

Ανάλυση Δεδομένων, Big Data και Ευκαιρίες στη Σύγχρονη Εποχή

Στα πλαίσια της ημερίδας

ΣΕΜΦΕ: ο παράξενος ελκυστής

Οι διπλωματούχοι "επιστρέφουν" στη Σχολή για να παρουσιάσουν τη δουλειά τους



Ευστρατία Χαριτίδου

Scientific Software Engineer

30 Μαΐου 2016



Σύνοψη

- 1** Προσωπική Πορεία
 - 2** IRI – Συνοπτική Παρουσίαση
 - 3** Η Επιστήμη των Big Data
 - 4** Ο Επιστήμονας Δεδομένων
-

Εν αρχή ην.. Οι Σπουδές

BSc ... Σ.Ε.Μ.Φ.Ε.

- Μαθηματικός εφαρμογών
- Ροές: Εφαρμοσμένη στατιστική και Μαθηματικά πληροφορικής

--- διάλειμμα ---

Erasmus Βιοστατιστική και Κλινική Έρευνα (MSc), Πανεπιστήμιο Paul-Sabatier, Τουλούζη, Γαλλία

MSc Statistics with Applications in Medicine, School of Mathematics, Πανεπιστήμιο του Southampton, Ηνωμένο Βασίλειο.

--- end ---

PhD Μπεϋζιανή Στατιστική, Τομέας Μαθηματικών, **Σ.Ε.Μ.Φ.Ε.**

- Θέμα: Μπεϋζιανή επιλογή οικογένειας μετασχηματισμών
- Επιβλέπων: Δημήτρης Φουσκάκης

Εργασιακή Δραστηριότητα

➤ **Στατιστικός Αναλυτής (free lance)**

Ενδεικτικοί συνεργάτες:

- *Frigoglass S.A.I.C.*,
- Πανεπιστημιακό Νοσοκομείο “*Ευαγγελισμός*”,
- ΜΚΟ *Ε.Μ.Ε.Ι.Σ.*, κ.α.

➤ **Εθνικό Μετσόβιο Πολυτεχνείο, Σ.Ε.Μ.Φ.Ε.**

Θέση: Υπεύθυνη εργαστηρίων στα μαθήματα

- “Ανάλυση Δεδομένων με Η/Υ” και
- “Μπευζιανή Στατιστική και MCMC”.

➤ **Νοσοκομείο ΡΕΑ, Κλινική Μαστού**

Θέση: Επιστήμονας δεδομένων και ερευνήτρια βιοστατιστικής σε θέματα καρκίνου μαστού.

Σήμερα..

IRi

- Scientific software engineer
- Innovations team (~ R&D)

Λέξεις – κλειδιά:

- R
- Big Data
- Διαχείριση δεδομένων
- Στατιστική μοντελοποίηση
- Cloud computing
- Προγνωστικά analytics - Forecasting
- Βελτιστοποίηση κώδικα
- Machine learning



IRi

Σύντομες Πληροφορίες



IRi
Growth delivered.

Λίγα λόγια για την εταιρεία

Η IRi αποτελεί έναν από τους κορυφαίους παρόχους:

- Διαχείρισης Big Data
 - Predictive Analytics
- με δραστηριότητα πάνω από 30 χρόνια.

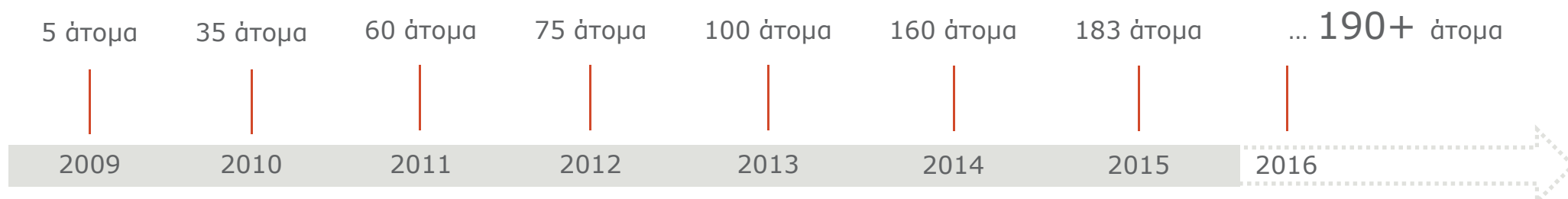
Βασικοί τομείς εμπειρογνωμοσύνης:

- Μέτρηση επιδόσεων της αγοράς και στρατηγική
- Συμπεριφορά καταναλωτών
- Analytics της αγοράς
- Μέσα επικοινωνίας και διαφήμισης
- Τεχνολογία

Βασικός στόχος:

Η παροχή λύσεων στην αγορά για τους μεγάλους κατασκευαστές και τους εμπόρους λιανικής.

