



Σεμινάριο:
“Business Analytics & Big Data, Μία Πρακτική Εισαγωγή”
Οικονομικό Πανεπιστήμιο Αθηνών
Σειρά 3^η, Μάρτιος-Απρίλιος-Μάιος 2016
Διάρκεια: 60 ώρες (10 εβδομάδες)

Η χρησιμοποίηση δεδομένων στη λήψη σωστών, έγκυρων και έγκαιρων αποφάσεων έχει αναχθεί σε «εκ των ουκ άνευ» παράγοντα επιτυχίας για τις περισσότερες σύγχρονες επιχειρήσεις και οργανισμούς. Ταυτόχρονα, τα τελευταία χρόνια, με την ανάπτυξη νέων τεχνολογιών και εφαρμογών, όπως η εξάπλωση των κοινωνικών δικτύων, η εκτεταμένη χρήση smart phones, η εξάπλωση των RFID συστημάτων, κ.α. ο όγκος και η μορφή των δεδομένων έχει αλλάξει δραματικά. Σαν αποτέλεσμα, έχουν αναπτυχθεί νέες μέθοδοι διαχείρισης και αξιοποίησης δεδομένων. Οι όροι Big Data, Business Analytics και Data Science βρίσκονται πλέον στο επίκεντρο των δραστηριοτήτων των IT τμημάτων μικρών και μεγάλων οργανισμών.

Τα τελευταία τρία χρόνια έχει αναδειχθεί ένας νέος ρόλος στις εταιρίες και τους οργανισμούς, με την ονομασία data scientist. Όπως αναφέρει το περιοδικό Economist «μία νέα ιδιότητα στελέχους έχει αναδειχθεί, ο επιστημονικός υπεύθυνος δεδομένων (the data scientist), ο οποίος συνδυάζει τα προσόντα ενός ικανού προγραμματιστή, στατιστικού και αφηγητή με σκοπό να ανακαλύψει τους ‘σβώλους χρυσού’ που υπάρχουν κάτω από τεράστιους όγκους δεδομένων». Επίσης, ο Tom Davenport σε πρόσφατο άρθρο του στο Harvard Business Review («Data Scientist: The Sexiest Job of the 21st Century») επιχειρηματολογεί για την αναγκαιότητα τέτοιων επαγγελματιών, ενώ επισημαίνει τη μεγάλη έλλειψη ανθρώπων σε αυτόν τον κλάδο. Μία πρόσφατη μελέτη της Gartner, προβλέπει 1,9 εκατ. θέσεις εργασίας για data scientists και άλλα 4 εκατ. για την υποστήριξη αυτών – μόνο στις Η.Π.Α - τα επόμενα τρία χρόνια.

Το συγκεκριμένο σεμινάριο έχει διάρκεια 60 ωρών, διαιρεμένο σε 10 εβδομάδες των δύο τριώρων διαλέξεων/εβδομάδα. Ο σκοπός του είναι να αποτελέσει για τον καταρτιζόμενο μία αρχική εισαγωγή στους βασικούς τομείς των business analytics & big data: διαχείριση δεδομένων, στατιστική, μηχανική εκμάθηση, εργαλεία και συστήματα. Ιδιαίτερη έμφαση δίνεται στην πρακτική διάσταση των παραπάνω τομέων και λιγότερο στις θεωρητικές έννοιες. Σαν αποτέλεσμα, τα θέματα του σεμιναρίου επικεντρώνονται σε εργαλεία και συστήματα που χρησιμοποιούνται σε εφαρμογές big data και την επίδειξη και κατανόηση του λεγόμενου data analysis life cycle.

