



**Πανεπιστήμιο Πειραιώς Τμήμα**

**Ψηφιακών Συστημάτων**

**Π.Μ.Σ. 2015-16**

*Διπλωματική εργασία*

**ΕΛΕΓΧΟΣ ΙΔΙΩΤΙΚΟΤΗΤΑΣ ΣΕ ΠΑΡΟΧΟΥΣ  
ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΗΡΕΣΙΩΝ**

**Νίκος Κυραμάς ΜΤΕ 1519**

**Επιβλέπων Καθηγητής: ΚΩΝΣΤΑΝΤΙΝΟΣ ΛΑΜΠΡΙΝΟΥΔΑΚΗΣ**

# Περιεχόμενα

<b>Κεφάλαιο 1: Big Data, Data mining και Profiling</b> .....	<b>4</b>
1.1 Τι είναι τα “Μεγάλα Δεδομένα”.....	4
Ορισμός.....	4
Βασικά Χαρακτηριστικά των Big Data.....	4
Παράγοντες Κλειδί στα BD.....	7
1.2 Τι είναι το Εξόρυξη Δεδομένων (Data mining) .....	8
Ορισμός.....	8
Στάδια της εξόρυξης δεδομένων .....	8
Κατηγορίες εξόρυξης δεδομένων.....	10
Τεχνικές εξόρυξης δεδομένων.....	11
Εφαρμογές εξόρυξης δεδομένων.....	13
1.3 Κατασκευή Προφίλ (Profiling) .....	14
Ορισμός.....	14
Profiling και Data Mining .....	15
Μηχανές Κατασκευής Προφίλ: .....	16
1.4 Εφαρμογές και Συνεισφορά των Big Data .....	17
Υγεία: .....	17
Mobile: .....	18
Έξυπνο δίκτυο:.....	18
Διαχείριση της κυκλοφορίας: .....	19
Λιανεμπόριο: .....	19
Πληρωμές.....	20
Διαδίκτυο.....	20
<b>Κεφάλαιο 2: Big Data στον Τραπεζικό Τομέα</b> .....	<b>22</b>
2.1 Τράπεζες και Information Technology.....	22
Τράπεζες και η επιστήμη των Πληροφοριακών Συστημάτων (Information Technology - IT):....	22
Χρηματοοικονομική Τεχνολογία ή FinTech .....	24
Η ανάγκη των Big Data στις τράπεζες .....	26
2.2 Εφαρμογές BD και DM στο Τραπεζικό Σύστημα.....	27
Διαχείριση κινδύνων .....	27
Διαχείριση χαρτοφυλακίου.....	29
Διαχείριση και Προφίλ Πελατών .....	31
Μάρκετινγκ .....	34
Ανίχνευση Απάτης στον τραπεζικό τομέα .....	36
2.3 Το Τραπεζικό Σύστημα στην εποχή των Big Data.....	38
Χρήση BD εκ των έσω .....	38
Μελέτες Περίπτωσης: .....	39

<b>Κεφάλαιο 3: Νομοθεσία στα Big Data και τον Τραπεζικό Τομέα.....</b>	<b>44</b>
3.1 Προβλήματα με τα Big Data .....	44
Ρίσκα που ανακύπτουν από την κατασκευή Προφίλ.....	44
Προβλήματα εφαρμογής .....	47
Προβλήματα από την χρήση στο Τραπεζικό σύστημα.....	53
3.2 Προστασία από τα Big Data.....	53
Νομικό πλαίσιο προστασίας Big Data, Data Mining και Profiling .....	53
The GDPR και ο Τραπεζικός τομέας .....	62
<i>Καταφέρνει το Νομικό πλαίσιο το στόχο του;</i> .....	66
Μελέτη περίπτωσης: Εξατομικευμένη τιμολόγηση υπό GDPR.....	71
Πως θα πρέπει να αντιμετωπίζεται οι Παραβιάσεις.....	74
Προτεινόμενες Λύσεις.....	83
Συμπεράσματα.....	88
Βιβλιογραφία.....	89

# Κεφάλαιο 1: Big Data, Data mining και Profiling

## 1.1 Τι είναι τα “Μεγάλα Δεδομένα”

### Ορισμός

Ο όρος **Big Data** είναι μια αόριστη έννοια για την οποία υπάρχει μεγάλος διάλογος ,συμφωνίες και διαφωνίες πάνω στον ορισμό της από ακαδημαϊκούς, επαγγελματίες της πληροφορικής ,ακόμα και κυβερνητικούς εκπροσώπους.

Σε γενικές γραμμές μπορούμε ορίσουμε σαν “Μεγάλα Δεδομένα” Τα μεγάλα δεδομένα προέκυψαν αρχικά ως όρος για την περιγραφή μεγάλων συνόλων δεδομένων που δεν μπορούσαν να ληφθούν, να αποθηκευτούν, να διαχειριστούν ούτε να αναλυθούν χρησιμοποιώντας παραδοσιακές βάσεις δεδομένων. Ωστόσο, ο ορισμός έχει διευρυνθεί σημαντικά τα τελευταία χρόνια. Τα μεγάλα δεδομένα δεν αφορούν μόνο τα ίδια τα δεδομένα αλλά και το σύνολο των τεχνολογιών που εκτελούν όλες τις προαναφερόμενες λειτουργίες, καθώς και ποικίλες συλλογές που επιλύουν περίπλοκα προβλήματα (.1) και καθιστούν την αξία ξεκλειδώματος από αυτά τα δεδομένα πιο οικονομικά(.2).

Ωστόσο αξίζει να σημειωθεί ότι ένα τεράστιο σύνολο ηλεκτρονικών δεδομένων αυτούσιο είναι πού μικρής αξίας και χρησιμότητας. Έχει αναπτυχθεί μια σειρά νέων τεχνολογιών γύρω από τη συλλογή, ταξινόμηση, ανάλυση, συνδυασμό και αξιοποίηση πληροφοριών που προκύπτουν από την επεξεργασία των δεδομένων αυτών οι οποίες συγκροτούν την **Επιστήμη των Big Data**.

Η **ουσιαστική αξία των Big Data** δεν βρίσκεται στο μέγεθος των δεδομένων αλλά στην συνδεσιμότητά τους με άλλα δεδομένα. Χάρη στις επιστήμες εξόρυξης (Data Mining), συσχετισμού και συνάθροισης τα Big Data μετατρέπονται ουσιαστικά σε ένα δίκτυο. Η αξία της προέρχεται από τις σχέσεις και τα μοτίβα που μπορεί να προκύψουν από την πραγματοποίηση των συνδέσεων μεταξύ των δεδομένων, για ένα άτομο, για το άτομο σε σχέση με μια ομάδα, για ομάδες ανθρώπων, ή απλά για τη δομή της ίδιας της πληροφορίας(.3).

### Βασικά Χαρακτηριστικά των Big Data

Ένας άλλος, πιο ακριβής ορισμός για τον όρο “Μεγάλα Δεδομένα ” είναι “δεδομένα τα οποία χαρακτηρίζονται από τα λεγόμενα **3Vs :Volume, velocity, variety**” δηλαδή δεδομένα που χαρακτηρίζονται από μεγάλο όγκο, ταχύτητα επεξεργασίας και ποικιλία δεδομένων. Στον διάλογο για τον ορισμό των Big Data τους έχουν αποδοθεί και άλλα χαρακτηριστικά όπως Μεταβλητότητα (Variability), Ακρίβεια (Veracity), Απεικόνιση (Visualization) και Αξία (Value). Στην κυριότητά τους όμως όλοι συμφωνούν στα 3Vs στα οποία θα μείνουμε και περισσότερο (.4).

**Όγκος :** είναι το πόσα δεδομένα που έχουμε. Τεράστιες ποσότητες δεδομένων είναι τώρα σε επεξεργασία με τη βασική μονάδα των «Big Data» να μεταβαίνει από τα Gigabytes και Terabytes σε

Exabytes ή ακόμα και Zettabytes(.5). Το Ίντερνετ των πραγμάτων, τα Δίκτυα Κοινωνικής Δικτύωσης, οι online συναλλαγές και η κίνηση του παγκοσμίου δικτύου συνολικά έχει δημιουργήσει εκθετική αύξηση των δεδομένων.

Εκτιμάται ότι σήμερα το ψηφιακό σύμπαν αποτελείται από περισσότερα από 1.200 δισεκατομμύρια δισεκατομμύρια ψηφιολέξεις, το 90% των οποίων φαίνεται να έχει παραχθεί τα τελευταία δύο χρόνια. Ο αριθμός αυτός, ο οποίος διπλασιάζεται ανά διετία, θα πρέπει να πολλαπλασιαστεί δεκαπλάσια έως το έτος 2020, φθάνοντας συνολικά 44 zettabytes ή 44 τρισεκατομμύρια gigabytes.

Η μετάφραση ή μάλλον η μεταγραφή από τα συστήματα ηλεκτρονικών υπολογιστών του φυσικού κόσμου και των κατοίκων της υπό τη μορφή μεταβληζόμενων δεδομένων δεν περιορίζεται πλέον ούτε περιορίζεται ουσιαστικά από τεχνική ή οικονομική δυσκολία. Ενώ η συλλογή, η μεταφορά και η αποθήκευση δεδομένων έχουν καθαρό άμεσο κόστος (.6), το κόστος αυτό μειώνεται σύμφωνα με το νόμο του Moore (ο διπλασιασμός της χωρητικότητας καταγραφής δεδομένων σε ένα τσιπ πυριτίου κάθε 18 μήνες(.7), αυξάνοντας έτσι την ικανότητα επεξεργασίας και συνεπώς την αποτελεσματικότητα καθιστώντας την είναι δυνατόν, χάρη στον πολλαπλασιασμό των τρανζίστορ, να εκτελείται ένας μεγάλος αριθμός σύνθετων λειτουργιών) και ο νόμος του Nielsen (η ταχύτητα σύνδεσης διπλασιάζεται κάθε 21 μήνες) (.8). Μπορεί επίσης να γίνει αναφορά στον νόμο του Kryder, ο οποίος το 2005 προέβλεπε ότι η πυκνότητα αποθήκευσης του μαγνητικού δίσκου θα διπλασιάστηκε κάθε 13 μήνες. Αν προσθέσουμε σε αυτό την εμφάνιση νέων μορφών αποθήκευσης όπως το SSD, καθίσταται σαφές ότι μπορούμε να αποθηκεύουμε όλο και περισσότερα δεδομένα και να τα αποκτήα όλο και πιο γρήγορα πρόσβαση. Συνεπώς, αυτό που βλέπουμε είναι μια εκθετική αύξηση της μεταφορικής ικανότητας (Moore), της χωρητικότητας αποθήκευσης (Kryder) και της ικανότητας επικοινωνίας (Nielsen).

Η εκθετική αύξηση των Big Data είναι ένα αποτέλεσμα της κατακράτησης όχι μόνο των άμεσα χρήσιμων δεδομένων (η χρησιμότητα(.9) της οποίας ορίζεται από την πραγματική χρήση για ένα δεδομένο σκοπό(.10)), αλλά επίσης και των δεδομένων των οποίων η χρησιμότητα έχει λήξει (και τα οποία είναι δεν είναι πλέον απαραίτητα για το σκοπό αυτό). Είναι η ποσότητα (ή όγκος) πολύ περισσότερο από την ποιότητα των δεδομένων που μπορούν να οδηγήσουν σε μια απροσδόκητη χρησιμότητα όλων των ειδών των δεδομένων, συμπεριλαμβανομένων εκείνων που εκ πρώτης όψεως έχουν το λιγότερο νόημα, που ξεχωριστά μεταφέρουν πολύ λίγες πληροφορίες(.11).

**Ταχύτητα:** Το δεύτερο σημαντικότερο χαρακτηριστικό των Big Data είναι η ταχύτητα με την οποία αυτά συλλέγονται, επεξεργάζονται και είναι προσβάσιμα σε «πραγματικό χρόνο». Ποτέ στο παρελθόν σύνολα δεδομένων ήταν σε θέση να συλλεχθούν τόσο ευέλικτα και γρήγορα. Αυτό

σημαίνει, επίσης, ότι η χρησιμότητα και η αξία των δεδομένων εξελίσσεται σε πραγματικό χρόνο, σύμφωνα με την εισροή νέων δεδομένων.

Δεν είναι μόνο η ταχύτητα με την οποία συγκεντρώνονται τα δεδομένα, αλλά και η ταχύτητα με την οποία επεξεργάζονται τα οποία παρακάμπτουν και απενεργοποιούν τις διαδικασίες ανθρώπινης αντίληψης και κατανόησης και τις διαδικασίες διακήρυξης κινήτρων. Για παράδειγμα, ο στόχος των συσκευών για την ανίχνευση, την ταξινόμηση και την πρόβλεψη της ανθρώπινης συμπεριφοράς και των τάσεων (είτε χρησιμοποιείται στον τομέα της ασφάλειας, της καταπολέμησης της τρομοκρατίας ή του μάρκετινγκ) δεν είναι η δήλωση προθέσεων ενός ατόμου ούτε τα πρώτα στάδια της υλοποίησης αυτών των προθέσεων, αλλά οι διαδικασίες που προηγούνται αυτών, συχνά σε προ-συνειδητό στάδιο. Κάποιος μπορεί να πιστεύει ότι όλα αυτά είναι επιστημονική φαντασία. Καθόλου. Π.χ. ο Eric Schmidt, CEO της Google, αναφέρει πως “ η τεχνολογία θα γίνει σύντομα τόσο αποτελεσματική πως θα είναι πολύ δύσκολο για τους ανθρώπους να δουν ή να καταναλώσουν κάτι που δεν έχει κατά κάποιο τρόπο προσαρμοστεί για αυτούς”(12).

**Ποικιλία:** Ένα άλλο χαρακτηριστικό των Big Data είναι η ποικιλία τους και περιγράφει μια από τις μεγαλύτερες προκλήσεις των μεγάλων δεδομένων. Η ποικιλία των δεδομένων μπορεί να έχει πολλές κατηγορίες. Οι βασικότερες είναι:

- 1) Η ποικιλία των **μορφών** (κείμενο, εικόνες, ήχος, γεωγραφικά δεδομένα, δεδομένα κινητών συσκευών, κλπ),
- 2) **Δομημένα ή αδόμητα** δεδομένα, όπου **δομημένα** είναι τα δεδομένα των οποίων όλοι οι δυνατές τιμές καθορίζονται και γνωστά εκ των προτέρων, ενώ τα **αδόμητα** όχι. Π.χ. σε μια βάση δεδομένων e-mail client, ο συγγραφέας και η ημερομηνία είναι δομημένα δεδομένα, αλλά το σώμα του μηνύματος είναι αδόμητη.
- 3) Από την προέλευσή τους ή το πού παράγονται: Τα **Hard data** είναι δεδομένα που παράγονται από ιδρύματα και δημόσιες διοικητικές αρχές (μητρώα παντός τύπου, καταγγελίες και δικαστικές αποφάσεις, εκλογικούς καταλόγους, άδειες όλων των ειδών, κ.λπ.) **Soft data** δεδομένα που παράγονται από φυσικά πρόσωπα, είτε εκ προθέσεως μέσω των blogs, κοινωνικά δίκτυα και φόρουμ συζητήσεων, είτε αθέλητα από τις δραστηριότητές και αλληλεπιδράσεις, τις online και offline κινήσεις τους, ψηφιακά «ίχνη» που συχνά συλλέγονται από προεπιλεγμένους μηχανισμούς.
- 4) Σε όλα αυτά προστίθενται τα **μεταδεδομένα (metadata)** τα οποία - με την γενικότερη έννοια του όρου - είναι «δεδομένα για τα δεδομένα», δημιουργούνται αυτόματα από τα ίδια τα συστήματα ηλεκτρονικών υπολογιστών, τα οποία καθιστούν δυνατή την περιγραφή και τη δομή άλλων στοιχείων, ανεξάρτητα από το περιεχόμενό τους. τα

μεταδεδομένα μπορούν, για παράδειγμα, είναι τα στοιχεία σχετικά με την τοποθεσία των δεδομένων, σχετικά με το είδος των διαθέσιμων στοιχείων, σχετικά με την πηγή των δεδομένων.

Σήμερα, όταν δουλεύουμε, καταναλώνουμε ή ταξιδεύουμε, αναπαράγουμε αναπόφευκτα δεδομένα, (.12) ταξινομούμε ως μαλακά δεδομένα τα δεδομένα που παράγονται από τα άτομα, είτε εσκεμμένα μέσω ιστολογίων, κοινωνικών δικτύων και φόρουμ συζητήσεων είτε ακούσια, καθώς ένα αυξανόμενο ποσοστό των δραστηριοτήτων τους, των αλληλεπιδράσεων, οι κινήσεις σε απευθείας σύνδεση και εκτός σύνδεσης αφήνουν ψηφιακά "αποτυπώματα" τα οποία συλλέγονται συχνά από μηχανισμούς για την παρακολούθηση επιγραμμικών κινήσεων, CCTV, παρακολούθηση GPS, παρακολούθηση ροής κυκλοφορίας, δορυφορικές εικόνες, καταγραφή τραπεζικών συναλλαγών κ.λπ. και αποθηκεύονται για διάφορους σκοπούς, σπάνια έγινε σαφής κατά τη στιγμή της συλλογής.

