



Πανεπιστήμιο Πειραιώς
Σχολή Τεχνολογιών Πληροφορικής και Επικοινωνιών
Εργαστήριο Επιστήμης των Δεδομένων - Data Science Lab.
<http://www.datastories.org>

Απολογισμός Δραστηριοτήτων
Εργαστηρίου Επιστήμης των Δεδομένων
(Data Science Lab.)

2016-2017



Πειραιάς, Ιούλιος 2017

ΠΡΟΛΟΓΟΣ

Το Εργαστήριο Επιστήμης των Δεδομένων (Data Science Lab. – DataStories) αποτελεί θεσμοθετημένο εργαστήριο της Σχολής Τεχνολογιών Πληροφορικής & Επικοινωνιών (ΤΠΕ) του Πανεπιστημίου Πειραιώς (ΦΕΚ 1819 τ.Β/21-6-2016). Κύριος σκοπός του εργαστηρίου είναι η προαγωγή της γνώσης και η ανάπτυξη της έρευνας στην Επιστήμη των Δεδομένων, όπως αυτά επιτυγχάνονται μέσω εκπαιδευτικών δραστηριοτήτων, συμμετοχών σε ερευνητικά προγράμματα, συνεργασιών κ.ά.

Στην παρούσα έκθεση καταγράφεται ένας συνοπτικός απολογισμός των δραστηριοτήτων του εργαστηρίου για το ακαδημαϊκό έτος 2016-17. Ειδικά επειδή η παρούσα έκθεση αποτελεί τον πρώτο επίσημο απολογισμό δραστηριοτήτων, για πληρότητα σε κάποιες ενότητες συμπεριλαμβάνονται δραστηριότητες πριν το 2016, καλύπτοντας το χρονικό διάστημα από την άτυπη σύσταση του εργαστηρίου ως ερευνητικής ομάδας τον 10.2014.

Περιεχόμενα:

1. Συνοπτική παρουσίαση σελ. 4
2. Ανθρώπινο δυναμικό / Υποδομές σελ. 5
3. Διδακτικό έργο / Διαλέξεις / Σεμινάρια σελ. 6
4. Επιστημονική – ερευνητική δραστηριότητα σελ. 8
5. Συμμετοχή σε χρηματοδοτούμενα ερευνητικά προγράμματασελ. 14
- Παράρτημα: Κατάλογος δημοσιεύσεωνσελ. 16

καθ. Γιάννης Θεοδωρίδης
Διευθυντής Εργαστηρίου Επιστήμης Δεδομένων

Πειραιάς, Ιούλιος 2017

1. ΣΥΝΟΠΤΙΚΗ ΠΑΡΟΥΣΙΑΣΗ

*Το **Εργαστήριο Επιστήμης των Δεδομένων** του Πανεπιστημίου Πειραιώς (Data Science Lab. – DataStories) δραστηριοποιείται ερευνητικά σε ένα πλήθος πεδίων της Επιστήμης Δεδομένων, όπως η διαχείριση μεγάλων δεδομένων, η αναλυτική δεδομένων, η μηχανική μάθηση, η εξόρυξη γνώσης από βάσεις δεδομένων, κειμένων και ήχων, η εξερεύνηση δεδομένων κίνησης, η σημασιολογική ολοκλήρωση και η ιδιωτικότητα δεδομένων. Το εργαστήριο έχει να επιδείξει σημαντική εμπειρία σε ερευνητικά έργα χρηματοδοτούμενα από Ευρωπαϊκούς και Εθνικούς πόρους, με πιο αξιοσημείωτα από τα έργα που βρίσκονται σε εξέλιξη, τα Ευρωπαϊκά H2020-ICT DATACRON (Big Data Analytics for Time Critical Mobility Forecasting, 2016-18) και H2020-SESAR DART (Data-driven Aircraft Trajectory Prediction Research, 2016-18). Στην τρέχουσα σύνθεσή του, το εργαστήριο αποτελείται από 9 μέλη ΔΕΠ (4 διαφορετικών Τμημάτων) του Πανεπιστημίου Πειραιώς καθώς και αρκετούς έμπειρους και νέους ερευνητές.*

2. ΑΝΘΡΩΠΙΝΟ ΔΥΝΑΜΙΚΟ / ΥΠΟΔΟΜΕΣ

Στις δραστηριότητες του Data Science Lab συμμετέχουν ή έχουν συμμετάσχει κατά την εν λόγω περίοδο οι:

- μέλη ΔΕΠ¹ (9): Αλέξανδρος Αρτίκης, Γιώργος Βούρος, Χρήστος Δουλκερίδης, Γιάννης Θεοδωρίδης, Μάρκος Κούτρας, Λευτέρης Κοφίδης, Νίκος Πελέκης, Άγγελος Πικράκης, Μαρία Χαλκίδη.
- μεταδιδάκτορες / έμπειροι ερευνητές (5): Ακριβή Βλάχου, Χάρης Γεωργίου, Σοφία Καραγιώργου, Κώστας Πατρούμπας, Γιώργος Σαντιπαντάκης.
- υποψήφιοι διδάκτορες / νέοι ερευνητές (11): Μάριος Βόντας, Απόστολος Γλένης, Φραγκίσκος Γρυλλάκης, Γιάννης Κοντούλης, Δέσποινα Κοπανάκη, Στράτος Μάνσαλης, Παναγιώτης Νικητόπουλος, Κυριάκος Πούτος, Στέλιος Σιδερίδης, Ευγενία Στουφή, Παναγιώτης Ταμπάκης.

καθώς και μεγάλος αριθμός μεταπτυχιακών και προπτυχιακών φοιτητών του Παν/μίου Πειραιώς.

Το Data Science Lab φιλοξενείται στην αίθ. 205, στο Κεντρικό Κτήριο του Παν/μίου Πειραιώς.

Στο χώρο του εργαστηρίου συντηρείται βιβλιοθήκη με ~150 βιβλία και 8 σειρές περιοδικών (Communications of the ACM, ACM Trans Database Systems, SIGMOD Record, SIGKDD Explorations, IEEE Computer, IEEE Trans Knowledge and Data Engineering, Data Mining & Knowledge Discovery, GeoInformatica) για τις ανάγκες των μελών του εργαστηρίου και των προπτυχιακών / μεταπτυχιακών φοιτητών.

¹ Σημείωση: βάσει του ΦΕΚ ίδρυσης του εργαστηρίου, πέρα από τα μέλη ΔΕΠ της Σχολής ΤΠΕ, μπορούν να συμμετέχουν στο εργαστήριο και μέλη ΔΕΠ εκτός Σχολής με την ιδιότητα του συνδεδεμένου μέλους.