Οι θεματικές ενότητες είναι οι εξής:

- **Modern Data Management, Business Intelligence and Big Data Systems:** *Fundamentals of modern data management:* data modeling, query languages, query processing, parallel and distributed systems, transactions. The 3Vs of big data. The data analysis lifecycle. *Business Intelligence:* Architectures, multi-dimensional modeling, data cubes, OLAP, indexing, applications. In-memory and column-oriented technologies. Business intelligence in the big data era. *Big Data Systems:* Hadoop and related technologies, NoSQL systems (Mongo, Redis, Cassandra, graph databases), stream engines.
- **Statistics for Business Analytics, An Overview:** An overview of statistical methods, techniques and principles used in predictive analytics along with examples.
- **Big Data Systems, A Practical Outline:** A brief presentation of five systems used in big data applications, along with representative applications: Hadoop/HDFS/Hive, Cassandra, MongoDB, Spark and Redis.
- **Mining Big Datasets:** An overview of data mining concepts such as clustering, classification, association rules, graph mining over large datasets, along with techniques, algorithms, examples.
- **Working with Python for Analytics:** An introduction to Python and examples on how to use it in analytics application.
- **An Introduction to R with Applications:** An introduction to Python and examples on how to use it in analytics application.
- **Content Analytics:** Methods, techniques, tools and systems to extract structure from unstructured data, such as text, audio, speech, images. Sentiment analysis.
- **Visual Analytics:** Advanced data visualization techniques and practices, specifically for big data interactive exploratory analysis, with live demos and applications.
- **Analytics Cases Studies:** Four representative end-to-end analytics case studies, presented by established consulting/software firms and ambitious startups alike. The case studies may be drawn from financial, energy, insurance, healthcare and other domains.

Το κόστος συμμετοχής είναι €720 και καλύπτεται από τον ΟΑΕΔ μέσω του προγράμματος ΛΑΕΚ για την επιμόρφωση εργαζομένων σε επιχειρήσεις. Δίνεται έκπτωση 10% για περισσότερα του ενός ατόμων από τον ίδιο οργανισμό. Θα δοθεί πιστοποιητικό παρακολούθησης και θα διανεμηθούν σημειώσεις. Δεδομένου του περιορισμένου αριθμού θέσεων, θα τηρηθεί σειρά προτεραιότητας.

Επιστημονικός Υπεύθυνος:
Δαμιανός Χατζηαντωνίου, Αναπληρωτής Καθηγητής, Οικονομικό Πανεπιστήμιο Αθηνών

Περσσότερες πληροφορίες:
Υπεύθυνος ΚΕΚ Ο.Π.Α. κ. Πέτρος Καλκάνης & Ευαγγελία Φλέσσουρα, Ευαγγελία Λυκούρη, 210-8203-912, -913, -914
<http://kek.aueb.gr/big-data-business-analytics/>

Πρόγραμμα Διαλέξεων

1^η εβδομάδα: Τρίτη 8/3/2016 & Τετάρτη 9/3/2016, 18.00-21.00	
Εισηγητής: Αναπ. Καθηγητής Δαμιανός Χατζηαντωνίου <i>Τμήμα Διοικητικής Επιστήμης & Τεχνολογίας, Ο.Π.Α.</i>	Modern Data Management, Business Intelligence and Big Data Systems
2^η εβδομάδα: Τρίτη 15/3/2016 & Τετάρτη 16/3/2016, 18.00-21.00	
Εισηγητής: Αναπ. Καθηγητής Δαμιανός Χατζηαντωνίου <i>Τμήμα Διοικητικής Επιστήμης & Τεχνολογίας, Ο.Π.Α.</i>	Modern Data Management, Business Intelligence and Big Data Systems
3^η εβδομάδα: Δευτέρα 21/3/2016 & Τετάρτη 23/3/2016, 18.00-21.00	
Εισηγητής: Αναπ. Καθηγητής Δημητρός Καρλής <i>Τμήμα Στατιστικής, Ο.Π.Α.</i>	Statistics for Business Analytics: An Overview
4^η εβδομάδα: Δευτέρα 28/3/2016 & Τετάρτη 30/3/2016, 18.00-21.00	
Εισηγητής: Αναπ. Καθηγητής Γιάννης Κοτίδης <i>Τμήμα Πληροφορικής, Ο.Π.Α.</i>	Mining Big Datasets
5^η εβδομάδα: Δευτέρα 4/4/2016 & Τετάρτη 6/4/2016, 18.00-21.00	
Εισηγητής: Παύλος Μητσούλης-Ντόμπος, MSc <i>Workable Inc.</i>	Big Data Systems: A Practical Outline
6^η εβδομάδα: Δευτέρα 11/4/2016 & Τετάρτη 13/4/2016, 18.00-21.00	
Εισηγητής: Rui Miguel Forte, Msc, PhD candidate <i>Workable & Computer Science Dept., University of Edinburgh</i>	An Introduction to R with Applications
7^η εβδομάδα: Δευτέρα 18/4/2016 & Τετάρτη 20/4/2016, 18.00-21.00	
Εισηγητής: Αναπ. Καθηγητής Πάνος Λουρίδας <i>Τμήμα Διοικητικής Επιστήμης & Τεχνολογίας, Ο.Π.Α.</i>	Working with Python for Analytics
8^η εβδομάδα: Δευτέρα 9/5/2016 & Τετάρτη 11/5/2016, 18.00-21.00	
Εισηγητής: Δρ. Ανδρέας Μανιάτης <i>CyberStream</i>	Visual Analytics
9^η εβδομάδα: Δευτέρα 16/5/2016 & Τετάρτη 18/5/2016, 18.00-21.00	
Εισηγητής: Δρ. Χάρης Παπαγεωργίου <i>Ινστιτούτο Επεξεργασίας Λόγου Ερευνητικό Κέντρο "Αθηνά" & Qualia</i>	Content Analytics: Text, Speech and Video Analytics
10^η εβδομάδα: Δευτέρα 23/5/2016 & Τετάρτη 25/5/2016, 18.00-21.00	
Εισηγητές: Accenture, EY, IBM, Oracle, QlikView, SAS, Targit (TBA)	Analytics Cases Studies