### Παράγοντες Κλειδί στα BD

Η τεχνολογία των Big Data έχει τέσσερις βασικούς παράγοντες(.13) (.14): Υποδομή, Αποθήκευση δεδομένων, Επεξεργασία/Διαχείριση και Ανάλυση Δεδομένων. Το 2012, οι πέντε πρώτες σε όλο τον κόσμο γεννήτριες σε έσοδα Big Data ήταν η IBM (USD 1,3 δις), η Hewlett-Packard (USD 664m), Teradata (USD 435), Dell (USD 425), και η Oracle (USD 415) (.15).

- **Υποδομές:** Το κλειδί για την υποδομή μεγάλων δεδομένων είναι η επεκτασιμότητα και η ευελιξία στο χειρισμό petabytes δεδομένων, έτσι το σύννεφο γίνεται μια φυσική επιλογή. Βασικά οι πάροχοι δημόσιων cloud περιλαμβάνουν: Amazon Web Services, GoGrid και Rackspace,
- **Αποθήκευση δεδομένων:** Παραδοσιακά, τα παραδοσιακά συστήματα και οι μέθοδοι αποθήκευσης είναι μη ιδανικά λόγω των περιορισμών των τιμών και της επεκτασιμότητας. Νέες μέθοδοι αποθήκευσης, ιδιαίτερα NoSQL και DFS (Distributed File System) αντιπροσωπεύουν το παραδειγματική στροφή στην αρένα αποθήκευσης. Μεταξύ αυτών, το Hadoop είναι η πιο συχνά χρησιμοποιούμενη μέθοδος αποθήκευσης για Big Data. Βασικοί παίκτες αποθήκευσης Big Data περιλαμβάνουν: MongoDB, Hadoop, Clustrix και Netezza
- **Επεξεργασία και Διαχείριση Δεδομένων:** Αυτό το πλαίσιο χρησιμοποιείται για την επεξεργασία μια αυξανόμενη μάζα των δεδομένων παράλληλα. Βασικά επεξεργαστές δεδομένων και διαχειριστές περιλαμβάνουν: Cloudera, Hadoop, και Greenplum.
- **Data Analytics:** Αυτή είναι η περιοχή που παρέχει οπτικοποίηση, προγνωστική ανάλυση και τις τεχνολογίες data analytics. Βασικά οι πάροχοι analytics

δεδομένων περιλαμβάνουν: Splunk, Clickfox, Rainstor, Pervasive, MapR και Progress dataDirect.

## 1.2 Τι είναι το Εξόρυξη Δεδομένων (Data mining)

### Ορισμός

Αφού έχουμε ορίσει και συγκεντρώσει τα “Μεγάλα” δεδομένα μας αυτό που μένει είναι να τα αξιοποιήσουμε, να τα επεξεργαστούμε και να εξάγουμε χρήσιμα συμπεράσματα και πορίσματα από αυτά. Ο ρόλος αυτός ανήκει σε έναν άλλο επιστημονικό κλάδο της πληροφορικής και της επιστήμης των δεδομένων, στην Εξόρυξη Δεδομένων .

Η Εξόρυξη Δεδομένων (Data Mining) κατά Han και Kamber (.16) είναι μια σειρά από τεχνικές που χρησιμοποιούνται με στόχο την εξαγωγή χρήσιμων μοτίβων και συμπερασμάτων από πολύ μεγάλα σετ δεδομένων. Αυτό επιτυγχάνεται συνδυάζοντας την επιστήμη της στατιστικής και της πληροφορικής, δηλαδή με τη χρησιμοποίηση μεθόδων που βασίζονται στην τεχνητή νοημοσύνη και εκμάθηση μηχανής.

Η Εξόρυξη Δεδομένων επιτρέπει στους χρήστες του τη δυνατότητα μέσω της εφαρμογής του να εξάγουν απαντήσεις σε πολύ δύσκολα ερωτήματα (στατιστικού και μη χαρακτήρα) αλλά και γενική γνώση πάνω στο υποκείμενο προς εξέταση. Γνώση την οποία δεν έχουν φανταστεί ότι μπορεί να προκύψει και απαντήσεις σε ερωτήματα που δεν έχουν ακόμα θέσει. Είναι χαρακτηριστικό ένα απόφθεγμα των Grey και Watson που συνοψίζει πλήρως την χρησιμότητα του "Το data mining επιτρέπει σε αναλυτές και διευθυντές να βρουν απαντήσεις χωρίς να έχουν καν ρωτήσει."

### Στάδια της εξόρυξης δεδομένων

Η ανάλυση των Big Data περιλαμβάνει πολλαπλές διακριτές φάσεις, καθεμία από τις οποίες εισάγει προκλήσεις. Πολλοί άνθρωποι εστιάζουν μόνο στη φάση ανάλυσης: ενώ η φάση είναι ζωτικής σημασίας, έχει μικρή χρησιμότητα χωρίς τις άλλες φάσεις του αγωγού ανάλυση των δεδομένων(.17).

**Συλλογή Δεδομένων:** Τα Big Data δεν προκύπτουν από ένα κενό, καταγράφονται από κάποια πηγή παραγωγής δεδομένων. Μεγάλο μέρος αυτών των δεδομένων δεν παρουσιάζει κανένα ενδιαφέρον, και μπορεί να φιλτραριστεί και να συμπιεστεί κατά τάξεις μεγέθους. Μία πρόκληση είναι να καθοριστούν αυτά τα φίλτρα με τέτοιο τρόπο ώστε να μην απορρίπτονται χρήσιμες πληροφορίες. Η δεύτερη μεγάλη πρόκληση είναι να δημιουργήσουν αυτόματα τα σωστά μεταδεδομένα για να περιγράψουν τα δεδομένα που καταγράφονται και πώς αυτό καταγράφεται και μετράτε.

**Απόσπαση και Καθαρισμός:** Συχνά, οι πληροφορίες που συλλέγονται δεν θα είναι σε μορφή έτοιμη για ανάλυση. Για παράδειγμα, ας δούμε τη συλλογή ηλεκτρονικών αρχείων υγείας σε ένα νοσοκομείο, που περιλαμβάνει μεταγραφόμενες υπαγορεύσεις από διάφορους γιατρούς, δομημένα



δεδομένα από αισθητήρες και μετρήσεις (ενδεχομένως με κάποια σχετική αβεβαιότητα) και δεδομένα εικόνας όπως οι ακτίνες X. Χρειαζόμαστε μια μέθοδο εκχύλισης πληροφοριών που τραβά έξω τις απαιτούμενες πληροφορίες από τις υποκείμενες πηγές και το εκφράζει αυτό σε μια δομημένη μορφή κατάλληλη για ανάλυση.

**Ενσωμάτωση, Άθροιση και Αντιπροσώπευση Δεδομένων:** Δεδομένης της ετερογένειας και της πλημμύρας των δεδομένων, δεν αρκεί απλώς καταγραφούν και να τα ρίξουμε σε ένα αποθετήριο. Η ανάλυση δεδομένων είναι πολύ πιο δύσκολο από τον απλά εντοπισμό, την αναγνώριση, την κατανόηση, και την αναφορά δεδομένων. Για την αποτελεσματική ανάλυση μεγάλης κλίμακας δεδομένων όλα αυτά πρέπει να συμβούν με ένα πλήρως αυτοματοποιημένο τρόπο. Αυτό απαιτεί διαφορές στη δομή των δεδομένων και σημασιολογία που πρόκειται να εκφραστεί σε μορφές που είναι κατανοητές από υπολογιστή, και στη συνέχεια «ρομποτικά» διαχωρίσιμο.

**Επεξεργασία Ερωτημάτων, Μοντελοποίηση και Ανάλυση Δεδομένων:** Οι μέθοδοι για την αναζήτηση και εξόρυξη Big Data είναι θεμελιωδώς διαφορετική από την παραδοσιακή στατιστική ανάλυση σε μικρά δείγματα. Τα Big Data είναι συχνά θορυβώδη, δυναμικά, ετερογενή, αλληλένδετα και αναξιόπιστα. Παρ'όλα αυτά, ακόμα και θορυβώδη Big Data μπορούν να είναι πιο πολύτιμα από μικροσκοπικά δείγματα επειδή τα γενικά στατιστικά στοιχεία που προέρχονται από συχνές μοτίβα και την ανάλυση συσχέτισης συνήθως εξουδετερώνουν τις ατομικές διακυμάνσεις και συχνά αποκαλύπτουν πιο αξιόπιστα, κρυμμένα μοτίβα και γνώση. Περαιτέρω, διασυνδεδεμένα Big Data σχηματίζουν μεγάλα ετερογενή δίκτυα πληροφοριών, με την οποία μπορούν να διερευνηθούν περιττές πληροφορίες για να αντισταθμίσουν τα δεδομένα που λείπουν, για να διασταυρώσει αντικρουόμενες περιπτώσεις, για την επικύρωση αξιόπιστες σχέσεις, να αποκαλύψει εγγενείς συστάδες, και να αποκαλύψει κρυφές σχέσεις και τα μοντέλα.

**Ερμηνεία:** Η ανάλυση των Big Data είναι περιορισμένης αξίας αν οι χρήστες δεν μπορούν να καταλάβουν την ανάλυση και εν ολίγης, τι λένε. Οι λαμβάνοντες αποφάσεων έπειτα της λήψης των αποτελεσμάτων της ανάλυσής τους πρέπει τελικά να ερμηνεύσει αυτά τα αποτελέσματα. Αυτή η ερμηνεία δεν μπορεί να συμβεί από το πουθενά. Συνήθως, πρόκειται για την εξέταση όλων των παραδοχών που προκύπτουν από εντοπίζονται πίσω στην ανάλυση. Επιπλέον υπάρχουν πολλές πιθανές πηγές λάθους: συστήματα ηλεκτρονικών υπολογιστών μπορεί να έχουν σφάλματα, τα μοντέλα έχουν σχεδόν πάντα παραδοχές και τα αποτελέσματα μπορεί να βασίζεται σε λανθασμένα στοιχεία. Για όλους αυτούς τους λόγους, ο χρήστης δεν μπορεί να παραχωρήσει την εξουσία του στο σύστημα. Μάλλον θα προσπαθήσει να καταλάβει, και να επαληθεύσει, τα αποτελέσματα που παράγονται από τον υπολογιστή. Το υπολογιστικό σύστημα πρέπει να είναι εύκολο για εκείνη την πράξη. Αυτό αποτελεί μια ιδιαίτερη πρόκληση για τα Big Data, λόγω της πολυπλοκότητάς τους. Υπάρχουν συχνά ζωτικής σημασίας υποθέσεις πίσω από τα δεδομένα που καταγράφονται.

## Κατηγορίες εξόρυξης δεδομένων

Οι αλγόριθμοι εξόρυξης δεδομένων προσδιορίζουν μια ποικιλία προβλημάτων που μπορούν να διαμορφωθούν και να λυθούν. Οι λειτουργίες εξόρυξης δεδομένων εμπίπτουν γενικά σε δύο κατηγορίες(.18):

1. Εποπτευόμενη μάθηση
2. Μη εποπτευόμενη μάθηση

Οι έννοιες της εποπτευόμενης και της μη εποπτευόμενης μάθησης προέρχονται από την επιστήμη της μηχανικής μάθησης, η οποία έχει ονομαστεί υποπεριοχή τεχνητής νοημοσύνης. Η τεχνητή νοημοσύνη σημαίνει την εφαρμογή και τη μελέτη συστημάτων που παρουσιάζουν αυτόνομη νοημοσύνη ή συμπεριφορά δική τους. Η μηχανική μάθηση ασχολείται με τεχνικές που επιτρέπουν στις συσκευές να μαθαίνουν από τις δικές τους επιδόσεις και να τροποποιούν τη δική τους λειτουργία. Η εξόρυξη δεδομένων χρησιμοποιεί έννοιες μηχανικής μάθησης στα δεδομένα.

**Μέθοδοι Εποπτευόμενης μάθησης (Supervised Methods)** : Η εποπτευόμενη μάθηση είναι επίσης γνωστή ως κατευθυνόμενη μάθηση. Η διαδικασία εκμάθησης κατευθύνεται από ένα προηγούμενος γνωστό εξαρτώμενο χαρακτηριστικό ή στόχο. Η κατευθυνόμενη εξόρυξη δεδομένων επιχειρεί να εξηγήσει τη συμπεριφορά του στόχου ως συνάρτηση ενός συνόλου ανεξαρτήτων χαρακτηριστικών ή προγνωστικών. Η εποπτευόμενη μάθηση γενικά οδηγεί σε προγνωστικά μοντέλα. Αυτό έρχεται σε αντίθεση με την μη επιτηρούμενη μάθηση όπου ο στόχος είναι η ανίχνευση προτύπων.

Η δημιουργία ενός εποπτευόμενου μοντέλου περιλαμβάνει εκπαίδευση, μια διαδικασία με την οποία το λογισμικό αναλύει πολλές περιπτώσεις όπου η τιμή-στόχος είναι ήδη γνωστή. Στη διαδικασία κατάρτισης, το μοντέλο "μαθαίνει" τη λογική για την πραγματοποίηση της πρόβλεψης. Για παράδειγμα, ένα μοντέλο που επιδιώκει να εντοπίσει τους πελάτες που ενδέχεται να ανταποκριθούν σε μια προσφορά πρέπει να εκπαιδευτεί αναλύοντας τα χαρακτηριστικά πολλών πελατών που είναι γνωστό ότι έχουν ανταποκριθεί ή δεν έχουν ανταποκριθεί σε κάποια προσφορά στο παρελθόν.

**Μέθοδοι Μη εποπτευόμενης μάθησης (Unsupervised Methods)** : Η μη επιτηρούμενη μάθηση δεν κατευθύνεται. Δεν υπάρχει διάκριση μεταξύ εξαρτημένων και ανεξάρτητων χαρακτηριστικών. Δεν υπάρχει προηγούμενος γνωστό αποτέλεσμα για την καθοδήγηση του αλγορίθμου κατά την κατασκευή του μοντέλου. Η μη επιτηρούμενη μάθηση μπορεί να χρησιμοποιηθεί για περιγραφικούς σκοπούς. Μπορεί επίσης να χρησιμοποιηθεί για να κάνει προβλέψεις.

## Τεχνικές εξόρυξης δεδομένων

Διάφοροι αλγόριθμοι και τεχνικές όπως (.19) η ταξινόμηση, η ομαδοποίηση, η παλινδρόμηση, η τεχνητή νοημοσύνη, τα νευρωνικά δίκτυα, οι κανόνες σύνδεσης, τα δέντρα αποφάσεων, ο γενετικός αλγόριθμος, η πλησιέστερη μέθοδος γειτονιάς κλπ. Χρησιμοποιούνται για την αποκάλυψη γνώσεων από βάσεις δεδομένων.

**Ταξινόμηση:** Η ταξινόμηση είναι η πιο συχνά εφαρμοζόμενη τεχνική εξόρυξης δεδομένων, η οποία χρησιμοποιεί μια σειρά από προ-ταξινομημένα παραδείγματα για την ανάπτυξη ενός μοντέλου που μπορεί να ταξινομήσει τον πληθυσμό των αρχείων γενικότερα. Αυτή η προσέγγιση χρησιμοποιεί συχνά αλγόριθμους ταξινόμησης βάσει δέντρων αποφάσεων ή νευρωνικού δικτύου. Η διαδικασία ταξινόμησης δεδομένων περιλαμβάνει μάθηση και ταξινόμηση. Στη μάθηση τα δεδομένα εκπαίδευσης αναλύονται με αλγόριθμο ταξινόμησης. Κατά την ταξινόμηση χρησιμοποιούνται δεδομένα δοκιμών για την εκτίμηση της ακρίβειας των κανόνων ταξινόμησης. Εάν η ακρίβεια είναι αποδεκτή, οι κανόνες μπορούν να εφαρμοστούν στις νέες πλειάδες δεδομένων.