3. ΔΙΔΑΚΤΙΚΟ ΕΡΓΟ / ΔΙΑΛΕΞΕΙΣ / ΣΕΜΙΝΑΡΙΑ

Το Data Science Lab συνεισφέρει διδακτικά σε πλήθος μαθημάτων, στα πλαίσια των ΠΠΣ των Τμημάτων Πληροφορικής και Ψηφιακών Συστημάτων της Σχολής ΤΠΕ καθώς και τεσσάρων (4) ΠΜΣ του Παν/μίου Πειραιώς. Συγκεκριμένα:

- Προπτυχιακά μαθήματα Τμήματος Πληροφορικής: Βάσεις Δεδομένων, Συστήματα Διαχείρισης Βάσεων Δεδομένων, Αποθήκες Δεδομένων και Εξόρυξη Γνώσης, Γεωγραφικά Πληροφοριακά Συστήματα.
- Προπτυχιακά μαθήματα Τμήματος Ψηφιακών Συστημάτων: Δομές Δεδομένων, Σχεδιασμός Βάσεων Δεδομένων, Βάσεις Δεδομένων, Αποθήκες και Εξόρυξη Δεδομένων, Ανάκτηση Πληροφοριών, Διαχείριση Δεδομένων από τον Παγκόσμιο Ιστό, Εργαστήριο Ανάπτυξης Εφαρμογών Βάσεων Δεδομένων, Εργαστήριο Συστημάτων Επεξεργασίας Πληροφοριών από τον Παγκόσμιο Ιστό, Αλγόριθμοι και Πολυπλοκότητα, Τεχνητή Νοημοσύνη, Συστήματα Ευφύων Πρακτόρων, Σηματολογικός Ιστός και Οντολογίες.
- Μεταπτυχιακά μαθήματα ΠΜΣ «Προηγμένα Συστήματα Πληροφορικής» Τμήματος Πληροφορικής: Διαχείριση Δεδομένων, Αναλυτική Δεδομένων, Εξόρυξη Γνώσης από Δεδομένα, Διαχείριση Γεωγραφικής Πληροφορίας.
- Μεταπτυχιακά μαθήματα ΠΜΣ «Πληροφορική» Τμήματος Πληροφορικής: Βάσεις Δεδομένων, Γεωπληροφορική, Ναυτιλιακή Πληροφορική.
- Μεταπτυχιακά μαθήματα ΠΜΣ «Ψηφιακά Συστήματα & Υπηρεσίες» Τμήματος Ψηφιακών Συστημάτων: Διαχείριση Δεδομένων, Μεγάλα Δεδομένα και Αναλυτική, Μεγάλα Δεδομένα και Αναλυτική: Τεχνικές και Εργαλεία I, Μεγάλα Δεδομένα και Αναλυτική: Τεχνικές και Εργαλεία II, Εφαρμογές Αναλυτικής Μεγάλων Δεδομένων, Εξόρυξη Δεδομένων και Προβλεπτική Αναλυτική.
- Μεταπτυχιακά μαθήματα ΠΜΣ «Εφαρμοσμένη Στατιστική» Τμήματος Στατιστικής & Ασφαλιστικής Επιστήμης: Στατιστικές Μέθοδοι Εξόρυξης Δεδομένων.

Επίσης, μέλη του εργαστηρίου προσφέρουν προσκεκλημένες διαλέξεις σε ΠΜΣ άλλων ιδρυμάτων, όπως του Ε.Μ. Πολυτεχνείου, του Χαροκόπειου Παν/μίου, του ΤΕΙ Πειραιά κ.ά.

Ήδη από την άτυπη σύστασή του (10.2014), το Data Science Lab οργανώνει σε περιοδική βάση Ερευνητικό Σεμινάριο με ομιλητές μέλη της ερευνητικής ομάδας αλλά και προσκεκλημένους ομιλητές, με στόχο τη διάχυση ερευνητικών αποτελεσμάτων προς τα μέλη της Πανεπιστημιακής κοινότητας. Συγκεκριμένα, η λίστα των ομιλιών μέχρι σήμερα έχει ως εξής [ημερομηνία διάλεξης (σε αντίστροφη χρονολογική σειρά, τίτλος, ομιλητής):

- 24/1/2017: “In silico analysis of miRNAs genes towards identification of pharmaceutical biomarkers”, Artemis HATZIGEORGIOU, Professor, Department of Electrical & Computer Engineering, University of Thessaly
- 14/12/2016: “Stream Processing in the Big Data Era: a blast from the past”, Yannis KOTIDIS, Assoc. Professor, Department of Informatics, Athens University of Economics and Business

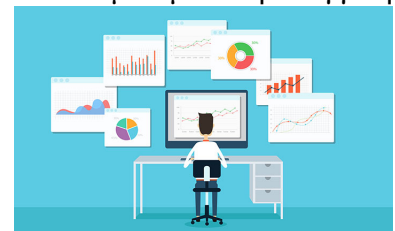
- 24/11/2016: “Advances in Feature Selection in Data Analytics”, Ioannis TSAMARDINOS, Assoc. Professor, Department of Computer Science, University of Crete
- 25/10/2016: “Graph Degeneracy for Graph Mining and applications”, Michalis VAZIRGIANNIS, Professor in LIX, École Polytechnique.
- 13/1/2016: “Introduction to linked geospatial data”, Manolis KOUBARAKIS, Professor in Department of Informatics and Telecommunications. University of Athens
- 22/6/2015: “Big Sequence Management”, Themis PALPANAS, Professor, Paris Descartes University of France
- 27/5/2015: “Recognising Complex Events in Noisy Data Streams”, Georgios PALIOURAS, Researcher at the IIT in the NCSR Demokritos in Athens, Greece.
- 1/4/2015: “A Tour of Sparsity-Aware Learning – Calling at: Online, Distributed, Robust And Dictionary Learning”, Sergios THEODORIDIS, Professor, University of Athens
- 10/3/2015: “Blocking Techniques for Web-Scale Entity Resolution”, George PAPADAKIS, Post-doctoral researcher, University of Athens
- 12/2/2015: “Data management challenges in the Medical Informatics Platform of the Human Brain Project”, Vasilis VASSALOS, Assoc. Professor, Dept. Informatics, AUEB
- 10/1/2015: “Big data infrastructures exploiting the power of big data”, Timos SELLIS, Professor, Computer Science & Info Tech, RMIT University
- 17/12/2014: “Visual Analytics in Practice”, Andreas MANIATIS, Commercial Manager, Cyberstream Technologies
- 17/12/2014: “The Emerging Landscape of Big Data Systems”, Damianos CHATZIANTONIOU, Assoc. Professor, Dept. Management Science & Technology, AUEB
- 5/11/2014: “Big Data and Cloud with Oracle”, Paul FLANNERY, Senior Director, Product Management Group - Servers, EMEA, Oracle

Υλικό σχετικό με τις ανωτέρω διαλέξεις (περίληψη, διαφάνειες κλπ.) υπάρχει διαθέσιμο στη σελίδα http://www.datastories.org/?page_id=3632.