Κύριοι Εισηγητές

Δρ. Δημήτρης Καρλής. Ο κ. Δημήτρης Καρλής είναι Αναπληρωτής Καθηγητής στο Τμήμα Στατιστικής του Οικονομικού Πανεπιστημίου Αθηνών (ΟΠΑ). Έχει λάβει πτυχίο Στατιστικής το 1992 από το Τμήμα Στατιστικής του ΟΠΑ και διδακτορικό τίτλο στη Στατιστική από το ίδιο Τμήμα το 1999. Έχει θητεύσει στο ίδιο Τμήμα ως προσωρινός διδάσκων (2000-2001) και ως Λέκτορας (2001-2004). Από τον Οκτώβριο του 2004 μέχρι τον Δεκέμβριο του 2010 διετέλεσε Επίκουρος Καθηγητής στο ίδιο Τμήμα. Έχει δημοσιεύσει περίπου 45 άρθρα σε διεθνή περιοδικά Στατιστικής με κριτές. Τα περιοδικά στα οποία έχει δημοσιεύσει μεταξύ άλλων περιλαμβάνουν Journal of the Royal Statistical Society Series A, Statistics in Medicine, Statistics and Computing, και αρκετά άλλα. Τα ερευνητικά του ενδιαφέροντα περιλαμβάνουν θέματα υπολογιστικής στατιστικής και αλγορίθμων, ανάλυση πολυμεταβλητών διακριτών δεδομένων, μείξεις κατανομών, στατιστικά μοντέλα για εφαρμογές στον αθλητισμό, μοντελοποίησης εξαρτημένων δεδομένων με τη χρήση συναρτήσεων copulas και άλλα. Είναι Associate editor στα περιοδικά Metron journal, Communications in Statistics (Theory and Methods and Computation and Simulation) και IMA Journal of Management Mathematics, ενώ έχει κάνει refereee σε περισσότερα από 85 άρθρα σε διάφορα περιοδικά Στατιστικής και συναφών αντικειμένων. Έχει συμμετάσχει σε πολλά διεθνή συνέδρια, σε μερικά εκ των οποίων ως προσκεκλημένο ομιλητή. Είναι μέλος διάφορων ενώσεων σχετικές με την επιστήμη της Στατιστικής, όπως της American Statistical Society, εκλεγμένο μέλος του International Statistical Institute, μέλος της International Association of Statistical Computing, υπεύθυνος δημοσιότητας της Eastern Mediterranean Region of the International Biometrics Society, καθώς και μέλος του Ελληνικού Στατιστικού Ινστιτούτου. Έχει, επίσης, εργασθεί σε μια σειρά από ερευνητικά προγράμματα που χρηματοδοτήθηκαν από την Ευρωπαϊκή Ένωση και τη EUROSTAT για έργα σχετικά με official statistics.