**Ομαδοποίηση:** Η ομαδοποίηση μπορεί να λεχθεί ως αναγνώριση παρόμοιων κατηγοριών αντικειμένων. Με τη χρήση τεχνικών ομαδοποίησης μπορούμε να εντοπίσουμε περαιτέρω πυκνές και αραιές περιοχές στον χώρο αντικειμένων και να ανακαλύψουμε το γενικό μοτίβο διανομής και τις συσχετίσεις μεταξύ των χαρακτηριστικών δεδομένων. Η προσέγγιση κατάταξης μπορεί επίσης να χρησιμοποιηθεί για αποτελεσματικά μέσα διαχωρισμού ομάδων ή κατηγοριών αντικειμένων, αλλά καθίσταται δαπανηρή, ώστε η ομαδοποίηση μπορεί να χρησιμοποιηθεί ως προσέγγιση προεπεξεργασίας για την επιλογή και την ταξινόμηση υποσύνολο χαρακτηριστικών. Για παράδειγμα, για να σχηματίσουν μια ομάδα πελατών με βάση τα πρότυπα αγορών, σε κατηγορίες γονιδίων με παρόμοια λειτουργικότητα.

**Πρόβλεψη:** Η τεχνική παλινδρόμησης μπορεί να προσαρμοστεί για την πρόβλεψη. Η ανάλυση παλινδρόμησης μπορεί να χρησιμοποιηθεί για να μοντελοποιήσει τη σχέση μεταξύ μίας ή περισσότερων ανεξάρτητων μεταβλητών και εξαρτημένων μεταβλητών. Στην εξόρυξη δεδομένων οι ανεξάρτητες μεταβλητές είναι γνωστά γνωρίσματα και οι μεταβλητές απόκρισης είναι αυτό που θέλουμε να προβλέψουμε. Δυστυχώς, πολλά προβλήματα πραγματικού κόσμου δεν είναι απλά πρόβλεψη. Για παράδειγμα, οι όγκοι των πωλήσεων, οι τιμές των μετοχών και τα ποσοστά αποτυχίας του προϊόντος είναι πολύ δύσκολο να προβλεφθούν επειδή μπορεί να εξαρτώνται από πολύπλοκες αλληλεπιδράσεις πολλαπλών μεταβλητών πρόβλεψης. Συνεπώς, ενδέχεται να απαιτούνται πιο σύνθετες τεχνικές (π.χ., λογιστική παλινδρόμηση, δέντρα αποφάσεων ή νευρικά δίκτυα) για την πρόβλεψη μελλοντικών τιμών.

**Κανόνας σύνδεσης:** Η συσχέτιση και η συσχέτιση είναι συνήθως η εύρεση συχνοτήτων που εντοπίζονται συχνά μεταξύ μεγάλων συνόλων δεδομένων. Αυτός ο τύπος εύρεσης βοηθά τις επιχειρήσεις να λάβουν ορισμένες αποφάσεις, όπως το σχεδιασμό του καταλόγου, το cross marketing και την ανάλυση της αγοραστικής συμπεριφοράς των πελατών. Οι αλγόριθμοι του κανόνα σύνδεσης πρέπει να είναι σε θέση να παράγουν κανόνες με τιμές εμπιστοσύνης μικρότερες από μία. Ωστόσο, ο αριθμός των πιθανών Κανόνων Σύνδεσης για ένα δεδομένο σύνολο δεδομένων είναι γενικά πολύ μεγάλος και ένα μεγάλο ποσοστό των κανόνων είναι συνήθως μικρής αξίας (αν υπάρχει).

**Νευρωνικά δίκτυα:** Το νευρικό δίκτυο είναι ένα σύνολο συνδεδεμένων μονάδων εισόδου / εξόδου και κάθε σύνδεση έχει ένα βάρος που υπάρχει μαζί του. Κατά τη διάρκεια της φάσης μάθησης, το δίκτυο μαθαίνει προσαρμόζοντας τα βάρη έτσι ώστε να είναι σε θέση να προβλέψει τις σωστές ετικέτες κλάσης των πλειάδων εισόδου. Τα νευρικά δίκτυα έχουν την αξιοσημείωτη ικανότητα να αντλούν νόημα από περίπλοκα ή ανακριβή δεδομένα και μπορούν να χρησιμοποιηθούν για την εξαγωγή μοτίβων και την ανίχνευση τάσεων που είναι πολύ περίπλοκες για να παρατηρηθούν είτε από τους ανθρώπους είτε από άλλες τεχνικές υπολογιστών. Αυτά είναι κατάλληλα για συνεχείς αποτιμήσεις εισροών και εκροών. Για παράδειγμα, η χειρόγραφη αναδιοργάνωση χαρακτήρων, για την εκπαίδευση ενός υπολογιστή για να προφέρει αγγλικό κείμενο και πολλά επιχειρηματικά προβλήματα σε πραγματικό κόσμο και έχει ήδη εφαρμοστεί με επιτυχία σε πολλές βιομηχανίες. Τα νευρωνικά δίκτυα είναι καλύτερα να εντοπίζουν πρότυπα ή τάσεις στα δεδομένα και είναι κατάλληλα για προβλέψεις ή προβλέψεις.

## Εφαρμογές εξόρυξης δεδομένων

Μερικά παραδείγματα εφαρμογών (.20) :

Πίνακας 1

Εταιρία	Πρόκληση	Αποτέλεσμα
FBTO Ολλανδικές Ασφαλιστικές Εταιρείες	<ul style="list-style-type: none"> <li>• Για να μειώσετε το κόστος του διαφημιστικού ταχυδρομείου.</li> <li>• Αύξηση την αποτελεσματικότητα των εκστρατειών μάρκετινγκ.</li> <li>• Αύξηση των διασταυρούμενων πωλήσεων σε υφιστάμενους πελάτες, χρησιμοποιώντας εισερχόμενα κανάλια, όπως το κέντρο πωλήσεων της εταιρείας και το διαδίκτυο, δοκιμή ενός έτους για την αποτελεσματικότητα της λύσης.</li> </ul>	<ul style="list-style-type: none"> <li>• Εφόσον η ομάδα μάρκετινγκ έχει τη δυνατότητα να προβλέψει την αποτελεσματικότητα των καμπανιών της.</li> <li>• Αυξήθηκε η αποτελεσματικότητα της δημιουργίας, βελτιστοποίησης και εκτέλεσης της καμπάνιας μάρκετινγκ.</li> <li>• Μειώθηκε το κόστος ταχυδρομείου κατά 35%.</li> <li>• Αυξημένα ποσοστά μετατροπής κατά 40%.</li> </ul>
Provident Financial's Home credit Division, Ηνωμένο Βασίλειο	<ul style="list-style-type: none"> <li>• Δεν υπάρχει σύστημα για την ανίχνευση και την πρόληψη της απάτης.</li> </ul>	<ul style="list-style-type: none"> <li>• Μειωμένη συχνότητα και έκταση της απάτης πράκτορα και πελάτη.</li> <li>• Εξοικονόμηση χρημάτων μέσω έγκαιρης ανίχνευσης απάτης.</li> <li>• Εξοικονόμισε τον χρόνο του ερευνητή και αύξησε το ποσοστό δίωξης.</li> </ul>
Εταιρείες παροχής αμοιβαίων χρηματοοικονομικών υπηρεσιών Standard Life	<ul style="list-style-type: none"> <li>• Προσδιορίστε τα βασικά χαρακτηριστικά των πελατών που προσελκύονται από την προσφορά υποθηκών τους.</li> <li>• Σταυροί πωλούν προϊόντα Standard Life Bank στους πελάτες άλλων εταιρειών της Standard Life.</li> <li>• Αναπτύξτε ένα μοντέλο επαναγοράς που θα μπορούσε να αναπτυχθεί στην ιστοσελίδα του ομίλου για να εξετάσει την αποδοτικότητα της επιχείρησης υποθηκών που γίνεται δεκτή από την Standard Life Bank.</li> </ul>	<ul style="list-style-type: none"> <li>• Ενσωματωμένο πρότυπο προσφοράς για την προσφορά υποθηκών της Standard Life Bank που προσδιορίζει τους βασικούς τύπους πελατών που μπορούν να εφαρμοστούν σε ολόκληρη την ομάδα προοπτικών ομάδας.</li> <li>• Ανακάλυψε τους βασικούς οδηγούς για την αγορά ενός προϊόντος remortgage.</li> <li>• Επιτεύχθηκε, με το μοντέλο, εννέα φορές μεγαλύτερη ανταπόκριση από αυτή που επιτυγχάνεται από την ομάδα ελέγχου.</li> <li>• Εγγυημένα έσοδα από υποθήκες υποψηφιότητας ύψους £ 33 εκατ. (Περίπου 47 εκατ.).</li> </ul>
Shenandoah Ασφαλιστική εταιρεία ζωής Ηνωμένες Πολιτείες Αμερικής.	<ul style="list-style-type: none"> <li>• Η διαδικασία έγκρισης πολιτικής ήταν βασισμένη σε χαρτί και δυσκίνητη.</li> <li>• Η δρομολόγηση αυτών των αντιγράφων σε διάφορα τμήματα, υπήρχαν καθυστερήσεις στην έγκριση.</li> </ul>	<ul style="list-style-type: none"> <li>• Εξουσιοδοτημένη διαχείριση με τρέχουσες πληροφορίες σχετικά με εκκρεμείς πολιτικές.</li> <li>• Μείωση του χρόνου που απαιτείται για την έκδοση ορισμένων πολιτικών κατά 20%.</li> <li>• Βελτιωμένες διαδικασίες αναδοχής και ελέγχου των επιδόσεων των εργαζομένων.</li> </ul>
Soft map Company Ltd., Τόκυο	<ul style="list-style-type: none"> <li>• Οι πελάτες δυσκολεύονταν να κάνουν αποφάσεις αγοράς υλικού και λογισμικού, οι οποίες εμπόδιζαν τις ηλεκτρονικές πωλήσεις.</li> </ul>	<ul style="list-style-type: none"> <li>• Οι προβολές σελίδας αυξήθηκαν κατά 67% το μήνα μετά την κυκλοφορία του μηχανισμού συστάσεων.</li> <li>• Τα κέρδη τριπλασιάστηκαν το 2001, καθώς οι πωλήσεις αυξήθηκαν κατά 18% σε σχέση με την αντίστοιχη περίοδο του προηγούμενου έτους.</li> </ul>

### 1.3 Κατασκευή Προφίλ (Profiling)

#### Ορισμός

Ο **όρος προφίλ** κυριολεκτικά είναι η πλάγια όψη του προσώπου, μεταφορικά αποτυπώνει το χαρακτήρα, το στυλ, τον τρόπο συμπεριφοράς. Η επιστήμη ωστόσο της κατασκευής προφίλ (profiling) αφορά την συγκέντρωση, ανάλυση και παρουσίαση (των αποτελεσμάτων ή των ιδίων) των χαρακτηριστικών (βιολογικών, ψυχολογικών, κοινωνικών, συμπεριφορικών κ.α.) ενός ατόμου ή ομάδας ατόμων π.χ. καταναλωτικό προφίλ, φυλετικό προφίλ κ.α. (.21)

Η κατασκευή “προφίλ” είναι τόσο παλιά όσο και η ίδια η ζωή. Είναι ένας μηχανισμός επιβίωσης των έμβιων όντων στο περιβάλλον τους, ένα ζωτικό σημάδι της ζωής. Στην ουσία είναι η αλληλεπίδραση μεταξύ παρακολούθησης και προσαρμογής. Για να επιβιώσει και να απολαύσει τη ζωή, οποιοσδήποτε οργανισμός πρέπει να προσαρμόζεται συνεχώς στις αλλαγές στο περιβάλλον του, ενώ παράλληλα μπορεί να προσαρμόσει το περιβάλλον του στις προτιμήσεις του. Αυτό συνιστά έναν επαναλαμβανόμενο προσδιορισμό των κινδύνων και των ευκαιριών από έναν οργανισμό στο περιβάλλον του (.22).

Μπορούμε να συμπεράνουμε ότι η κατασκευή προφίλ δεν είναι τυπικά ανθρώπινο χαρακτηριστικό αν και έχουμε αναπτύξει δικό μας είδος δημιουργίας και εφαρμογής προφίλ. Αυτό που είναι ιδιαίτερο για τον άνθρωπο είναι η ικανότητά μας να σκεφτόμαστε τα προφίλ που δημιουργούν οι εγκέφαλοί μας, η ικανότητά μας για συνειδητό προβληματισμό σχετικά με τα προφίλ που δημιουργήσαμε ασυνείδητα, διότι αυτό μας δίνει την ελευθερία να τα συζητάμε, να τα απορρίπτουμε, να τα ενισχύουμε και να τα εφαρμόζουμε σκόπιμα. Είναι προϋπόθεση για να χαρακτηρίζονται οι ενέργειές μας ως αποτέλεσμα της ελευθερίας βούλησης: να μπορούμε να γνωρίσουμε τα πρότυπα που ρυθμίζουν τις πράξεις μας και να τα αναθεωρήσουμε για να αλλάξουμε τις συνήθειες μας(.23).

**Στην επιστήμη της πληροφορικής**, η κατασκευή προφίλ (profiling) αναφέρεται στη διαδικασία δημιουργίας και εφαρμογής των προφίλ των χρηστών που παράγονται από την ανάλυση των ηλεκτρονικών δεδομένων. Αυτό που αλλάζει με τις νέες τεχνολογίες είναι το μέγεθος και το εύρος των δυνατοτήτων σε σχέση με αυτά που είχαμε μέχρι τώρα(.21). Τα Big Data προσφέρουν ασύγκριτα περισσότερα δεδομένα (volume), μεγαλύτερη ποικιλία (variety) και ταχύτητα επεξεργασίας (velocity). Ενώ οι τεχνολογίες του Data mining προσφέρουν την ανακάλυψη μοτίβων ή συσχετισμών στις μεγάλες αυτές ποσότητες δεδομένων.

Η κατασκευή προφίλ **μπορεί να χρησιμοποιηθεί**, για να ταξινομήσει, να περιγράψει και να αναλύσει “τι συνέβη”, κάτι που δεν είναι ιδιαίτερα νέο ή προβληματικό. Αυτός ο τύπος ανάκτησης των αποθηκευμένων πληροφοριών ονομάζεται ερώτημα (query). Σε τέτοιες περιπτώσεις η δημιουργία προφίλ επιτρέπει μια δομή του ήδη γνωστού. Από την άλλη χρησιμοποιείται για τη συγκέντρωση δεδομένων με τέτοιο τρόπο ώστε να παραχθούν πληροφορίες και να προταθούν προβλέψεις ή

προσδοκίες. Αυτή η δραστηριότητα (δημιουργίας προφίλ) δημιουργεί έτσι μια συγκεκριμένη γνώση, μέσω της διαδικασίας που είναι γνωστή ως αποκάλυψη γνώσεων σε βάσεις δεδομένων (KDD). Οι γνώσεις που παράγονται είναι μη αντιπροσωπευτικές: δεν αντιπροσωπεύουν μια τρέχουσα κατάσταση υποθέσεων.

Τα προφίλ είναι πρότυπα που προκύπτουν από μια πιθανοτική επεξεργασία δεδομένων. Δεν περιγράφουν την πραγματικότητα, αλλά ανιχνεύονται σε βάσεις δεδομένων με τη διαδικασία της συνάθροισης, της εξόρυξης και τον καθαρισμό δεδομένων. Ωστόσο, ακόμη και αν η διαδικασία δημιουργίας προφίλ δείχνει ότι ένα πρότυπο συμβαίνει κάθε φορά που πληρούνται ορισμένες προϋποθέσεις, δεν μπορεί κανείς να είναι 100% σίγουρος ότι θα συμβεί σήμερα και αύριο επίσης(.21).

## Profiling και Data Mining

Όπως είδαμε η κατασκευή προφίλ στη σύγχρονη πληροφορική είναι άρρηκτα συνδεδεμένη με το Data mining. Όταν μελετάμε τη χρήση της εξόρυξης δεδομένων για σκοπούς δημιουργίας προφίλ, μπορούμε να διακρίνουμε δύο διαφορετικές προσεγγίσεις εξόρυξης δεδομένων, την "περιγραφική εξόρυξη δεδομένων" και την "προγνωστική εξόρυξη δεδομένων".

**Περιγραφική εξόρυξη δεδομένων:** Ο στόχος της περιγραφικής εξόρυξης δεδομένων είναι να ανακαλυφθούν *άγνωστες σχέσεις* μεταξύ διαφορετικών αντικειμένων δεδομένων σε μια βάση δεδομένων. Οι αλγόριθμοι εξόρυξης περιγραφικών δεδομένων προσπαθούν να ανακαλύψουν τη γνώση σχετικά με ένα συγκεκριμένο τομέα προσδιορίζοντας κοινά σημεία και χαρακτηριστικά μεταξύ διαφορετικών αντικειμένων. Ανακαλύπτοντας συσχετίσεις μεταξύ αντικειμένων δεδομένων σε ένα σύνολο δεδομένων που αντιπροσωπεύει ένα συγκεκριμένο τομέα, μπορούμε να τον καταλάβουμε.