4. ΕΠΙΣΤΗΜΟΝΙΚΗ – ΕΡΕΥΝΗΤΙΚΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑ

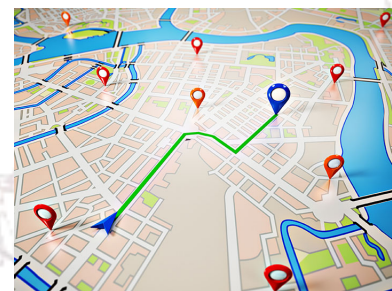
Η ερευνητική δραστηριότητα του Data Science Lab εστιάζει στις νέες προκλήσεις που φέρνουν τα 'Μεγάλα Δεδομένα' (Big Data) στο χώρο της διαχείρισης δεδομένων, της ανάλυσης πληροφορίας κλπ. Συνοπτικά:

- ο **Διαχείριση Μεγάλων Δεδομένων (Big Data Management)**. Τα Μεγάλα Δεδομένα αφορούν σε δεδομένα τα οποία είναι τεράστια σε όγκο, ετερογενούς φύσης, προέρχονται από διαφορετικές πηγές, παράγονται με υψηλούς ρυθμούς, και συνεπώς δεν μπορούμε να τα διαχειριστούμε με παραδοσιακά συστήματα διαχείρισης δεδομένων. Ως εκ τούτου πρέπει να επανασχεδιαστεί ολόκληρος ο κύκλος ζωής της διαχείρισης δεδομένων, ώστε να αντιμετωπιστούν οι ιδιαίτερες απαιτήσεις που προκύπτουν. Η ερευνητική μας ομάδα ασχολείται με νέες προσεγγίσεις επεξεργασίας χωρικών (Doulkeridis et al. 2017), χωρο-χρονικών (Nikitoroulos et al. 2016) και πολυδιάστατων δεδομένων (Saouk et al. 2016) μεγάλης κλίμακας, σε πλατφόρμες όπως το Hadoop και το Spark, όπου η έμφαση είναι σε θέματα απόδοσης και κλιμάκωσης με το μέγεθος των δεδομένων εισόδου.
- ο **Αναλυτική Κειμένου & Ήχου (Text & Audio Analytics)**. Ο πρωτοφανής όγκος δεδομένων κειμένου που παράγεται καθημερινά σε κοινωνικά δίκτυα και εν γένει στον Παγκόσμιο Ιστό, αλλά και με τη μορφή Ανοικτών Δεδομένων, απαιτεί νέες τεχνικές ανάλυσης κειμένων που είναι κλιμακούμενες και ανακαλύπτουν πολύτιμη γνώση που μπορεί να κρύβεται στα δεδομένα. Τα τελευταία χρόνια έχει δοθεί ιδιαίτερο ενδιαφέρον στην ανάλυση του συναισθήματος που μπορεί να κρύβεται σε ένα κείμενο, και αναδεικνύει εάν το κείμενο περιγράφει κάτι με θετικό, αρνητικό ή ουδέτερο τρόπο. Η ερευνητική μας ομάδα ασχολείται με θέματα ανάλυσης συναισθήματος σε μεγάλες συλλογές κειμένων από το Twitter με χρήση νευρωνικών δικτύων (Baziotis et al. 2017a) αλλά και με παραδοσιακούς ταξινομητές κειμένων (Karanasou et al. 2015) σε περιβάλλον ρευμάτων δεδομένων (Karanasou et al. 2016). Εστιάζοντας στις δύο τελευταίες εργασίες, στην πρώτη προτείνουμε μια προσέγγιση για ανάλυση συναισθήματος σε tweets με μεταφορικό και μη μεταφορικό λόγο. Η προτεινόμενη προσέγγιση βασίζεται στην εξαγωγή χαρακτηριστικών από τα tweets, χρησιμοποιώντας μορφολογικά, συντακτικά χαρακτηριστικά και σημασιολογικές πληροφορίες. Ως συνέχεια της πρώτης, στη δεύτερη εργασία παρουσιάζουμε ένα σύστημα για ανάλυση σε πραγματικό χρόνο του συναισθήματος των μηνυμάτων από Twitter. Για την ανάλυση του συναισθήματος υιοθετούμε μια προσέγγιση εποπτευόμενη μάθησης, όπου εκπαιδευούμε διάφορους κατηγοριοποιητές με βάση τα χαρακτηριστικά που έχουμε εξάγει από τα tweets. Επίσης υιοθετούμε μια αρχιτεκτονική βασισμένη στο Storm, προκειμένου να διαχειριστούμε μεγάλης κλίμακας και με γρήγορο ρυθμό άφιξης δεδομένα. Από την άλλη, στο χώρο του ήχου, η συνεχώς αυξανόμενη διαθεσιμότητα ροών ήχου μέσω



διαφόρων διαύλων διανομής, τόσο για εμπορικούς όσο και για μη κερδοσκοπικούς / ερευνητικούς σκοπούς, έχει αναδείξει την ανάγκη ύπαρξης συστημάτων που είναι ικανά να αναλύουν τους ήχους, προκειμένου να εξάγουν χρήσιμη γνώση, δηλαδή μεταδεδομένα σε διάφορα επίπεδα αφαίρεσης, τα οποία μπορούν να καταναλωθούν από τους χρήστες ή και να γίνουν αντικείμενο περαιτέρω επεξεργασίας από άλλα υπολογιστικά συστήματα. Ανάλογα με το εκάστοτε πρότυπο διαμόρφωσης για την αποθήκευση ή διανομή του ψηφιακού ήχου, τις κατηγορίες ήχων, τη συνύπαρξη του ήχου με άλλους τύπους μέσων (π.χ. κείμενο και κινούμενη εικόνα), τις απαιτήσεις των χρηστών, τον όγκο των δεδομένων, το πλαίσιο εφαρμογής και μια πληθώρα άλλων παραμέτρων, συνεχίζουν να αναπτύσσονται διάφορες εφαρμογές και ερευνητικές τάσεις, προκειμένου να αντιμετωπιστούν οι προκύπτουσες τεχνολογικές προκλήσεις στο χώρο της αναγνώρισης ομιλίας, της ταυτοποίησης / επιβεβαίωσης ταυτότητας ομιλητή, της ανάκτησης μουσικής πληροφορίας, του εντοπισμού ηχητικών συμβάντων, της ταξινόμησης ηχητικής σκηνής, της αναγνώρισης συναισθήματος και της πολυτροπικής ανάλυσης οπτικοακουστικού υλικού. Η σπουδαιότητα των εξαγόμενων μεταδεδομένων έγκειται στο ότι επιτρέπουν την περαιτέρω ανάπτυξη μεθόδων δεικτοδότησης και γρήγορης πλοήγησης, οι οποίες βρίσκονται στον πυρήνα των πολυμεσικών συστημάτων νέας γενιάς. Στο εργαστήριο διεξάγουμε έρευνα αιχμής σε μεθόδους ανάλυσης ήχων, με έμφαση σε νευρωνικά δίκτυα μεγάλου βάθους, τεχνικές εκμάθησης λεξικού και κρυφά Markov μοντέλα.