Ανάπτυξη της IRi στην Ελλάδα





Η Επιστήμη των Big Data

Στο μεταίχμιο της επανάστασης των *Analytics*



Big Data: Εισαγωγικά

Big Data:

- Θεωρούνται ο νέος “χρυσός”.
- Καταλύτης για τη σύγχρονη βιομηχανία.
- ΟΜΩΣ προϋπήρχαν εξ αρχής!
- ΣΗΜΕΡΑ: είναι πολύ φθηνότερη η αποθήκευσή τους παρά η απόρριψή τους.

Τι είναι τα Big Data:

- Δεδομένα που δε χωρούν σε ένα αρχείο Excel?
- Δεδομένα που δε χωρούν σε ένα τυπικό PC?
- Όλα αυτά, αλλά και δεδομένα με μεγάλη χρησιμότητα...

Μέγεθος:

- Από μερικά **Terabytes** ως και **Petabytes** (10^{15} bytes).
- Π.χ. Αρχείο εκατοντάδων εκατομμυρίων γραμμών δεδομένων.

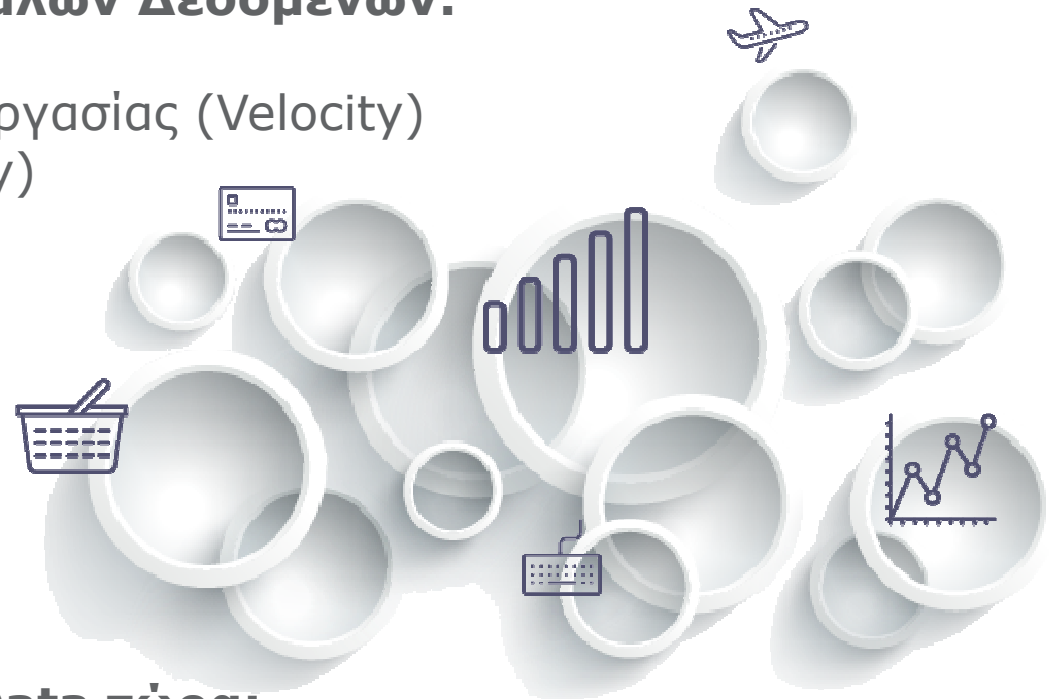


Τα Χαρακτηριστικά των Big Data: Η Θεωρία των 3V

Πότε έχουμε ένα πρόβλημα Μεγάλων Δεδομένων:

- Πάρα πολλά bytes (Volume)
- Πάρα πολύ υψηλός ρυθμός επεξεργασίας (Velocity)
- Πάρα πολλές πηγές/τύποι (Variety)
- Δύσκολη ανάλυση (non scalable)
- Δύσκολη οπτικοποίηση

- Τα πληροφοριακά δεδομένα υπερδιπλασιάζονται κάθε δύο χρόνια.



Γιατί υπάρχει η ανάγκη για Big Data τώρα:

- Αφθονία δεδομένων σε πάρα πολλούς κλάδους (υγεία, εμπόριο, κλπ).
- Έκρηξη δύναμικής των υπολογιστικών συστημάτων.