Η περιγραφική εξόρυξη δεδομένων είναι ενδιαφέρουσα επειδή η καλύτερη γνώση ενός συγκεκριμένου τομέα επιτρέπει τον καλύτερο προγραμματισμό και τη διάθεση των πόρων σε αυτόν τον τομέα. Για παράδειγμα, η περιγραφική εξαγωγή δεδομένων σχετικά με τα έντυπα φορολογικής δήλωσης μπορεί να διαπιστώσει μια φυσική κατανομή των ατόμων σε περιορισμένο αριθμό ομάδων με βάση την οικονομική δραστηριότητα, την οικονομική κατάσταση και την κατοχή τους. Επίσης, η περιγραφική εξόρυξη δεδομένων θα μπορούσε να εκθέσει τις σχέσεις μεταξύ των διαφόρων πεδίων των εντύπων.

Είναι σημαντικό να παρατηρήσουμε εδώ ότι στην περιγραφική εξόρυξη δεδομένων δεν έχει δοθεί "στόχος". Επομένως, η περιγραφική εξόρυξη δεδομένων συχνά αποκαλείται επίσης "εξόρυξη δεδομένων". Επιπλέον, η περιγραφική εξόρυξη δεδομένων περιγράφει τις σχέσεις και δεν τις εξηγεί. Η ύπαρξη μιας συσχέτισης σε ένα σύνολο δεδομένων δεν σημαίνει απαραίτητα ότι αυτή η σχέση θα εμφανίζεται πάντα στον πραγματικό κόσμο, ούτε εξηγεί γιατί υπάρχει ο συσχετισμός. Με άλλα λόγια,

είναι ζωτικής σημασίας να μην συγχέεται η συσχέτιση με την αιτιώδη συνάφεια όταν πρόκειται για περιγραφική εξόρυξη δεδομένων.

**Προγνωστική** εξόρυξη δεδομένων: Όπως υποδηλώνει το όνομα, ο στόχος της εξόρυξης δεδομένων είναι να κάνει προβλέψεις πάνω σε γεγονότα βάση μοτίβων που προσδιορίστηκαν χρησιμοποιώντας γνωστές πληροφορίες. Όταν πρόκειται για τη δημιουργία προφίλ, αυτό σημαίνει ότι πληροφορίες σχετικά με ένα άτομο εξορυσσονται προκειμένου να προσδιοριστεί αν ταιριάζει με το προηγούμενο προφίλ. Επομένως, τα δεδομένα εκπαίδευσης στην προγνωστική εξόρυξη συνίστανται συνήθως από μια συλλογή σχολιασμένων αντικειμένων ή ατόμων. Π.χ., τα δημογραφικά στοιχεία ενός συνόλου ατόμων μαζί με το σχόλιο ότι είναι γνωστοί τρομοκράτες. Ως εκ τούτου, η προγνωστική αναζήτηση δεδομένων ονομάζεται συχνά "εξόρυξη δεδομένων Με επίβλεψη".

Μια συχνά χρησιμοποιούμενη μέθοδος στην προγνωστική εξόρυξη δεδομένων είναι η ταξινόμηση. Χρησιμοποιώντας αλγόριθμους εξόρυξης δεδομένων που ονομάζουμε ταξινομητές, μπορούμε να διαπιστώσουμε αν ένα νέο αντικείμενο ταιριάζει με μια προηγουμένως καθιερωμένη κλάση. Οι κλάσεις βασίζονται σε πεδία εισαγωγής που περιέχουν διαφορετικά χαρακτηριστικά που σχετίζονται με την κλάση. Όταν ένα αντικείμενο μοιράζεται χαρακτηριστικά με τα αντικείμενα δεδομένων που ανήκουν στην κλάση, είναι πιθανό να ανήκει και σε αυτήν την κλάση. Για παράδειγμα, η ετικέτα κλάσης "καναρίνι" θα αποτελείται από χαρακτηριστικά όπως {κίτρινο, ράμφος, φτερά, ουρά}. Όταν ένα αντικείμενο έχει ένα ορισμένο ποσό αυτών των χαρακτηριστικών, γίνεται όλο και πιο πιθανό ότι είναι πράγματι ένα καναρίνι.

Είναι σημαντικό να σημειωθεί ότι η ταξινόμηση καθορίζει μόνο την πιθανότητα να ανήκει σε κάποια κατηγορία. Όταν η ταξινόμηση βασίζεται σε λίγους καθοριστικούς παράγοντες (π.χ. κάτι είναι είτε 1 είτε δεν είναι 1), η ταξινόμηση μπορεί να είναι αρκετά ακριβής. Ωστόσο, όταν ένα συγκεκριμένο γεγονός ή συμβάν εξαρτάται από πολλούς παράγοντες, η ακριβής ταξινόμηση γίνεται πιο δύσκολη, αν όχι αδύνατη. Επίσης, πολλές φορές οι παράγοντες που καθορίζουν την κλάση του αντικειμένου δεν υπάρχουν στο σύνολο δεδομένων. Για παράδειγμα, αν κάποιος θα μεγαλώσει ή όχι ως εγκληματίας εξαρτάται από πολλούς παράγοντες, η ακριβής ταξινόμηση (δηλαδή, ποινική ή μη ποινική) είναι αδύνατη. Ως εκ τούτου, το αποτέλεσμα ενός ταξινομητή σε ένα νέο αντικείμενο θα είναι πάντα πιθανολογικού χαρακτήρα.

### Μηχανές Κατασκευής Προφίλ:

Στην πληροφορική πλέον μιλάμε για κάτι νέο: την **αυτοματοποιημένη κατασκευή προφίλ**. Η αυτοματοποιημένη κατασκευή προφίλ χρησιμοποιεί τη διαδικασία της ανακάλυψης γνώσης σε βάσεις δεδομένων - KDD και μάθηση μηχανής (machine learning).

Η αυτό αυτοματοποίηση αυτή είναι νέα για τρεις λόγους. **Πρώτον**, δεν μιλάμε για τον προσδιορισμό προφίλ από οργανισμούς, αλλά από μηχανές. Οι μηχανές αυτές είναι εκπαιδευμένα



προγράμματα λογισμικού για την ανάκτηση απροσδόκητων συσχετισμών σε μάζες δεδομένων.

**Δεύτερον**, δεν μιλάμε για την αναζήτηση ερωτημάτων σε βάσεις δεδομένων, για να συνοψίσουμε τα χαρακτηριστικά προκαθορισμένων κατηγοριών, αλλά για να ανακαλύψουμε γνώσεις που δεν γνωρίζουμε και είναι «κρυμμένες» στα δεδομένα. **Τρίτον**, σε αυτό το σημείο δεν μπορούμε να σκεφτούμε τον τρόπο με τον οποίο τα προφίλ επηρεάζουν τις ενέργειές μας, διότι ουσιαστικά δεν έχουμε πρόσβαση στον τρόπο παραγωγής και χρήσης τους. Αυτή η τελευταία διαφορά υποδηλώνει ότι η δημιουργία προφίλ εμποδίζει την ελευθερία μας να ενεργούμε αυτόνομα.

#### 1.4 Εφαρμογές και Συνεισφορά των Big Data

Τα Big Data είναι μια μεγάλη βιομηχανία. Μια έρευνα στο Ινστιτούτο Τεχνολογίας της Μασαχουσέτης δείχνει ότι οι εταιρείες που βασίζουν τη λήψης αποφάσεων στην επιστήμη των δεδομένων, απολαμβάνουν μια αύξηση της παραγωγικότητας στο 5-6%. Οι εταιρείες που χρησιμοποιούν τα δεδομένα πιο αποτελεσματικά ξεχωρίζουν από τις υπόλοιπες. Μια έκθεση του Ινστιτούτου McKinsey Global Institute (MGI) καταδεικνύει τη μετασχηματιστική επίδραση των big data σε ολόκληρους τομείς που κυμαίνονται από την υγειονομική περίθαλψη στο λιανικό εμπόριο έως και την κατασκευή πολιτικών εκστρατειών (.25).

Ακριβώς όπως βοηθά τις επιχειρήσεις να αυξήσουν την παραγωγικότητα τους, τα big data επιτρέπουν στις κυβερνήσεις να βελτιώσουν τη διαχείριση του δημόσιου τομέα και βοηθά παγκόσμιους οργανισμούς στην ανάλυση των πληροφοριών για την εκπόνηση στρατηγικού σχεδιασμού. Η ζήτηση για μεγάλα στοιχεία επιταχύνεται.

Σε αυτό το κεφάλαιο, παρουσιάζουμε μερικά τυπικά παραδείγματα με τα οφέλη που τα big data έχουν επιφέρει.

##### Υγεία:

Ο Δρ Russ Altman, καθηγητής της ιατρικής και βιοτεχνολογίας στο Πανεπιστήμιο του Στάνφορντ, και οι συνεργάτες του έκαναν μια πρωτοποριακή με τη χρήση των Big Data. Βρήκαν ότι όταν λαμβάνοντας μαζί, Paxil, το πολύ γνωστό αντικαταθλιπτικό που παρέχονται σε εκατομμύρια Αμερικανών, και το Pravachol, ένα πολύ δημοφιλές φάρμακο μείωσης της χοληστερόλης, έχουν φοβερές παρενέργειες. Αυξάνουν τα επίπεδα γλυκόζης στο αίμα των ασθενών σε επίπεδα διαβητικών. Κάθε φάρμακο που λαμβάνεται μόνο του δεν έχει τις διαβητική παρενέργειες, ως εκ τούτου, η Υπηρεσία Τροφίμων και Φαρμάκων (Food and Drug Administration FDA) ενέκρινε τα φάρμακα για χρήση. Σίγουρα ο FDA δεν μπορούσε να δοκιμάσει κάθε φάρμακο για κάθε πιθανή αλληλεπίδραση.

Ο Altman και η ομάδα του έκαναν την ανακάλυψή τους κατά την επιδίωξη στατιστικών αναλύσεων και τεχνικές εξόρυξης δεδομένων για την αναγνώριση προτύπων σε μεγάλα σύνολα

δεδομένων. Σχεδόν 15 εκατομμύρια Αμερικανοί χρησιμοποιούν είτε Paxil ή Pravachol. Υπολογίζεται ότι ένα εκατομμύριο χρησιμοποιούσε και τα δύο. Για αυτούς τους χρήστες, το έργο του Altman και των συνεργατών του ήταν δυνητικά σωτήριο. (.26)

Τα ευρήματα των Altman και της ομάδα του δεν είναι η μόνη σημαντική ανακάλυψη βασίζεται σε ανάλυση των big data στον τομέα της υγείας. Η ανακάλυψη των ανεπιθύμητων παρενεργειών του φαρμάκου Vioxx, το οποίο οδήγησε στην απόσυρσή του από την αγορά, κατέστη δυνατή με την ανάλυση των κλινικών και οικονομικών δεδομένων που συλλέχθηκαν από το Kaiser Permanente, την κοινοπραξία διαχειρίζεται περίθαλψης με βάση την Καλιφόρνια (.27). Αν η Kaiser Permanente δεν άθροιζε κλινικά και οικονομικά δεδομένα, οι ερευνητές μπορεί να μην ήταν σε θέση να αποδώσουν 27.000 καρδιακούς θανάτους από ανακοπή που συνέβησαν μεταξύ 1999 και 2003, με τη χρήση του φαρμάκου.

Ένα άλλο συχνά αναφερόμενο παράδειγμα είναι το Google Flu Trends, το οποίο προβλέπει και εντοπίζει εστίες γρίπης και λοιπών λοιμώξεων της κάνοντας χρήση πληροφοριών – συσχέτιση ερωτημάτων αναζήτησης – τα οποία αρχικά δεν είχαν συλλεχθεί για αυτή την καινοτόμο εφαρμογή. Φυσικά, η έγκαιρη ανίχνευση της δραστηριότητας της νόσου, όταν ακολουθείται από ταχεία αντίδραση, μπορεί να μειώσει τον αντίκτυπο της εποχικής και της πανδημικής γρίπης. (.28)

#### Mobile:

Ο τομέας της υγείας δεν είναι ο μοναδικός χώρος για την πρωτοποριακή χρήση των big data. Μια ομάδα επιστημόνων που εργάζονται σε ένα συνεργατικό έργο στο MIT, το Harvard και αντίστοιχα ερευνητικά πανεπιστήμια αναλύουν σε αυτή τη φάση δεδομένα επικοινωνιών κινητής τηλεφωνίας για να κατανοήσουν καλύτερα τις ανάγκες του 1 δισεκατομμυρίου ανθρώπων που ζουν σε οικισμούς ή φτωχογειτονιές στις αναπτυσσόμενες χώρες (.29). Ερευνούν τρόπους για την πρόβλεψη ελλείψεων τροφίμων χρησιμοποιώντας μεταβλητές όπως οι τιμές της αγοράς, η ξηρασία, οι μεταναστεύσεις, η προηγούμενη περιφερειακή παραγωγή και εποχιακές διακυμάνσεις(.30). Για να ποσοτικοποιήσουν τα κύματα εγκληματικότητας, παρακολουθώντας το χρόνο, τον τόπο και τη φύση των εγκληματικών δραστηριοτήτων σε τοποθεσίες σε όλη την πόλη(.31). Και για να αποφασίσουν ποια παρέμβαση λειτουργεί καλύτερα για τη βελτίωση των μαθησιακών αποτελεσμάτων στην ανάπτυξη των σχολείων της χώρας(.32).

#### Έξυπνο δίκτυο:

Ένα άλλο παράδειγμα είναι το «έξυπνο δίκτυο», το οποίο αναφέρεται στον εκσυγχρονισμό του ισχύοντος ηλεκτρικού δικτύου για να εισαγάγει μια αμφίδρομη ροή πληροφοριών της ηλεκτρικής ενέργειας(.33). Το έξυπνο δίκτυο έχει σχεδιαστεί για να επιτρέπει στους φορείς παροχής υπηρεσιών

ηλεκτρικής ενέργειας, στους χρήστες(καταναλωτές ηλ. ενεργείας), και τρίτους φορείς να παρακολουθούν και να ελέγχουν τη χρήση ηλεκτρικής ενέργειας. Οι βοηθητικές εφαρμογές βλέπουν το έξυπνο δίκτυο ως έναν τρόπο για να εντοπίσετε με ακρίβεια διακοπές ρεύματος ή άλλα προβλήματα, συμπεριλαμβανομένων των κυβερνο-επιθέσεων ή φυσικών καταστροφών, έτσι ώστε τεχνικοί να αποστέλλονται για την επίλυση των προβλημάτων. Οι καταναλωτές επωφελούνται από περισσότερες επιλογές για το πώς, πότε και πόση ηλεκτρική ενέργεια χρησιμοποιούν(.34). Οι φίλο-περιβαλλοντικοί φορείς βλέπουν το έξυπνο δίκτυο ως ένα κλειδί για την παροχή καλύτερης ποιότητας ισχύος, και την πιο αποτελεσματική κίνηση προς τις ανανεώσιμες πηγές ενέργειας.

### Διαχείριση της κυκλοφορίας:

Μια παρακείμενη περιοχή για τα δεδομένα με γνώμονα την περιβαλλοντική καινοτομία είναι η διαχείριση και ο έλεγχος της κυκλοφορίας. Οι κυβερνήσεις ανά τον κόσμο εγκαθιστούν ηλεκτρονικά συστήματα τιμολόγησης των διοδίων, τα οποία καθορίζουν χρεώσεις και πληρωμές με βάση την κινητικότητα και την κυκλοφοριακή συμφόρηση(.35). Οι χρήστες πληρώνουν ανάλογα με τη χρήση των οχημάτων τους και των δρόμων. Οι πολεοδόμοι επωφελούνται από την ανάλυση των δεδομένων τοποθεσία για αποφάσεις που αφορούν την κατασκευή δρόμων και τις μαζικές διελεύσεις, τον μετριασμό της κυκλοφοριακής συμφόρησης, καθώς και το σχεδιασμό για την ανάπτυξη υψηλής πυκνότητας ροής οχημάτων(.36). Οι αποφάσεις αυτές μπορούν όχι μόνο να μειώσουν τη συμφόρηση, αλλά και να μετριάσουν τις εκπομπές ρύπων(.37). Παράλληλα, επωφελούνται οι οδηγοί ατομικά από έξυπνες δρομολόγησης, βασιζόμενες σε πληροφορίες για την κυκλοφορία σε πραγματικό χρόνο, συμπεριλαμβανομένων των εκθέσεων ατυχημάτων και πληροφορίες σχετικά με προγραμματισμένα οδικά έργα και περιοχές υψηλής κυκλοφοριακής συμφόρησης(.38).