- ο **Αναλυτική Δεδομένων Κίνησης (Mobility Data Analytics)**. Η ομάδα μας σταθερά τα τελευταία χρόνια ασχολείται με τη μελέτη τεχνικών ανάλυσης και εξόρυξης γνώσης από βάσεις δεδομένων κινούμενων αντικειμένων (moving object databases), τα οποία αντικείμενα μπορεί να είναι πεζοί, οχήματα κ.ά. τα οποία κινούνται στο 2-διάστατο (x-, y-) ή 3-διάστατο χώρο (x-, y-, z-). Ενδεικτικά, έχουμε αναπτύξει τεχνικές συσταδοποίησης τροχιών (Pelekis et al. 2017a; Pelekis et al. 2017b) και, ευρύτερα, το σύστημα Hermes (Pelekis et al. 2015a) όπου οι αναλυτικές εργασίες υλοποιούνται ως ερωτήματα σε ένα κατάλληλα σχεδιασμένο σύστημα βάσεων δεδομένων. Επίσης, πρόσφατα έχει αναδυθεί η ερευνητική κατεύθυνση της διαχείρισης και ανάλυσης σημασιολογικών τροχιών, δηλαδή τροχιών κινούμενων αντικειμένων με επισημειώσεις για το πώς και γιατί μιας κίνησης («από το σπίτι στο γραφείο» κλπ.). Σε αυτό το πλαίσιο, έχουμε αναπτύξει κατάλληλες τεχνικές επεξεργασίας ερωτημάτων και ανάλυσης της σχετικής πληροφορίας (Sideridis et al. 2016) καθώς και το λογισμικό Hermapolis, μια γεννήτρια παραγωγής ρεαλιστικών δεδομένων που ικανοποιούν κάποια μοτίβα κίνησης (Pelekis et al. 2016).
- ο **Αναγνώριση Συμβάντων από Ρεύματα Δεδομένων (Stream Reasoning)**. Η αναγνώριση συμβάντων αναφέρεται στην εξόρυξη κρίσιμης πληροφορίας από (γεωγραφικά) κατανεμημένα ρεύματα δεδομένων. Η ερευνητική μας ομάδα έχει αναπτύξει τεχνικές αναγνώρισης συμβάντων, όπου χωρο-χρονικά πρότυπα αντιστοιχίζονται σε ρεύματα σε πραγματικό χρόνο (Alevizos et al. 2017; Patroumpas et al. 2017; Alevizos et al. 2015; Artikis et al. 2015; Patroumpas et al. 2015). Επιπλέον, έχουμε αναπτύξει τεχνικές μηχανικής μάθησης για την αυτοματοποιημένη κατασκευή



και βελτίωση χωρο-χρονικών προτύπων (Katzouris et al. 2017; Micheloudakis et al. 2016).

- **Ανάλυση Κοινωνικών Δικτύων (Social Network Analysis).** Στην εργασία (Vahabi et al. 2015) παρουσιάζουμε ένα νέο σύστημα συστάσεων που στοχεύει στην αύξηση της συνολικής συμμετοχής των χρηστών σε διαδικτυακές κοινωνικές πλατφόρμες, λαμβάνοντας υπόψη τη διαδικασία διάχυσης πληροφορίας σε ένα κοινωνικό δίκτυο. Έχουμε σχεδιάσει ένα νέο αλγόριθμο συστάσεων, DifRec, διατυπώνοντας το πρόβλημα της μεγιστοποίησης της συνολικής συμμετοχής του χρήστη ως πρόβλημα αναθέσεων σε ένα κατάλληλα ορισμένο γράφο γειτνίασης. Ο γράφος αυτός μοντελοποιεί τις συγκρούσεις που προκύπτουν από το να συστήσει κανείς σε ένα χρήστη αντικείμενα που θα τα λάβει ούτως ή άλλως μέσω της κοινωνικής διάχυσης.



- **Σημασιολογική Ολοκλήρωση / Ενοποίηση (Semantic Integration).** Η σημασιολογική ολοκλήρωση δεδομένων από ετερογενείς πηγές δεδομένων αναφέρεται στη σύνδεση των δεδομένων με τρόπο που να διατηρείται το νόημα αυτών. Για την επίτευξη του στόχου αυτού απαιτούνται ρητά μοντέλα των δεδομένων (οντολογίες), σύνδεση των μοντέλων αυτών και, τέλος, σύνδεση των οντοτήτων που αναπαρίστανται. Το πρόβλημα αυτό, ενώ είναι από μόνο του ιδιαίτερα δύσκολο στη γενικότητά του, αποκτά νέες διαστάσεις όταν συνδέεται με big data χαρακτηριστικά.



- **Ανάλυση Μεγάλων Δεδομένων με Μοντέλα Χαμηλής Τάξης (Low-rank Models for Big Data Analysis).** Συχνά οι (γνωστοί ή κρυφοί (latent)) μηχανισμοί παραγωγής των δεδομένων επιδέχονται περιγραφής με μοντέλα χαμηλής (πολύ χαμηλότερης από τις διαστάσεις τους) τάξης, κάτι που επιτρέπει την αποδοτική αποθήκευση, επεξεργασία, κι ανάλυση όγκου δεδομένων κατά τα άλλα μη διαχειρίσιμων. Έμφαση δίνεται σε μοντέλα βασισμένα σε πολύ-διάστατους πίνακες (tensors) και σε τεχνικές υπολογισμού τους για ακριβέστερη αναπαράσταση (κι ερμηνεία) πολλαπλών σχέσεων «θαμμένων» στα δεδομένα. Παραδείγματα εφαρμογών: (α) τηλεπικοινωνιακά συστήματα πολλαπλών κεραιών εκπομπής-λήψης, (β) ανάλυση μετρήσεων λειτουργικής μαγνητικής τομογραφίας (fMRI).