Δρ. Γιάννης Κωτίδης. Ο Γιάννης Κωτίδης είναι Αναπληρωτής Καθηγητής στο Τμήμα Πληροφορικής του Οικονομικού Πανεπιστημίου Αθηνών. Έχει λάβει πτυχίο από το Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου το 1995 και τίτλο μεταπτυχιακών σπουδών (1997) και διδακτορικό (2000) στην Επιστήμη των Υπολογιστών από το University of Maryland. Μέχρι τον Ιανουάριο του 2006 δούλεψε σαν ανώτερο στέλεχος της ομάδας διαχείρισης βάσεων δεδομένων στο ερευνητικό κέντρο της AT&T Labs – Research. Τα ερευνητικά του ενδιαφέροντα περιλαμβάνουν συστήματα, εφαρμογές και εργαλεία διαχείρισης πολύ μεγάλων βάσεων δεδομένων. Ο κ. Κωτίδης έχει αναπτύξει έντονο ερευνητικό έργο έχοντας περισσότερα από 80 δημοσιευμένα άρθρα σε κορυφαία περιοδικά και συνέδρια της περιοχής διαχείρισης δεδομένων και είναι κάτοχος 5 διπλωμάτων ευρεσιτεχνίας (Αμερική). Έχει λάβει σημαντικά βραβεία για τις δημοσιεύσεις του (ACM SIMGOD'99, IEEE ICDE'09) όπως το βραβείο επιστημονικής αριστείας από τα AT&T Labs για την εργασία του σε θέματα ολοκλήρωσης πληροφοριακών συστημάτων.

Δρ. Πάνος Λουριδάς. Ο Πάνος Λουριδάς είναι Αναπληρωτής Καθηγητής στο Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας, στο Οικονομικό Πανεπιστήμιο Αθηνών. Έχει εργαστεί επί σειρά ετών στο σχεδιασμό και στην ανάπτυξη μεγάλων συστημάτων λογισμικού και στην επεξεργασία και αξιοποίηση δεδομένων μεγάλου όγκου, τόσο ως ερευνητής όσο και ως επαγγελματίας προγραμματιστής. Έχει δημοσιεύσει εργασίες σε πολλούς διαφορετικούς τομείς, από την ανάλυση της Wikipedia μέχρι θεωρία σχεδίου, και από συστήματα ασφαλείας μέχρι τεχνολογία λογισμικού. Με το ρόλο του συμβούλου στο Εθνικό Δίκτυο Έρευνας και Τεχνολογίας (ΕΔΕΤ) είναι υπεύθυνος για το νεοϋπολογιστικό σύστημα ~okeanos (<http://okeanos.grnet.gr>), το οποίο παρέχει χιλιάδες ιδεατές μηχανές στην ελληνική ερευνητική και ακαδημαϊκή κοινότητα, και για το σύστημα ηλεκτρονικών εκλογών Ζεός (<http://zeus.grnet.gr>), το οποίο χρησιμοποιείται για τη διενέργεια νομικά έγκυρων ηλεκτρονικών εκλογών. Παπαγεωργίου Χάρης είναι ερευνητής Β' βαθμίδος στο Ινστιτούτο Επεξεργασίας του Λόγου-ΙΕΛ/Ερευνητικό Κέντρο «ΑΘΗΝΑ» και συνιδρυτής της spin off εταιρίας Qualia [www.qualia.gr], που δραστηριοποιείται στο χώρο της επιχειρηματικής ευφυΐας και αναλυτικής δεδομένων και η οποία συστήθηκε το 2004. Απόφοιτος της Σχολής