### Λιανεμπόριο:

Τα big data μετασχηματίζουν και την λιανική αγορά. Το σύστημα απογραφής διαχείρισης της γνωστής αλυσίδας λιανικής πώλησης Wal-Mart, που ονομάζεται Retail Link, ήταν το πρώτο το οποίο πρωτοστάτησε στην εποχή των big data, δίνοντας στους προμηθευτές τη δυνατότητα να βλέπουν το ακριβή αριθμό των προϊόντων τους ,σε κάθε ράφι , σε κάθε κατάσταση, για κάθε συγκεκριμένη χρονική στιγμή. Πολλοί από εμάς χρησιμοποιούν «Οι πελάτες που αγόρασαν αυτό επίσης αγόρασαν» χαρακτηριστικό της Amazon (και πλέον σήμερα χρησιμοποιείτε από πολλά e-shop), προτρέποντας τους χρήστες να εξετάσει την αγορά επιπλέον προϊόντων που επιλέγονται από ένα συνεργατικό εργαλείο φιλτραρίσματος(.39). Το βασικό επιχειρηματικό μοντέλο του Διαδικτύου χρηματοδοτεί προϊόντα και υπηρεσίες με στοχευμένες διαφημίσεις η αξία των οποίων συσχετίζεται άμεσα με την ποσότητα των πληροφοριών που συλλέγονται από τους χρήστες(.40). Οι επιχειρήσεις δεν ενδιαφέρονται τόσο πολύ για την ταυτότητα του κάθε χρήστη, αλλά μάλλον για τα χαρακτηριστικά

του προφίλ τους, το οποίο καθορίζει τη φύση των διαφημίσεων που του εμφανίζονται. Τα Analytics μπορεί επίσης να χρησιμοποιηθούν σε offline περιβάλλον για τη μελέτη της συμπεριφοράς των πελατών μέσα στο κατάστημα για τη βελτιώσει της διάταξης των καταστημάτων, το φάσμα προϊόντων, και την τοποθέτηση στο ράφι. Ο McKinsey αναφέρει σαφώς ότι «πρόσφατες καινοτομίες έχουν επιτρέψει στους λιανοπωλητές να παρακολουθείτε τα αγοραστικά πρότυπα των πελατών (π.χ., μονοπάτι και το χρόνο που δαπανάται σε διάφορα μέρη του ένα κατάστημα), αντλώντας δεδομένα θέσης πραγματικού χρόνου από το smartphone εφαρμογές (π.χ., Shopkick), αναμεταδότες από το καλάθι αγορών, ή παθητικά την παρακολούθηση της θέσης των κινητών τηλεφώνων μέσα σε ένα περιβάλλον λιανικής πώλησης.» (.41). Ολοένα και περισσότερο, οι οργανισμοί επιδιώκουν να συνδέσουν την ηλεκτρονική δραστηριότητα με τη συμπεριφορά εκτός σύνδεσης, τόσο για να αξιολογήσουν την αποτελεσματικότητα των διαφημιστικών καμπανιών στο διαδίκτυο, όπως κρίνονται με τη μετατροπή σε αγορές στο παρελθόν, και να προωθήσουν τους πελάτες τους με διαφημίσεις στο διαδίκτυο.

### Πληρωμές

Ένα ακόμα σημαντικός χώρος για την αξιοποίηση των big data είναι η ανίχνευσης της απάτης στον τομέα των πληρωμών με κάρτα. Με το ηλεκτρονικό εμπόριο να κατακτά ένα ολοένα και μεγαλύτερο μερίδιο της αγοράς λιανικής πώλησης, οι έμποροι, οι οποίοι φέρουν την τελική ευθύνη για τη δόλια και κακόβουλη χρήση πληρωμών με κάρτα (.42), πρέπει να εφαρμόσουν ισχυρούς μηχανισμούς για τον εντοπισμό ύποπτων συναλλαγών που συχνά πραγματοποιούνται “πρωτάρηδες” πελάτες. Με το σκοπό αυτό, οι εταιρείες ανέπτυξαν λύσεις για τους εμπόρους συστήματα πρόγνωσης απάτης «Συναλλαγή Μη Παρούσας Κάρτας», προκειμένου να μετρηθεί σε πραγματικό χρόνο η πιθανότητα μιας συναλλαγής να είναι ύποπτη ως απατηλή. Για να γίνει αυτό, οι υπηρεσίες αναλύουν το ιστορικό του αγοραστή και να παρέχει αξιολογήσεις, σαν μια συνοπτική λίστα των αναφορών, αλλά με τη μορφή μιας ενιαίας βαθμολογίας.

### Διαδίκτυο

Τέλος, σκεφτείτε τα ογκώδη δεδομένα των σιλό που διατηρούν οι online giants της τεχνολογίας, το Google, το Facebook, η Microsoft, η Apple και το Amazon. Οι εταιρείες αυτές συγκεντρώνουν προηγουμένως αδιανόητες ποσότητες προσωπικών δεδομένων. Το Facebook, για παράδειγμα, έχει πάνω από 900 εκατομμύρια χρήστες που ανεβάζουν πάνω από 250 εκατομμύρια φωτογραφίες και κάνουν κλικ στο κουμπί "όπως" περισσότερο από 2,5 δισεκατομμύρια φορές την ημέρα (.43). Η Google προσφέρει μια πληθώρα προϊόντων και υπηρεσιών με μεγάλη ένταση δεδομένων, όπως η πανταχού παρούσα μηχανή αναζήτησης, το λειτουργικό σύστημα Android, το πρόγραμμα περιήγησης Chrome, το Gmail, το Youtube, οι Χάρτες Google, το Google Plus, το Google Analytics, το Google Apps και πολλά άλλα(.44). Επιπλέον, η Google κατέχει τη μεγαλύτερη εταιρεία

προβολής διαφημίσεων σε απευθείας σύνδεση DoubleClick, την οποία αγόρασε το 2007, σε μεγάλο βαθμό για τη συντριβή των υπερασπιστών ιδιωτικού απορρήτου(.45), καθώς και για την AdMob, την κορυφαία εταιρεία διαφήμισης για κινητά. Ως αποτέλεσμα, η Google έχει πλέον παρουσία σε πάνω από το 70% των ιστότοπων τρίτων(.46). Η Amazon και το Yahoo αναζητούν νέους τρόπους για να εκμεταλλευτούν και να αποκομίσουν έσοδα από το θησαυρό των δεδομένων των πελατών τους(.47). Η Apple και η Microsoft καταστήσουν λειτουργικά συστήματα καθώς και προγράμματα περιήγησης, τα οποία αποτελούν σημαντικό σημείο αναφοράς για τη συλλογή πληροφοριών χρηστών και σε απευθείας σύνδεση.

# Κεφάλαιο 2: Big Data στον Τραπεζικό Τομέα

## 2.1 Τράπεζες και Information Technology

Τράπεζες και η επιστήμη των Πληροφοριακών Συστημάτων (Information Technology - IT):

Η επιστήμη των Πληροφοριακών Συστημάτων είναι άρρηκτα συνδεδεμένη με την σύγχρονη ιστορία του τραπεζικού συστήματος. Η αλλαγή που προκάλεσε η καινοτομία της τεχνολογίας της Πληροφορικής και των Τηλεπικοινωνιών μείωσε τις διαφορές τιμών στις γεωγραφικά απομακρυσμένες αγορές. Κατόπιν παρουσίασε την εμφάνιση συγκεκριμένων εφαρμογών πληροφορικής που τροποποίησαν τις σχέσεις μεταξύ των απομονωμένων τμημάτων τραπεζικών οργανισμών. Με την πάροδο του χρόνου, οι εφαρμογές πληροφορικής παρήγαγαν τροποποιήσεις σε ολόκληρη την οργανωτική δομή όσον αφορά τις εσωτερικές λειτουργίες και τις σχέσεις μεταξύ τραπεζών και πελατών(.48).

Μια σύντομη ιστορική αναδρομή των τραπεζών και της επιστήμης των Π.Σ.

### Περίοδος 1864 -1945:

- Η εισαγωγή του τηλέγραφου το 1846 μείωσε τις διαφορές των τιμών μετοχών μεταξύ των αγορών.
- Σύνδεση και συντονισμός κεντρικών γραφείων με υποκαταστήματα σε διαφορετικές τοποθεσίες.
- Μείωση των διαφορών τιμών μεταξύ των αγορών.
- Charga-Plate (1928) και Air Travel Card (1930), προκάτοχοι της χρεωστικής κάρτας.

### Περίοδος 1945 – 65:

- Μετατόπιση βάρους από υποκαταστήματα σε τραπεζικές σχέσεις.
- Αυτοματοποιημένες τραπεζικές δηλώσεις.
- Εγγυητήριες κάρτες επιταγών (προκάτοχος χρεωστικής/πιστωτικής κάρτας).
- Μείωση του κόστους των δραστηριοτήτων έντασης εργασίας (π.χ. σύστημα εκκαθάρισης).

### Περίοδος 1965-80:

- Αύξηση των διασυννοριακών πληρωμών.
- Εισαγωγή του ATM στο ευρύ κοινό.
- Αυτοματοποίηση λογιστικής υποκαταστημάτων.
- Ο έλεγχος πραγματικού χρόνου αρχίζει

### Περίοδος 1980-95:

- Προμήθεια προϊόντων μη πληρωμής όπως ασφάλειες, στεγαστικά δάνεια και συντάξεις.
- Ανάπτυξη εναλλακτικών καναλιών διανομής, όπως η τηλεφωνική τραπεζική.
- Εισαγωγή της ηλεκτρονικής τραπεζικής 1980 στη Νέα Υόρκη (ΗΠΑ). Τέσσερις μεγάλες τράπεζες (Citibank, Chase Bank, Chemical Bank και Manufactures Hanover) προσέφεραν τραπεζικές υπηρεσίες στο σπίτι, μέσω τερματικού.

### **Περίοδος 1995–2001:**

- Τηλεφωνική τραπεζική (ΗΠΑ, Ηνωμένο Βασίλειο). Λιανικοί διαμεσολαβητές χωρίς υποκαταστήματα. Πολυκαναλικό σύστημα διανομής για τις τράπεζες βάσει ολοκληρωμένου λογαριασμού πελατών και συστήματος πληροφοριών.
- Κάρτες Mondex (UK) Χρεωστική κάρτα με μικρο-τσιπ. Διευκόλυνση των συναλλαγών λιανικής μικρής αξίας με δυνατότητα υποκατάστασης των χαρτονομισμάτων και κερμάτων που εκδίδει η κεντρική τράπεζα. Επισημοποιήστε τους τρόπους συλλογής ευρέων.
- DigiCash (Ολλανδία ) Ηλεκτρονικό μέσο ανταλλαγής και λογιστικής μονάδας. Συστήματα πληρωμών και προϊόντα που εξαρτώνται αποκλειστικά από επικοινωνίες υψηλής ταχύτητας που πραγματοποιούνται μέσω υπολογιστών.
- Διαδραστική τραπεζική στο Διαδίκτυο. Το 1995, η Wells Fargo ήταν η πρώτη τράπεζα των ΗΠΑ που προσέθεσε υπηρεσίες λογαριασμού στον ιστότοπό της, ενώ άλλες τράπεζες ακολούθησαν σύντομα. Την ίδια χρονιά, η Προεδρική τράπεζα των ΗΠΑ έγινε η πρώτη τράπεζα των ΗΠΑ για να ανοίξει τραπεζικούς λογαριασμούς μέσω του Διαδικτύου. Η τεχνολογία ανοίγει νέες ευκαιρίες για την ανάπτυξη των τραπεζών και προσφέρει στους διαχειριστές των τραπεζών δυνατότητες να επιτύχουν υψηλή οργανωτική ευελιξία.

### **Περίοδος 2001- σήμερα:**

- Mobile Banking, οι υπολογιστικές δυνατότητες των κινητών συσκευών σε συνδυασμό με την έξαρση των δυνατοτήτων και έκτασης του διαδικτύου η τραπεζική γίνεται δυνατή από τα προσωπικά τεματικά/κινητές συσκευές.
- Η Fintech είναι μια νέα χρηματοπιστωτική βιομηχανία που εφαρμόζει τεχνολογία για τη βελτίωση των οικονομικών δραστηριοτήτων.
- Εισαγωγή του Bitcoin, ένα peer-to-peer σύστημα πληρωμών και ένα ψηφιακό συνάλλαγμα ανοιχτού κώδικα.

## Χρηματοοικονομική Τεχνολογία ή FinTech

Η FinTech είναι συντομογραφία για τη Χρηματοοικονομική Τεχνολογία και υποδηλώνει τη χρήση τεχνολογίας για την υποστήριξη, την ενίσχυση και την καινοτομία των υπηρεσιών και των προϊόντων που προσφέρονται παραδοσιακά από τους εγκατεστημένους χρηματοπιστωτικούς οργανισμούς, συμπεριλαμβανομένων τραπεζών, ασφαλιστικών εταιρειών και επενδυτικών κεφαλαίων.

Αρχικά, ο όρος εφαρμόζοταν στην τεχνολογία που χρησιμοποιείτο στο πίσω μέρος των καθιερωμένων χρηματοπιστωτικών ιδρυμάτων καταναλωτών και εμπορίου. Σε ένα επιστημονικό άρθρο, όπου αναλύει τα μοντέλα για τον τρόπο με τον οποίο είχε αναλύσει και επιλύσει τα καθημερινά τραπεζικά προβλήματα που αντιμετώπισαν στην τράπεζα Manufactures Hanover Trust, ο αντιπρόεδρος της τράπεζας, ο κ. Abraham Leon Bettinger (Prabook, 2016) όρισε σαν FinTech ή αλλιώς ένα ακρωνύμιο Financial Technologies που σημαίνει χρηματοοικονομική τεχνολογία, “συνδυάζοντας την τεχνογνωσία των τραπεζών με τις σύγχρονες τεχνικές διαχείρισης της επιστήμης και τον υπολογιστή.”(.49). Από το τέλος της πρώτης δεκαετίας του 21ου αιώνα, ο όρος έχει επεκταθεί ώστε να περιλαμβάνει οποιαδήποτε τεχνολογική καινοτομία στον χρηματοπιστωτικό τομέα, συμπεριλαμβανομένων των καινοτομιών στον τομέα της χρηματοοικονομικής παιδείας, της λιανικής τραπεζικής, των επενδύσεων και ακόμη και των κρυπτογραφικών νομισμάτων.

Η Fintech είναι ένας φορέας χρηματοοικονομικής τεχνολογίας που περιγράφει έναν αναδυόμενο τομέα χρηματοπιστωτικών υπηρεσιών τον 21ο αιώνα. Οργανισμοί FinTech είναι ένα σύνολο επιχειρήσεων τραπεζικών και/ή χρηματοπιστωτικών υπηρεσιών που το μοντέλο τους δεν είναι απλά να χρησιμοποιούν τις σύγχρονες ηλεκτρονικές και τεχνολογικές υπηρεσίες αλλά στήνονται με βάση και γύρω αυτές, προκαλώντας του υπάρχων επιχειρησιακό status quo (.50). Η FinTech είναι συντομογραφία για τη Χρηματοοικονομική Τεχνολογία και υποδηλώνει τη χρήση τεχνολογίας για την υποστήριξη, την ενίσχυση και την καινοτομία των υπηρεσιών και των προϊόντων που προσφέρονται παραδοσιακά από τους εγκατεστημένους χρηματοπιστωτικούς οργανισμούς, συμπεριλαμβανομένων τραπεζών, ασφαλιστικών εταιρειών και επενδυτικών κεφαλαίων.

Η FinTech έχει κάνει τις τραπεζικές συναλλαγές φιλικότερες προς τον χρήστη, αποτελεσματικότερες, φθηνότερες και γρηγότερες, δηλαδή ψηφιακές (Dorfleitner, Hornhuf, Schmitt, & Weber, 2017) και δημιούργησε νέα επιχειρηματικά μοντέλα, υπηρεσίες και προϊόντα.

Είναι ενδεικτικό να δούμε το πώς οι εταιρίες αυτές έχουν μεγαλώσει και έχουν πρόσφατα αποκτήσει τη δυνατότητα να απειλήσουν τις θέσεις των τραπεζών που κατέχουν μακρόχρονη θέση από τα συστατικά στοιχεία των αλυσίδων αξίας των τραπεζών, να τα εξευρίσκει και να τα βελτιστοποιούν και να προσφέρουν προτάσεις αξίας για συγκεκριμένες κατηγορίες που να ανταποκρίνονται στις μεταβαλλόμενες καταναλωτικές προτιμήσεις και οικονομικές ανάγκες. Καθώς η



ανάπτυξη της FinTech αυξάνεται στην ποιότητα και τη διάδοση, το ίδιο ισχύει και για τις δυνατότητες της ουσιαστικής αλλαγής τραπεζικών συναλλαγών όπως την ξέρουμε.