- **Ιδιωτικότητα Δεδομένων (Data Privacy).** Η αυξανόμενη διαθεσιμότητα πληροφορίας λόγω της έκρηξης της σχετικής τεχνολογίας έχει οδηγήσει στην ανάπτυξη αποτελεσματικών τεχνικών διαχείρισης και ανάλυσης δεδομένων, κάτι που όμως μπορεί να οδηγήσει σε σημαντικούς κινδύνους όσον αφορά την προστασία των προσωπικών δεδομένων και της ιδιωτικότητας των ατόμων. Στη εργασία μας (Koutsopoulos et al. 2015) σχεδιάζουμε μια αγορά ιδιωτικών δεδομένων

προσανατολισμένη στη διαδικασία εκπαίδευσης ενός κατηγοριοποιητή. Κάθε κάτοχος δεδομένων έχει ιδιωτικό κόστος που αναφέρεται στη δυσκολία του για παροχή δεδομένων στον εκπαιδευόμενο κατηγοριοποιητή. Επίσης, σε κάθε στάδιο της μάθησης, κάθε ιδιοκτήτης δεδομένων χαρακτηρίζεται από ένα σκορ χρησιμότητας, το οποίο εκφράζει τη χρησιμότητα των δεδομένων του για τη διαδικασία της εκπαίδευσης. Ο εκπαιδευόμενος ξεκινά ζητώντας δεδομένα και οι ενδιαφερόμενοι ιδιοκτήτες δεδομένων αποκρίνονται δηλώνοντας το κόστος τους. Ο εκπαιδευόμενος υπολογίζει το σκορ χρησιμότητας από κάθε κάτοχο δεδομένων. Στη συνέχεια, επιλέγει έναν κάτοχο για να αγοράσει δεδομένα και κάνει τη σχετική πληρωμή. Σε μια διαφορετική κατεύθυνση, αυτή της ιδιωτικότητας δεδομένων θέσεων και κίνησης ανθρώπων (τροχιάς σε 3-διάστατο χώρο), έχουμε αναπτύξει τα συστήματα Hermes++ και Private-Hermes (Koranaki et al. 2017), τα οποία διατηρούν τα δεδομένα εσωτερικά στον οργανισμό φιλοξενίας και παρέχουν δυνατότητα ερωταποκρίσεων όντας υπεύθυνα για την ασφαλή παραχώρηση πληροφορίας προς τον έξω κόσμο, καθώς και προηγμένες τεχνικές ανωνυμοποίησης όπου ο κάθε χρήστης εξατομικευμένα δηλώνει το επίπεδο ανωνυμοποίησης – άρα γενίκευσης της τροχιάς του – που επιθυμεί (Koranaki et al. 2016) καθώς και ανωνυμοποίησης της τροχιάς με αξιοποίηση σημασιολογικής πληροφορίας (Koranaki & Pelekis 2017).

Τα προϊόντα της έρευνας που πραγματοποιείται μπορούν να συνοψιστούν στη διαρκή ανάπτυξη εργαλείων λογισμικού και τη συντήρηση portals γενικότερου επιστημονικού ενδιαφέροντος που περιλαμβάνουν σύνολα δεδομένων και αλγορίθμους διαθέσιμους προς γενική χρήση. Συγκεκριμένα:

- **Hermes MOD engine:** μηχανή διαχείρισης βάσεων δεδομένων κινούμενων αντικειμένων (σε 2 εκδοχές: πάνω σε Oracle ή PostgreSQL). URL: <http://infolab.cs.unipi.gr/hermes/>.
- **Hermoupolis:** γεννήτρια συνθετικών δεδομένων τροχιών κινούμενων αντικειμένων πάνω σε οδικό δίκτυο. URL: <http://infolab.cs.unipi.gr/hermoupolis/>.
- **ChoroChronos.org:** συλλογή πραγματικών συνόλων δεδομένων τροχιών κινούμενων αντικειμένων και αλγορίθμων επεξεργασίας τους. URL: <http://chorochronos.datastories.org>.
- **UniPi-AIS:** σταθμός βάσης εγκατεστημένος στο Παν/μιο Πειραιώς για τη συλλογή και αποθήκευση μηνυμάτων τύπου AIS που εκπέμπουν τα πλοία που βρίσκονται εντός της εμβέλειας της κεραίας του σταθμού. Ο σταθμός είναι συνδεδεμένος με το παγκοσμίου κάλυψης δίκτυο MarineTraffic. URL: <http://infolab.cs.unipi.gr/UniPi-AIS/> (απεικόνιση πληροφορίας AIS σε πραγματικό χρόνο), <https://www.marinetraffic.com/en/ais/details/stations/2782> (ο σταθμός στο δίκτυο MarineTraffic).
- **UniPi-ADSB:** σταθμός βάσης εγκατεστημένος στο Παν/μιο Πειραιώς για τη συλλογή και αποθήκευση μηνυμάτων τύπου ADS-B που εκπέμπουν τα αεροσκάφη που βρίσκονται εντός της εμβέλειας της κεραίας του σταθμού. URL: [υπό κατασκευή].

Η ερευνητική δραστηριότητα του εργαστηρίου έχει οδηγήσει σε πλήθος δημοσιεύσεων σε διεθνή περιοδικά, συλλογικούς τόμους και πρακτικά διεθνών συνεδρίων με κρίση στο πλήρες κείμενο. Συνοπτικά, η παραγωγή δημοσιεύσεων της τελευταίας 3-ετίας για τα μέλη του εργαστηρίου έχει ως εξής:

	2017	2016	2015
Άρθρα σε επιστημονικά περιοδικά με κριτές	4	3	11
Εργασίες σε πρακτικά διεθνών συνεδρίων (κρίση στο πλήρες κείμενο)	10	11	16
Κεφάλαια σε συλλογικούς τόμους	1	-	1

Αναλυτικά, ο πλήρης κατάλογος δημοσιεύσεων για την εν λόγω περίοδο καταγράφεται στο Παράρτημα Ι.

Επίσης, το εργαστήριο έχει συμμετάσχει στη διοργάνωση επιστημονικών συνεδρίων, παρουσίαση tutorials σε επιστημονικά συνέδρια, επιμέλεια ειδικών τευχών περιοδικών κλπ. Συγκεκριμένα:

2017

- 1st Int'l VLDB Workshop on Mobility Analytics for Spatio-temporal and Social Data (MATES'17), Munich, Germany, Sep. 1, 2017. General co-chair: Christos Doulkeridis; PC co-chair: George Vouros (<http://ai-group.ds.unipi.gr/mates17>).
- 1st Int'l EDBT Workshop on Big Geo Data Quality and Privacy (BigQP'17), Venice, Italy, Mar. 21, 2017. Workshop co-chair: Yannis Theodoridis (<http://www-etis.ensea.fr/BigGeoQ-UP/BIGQP2017>).
- Giatrakos N, Artikis A, Deligiannakis A, Garofalakis M (2017) Complex event recognition in the big data era. Tutorial at VLDB 2017.
- Artikis A, Margara A, Ugarte M, Vansummeren S, Weidlich M (2017) Complex event recognition languages. Tutorial at DEBS'2017.