Δρ. Χάρης Παπαγεωργίου. Ο Παπαγεωργίου Χάρης είναι ερευνητής Β' βαθμίδος στο Ινστιτούτο Επεξεργασίας του Λόγου-ΙΕΛ/Ερευνητικό Κέντρο «ΑΘΗΝΑ» και συνιδρυτής της spin off εταιρίας Qualia [www.qualia.gr], που δραστηριοποιείται στο χώρο της επιχειρηματικής ευφυΐας και αναλυτικής δεδομένων και η οποία συστήθηκε το 2004. Απόφοιτος της Σχολής Ηλεκτρολόγων Μηχανικών του ΕΜΠ το 1992. Το 1996 πήρε τον τίτλο του Διδάκτορα Μηχανικού από το ΕΜΠ. Το 1989 συμμετείχε στην ομάδα Eurotra στο ΙΕΛ. Από το 1992 είναι ερευνητής στο ΙΕΛ στο χώρο της γλωσσικής τεχνολογίας και της αναλυτικής δεδομένων. Έχει συμμετάσχει σε πλήθος ευρωπαϊκών και εθνικών προγραμμάτων με ιδιαίτερη έμφαση στην πολυμεσική, πολυτροπική και πολυγλωσσική ανάλυση (ενδεικτικά Eurotra, LRE/Translearn, Esprit-Naplus, LE-Aventinus, IST-CIMWOS, IST-Reveal_This, eBusiness-MUSE, COST278, HEG-TV++, COST2102, IST-METANET). Έχει παρουσιάσει παραπάνω από 60 εκτενείς δημοσιεύσεις σε διεθνή περιοδικά καθώς και συνέδρια. Κατέχει δίπλωμα Ευρεσιτεχνίας σχετικά με την ανάκτηση πολύγλωσσων κειμένων.

Δρ. Δαμιανός Χατζηαντωνίου (Επιστημονικός Υπεύθυνος). Ο κ. Δαμιανός Χατζηαντωνίου είναι Αναπληρωτής Καθηγητής στο Τμήμα Διοικητικής Επιστήμης και Τεχνολογίας του Οικονομικού Πανεπιστημίου Αθηνών με αντικείμενο Διαχείριση Δεδομένων και Επιχειρηματική Ευφυΐα. Είναι επίσης Διευθυντής του διεθνούς μεταπτυχιακού προγράμματος "Business Analytics" στο Ο.Π.Α. Έχει πτυχίο Εφαρμοσμένων Μαθηματικών του Πανεπιστημίου Αθηνών, τίτλο μεταπτυχιακών σπουδών στην Επιστήμη Υπολογιστών από το Courant Institute of Mathematical Sciences του Πανεπιστημίου της Νέας Υόρκης (1993) και διδακτορικό από το Columbia University (1997). Τα ερευνητικά του ενδιαφέροντα περιλαμβάνουν θέματα βάσεων δεδομένων και επιχειρηματικής ευφυΐας, όπως για παράδειγμα αποθήκες δεδομένων, ανάλυση OLAP, συστήματα ροών δεδομένων, επεξεργασία ερωτημάτων και large-scale analytics. Έχει δημοσιεύσει περισσότερα από 30 άρθρα σε κορυφαία διεθνή περιοδικά και συνέδρια, όπως VLDB, ICDE, EDBT, SIGMOD, SIGMOD Record, Information Systems Journal, DKE Journal και άλλου, ενώ υπάρχουν περισσότερες των 200 ετεροαναφορών στις δημοσιεύσεις αυτές, περιλαμβανομένων και εννέα διεθνών textbooks. Τα αποτελέσματα της ερευνητικής εργασίας του κ. Χατζηαντωνίου έχουν χρησιμοποιηθεί από τα συστήματα Microsoft SQL Server (query processor), Oracle 8i και 9i (Analytic Functions), έχουν επηρεάσει τη διαμόρφωση του ANSI SQL standard (OLAP amendment) και αναφέρονται σαν πρωτεύουσες αναφορές σε περισσότερες από έντεκα πατέντες. Έχει διατελέσει Επίκουρος Καθηγητής στο Stevens Institute of Technology και έχει συνεργαστεί με τα AT&T Research Labs και Columbia Medical Informatics Department. Εκτός των ακαδημαϊκών δραστηριοτήτων, ήταν ιδρυτικό μέλος δύο εταιριών πληροφορικής (start ups), μία με βάση τη Νέα Υόρκη (Panakea Software Inc., DW&OLAP tools & consulting) και μία στην Αθήνα (Voiceweb, speech & telecom applications), ενώ ήταν senior research consultant το 2007-2008 στην Aster Data με βάση το San Francisco (πλέον Aster Teradata).