Η KPMG παρέχει έναν πιο ενημερωμένο κατάλογο των 100 παγκόσμιων κορυφαίων εταιρειών FinTech του 2016 (KPMG, 2016). Η KPMG είναι μια επαγγελματική εταιρεία παροχής υπηρεσιών και ένας από τους Big Four ελεγκτές, μαζί με την Deloitte, την Ernst & Young (EY) και την Price waterhouse Coopers (PwC).

Ο κατάλογός τους εξετάζει την ποιότητα κάθε επιχείρησης εξετάζοντας τα αντληθέντα κεφάλαια, τον ρυθμό άντλησης κεφαλαίων, τη γεωγραφική και τομεακή πολυμορφία και την καινοτομία. Ο αριθμός των εταιρειών ανά χώρα στον κατάλογο αυτό μπορεί να είναι πιο σημαντικός για την ποιότητα του τομέα FinTech. Ωστόσο, αυτός ο κατάλογος δεν έχει επίσης εξετάσει τη σχετική σχέση μεταξύ μεγέθους χώρας και δραστηριότητας της FinTech: Περισσότεροι άνθρωποι σημαίνουν περισσότερο διαθέσιμο κεφάλαιο και περισσότερες νέες επιχειρήσεις. Για να το προσαρμόσουμε και να μετρήσουμε την κατά κεφαλήν ποιότητα, επιλέγουμε να συμπεριλάβουμε και άλλα τρία στατιστικά στοιχεία: Αριθμός νεοεμφανιζόμενων επιχειρήσεων που βρέθηκαν από το Haddad & Hornuf ανά εκατομμύριο κατοίκους, ποσοστό υιοθεσίας ορίζεται ως χρήστες του FinTech ως ποσοστό του ψηφιακά ενεργού πληθυσμού (EY, 2017) και το ποσοστό των πελατών που χρησιμοποιούν τουλάχιστον μία μη παραδοσιακή επιχείρηση για χρηματοπιστωτικές υπηρεσίες (Cargemini, 2017). Πιστεύουμε ότι οι δείκτες αυτοί συμπληρώνουν την ένδειξη της πολυπλοκότητας του τομέα FinTech των χωρών. Με βάση αυτές τις πέντε στατιστικές, παρουσιάζουμε τις ηγετικές χώρες ανά περιφέρεια στο Πίνακα 2 παρακάτω (.77).

<i>ΧΩΡΑ</i>	<i>Εταιρίες στους Fintech100: Παγκόσμιοι Ηγέτες Καινοτομίας (KPMG, 2016)</i>	<i>Αριθμός από startups το 2014 (Haddad &amp; Hornhuf, 2016)</i>	<i>Αριθμός από startups; ανά M capita (κατά Haddad &amp; Hornhuf, 2016)</i>	<i>Υιοθέτηση : Fin.tech χρήστες σε % ψηφιακά ενεργού πληθυσμού (EY, 2017)</i>	<i>% καταναλωτών που χρησιμοποιούν τουλάχιστον μια μη παραδοσιακή επωνυμία από Fin.services (Capgemini)</i>
<i>Η.Π.Α.</i>	25	235	0,721	33 %	46 %
<i>Καναδάς</i>	6			18 %	40 %
<i>Ηνωμένο Βασίλειο</i>	12	231	3,519	42 %	49 %
<i>Γερμανία</i>	4	54	0,652	35 %	
<i>Γαλλία</i>	3	53	0,789	27 %	36 %
<i>Αυστραλία</i>	9			37 %	43 %
<i>Κίνα</i>	8			69 %	84 %
<i>Ινδία</i>	3			52 %	77 %
<i>Σιγκαπούρη</i>	1			23 %	53 %
<i>Βραζιλία</i>	3			40 %	
<i>Μεξικό</i>	1			36 %	

### Η ανάγκη των Big Data στις τράπεζες

Η **τραπεζική βιομηχανία** είναι μία από τις πλέον βασισμένες στα δεδομένα βιομηχανίες. Με την πρόοδο της τεχνολογίας, οι συσκευές που μπορούν να χρησιμοποιήσουν οι καταναλωτές για την έναρξη συναλλαγών πολλαπλασιάζονται εκτοξεύοντας τον αριθμό των συναλλαγών. Αποτέλεσμα η ταχεία αύξηση των δεδομένων (.12). Στα τέλη του 2012, εκτιμήθηκε ότι οι χρηματοπιστωτικοί οργανισμοί και οι οργανισμοί χρεογράφων διαχειρίζονται 3.8 petabytes δεδομένων ανά επιχείρηση.

#### (Όγκος)

Και αυτό αφορά μόνο το πρόβλημα του όγκου των δεδομένων, καθώς προστίθεται και πρόβλημα της ποικιλότητας των δεδομένων. Σύμφωνα με πρόσφατη έρευνα της Forrester, οι επιχειρήσεις αξιοποιούν μόλις το 25% των αδόμητων πληροφοριών για δημιουργία γνώσης, ενίσχυση της ικανότητας αναγνώρισης προβλημάτων και λήψη επιχειρηματικών αποφάσεων. Το 90% σχεδόν όλων των δεδομένων που συγκεντρώνει μία εταιρεία συνιστούν αδόμητη πληροφορία, η οποία αποτελεί τον μεγαλύτερο όγκο των δεδομένων της επιχείρησης. Τα αδόμητα δεδομένα (unstructured data) εξακολουθούν να αυξάνουν ταχύτητα σε όγκο και ποικιλία, δημιουργώντας ακόμα μεγαλύτερη - αλλά κρυμμένη- αξία στους οργανισμούς. Οι επιχειρήσεις συλλέγουν σε καθημερινή βάση τεράστιο όγκο πληροφορίας, η οποία βασίζεται -κυρίως- σε κείμενο. Η συλλογή γίνεται με πολλαπλούς τρόπους: ανατροφοδότηση από τον πελάτη (customer feedback), e-mails, ηλεκτρονικά αρχεία

κειμένου, blogs, έρευνες (π.χ. έρευνα ικανοποίησης πελάτη), αναφορές και αιτήματα, social media, επικοινωνίες μέσω Call Centers και γραμμές εξυπηρέτησης πελατών, και πολλά άλλα (.51).

### (Ποικιλότητα)

Τα σύνολα δεδομένων έχουν αυξηθεί πάρα πολύ από την άποψη του μεγέθους, του τύπου και της πολυπλοκότητας και είναι δύσκολο να συνεργαστούν με τα παραδοσιακά εργαλεία διαχείρισης βάσεων δεδομένων. Πολλά μεγάλα χρηματοπιστωτικά και τραπεζικά ιδρύματα φθάνουν στα ανώτερα όρια των συστημάτων τους και τώρα αναζητούν νέες αναλύσεις και λύσεις-πλαίσιο. (Ταχύτητα)

Η ανάγκη για καλύτερη αξιοποίηση των υφιστάμενων τεχνολογιών για την απόκτηση, οργάνωση, ενσωμάτωση και ανάλυση των δεδομένων είναι θέμα επιβίωσης για τους τραπεζικούς οργανισμούς. Και όπως είδαμε παραπάνω πληρεί σε μεγάλο βαθμό τα ελάχιστα κριτήρια για τον κλάδο των **Big Data**: Όγκο, Ποικιλότητα, Ταχύτητα.

**Στον γενικότερο επιχειρηματικό κόσμο**, η ανάγκη αυτή παρατηρείται καθώς εννέα από τους 10 ηγέτες επιχειρήσεων θεωρούν τα δεδομένα ως τον τέταρτο παράγοντα παραγωγής, όπως η γη, η εργασία και το κεφάλαιο. Σύμφωνα με την έρευνα, η χρήση μεγάλων δεδομένων έχει βελτιώσει τις επιδόσεις των επιχειρήσεων κατά μέσο όρο 26% και ο αντίκτυπος εκτιμάται ότι θα αυξηθεί στο 41% κατά τα επόμενα τρία χρόνια. Οι εταιρείες που θέτουν τα δεδομένα ως την κορυφαία προτεραιότητα για τις αποφάσεις μάρκετινγκ και πωλήσεων αυξάνουν την απόδοση των επενδύσεων μάρκετινγκ κατά 15-20% (.12).

Οι επενδύσεις Big Data εκτιμάται ότι αντιπροσωπεύουν μόνο 30 δισεκατομμύρια δολάρια μόνο το 2014 και οι επενδύσεις αυτές αναμένεται να αυξηθούν σε αθροιστικό ετήσιο ποσοστό αύξησης κατά 17% τα επόμενα έξι χρόνια και να φθάσουν τα 76 δισ. Δολάρια μέχρι το τέλος του 2020. (.12)

Σήμερα, αυτό που έχει σημασία για τις επιχειρήσεις και τους εμπόρους δεν είναι πλέον το ερώτημα εάν η επένδυση σε εργαλεία διαχείρισης δεδομένων είναι μια καλή συμφωνία ή όχι - αλλά πόσο αποτελεσματικά μπορούν οι επιχειρήσεις να ενεργήσουν στο σύνθετο υλικό που έχουν;

## 2.2 Εφαρμογές BD και DM στο Τραπεζικό Σύστημα

### Διαχείριση κινδύνων

Η διαχείριση και η μέτρηση του κινδύνου αποτελεί τον πυρήνα κάθε χρηματοπιστωτικού ιδρύματος. Η μεγάλη πρόκληση στον τραπεζικό και ασφαλιστικό κόσμο σήμερα είναι η εφαρμογή συστημάτων διαχείρισης κινδύνων προκειμένου να προσδιοριστεί, να μετρηθεί και να ελεγχθεί η έκθεση των επιχειρήσεων. Εδώ ο πιστωτικός κίνδυνος και ο κίνδυνος χρηματοοικονομικής αγοράς αποτελούν την κεντρική πρόκληση, μπορεί κανείς να παρατηρήσει μια σημαντική αλλαγή στον τομέα

του τρόπου μέτρησης και αντιμετώπισής τους, με βάση την εμφάνιση της προηγμένης τεχνολογίας βάσεων και εξόρυξης δεδομένων. (Άλλοι τύποι κινδύνων αποτελούν ο κίνδυνος ρευστότητας, λειτουργικός κίνδυνος ή ο κίνδυνος συγκέντρωσης)(.52).

Σήμερα, η ολοκληρωμένη μέτρηση των διαφόρων ειδών κινδύνου (δηλαδή, της αγοράς και του πιστωτικού κινδύνου) κινείται στο επίκεντρο της προσοχής. Όλα αυτά βασίζονται σε μοντέλα που αντιπροσωπεύουν μόνο χρηματοπιστωτικά μέσα ή παράγοντες κινδύνου, τη συμπεριφορά τους και την αλληλεπίδρασή τους με το σύνολο της αγοράς, καθιστώντας αυτό το πεδίο πολύ σημαντικό θέμα της έρευνας.

**Κίνδυνος Χρηματοοικονομικής αγοράς:** Για τα μεμονωμένα χρηματοπιστωτικά μέσα, δηλαδή τους χρηματιστηριακούς δείκτες, τα επιτόκια ή τα νομίσματα, η μέτρηση του κινδύνου αγοράς βασίζεται σε μοντέλα που εξαρτώνται από ένα σύνολο υποκείμενων συντελεστών κινδύνου, όπως τα επιτόκια, οι χρηματιστηριακοί δείκτες ή η οικονομική ανάπτυξη. Κάποιος ενδιαφέρεται για μια λειτουργική μορφή μεταξύ της τιμής του μέσου ή του κινδύνου και των υποκείμενων παραγόντων κινδύνου, καθώς και της λειτουργικής εξάρτησης των ίδιων των παραγόντων κινδύνου.

Σήμερα υπάρχουν διαφορετικές προσεγγίσεις μέτρησης κινδύνου αγοράς. Όλοι αυτοί βασίζονται σε μοντέλα που αντιπροσωπεύουν το ενιαίο όργανο, τη συμπεριφορά τους και την αλληλεπίδραση με τη συνολική αγορά. Πολλά από αυτά μπορούν να κατασκευαστούν μόνο με τη χρήση διαφόρων τεχνικών εξόρυξης δεδομένων στα δεδομένα ιδιοκτησιακού χαρτοφυλακίου, δεδομένου ότι τα δεδομένα δεν είναι διαθέσιμα στο κοινό και απαιτούν συνεπή εποπτεία.

**Πιστωτικός κίνδυνος:** Η αξιολόγηση του πιστωτικού κινδύνου αποτελεί βασική συνιστώσα της διαδικασίας εμπορικού δανεισμού. Χωρίς αυτό, ο δανειστής δεν θα ήταν σε θέση να κάνει μια αντικειμενική κρίση του καιρού για να δανείσει στον υποψήφιο δανειολήπτη, ή αν πόση χρέωση για το δάνειο. Η διαχείριση του πιστωτικού κινδύνου μπορεί να ταξινομηθεί σε δύο βασικές ομάδες:

1. Πιστωτική βαθμολογία / αξιολόγηση πιστοληπτικής ικανότητας. Ανάθεση ενός πελάτη ή ενός προϊόντος σε επίπεδο κινδύνου (δηλ. Έγκριση πίστωσης)
2. Αξιολόγηση της συμπεριφοράς και ανάλυση μετανάστευσης. Αξιολόγηση της πιθανότητας ενός πελάτη ή του προϊόντος να μεταβάλει το επίπεδο κινδύνου μέσα σε δεδομένη χρονική στιγμή (δηλ., Μεταβλητότητα της συναλλαγματικής ισοτιμίας)

Στον εμπορικό δανεισμό, η εκτίμηση κινδύνου είναι συνήθως μια προσπάθεια ποσοτικοποίησης του κινδύνου απώλειας για τον δανειστή όταν κάνει μια συγκεκριμένη απόφαση δανεισμού. Ο πιστωτικός κίνδυνος μπορεί να προσδιοριστεί ποσοτικά με τις μεταβολές της αξίας ενός πιστωτικού προϊόντος ή ενός χαρτοφυλακίου ολόκληρου του πιστωτικού πελάτη, το οποίο βασίζεται σε μεταβολή του επιτοκίου του μέσου, πιθανότητα αθέτησης και ποσοστό ανάκτησης του μέσου σε περίπτωση αθέτησης. Τα περαιτέρω αποτελέσματα διαφοροποίησης επηρεάζουν το αποτέλεσμα σε επίπεδο χαρτοφυλακίου. Έτσι, ένα σημαντικό μέρος της εφαρμογής και φροντίδας του συστήματος

διαχείρισης πιστωτικού κινδύνου θα είναι ένα τυπικό πρόβλημα εξόρυξης δεδομένων: η μοντελοποίηση της αξίας του πιστωτικού μέσου μέσω των πιθανών αδυναμιών, των μεταβολών αξιολόγησης και των ποσοστών ανάκτησης.

### Διαχείριση χαρτοφυλακίου

Οι προσεγγίσεις μέτρησης κινδύνου σε συγκεντρωτικό επίπεδο χαρτοφυλακίου ποσοτικοποιούν τον κίνδυνο ενός συνόλου μέσου ή πελάτη, συμπεριλαμβανομένων των αποτελεσμάτων διαφοροποίησης. Από την άλλη πλευρά, τα μοντέλα πρόβλεψης δίνουν την πρόβλεψη της αναμενόμενης απόδοσης ή της τιμής ενός χρηματοπιστωτικού μέσου. Και τα δύο καθιστούν δυνατή την ενεργή διαχείριση του ευρύτερου χαρτοφυλακίου της επιχείρησης με αποδοτικό τρόπο ανάληψης κινδύνου / απόδοσης. Η εφαρμογή της σύγχρονης θεωρίας κινδύνων είναι επομένως μέσα στη θεωρία του χαρτοφυλακίου, ένα σημαντικό μέρος της διαχείρισης του χαρτοφυλακίου.

Με τις τεχνικές εξόρυξης δεδομένων και βελτιστοποίησης, οι επενδυτές είναι σε θέση να καταναείμουν κεφάλαια μεταξύ εμπορικών δραστηριοτήτων για να μεγιστοποιήσουν τα κέρδη ή να ελαχιστοποιήσουν τον κίνδυνο. Αυτή η δυνατότητα υποστηρίζει την ικανότητα δημιουργίας εμπορικών συστάσεων και δομών χαρτοφυλακίου από τις απαιτήσεις χρήσης και κινδύνου που παρέχονται από τον χρήστη.