2016

- Artikis A (2016) Complex event recognition. Tutorial at ECAI'2016.

2015

- 13th European Conference on Multi-Agent Systems (EUMAS) & 3rd Int'l Conference on Agreement Technologies (AT), Athens, Greece, Dec. 17-18, 2015. Conference co-chair: George Vouros (<http://ai-group.ds.unipi.gr/eumas-at2015/eumas2015>).
- Artikis A, Weidlich M (2015) Complex event recognition in multi-agent systems. Tutorial at AAMAS'2015.

Επίσης, στην εν λόγω περίοδο έχουν προκύψει οι εξής διακρίσεις σε διεθνές επίπεδο:

2017

- (Baziotis et al. 2017a) 1st place in Int'l Workshop on Semantic Evaluation (SemEval-2017), Task 4 on “Sentiment Analysis in Twitter” (<http://alt.qcri.org/semeval2017/task4/>).
- (Baziotis et al. 2017b) 2nd place in Int'l Workshop on Semantic Evaluation (SemEval-2017), Task 6 on “#HashtagWars: Learning a Sense of Humor” (<http://alt.qcri.org/semeval2017/task6/>).

2016

- (Nikitopoulos et al. 2016) 3rd place in ACM SIGSPATIAL Cup 2016 (<http://sigspatial2016.sigspatial.org/giscup2016/>).

Τέλος, μέλη του εργαστηρίου συμμετέχουν σε επιστημονικές επιτροπές, συμβούλια κ.α. Ενδεικτικά: Γενική Συνέλευση Ελληνικού Ιδρύματος Έρευνας και Καινοτομίας (ΕΛΙΔΕΚ), Επιστημονικό Συμβούλιο Ινστιτούτου Πληροφοριακών Συστημάτων Ε.Κ. «Αθηνά» (ΙΠΣΥ/Αθηνά), Editorial Board - ACM Computing Surveys, Editorial board – IEEE Transactions on Signal Processing, Editorial Board - Int'l Journal of Data Warehousing and Mining (IJDWM), Endowment - Int'l Symposium on Spatial and Temporal Databases (SSTD).

Περισσότερες πληροφορίες για την ερευνητική δραστηριότητα του εργαστηρίου και τα εργαλεία που έχουν αναπτυχθεί είναι διαθέσιμες στις σελίδες http://www.datastories.org/?page_id=3333 και http://www.datastories.org/?page_id=3372, αντίστοιχα.

5. ΣΥΜΜΕΤΟΧΗ ΣΕ ΧΡΗΜΑΤΟΔΟΤΟΥΜΕΝΑ ΕΡΕΥΝΗΤΙΚΑ ΠΡΟΓΡΑΜΜΑΤΑ

Το Data Science Lab συμμετέχει σε ικανό αριθμό ανταγωνιστικών ερευνητικών προγραμμάτων, με Εθνική ή/και Κοινοτική χρηματοδότηση. Παρατίθεται στη συνέχεια αναλυτικός κατάλογος των έργων που ολοκληρώθηκαν πρόσφατα ή βρίσκονται σε εξέλιξη:

- (2016.07-2017.06) “DATASIM – Data Science for Simulating the Era of Electric Vehicles / Εθνική Συμμετοχή”. Ερευνητικό έργο χρηματοδοτούμενο από τη ΓΓΕΤ. Συντονιστής ομάδας: Γιάννης Θεοδωρίδης. Προϋπολογισμός ομάδας: 20.188 €.
- (2016.07-2017.06) “SEEK – Semantic Enrichment of Trajectory Knowledge Discovery / Εθνική Συμμετοχή”. Ερευνητικό έργο χρηματοδοτούμενο από τη ΓΓΕΤ. Συντονιστής ομάδας: Γιάννης Θεοδωρίδης. Προϋπολογισμός ομάδας: 5.894 €.
- (2016.06-2018.06) “DART – Data-driven Aircraft Trajectory Prediction Research”. Ερευνητικό έργο χρηματοδοτούμενο από το Ευρωπαϊκό Πρόγραμμα Horizon 2020 / SESAR. Συντονιστής ομάδας: Γιώργος Βούρος. Προϋπολογισμός ομάδας: 170.000 €. Περισσότερες πληροφορίες στο [dart-project.eu].
- (2016.01-2018.12) “DATACRON – Big Data Analytics for Time Critical Mobility Forecasting”. Ερευνητικό έργο χρηματοδοτούμενο από το Ευρωπαϊκό Πρόγραμμα Horizon 2020 / ICT. Συντονιστής ομάδας: Γιώργος Βούρος. Προϋπολογισμός ομάδας: 730.625 €. Περισσότερες πληροφορίες στο [datacron-project.eu].
- (2014.02-2015.10) “ROADRUNNER: Scalable and Efficient Big Data Analytics”. Ερευνητικό έργο χρηματοδοτούμενο από τη ΓΓΕΤ, Πρόγραμμα «ΑΡΙΣΤΕΙΑ II». Συντονιστής ομάδας: Χρήστος Δουλκερίδης. Προϋπολογισμός ομάδας: 209.911 €. Περισσότερες πληροφορίες στο [platforms.gr/roadrunner/].

Από τα παραπάνω στοιχεία, αφού αφαιρεθεί το ποσοστό παρακράτησης του ΕΛΚΕ του Παν/μίου Πειραιά, προκύπτει ένας μέσος προϋπολογισμός για το εργαστήριο της τάξης των 330.000 € ανά έτος (για αμοιβές, ταξίδια, εξοπλισμό κλπ.). Κύρια πηγή χρηματοδότησης (90% του συνόλου) αποτελεί το Ευρωπαϊκό Πρόγραμμα Horizon 2020, ακολουθούμενη (10% του συνόλου) από τα προγράμματα της ΓΓΕΤ.

Μέσω των ερευνητικών κοινοπραξιών στις οποίες συμμετέχει το εργαστήριο έχει αναπτυχθεί σταθερή συνεργασία σε ερευνητικό επίπεδο με Πανεπιστήμια και Ερευνητικά Κέντρα της Ελλάδας και του εξωτερικού. Ενδεικτικά:

- Boeing Research & Technology Europe, Madrid, Spain
- Fraunhofer Institute IAIS, Bonn, Germany
- Naval Academy Research Institute, Brest, France
- Norwegian University of Science and Technology NTNU, Trondheim, Norway
- ΕΚΕΦΕ «Δημόκριτος», Ελλάδα

Στο πλαίσιο των ανωτέρω συνεργασιών, έχουν φιλοξενηθεί στο εργαστήριο αρκετά μέλη ΔΕΠ και νέοι ερευνητές ιδρυμάτων του εξωτερικού, σε χρονικά διαστήματα μερικών εβδομάδων έως μερικών μηνών.