Πού Τοποθετείται ο Επιστήμονας Δεδομένων

Μηχανική

Μαθηματικά/
Στατιστική

Κοινωνικές
Επιστήμες

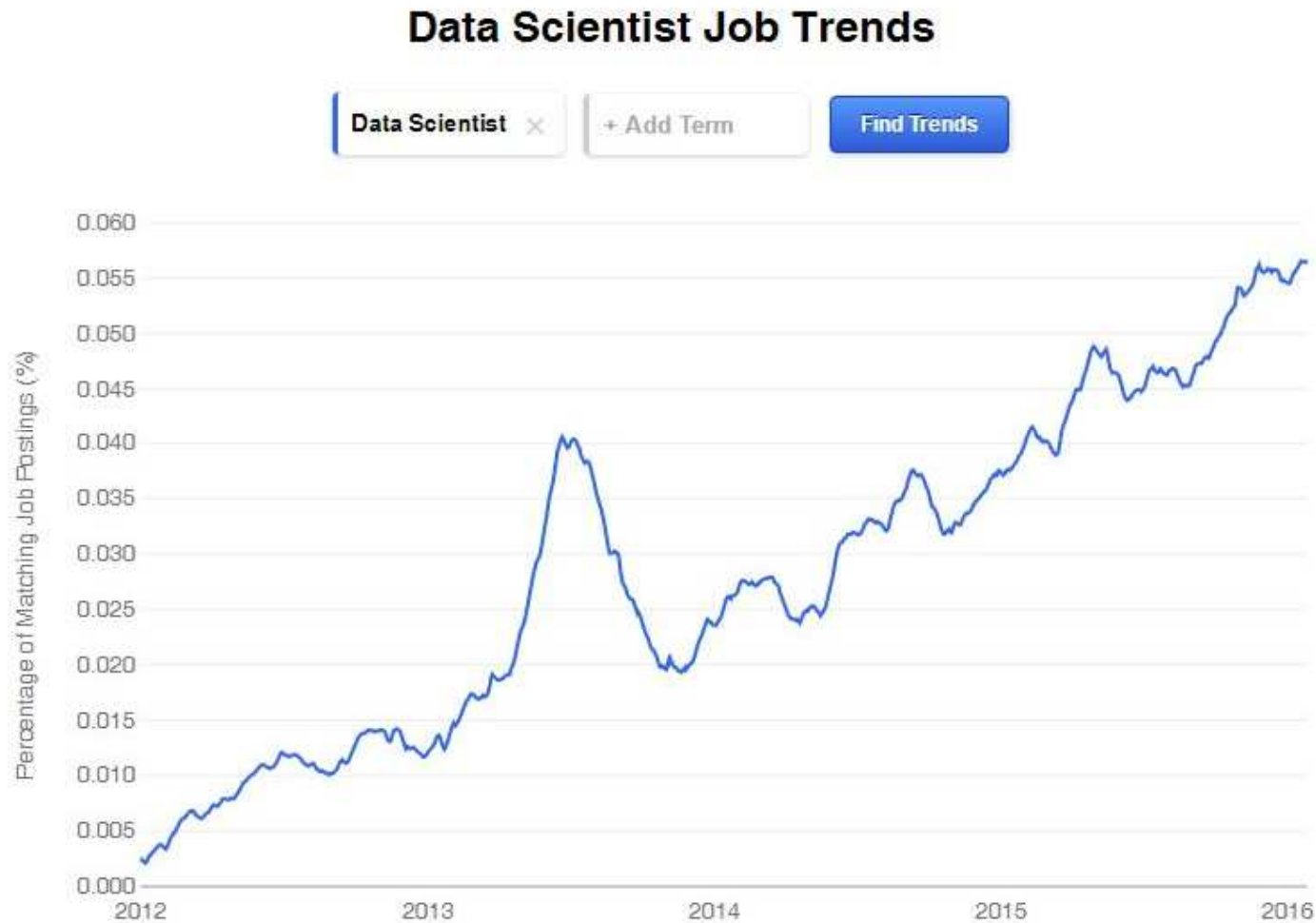
Επιστήμονας Δεδομένων

Επικοινωνία

Εν Τέλει... Με τι Ασχολούνται οι Επιστήμονες Δεδομένων;

- **Προγραμματισμός:** Back-end, Front-end, Διαχείριση Συστημάτων
- **Στατιστική:** Κλασσική, Χωρική, Διαχείριση Δεδομένων, Οπτικοποίηση
- **Μαθηματικά:** Αλγόριθμοι, Βελτιστοποίηση, Γραφικά Μοντέλα
- **Business:** Ανάπτυξη Προϊόντων ή/και Υπηρεσιών
- **Machine Learning/Big Data:** Δομημένα ή μη

Επιστήμονας Δεδομένων: Σημαντική Αύξηση Ζήτησης



Και στην Ελλάδα έχουν ξεπηδήσει εταιρείες/start-ups που αναζητούν εργαζόμενους με ανάλογο προφίλ.

Πηγή: Indeed.com

Ευχαριστώ!



IRi
Growth delivered.