Με τις τεχνικές εξόρυξης δεδομένων είναι δυνατόν να παρέχονται εκτεταμένες δυνατότητες ανάλυσης σεναρίων σχετικά με τις αναμενόμενες τιμές ή αποδόσεις των περιουσιακών στοιχείων και τον σχετικό κίνδυνο. Με αυτή τη λειτουργικότητα, τι γίνεται αν οι προσομοιώσεις των διαφορετικών συνθηκών της αγοράς, π.χ. αλλαγές επιτοκίων και συναλλαγματικές ισοτιμίες), προκειμένου να εκτιμηθεί ο αντίκτυπος στην αξία ή / και στον κίνδυνο που συνδέεται με το χαρτοφυλάκιο, τον αντισυμβαλλόμενο της επιχειρηματικής μονάδας ή το γραφείο συναλλαγών. Διάφορα αποτελέσματα σεναρίων μπορούν να θεωρηθούν λαμβάνοντας υπόψη τις πραγματικές συνθήκες της αγοράς. Οι αναλύσεις κερδών και ζημιών επιτρέπουν στους χρήστες να έχουν πρόσβαση σε μια κατηγορία περιουσιακών στοιχείων, μια περιοχή, ένας αντισυμβαλλόμενος ή ένα προσαρμοσμένο υπο-χαρτοφυλάκιο μπορεί να συγκριθεί με κοινά διεθνή κριτήρια αναφοράς.

Τα τελευταία χρόνια ένα σημαντικό θέμα της έρευνας ήταν η δημιουργία εργαλείων ποσοτικού εμπορίου που χρησιμοποιούν μεθόδους εξόρυξης δεδομένων βασισμένες σε παρελθόντα δεδομένα ως εισροές για την πρόβλεψη βραχυπρόθεσμων κινήσεων σημαντικών νομισμάτων, επιτοκίων ή μετοχών.

Ο στόχος αυτής της τεχνικής είναι να εντοπίζονται οι χρόνοι όταν οι αγορές είναι φτηνές ή ακριβές, προσδιορίζοντας τον παράγοντα που είναι σημαντικός για τον προσδιορισμό των αποδόσεων της αγοράς. Το σύστημα διαπραγμάτευσης εξετάζει τη σχέση μεταξύ των σχετικών πληροφοριών και

ενός τμήματος χρηματοοικονομικών περιουσιακών στοιχείων και σας δίνει τη δυνατότητα να αγοράσετε ή να πωλήσετε συστάσεις όταν υποψιάζεστε ότι υπάρχει υποτίμηση ή υπερεκτίμηση. Επομένως, ακόμη και αν ορισμένοι έμποροι θεωρούν ότι η προσέγγιση εξόρυξης δεδομένων είναι πολύ μηχανική ή υπερβολικά επικίνδυνη για να χρησιμοποιηθεί συστηματικά, μπορεί να την χρησιμοποιήσουν επιλεκτικά ως περαιτέρω γνώμη.

Η τιμή ενός χρηματοοικονομικού περιουσιακού στοιχείου επηρεάζεται από διάφορους παράγοντες που μπορούν να ταξινομηθούν ευρέως ως οικονομικοί, πολιτικοί και παράγοντες της αγοράς. Οι συμμετέχοντες σε μια αγορά παρατηρούν τη σχέση μεταξύ αυτών των παραγόντων και της τιμής ενός περιουσιακού στοιχείου, υπολογίζουν την τρέχουσα αξία αυτών των παραγόντων και προβλέπουν τις μελλοντικές αξίες για να φθάσουν τελικά στη μελλοντική αξία του περιουσιακού στοιχείου και του εμπορίου ανάλογα. Πολύ συχνά από τη στιγμή που ένα εκπαιδευμένο μάτι ανιχνεύει αυτούς τους ευνοϊκούς παράγοντες, πολλοί άλλοι ίσως έχουν ανακαλύψει την ευκαιρία, μειώνοντας πιθανά τα έσοδα αλλιώς. Επίσης, αυτοί οι παράγοντες με τη σειρά τους μπορεί να σχετίζονται με πολλούς άλλους παράγοντες που καθιστούν δύσκολη την πρόβλεψη.

Οι τεχνικές εξόρυξης δεδομένων χρησιμοποιούνται για την ανεύρεση κρυφών γνώσεων, άγνωστων προτύπων και νέων κανόνων από μεγάλα σύνολα δεδομένων, τα οποία ίσως είναι χρήσιμα για μια ποικιλία δραστηριοτήτων λήψης αποφάσεων. Με την αυξανόμενη οικονομική παγκοσμιοποίηση και τη βελτίωση της τεχνολογίας της πληροφορίας, δημιουργούνται και αποθηκεύονται μεγάλα ποσά χρηματοοικονομικών στοιχείων. Αυτά μπορούν να υποβληθούν σε τεχνικές εξόρυξης δεδομένων για να ανακαλύψουν κρυμμένα μοτίβα και να αποκτήσουν προβλέψεις για τις τάσεις στο μέλλον και τη συμπεριφορά των χρηματοπιστωτικών αγορών. Με την αμεσότητα που προσφέρεται από την εξόρυξη δεδομένων, τα τελευταία δεδομένα μπορούν να εξορύσσονται για να αποκτήσουν κρίσιμες πληροφορίες το νωρίτερο. Αυτό με τη σειρά του θα είχε ως αποτέλεσμα μια βελτιωμένη ανταπόκριση στην αγορά και ευαισθητοποίηση που θα οδηγούσε σε μείωση του κόστους και αύξηση των εσόδων.

Οι προόδους που έγιναν στην τεχνολογία επέτρεψαν τη δημιουργία ταχύτερων και καλύτερων συστημάτων πρόβλεψης. Αυτά τα συστήματα βασίζονται σε ένα συνδυασμό τεχνικών εξόρυξης δεδομένων και μεθόδων τεχνητής νοημοσύνης, όπως είναι η υπόθεση βασισμένη στη συλλογιστική (CBR) και τα νευρωνικά δίκτυα (NN). Ο συνδυασμός ενός τέτοιου συστήματος πρόβλεψης μαζί με μια καλή εμπορική στρατηγική προσφέρει τεράστιες ευκαιρίες για μαζικές αποδόσεις.

Η αξία ενός χρηματοοικονομικού περιουσιακού στοιχείου εξαρτάται τόσο από τις μακροοικονομικές όσο και από τις μικροοικονομικές μεταβλητές και τα δεδομένα αυτά είναι διαθέσιμα σε ποικίλες διαφορετικές μορφές. Η εξόρυξη δεδομένων έρχεται εδώ, επειδή βοηθά στην ανακάλυψη πληροφοριών και κρυφών μοτίβων από μεγάλα σύνολα δεδομένων και πηγές δεδομένων

σε διαφορετικές μορφές. Οι τεχνικές NN και CBR μπορούν να εφαρμοστούν εκτεταμένα για την πρόβλεψη αυτών των οικονομικών μεταβλητών.

NN χαρακτηρίζονται από δυνατότητες μάθησης και τη δυνατότητα βελτίωσης της απόδοσης με την πάροδο του χρόνου. Επίσης, το NN μπορεί να γενικεύσει δηλ. Να αναγνωρίσει νέα αντικείμενα που μπορεί να είναι παρόμοια αλλά όχι ακριβώς ίδια με προηγούμενα αντικείμενα. NN με την ικανότητά τους να αντλούν νόημα από ασαφή δεδομένα μπορούν να χρησιμοποιηθούν για την ανίχνευση προτύπων τα οποία κατά τα άλλα είναι πολύ περίπλοκα για να ανιχνευθούν από τον άνθρωπο. NN ενεργούν ως εμπειρογνώμονες στην περιοχή που έχουν εκπαιδευτεί να εργαστούν. Αυτά μπορούν να χρησιμοποιηθούν για την παροχή προβλέψεων για νέες καταστάσεις και την εργασία σε πραγματικό χρόνο. Έτσι, τα διαθέσιμα ιστορικά δεδομένα σχετικά με τις χρηματοπιστωτικές αγορές και τις διάφορες μεταβλητές μπορούν να χρησιμοποιηθούν για την κατάρτιση NN για την προσομοίωση της αγοράς. Με βάση την εισαγωγή των σημερινών τιμών των μεταβλητών της αγοράς, το NN μπορεί να προβλέψει την κατάσταση την επόμενη ημέρα και μπορεί να χρησιμοποιηθεί για να δώσει μια σύσταση αγοράς / πώλησης.

Η μεθοδολογία CBR βασίζεται σε συλλογιστική από προηγούμενες επιδόσεις. Χρησιμοποιεί ένα μεγάλο αποθετήριο δεδομένων που αποθηκεύονται ως περιπτώσεις που θα περιλαμβάνουν όλες τις μεταβλητές της αγοράς στην περίπτωση αυτή. Όταν μια νέα περίπτωση τροφοδοτείται (με τη μορφή μιας περίπτωσης που περιέχει τις σχετικές μεταβλητές), ο αλγόριθμος CBR προβλέπει την απόδοση / το αποτέλεσμα αυτής της περίπτωσης με βάση τις περιπτώσεις που έχει στην αποθήκη του. Οι τεχνικές εξόρυξης δεδομένων μπορούν να χρησιμοποιηθούν για την ανίχνευση κρυφών μοτίβων σε αυτές τις περιπτώσεις, οι οποίες στη συνέχεια μπορούν να χρησιμοποιηθούν για περαιτέρω λήψη αποφάσεων. Οι μέθοδοι CBR μπορούν να χρησιμοποιηθούν σε πραγματικό χρόνο, γεγονός που καθιστά την ανάλυση πραγματικά γρήγορη και βοηθάει σε λήψη αποφάσεων σε πραγματικό χρόνο, με αποτέλεσμα να προκύψουν άμεσα κέρδη.

Έτσι, οι τεχνικές εξόρυξης δεδομένων και επιχειρηματικής πληροφόρησης (CBR και NN) μπορούν να χρησιμοποιηθούν σε συνδυασμό στις χρηματοπιστωτικές αγορές για να προβλέψουν τη συμπεριφορά της αγοράς και να αποκτήσουν συμπεριφορά με πρότυπο για να επηρεάσουν τη λήψη αποφάσεων.

### Διαχείριση και Προφίλ Πελατών

Οι τράπεζες διαθέτουν πολλές και τεράστιες βάσεις δεδομένων που περιέχουν συναλλαγές και άλλες λεπτομέρειες των πελατών της. Από αυτές τις αποθήκες δεδομένων μπορούν να εξαχθούν πολύτιμες επιχειρηματικές πληροφορίες. Ωστόσο, είναι ανέφικτη η υποστήριξη της ανάλυσης και της λήψης αποφάσεων χρησιμοποιώντας παραδοσιακές γλώσσες ερωτήσεων. γιατί η ανθρώπινη ανάλυση χωρίζεται με τον όγκο και τη διάσταση. Οι παραδοσιακές στατιστικές μέθοδοι δεν έχουν την

ικανότητα και την κλίμακα να αναλύουν αυτά τα δεδομένα και συνεπώς οι σύγχρονες μέθοδοι και εργαλεία εξόρυξης δεδομένων χρησιμοποιούνται όλο και περισσότερο στη διαδικασία λήψης αποφάσεων όχι μόνο σε τραπεζικά και χρηματοπιστωτικά ιδρύματα αλλά και σε όλες τις βιομηχανίες.

Η δημιουργία προφίλ πελατών είναι μια διαδικασία εξόρυξης δεδομένων που δημιουργεί προφίλ πελατών διαφορετικών ομάδων από την υπάρχουσα βάση δεδομένων πελατών της εταιρείας. Οι πληροφορίες που λαμβάνονται από αυτή τη διαδικασία μπορούν να χρησιμοποιηθούν για διάφορους σκοπούς, όπως η κατανόηση των επιχειρηματικών επιδόσεων, η ανάληψη νέων πρωτοβουλιών μάρκετινγκ, ο κατακερματισμός της αγοράς, η ανάλυση κινδύνου και η αναθεώρηση των πολιτικών των πελατών της εταιρείας. Το πλεονέκτημα της εξόρυξης δεδομένων είναι ότι μπορεί να χειριστεί μεγάλα ποσά δεδομένων και να μάθει εγγενείς δομές και πρότυπα στα δεδομένα. Μπορεί να δημιουργήσει κανόνες και μοντέλα που είναι χρήσιμα για τη λήψη αποφάσεων που μπορούν να εφαρμοστούν σε μελλοντικές περιπτώσεις.

Η Μοντελοποίηση συμπεριφοράς πελατών (Customer Behavior Modeling - CBM) ή ο χαρακτηρισμός πελατών είναι ένα εργαλείο για την πρόβλεψη της μελλοντικής αξίας ενός ατόμου και της κατηγορίας κινδύνου στην οποία ανήκει με βάση τα δημογραφικά του χαρακτηριστικά, τον τρόπο ζωής και την προηγούμενη συμπεριφορά. Αυτό βοηθά να επικεντρωθεί στη διατήρηση των πελατών. Τα δύο σημαντικά γεγονότα που έχουν σημαντικές επιπτώσεις στην επιλογή των μεθόδων χαρακτηριστικών πελατών είναι:

- Οι πληροφορίες προφίλ μπορούν να αποτελούνται από πολλές μεταβλητές (ή δεκάδες από αυτές).
- Τα περισσότερα από αυτά είναι κατηγορικές μεταβλητές (ή μη αριθμητικές μεταβλητές ή ονομαστικές μεταβλητές).

Ο χαρακτηρισμός του πελάτη πρέπει να χαρακτηρίζει χαρακτηριστικά των ειδικών ομάδων πελατών. Πολλές τεχνικές εξόρυξης δεδομένων αναζητούν προφίλ ειδικών ομάδων πελατών χρησιμοποιώντας συστηματικά τεχνικές τεχνητής νοημοσύνης. Δημιουργούν ακριβή προφίλ βασισμένα στην αναζήτηση δέσμης και τις τεχνικές εκμάθησης βαθμίδων.

Ο χαρακτηρισμός των πελατών χρησιμοποιεί επίσης πολλές μεθόδους πρόβλεψης μοντέλων. Οι εφαρμόσιμες τεχνικές προγνωστικής μοντελοποίησης μπορούν να κατηγοριοποιηθούν σε δύο ευρείες προσεγγίσεις. Εξαρτάται από τον τύπο των προβλεπόμενων πληροφοριών ή μεταβλητών, που ονομάζονται επίσης μεταβλητές στόχου. Αν ο τύπος των προβλεπόμενων τιμών είναι κατηγορηματικός, προτιμώνται οι τεχνικές ταξινόμησης.

**Μέθοδοι Ταξινόμησης:** Στην προσέγγιση αυτή, τα επίπεδα κινδύνου είναι οργανωμένα σε δύο κατηγορίες με βάση το παρελθόν ιστορικό προεπιλογής. Για παράδειγμα, οι πελάτες με παρελθόν ιστορικό προεπιλογής μπορούν να ταξινομηθούν σε "επικίνδυνη" ομάδα, ενώ οι υπόλοιποι



τοποθετούνται ως "ασφαλείς" ομάδες. Χρησιμοποιώντας αυτές τις πληροφορίες κατηγοριοποίησης ως στόχο της πρόβλεψης, οι τεχνικές Tree Decision και Induction Rule μπορούν να χρησιμοποιηθούν για την κατασκευή μοντέλων που μπορούν να προβλέψουν επίπεδα προεπιλογής κινδύνου για νέες αιτήσεις δανείων.

**Μέθοδοι Πρόβλεψης Αξίας:** Στη μέθοδο αυτή, για παράδειγμα, αντί να ταξινομεί νέες αιτήσεις δανείων, προσπαθεί να προβλέψει τα αναμενόμενα προεπιλεγμένα ποσά για νέες αιτήσεις δανείων. Οι προβλεπόμενες τιμές είναι αριθμητικές και συνεπώς απαιτούν τεχνικές μοντελοποίησης που μπορούν να λάβουν αριθμητικά δεδομένα ως μεταβλητές στόχους (ή προβλεπόμενες). Το νευρικό δίκτυο και η παλινδρόμηση χρησιμοποιούνται για το σκοπό αυτό. Οι πιο συνηθισμένες μέθοδοι εξόρυξης δεδομένων που χρησιμοποιούνται για τη δημιουργία προφίλ πελατών είναι:

Οι πιο συνηθισμένες μέθοδοι εξόρυξης δεδομένων που χρησιμοποιούνται για τη δημιουργία προφίλ πελατών είναι:

- Ομαδοποίηση (περιγραφικό)
- Ταξινόμηση (πρόβλεψη) και παλινδρόμηση (πρόβλεψη)
- Η ανακάλυψη κανόνα σύνδεσης (περιγραφική) και η ανακάλυψη διαδοχικών προτύπων (πρόβλεψη)

Στο Σύστημα Διαχείρισης Πελατών (Customer Relation Management ή CRM), η εξόρυξη δεδομένων χρησιμοποιείται συχνά για την εκχώρηση βαθμολογίας σε συγκεκριμένο πελάτη ή προοπτική που υποδεικνύει την πιθανότητα ότι το άτομο θα συμπεριφερθεί με συγκεκριμένο τρόπο. Για παράδειγμα, μια βαθμολογία θα μπορούσε να μετρήσει την τάση να ανταποκριθεί σε μια συγκεκριμένη προσφορά ασφάλισης ή πιστωτικής κάρτας ή να μεταβεί σε προϊόν ανταγωνιστή.