6. ΠΑΡΑΡΤΗΜΑ: ΚΑΤΑΛΟΓΟΣ ΔΗΜΟΣΙΕΥΣΕΩΝ

ΕΡΓΑΣΙΕΣ ΣΕ ΕΠΙΣΤΗΜΟΝΙΚΑ ΠΕΡΙΟΔΙΚΑ ΜΕ ΚΡΙΤΕΣ

2017

- Georgoulas K, Vlachou A, Doulkeridis C, Kotidis Y (2017) User-centric similarity search. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 29(1): 200-213.
- Gkorgkas O, Vlachou A, Doulkeridis C, Nørnvåg K (2017) Exploratory product search using top-k join queries. *Information Systems*, 64: 75-92.
- Patroumpas K, Alevizos E, Artikis A, Vodas M, Pelekis N, Theodoridis Y (2017) Online event recognition from moving vessel trajectories. *Geoinformatica*, 21(2): 389-427.
- Pelekis N, Tampakis P, Vodas M, Doulkeridis C, Theodoridis Y (2017b) On temporal-constrained sub-trajectory cluster analysis. *Data Mining and Knowledge Discovery*, online first.

2016

- Katzouris N, Artikis A, Paliouras G (2016) Online learning of event definitions. *Theory and Practice of Logic Programming*, 16(5-6).
- Pelekis N, Sideridis S, Tampakis P, Theodoridis Y (2016) Simulating our LifeSteps by example. *ACM Transactions on Spatial Algorithms and Systems*, 2(3), article no. 11.
- Sideridis S, Pelekis N, Theodoridis Y (2016) On querying and mining semantic-aware mobility timelines. *Journal of Data Science & Analytics*, 2(1): 29-44.

2015

- Artikis A, Sergot M, Paliouras G (2015) An event calculus for event recognition. *IEEE Transactions on Knowledge and Data Engineering*, 27(4): 895-908.
- Fileto R, May C, Renso C, Pelekis N, Klein D, Theodoridis Y (2015) The *Baquara*² knowledge-based framework for semantic enrichment and analysis of movement data. *Data and Knowledge Engineering*, 98: 104-122.
- Katzouris N, Artikis A, Paliouras G (2015) Incremental learning of event definitions with inductive logic programming. *Machine Learning*, 100(2-3): 555-585.
- Konstantinidis, A., Chatzimilioudis, G., Zeinalipour-Yazti, D., Mpeis, P., Pelekis, N., Theodoridis, Y. (2015): Privacy-preserving indoor localization on smartphones. *IEEE Transactions on Knowledge and Data Engineering*, 27(11): 3042-3055.
- Pelekis, N., Frentzos, E., Giatrakos, N., Theodoridis, Y. (2015a): HERMES: a trajectory DB engine for mobility-centric applications. *Int'l Journal of Knowledge-based Organizations*, 5(3).
- Pertesis D, Doulkeridis C (2015) Efficient skyline query processing in SpatialHadoop. *Information Systems*, 54: 325-335.
- Pitropakis N, Pikrakis A, Lambrinouidakis C (2015) Behaviour reflects personality: detecting co-residence attacks on Xen-based cloud environments. *International Journal of Information Security*, 14(4): 299-305.

- Santipantakis GM, Vouros GA (2015a) Decomposing ontologies for the construction of distributed knowledge bases: The mONTul method. *International Journal on Artificial Intelligence Tools*, 24(5).
- Santipantakis GM, Vouros GA (2015b) Distributed reasoning with coupled ontologies: the E-SHIQ representation framework. *Knowledge and Information Systems*, 45(2): 491-534.
- Skarlatidis A, Paliouras G, Artikis A, Vouros G (2015) Probabilistic event calculus for event recognition. *ACM Transactions on Computational Logic*, 16(2): 11:1-11:37.
- Vouros GA (2015) Decentralized semantic coordination of interconnected entities via belief propagation. *AI Communications*, 28(4): 617-634.

ΔΗΜΟΣΙΕΥΣΕΙΣ ΣΕ ΠΡΑΚΤΙΚΑ ΔΙΕΘΝΩΝ ΣΥΝΕΔΡΙΩΝ (ΚΡΙΣΗ ΣΤΟ ΠΛΗΡΕΣ ΚΕΙΜΕΝΟ)

2017

- Alevizos E, Artikis A, Paliouras G (2017) Event forecasting with pattern Markov chains. *Proceedings of DEBS*, Barcelona, Spain.
- Artikis A, Katzouris N, Correia I, Baber C, Morar N, Skarbovsky I, Fournier F, Paliouras G (2017) A prototype for credit card fraud management. *Proceedings of DEBS*, Barcelona, Spain.
- Baziotis C, Pelekis N, Doulkeridis C (2017a) DataStories at SemEval-2017 Task 4: Bidirectional LSTM with Attention for Message-level and Topic-based Sentiment Analysis. *Proceedings of SemEval*, Vancouver, Canada.
- Baziotis C, Pelekis N, Doulkeridis C (2017b) DataStories at SemEval-2017 Task 6: Siamese Bidirectional LSTM with Attention for Humorous Text Comparison. *Proceedings of SemEval*, Vancouver, Canada.
- Claramunt C, Ray C, Salmon L, Camossi E, Hadzagic M, Joussemme AL, Andrienko G, Andrienko N, Theodoridis Y, Vouros GA (2017) Maritime data integration and analysis: recent progress and research challenges [vision paper]. *Proceedings of EDBT*, Venice, Italy.
- Doulkeridis C, Pelekis N, Theodoridis Y, Vouros GA (2017a) Big data management and analytics for mobility forecasting in datAcron. *Proceedings of EDBT/ICDT Workshops*, Venice, Italy.
- Doulkeridis C, Vlachou A, Mpeostas D, Mamoulis N (2017b) Parallel and distributed processing of spatial preference queries using keywords. *Proceedings of EDBT*, Venice, Italy.
- Kopanaki D, Pelekis N (2017) Privacy preservation of semantic trajectory databases using query auditing techniques. *Proceedings of EDBT/ICDT Workshops*, Venice, Italy.
- Pelekis N, Tampakis P, Vodas M, Panagiotakis C, Theodoridis Y (2017a) In-DBMS sampling-based sub-trajectory clustering. *Proceedings of EDBT*, Venice, Italy.
- Vouros GA (2017) Learning conventions via social reinforcement learning in complex and open settings. *Proceedings of AAMAS*, São Paulo, Brazil.

2016

- Karamanolakis G, Iosif E, Zlatintsi A, Pikrakis A, Potamianos A (2016) Audio-based Distributional Representations of Meaning Using a Fusion of Feature Encodings. *Proceedings of Interspeech*, San Francisco – CA, USA.
 - Karanasou M, Ampla A, Doulkeridis C, Halkidi M (2016) Scalable and real-time sentiment analysis on Twitter data. *Proceedings of SENTIRE*, Barcelona, Spain.
 - Konstantinidis A, Chatzimilioudis G, Zeinalipour-Yazti D, Mpeis P, Pelekis N, Theodoridis Y (2016) Privacy-preserving indoor localization on smartphones [extended abstract]. *Proceedings of ICDE*, Helsinki, Finland.
 - Kopanaki D, Theodossopoulos V, Pelekis N, Kopanakis I, Theodoridis Y (2016) Who cares about others' privacy: personalized anonymization of moving object trajectories. *Proceedings of EDBT*, Bordeaux, France.
 - Michelioudakis E, Artikis A, Paliouras G (2016) Online structure learning for traffic management. *Proceedings of ILP*, London, UK.
 - Micheloudakis V, Skarlatidis A, Paliouras G, Artikis A (2016) OSLa: Online structure learning using background knowledge axiomatization, *Proceedings of ECML*, Riva del Garda, Italy.
 - Nikitopoulos P, Paraskevopoulos AI, Doulkeridis C, Pelekis N, Theodoridis Y (2016) BigCAB: distributed hot spot analysis over big spatio-temporal data using Apache Spark [GIS Cup]. *Proceedings of ACM SIGSPATIAL GIS*, San Francisco – CA, USA.
 - Pikrakis A, Kopsinis Y, Kroher N, Diaz-Banez JM (2016) Unsupervised singing voice detection using dictionary learning. *Proceedings of EUSIPCO*, Budapest, Hungary.
 - Saouk M, Doulkeridis C, Vlachou A, Nørnvåg K (2016) Efficient Processing of Top-k Joins in MapReduce. *Proceedings of IEEE Big Data*, Washington, USA.
 - Sfyris GA, Fragkos N, Doulkeridis C (2016) Profile-based selection of expert groups. *Proceedings of TPD*, Hannover, Germany.
 - Vlassopoulos C, Kontopoulos I, Apostolou M, Artikis A, Vogiatzis D (2016) Dynamic graph management for streaming social media analytics. *Proceedings of DEBS*, Irvine – CA, USA.
- 2015
- Alevizos E, Artikis A, Patroumpas K, Vodas M, Theodoridis Y, Pelekis N (2015) How not to drown in a sea of information: An event recognition approach [short paper]. *Proceedings of IEEE Big Data*, Santa Clara – CA, USA.
 - Costa CF, Nascimento M, Macedo J, Theodoridis Y, Pelekis N, Machado J (2015) Optimal time-dependent sequenced route queries in road networks. *Proceedings of ACM SIGSPATIAL GIS*, Seattle - Washington, USA.
 - Gkorgkas O, Vlachou A, Doulkeridis C, Nørnvåg K (2015a) Finding the most diverse products using preference queries. *Proceedings of EDBT*, Brussels, Belgium.
 - Gkorgkas O, Vlachou A, Doulkeridis C, Nørnvåg K (2015b) Maximizing influence of spatio-textual objects based on keyword selection. *Proceedings of SSTD*, Seoul, South Korea.
 - Karanasou M, Doulkeridis C, Halkidi M (2015) DsUniPi: an SVM-based approach for sentiment analysis of figurative language on Twitter. *Proceedings of SemEval*, Denver – CO, USA.

- Koutsopoulos I, Gionis A, Halkidi M (2015) Auctioning data for learning. *Proceedings of MoDAT (in conjunction with IEEE ICDM)*, Atlantic City – NJ, USA.
- Kroher N, Pikrakis A, Moreno J, Diaz-Banez JM (2015) Discovery of repeated vocal patterns in polyphonic audio: A case study on flamenco music. *Proceedings of EUSIPCO*, Nice, France.
- Patroumpas K, Artikis A, Katzouris N, Vodas M, Theodoridis Y, Pelekis N (2015) Event recognition for maritime surveillance. *Proceedings of EDBT*.
- Pelekis N, Sideridis S, Theodoridis Y (2015b) Hermes^{sem}: a semantic-aware framework for the management and analysis of our lifesteps. *Proceedings of IEEE DSAA*, Paris, France.
- Pikrakis A, Kopsinis Y, Chouvardas S, Theodoridis S (2015) Pattern classification formulated as a missing data task: the audio genre classification case. *Proceedings of IEEE ICASSP*, Brisbane, Queensland, Australia.
- dos Santos AA, Furtado AS, Alvares LO, Pelekis N, Bogorny V (2015) Inferring relationships from trajectory data. *Proceedings of GeoInfo*, Campos do Jordão, São Paulo, Brazil.
- Vahabi H, Koutsopoulos I, Gullo F, Halkidi M (2015) Optimal user engagement through social-diffusion-aware recommender systems, *Proceedings of CIKM*, Melbourne, Australia.
- Vouros GA (2015) The emergence of norms via contextual agreements in open societies. *Proceedings of CARE/MFSC@AAMAS*, Istanbul, Turkey.
- Santipantakis GM, Kotis KI, Vouros GA (2015a) Accessing and reasoning with data from disparate data sources using modular ontologies and OBDA. *Proceedings of SEMANTICS*, Vienna, Austria.
- Santipantakis GM, Kotis KI, Vouros GA (2015b) Ontology-based data sources' integration for maritime event recognition. *Proceedings of IISA*, Corfu, Greece.
- Santipantakis GM, Kotis KI, Vouros GA (2015c) Ontology-based data integration for event recognition in the maritime domain. *Proceedings of WIMS*, Larnaca, Cyprus.

ΚΕΦΑΛΑΙΑ ΣΕ ΣΥΛΛΟΓΙΚΟΥΣ ΤΟΜΟΥΣ, ΜΕΤΑ ΑΠΟ ΚΡΙΣΗ

2017

- Kopanaki D, Pelekis N, Theodoridis Y (2017) Systems for privacy-preserving mobility data management. In Gkoulalas-Divanis A, Bettini C (eds.) *First Handbook on Mobile Data Privacy*, Springer.

2016

- ---

2015

- Halkidi M, Vazirgiannis M, Hennig CM (2015) Method-independent indices for cluster validation and estimating the number of clusters. In Hennig C, Meila M, Murtagh F, Rocci R (eds.) *Handbook of Cluster Analysis*. Chapman and Hall/CRC.

Data Science Lab.
@
University of Piraeus

80 Karaoli & Dimitriou Str.
GR-18534 Piraeus, Greece

www.datastories.org



<https://www.facebook.com/pages/DataStories/845396452149334>



<https://www.linkedin.com/groups?home=&gid=8207219>



<https://twitter.com/UnipiDataSciLab>



<https://www.youtube.com/channel/UCVI4IqB91GJhccwHzdP-YjA>