Η εξόρυξη δεδομένων μπορεί να είναι χρήσιμη και στις τρεις φάσεις ενός κύκλου σχέσεων με τους πελάτες: απόκτηση πελατών, αύξηση της αξίας του πελάτη και διατήρηση πελατών. Για παράδειγμα, μια τυπική τραπεζική εταιρεία ας πούμε ότι στέλνει 1 εκατομμύριο άμεσα μηνύματα για την απόκτηση πελατών από πιστωτικές κάρτες. Οι παλαιότερες έρευνες έχουν δείξει ότι συνήθως το 6% αυτών των πελατών-στόχων ανταποκρίνεται σε αυτά τα άμεσα μηνύματα. Οι τράπεζες χρησιμοποιούν τα μοντέλα πιστωτικού κινδύνου για να ταξινομήσουν αυτούς τους ερωτηθέντες σε καλές κατηγορίες πιστωτικού κινδύνου και κακών πιστωτικών κινδύνων. Το ποσοστό των καλών ερωτηθέντων πιστωτικού κινδύνου είναι μόνο 16% από το σύνολο των ερωτηθέντων. Έτσι, ως καθαρό αποτέλεσμα, περίπου το 1% των συνολικών στοχευμένων πελατών μετατρέπονται σε πελάτες της πιστωτικής κάρτας μέσω απευθείας ταχυδρομείου. Λαμβάνοντας υπόψη το τεράστιο κόστος και την προσπάθεια που συνεπάγεται μια τέτοια διαδικασία μάρκετινγκ, οι τεχνικές εξόρυξης δεδομένων μπορούν να βελτιώσουν σημαντικά το ποσοστό μετατροπής των πελατών με πιο εστιασμένο μάρκετινγκ. Χρησιμοποιώντας ένα πρότυπο δοκιμαστικό μοντέλο χρησιμοποιώντας τεχνικές δέντρων

αποφάσεων όπως CHAID (Ανίχνευση αυτόματης αλληλεπίδρασης Chi-squared), CART (ταξινόμηση και regression δέντρα), Quest και C5.0. μπορεί να αναλυθεί ποιοι πελάτες είναι πιο πιθανό να απαντήσουν. Και χρησιμοποιώντας αυτό με το μοντέλο κινδύνου χρησιμοποιώντας τεχνικές όπως το νευρικό δίκτυο μπορεί να βοηθήσει στην κατασκευή ενός μοντέλου δοκιμής.

Ο τρόπος εξόρυξης δεδομένων μπορεί στην πραγματικότητα να ενσωματωθεί στην εφαρμογή CRM καθορίζεται από τη φύση της αλληλεπίδρασης των πελατών. Η αλληλεπίδραση του πελάτη μπορεί να είναι εισερχόμενη (όταν ο πελάτης επικοινωνεί με την εταιρεία) ή εξερχόμενη (όταν η εταιρεία επικοινωνεί με πελάτες). Οι απαιτήσεις ανάπτυξης είναι εντελώς διαφορετικές. Οι εξερχόμενες αλληλεπιδράσεις, όπως η διαφημιστική εκστρατεία άμεσης αλληλογραφίας, περιλαμβάνουν την επιλογή εταιρειών από τους οποίους θα αποσταλούν ταχυδρομικώς, εφαρμόζοντας το μοντέλο δοκιμής στη βάση δεδομένων των πελατών. Σε άλλες εξερχόμενες εκστρατείες όπως η διαφήμιση, το προφίλ των καλών προοπτικών που δείχνει το δοκιμαστικό μοντέλο πρέπει να ταιριάζει με το προφίλ των ανθρώπων που θα φτάσει η διαφήμιση.

Για τις εισερχόμενες συναλλαγές, όπως η τηλεφωνική ή η διαδικτυακή εντολή, η εφαρμογή πρέπει να ανταποκρίνεται σε πραγματικό χρόνο. Επομένως, το μοντέλο εξόρυξης δεδομένων είναι ενσωματωμένο στην εφαρμογή και συνιστά ενεργά μια ενέργεια. Και στις δύο περιπτώσεις, ένα από τα βασικά ζητήματα στην εφαρμογή ενός μοντέλου σε νέο σύνολο δεδομένων είναι οι μετασχηματισμοί που γίνονται στην κατασκευή του μοντέλου. Η ευκολία με την οποία ενσωματώνονται αυτές οι αλλαγές στο μοντέλο καθορίζει την παραγωγικότητα της ανάπτυξης αυτών των εργαλείων.

## Μάρκετινγκ

Επειδή οι υψηλοί ανταγωνισμοί στη βιομηχανία χρηματοδότησης, οι έξυπνες επιχειρηματικές αποφάσεις στο μάρκετινγκ είναι πιο σημαντικές από ποτέ για καλύτερη στόχευση, εξαγορά, διατήρηση και σχέση πελατών. Υπάρχει ανάγκη να υπάρχουν στρατηγικές φροντίδας πελατών και μάρκετινγκ για την επιτυχία και την επιβίωση της επιχείρησης. Με τη βοήθεια της εξόρυξης δεδομένων και των προγνωστικών αναλύσεων είναι δυνατόν να γίνουν τέτοιες στρατηγικές.

Τα χρηματοπιστωτικά ιδρύματα δυσκολεύονται να εντοπίσουν νέους προηγουμένως ανεπιθύμητους αγοραστές και ως εκ τούτου εφαρμόζουν επιθετικό πρόγραμμα μάρκετινγκ για να αποκτήσουν νέους πελάτες από τους ανταγωνιστές τους. Οι αβεβαιότητες του αγοραστή καθιστούν σχεδόν αδύνατο τον σχεδιασμό νέων υπηρεσιών και τη χρήση των μέσων. Η κλασική λύση είναι να εφαρμοστούν οι υποκειμενικές ανθρώπινες εμπειρογνωμοσύνη ως βασικοί κανόνες. Μέχρι πρόσφατα, η αντικατάσταση του ανθρώπινου εμπειρογνώμονα από την τεχνολογία των υπολογιστών ήταν δύσκολη.

Ένα ενδιαφέρον εργαλείο που διατίθεται στο μάρκετινγκ και το χρηματοπιστωτικό ίδρυμα είναι η ανάλυση των δεδομένων του πελάτη. Αυτό επιτρέπει την ανάλυση και τον υπολογισμό βασικών δεικτών που βοηθούν την τράπεζα να εντοπίσει παράγοντες που επηρέασαν τη ζήτηση του πελάτη στο παρελθόν και την ανάγκη του πελάτη στο μέλλον.

Οι πληροφορίες σχετικά με τα προσωπικά δεδομένα του πελάτη μπορούν επίσης να δώσουν ενδείξεις που επηρεάζουν τη μελλοντική ζήτηση. Σε περίπτωση ανάλυσης των οφειλετών λιανικής και των μικρών εταιρειών, τα καθήκοντα μάρκετινγκ συνήθως περιλαμβάνουν παράγοντες σχετικά με τον ίδιο τον πελάτη, την πιστοληπτική του ικανότητα και τη βαθμολόγησή του από εξωτερικούς οργανισμούς αξιολόγησης πιστοληπτικής ικανότητας.

Με την εμφάνιση εργαλείων εξόρυξης δεδομένων και επιχειρηματικών πληροφοριών, οι τράπεζες κατέστησαν δυνατή την ενίσχυση της απόκτησης πελάτη μέσω άμεσου μάρκετινγκ και την καθιέρωση πολυκαναλικών επαφών για τη βελτίωση της ανάπτυξης των πελατών μέσω της πολλαπλής πώλησης και της πώλησης προϊόντων και την αύξηση της διατήρησης των πελατών διαχείριση συμπεριφοράς. Οι τράπεζες μπορούν να χρησιμοποιήσουν τα διαθέσιμα δεδομένα για να διατηρήσουν τους καλύτερους πελάτες τους και να εντοπίσουν ευκαιρίες για να τους πουλήσουν πρόσθετες υπηρεσίες. Το προφίλ όλων των πολύτιμων λογαριασμών μπορεί να γίνει και τα κορυφαία λένε 5-10% μπορούν να ανατεθούν στους Managers Σχέσεων, των οποίων η δουλειά θα είναι να εντοπίσουν νέες ευκαιρίες πώλησης με αυτούς τους πελάτες.

Είναι επίσης δυνατή η δέσμευση διαφόρων προτάσεων για την κάλυψη της ανάγκης των πελατών. Η εξόρυξη δεδομένων μπορεί επίσης να βοηθήσει τις τράπεζες να προσαρμόσουν τις διάφορες προωθητικές προσφορές. Για παράδειγμα, τα άμεσα μηνύματα μπορούν να προσαρμοστούν σύμφωνα με το τμήμα των κατόχων λογαριασμών στην τράπεζα. Είναι επίσης δυνατό οι τράπεζες να εντοπίσουν το πρόβλημα των πελατών που μπορεί να είναι αχρεωστήτως στο μέλλον, από τα προηγούμενα αρχεία πληρωμών τους και το προφίλ και τα μοντέλα δεδομένων που είναι διαθέσιμα. Αυτό μπορεί επίσης να βοηθήσει τις τράπεζες να προσαρμόσουν τη σχέση τους με αυτούς τους πελάτες έτσι ώστε η απώλεια στο μέλλον να διατηρηθεί στο ελάχιστο.

Η εξόρυξη δεδομένων μπορεί να βελτιώσει τα ποσοστά απόκρισης στις εκστρατείες άμεσης αλληλογραφίας, καθώς ο χρόνος που απαιτείται για την ταξινόμηση των πελατών θα μειωθεί, γεγονός που με τη σειρά του θα αυξήσει τα έσοδα, θα βελτιώσει την αποδοτικότητα των πωλήσεων από την ομάδα στόχο. Η εξόρυξη δεδομένων βοηθά τις τράπεζες να βελτιστοποιήσουν το χαρτοφυλάκιο των υπηρεσιών τους, τα κανάλια παράδοσης. Ένα αρχείο παλαιών συναλλαγών μπορεί να δώσει χρήσιμη γνώση στην τράπεζα και διαφορετικές τοποθεσίες / υποκαταστήματα του ίδιου υποκαταστήματος μπορούν επίσης να ακολουθήσουν ορισμένα μοτίβα που όταν παρατηρηθούν μπορούν να χρησιμοποιηθούν ως παλαιότερα αρχεία για να μάθουν από και να βασίσουν τις μελλοντικές ενέργειες επάνω.

Οι τεχνικές εξόρυξης δεδομένων μπορούν να βοηθήσουν τεράστιες τις τράπεζες και τα χρηματοπιστωτικά ιδρύματα σε αυτόν τον τομέα για καλύτερη στόχευση και απόκτηση νέων πελατών, ανίχνευση απάτης σε πραγματικό χρόνο, παροχή προϊόντων με βάση το τμήμα για καλύτερη στόχευση των πελατών, ανάλυση των προτύπων αγοράς πελατών χρόνο για καλύτερη συγκράτηση και σχέση, ανίχνευση αναδυόμενων τάσεων για την ενεργητική στάση σε μια άκρως ανταγωνιστική αγορά προσθέτοντας πολύ περισσότερη αξία σε υπάρχοντα προϊόντα και υπηρεσίες και εκτόξευση νέων δεσμών προϊόντων και υπηρεσιών.

Ο Rajanish Dass (2006) (.53) πρότεινε ότι οι τεχνικές εξόρυξης δεδομένων μπορούν να βοηθήσουν σε μεγάλο βαθμό τις τράπεζες και τα χρηματοπιστωτικά ιδρύματα σε αυτόν τον τομέα για καλύτερη στόχευση και απόκτηση νέων πελατών, ανίχνευση απάτης σε πραγματικό χρόνο, παροχή προϊόντων με βάση το τμήμα για καλύτερη στόχευση των πελατών, Ανάλυση των αγοραστικών προτύπων των πελατών με την πάροδο του χρόνου για καλύτερη συγκράτηση και σχέση, ανίχνευση αναδυόμενων τάσεων για προληπτική προσέγγιση σε μια άκρως ανταγωνιστική αγορά προσθέτοντας πολύ περισσότερη αξία σε υπάρχοντα προϊόντα και υπηρεσίες και δρομολόγηση νέων πακέτων προϊόντων και υπηρεσιών.

### Ανίχνευση Απάτης στον τραπεζικό τομέα

Μερικές φορές η δεδομένη δημογραφική περιγραφή και το ιστορικό συναλλαγών των πελατών είναι πιθανό να εξαπατήσουν την τράπεζα. (.18) Η τεχνική εξόρυξης δεδομένων συμβάλλει στην ανάλυση τέτοιων προτύπων και συναλλαγών που οδηγούν σε απάτη. Ο τραπεζικός τομέας καταβάλλει περισσότερες προσπάθειες για την ανίχνευση της απάτης. Η διαχείριση της απάτης είναι μια δραστηριότητα έντασης γνώσης. Είναι τόσο σημαντικό για την ανίχνευση της απάτης ότι η διαπίστωση ποιων από τις συναλλαγές δεν είναι εκείνες που θα κάνει ο χρήστης. Έτσι, το σύστημα εξόρυξης ελέγχει ποιες από τις συναλλαγές δεν ταιριάζουν σε μια συγκεκριμένη κατηγορία ή δεν είναι τυποποιημένες επαναλαμβανόμενες συναλλαγές.

Απαιτείται να αποφασίσει ποιες από τις ενέργειες του χρήστη αντιστοιχούν στη φυσική συμπεριφορά του και οι οποίες είναι εξαιρεση, χωρίς καμία βοήθεια. Με τη βοήθεια μοναδικών αλγορίθμων, είναι σε θέση να ανιχνεύει ύποπτη δραστηριότητα μέσα στα δεδομένα με μη δεσμευτικό τρόπο. Ενώ το σύστημα παρατηρεί τις συναλλαγές του χρήστη, ανακαλύπτει κοινά πρότυπα συμπεριφοράς συγκεντρώνοντας ανάλογες συναλλαγές μαζί.

Για να ανακαλύψουν ανώμαλες συναλλαγές, συγκρίνονται νέες συναλλαγές με τα κοινά πρότυπα συμπεριφοράς του χρήστη. Μια συναλλαγή που δεν αντιστοιχεί σε ένα από αυτά τα πρότυπα θα αντιμετωπίζεται ως ύποπτη δραστηριότητα και θα ενεργοποιεί ανάλογα τα μέτρα προφύλαξης.

Ένα σημαντικό πρώιμο βήμα στην ανίχνευση της απάτης είναι ο εντοπισμός παραγόντων που μπορούν να οδηγήσουν σε απάτη. Ποια συγκεκριμένα φαινόμενα συμβαίνουν συνήθως πριν, κατά τη διάρκεια ή μετά από ένα δόλιο γεγονός; Ποια άλλα χαρακτηριστικά θεωρούνται γενικά με την απάτη; Όταν εντοπίζονται αυτά τα φαινόμενα και τα χαρακτηριστικά, η πρόβλεψη και η ανίχνευση της απάτης γίνεται ένα πολύ πιο εύχρηστο έργο. Λόγω της φύσης των δεδομένων, οι παραδοσιακές τεχνικές μηχανικής μάθησης δεν είναι κατάλληλες. Οι παραδοσιακές τεχνικές μπορούν να ανιχνεύσουν δόλιες ενέργειες παρόμοιες με αυτές που έχουν ήδη αναγνωριστεί ως απάτη, αλλά σπάνια εντοπίζουν δόλιες δραστηριότητες που δεν είχαν μάθει εκ των προτέρων. Το μοντέλο συγκέντρωσης και οι μέθοδοι εκτίμησης πυκνότητας πιθανότητας, που περιγράφονται στις επόμενες ενότητες, μπορούν να χρησιμοποιηθούν καλά για την ανίχνευση της απάτης στον τραπεζικό τομέα.

**1) Το μοντέλο της Συλλογής:** Η συλλογή βοηθά στην ομαδοποίηση των δεδομένων σε όμοια συμπλέγματα που βοηθούν στην απλή ανάκτηση δεδομένων. Η ανάλυση συμπλέγματος είναι μια τεχνική για το σπάσιμο των δεδομένων σε συναφή στοιχεία με τέτοιο τρόπο ώστε τα σχέδια και η σειρά να γίνουν ορατά. Αυτό το μοντέλο βασίζεται στη χρήση των περιοχών συγκέντρωσης δεδομένων των παραμέτρων.

**2) Μέθοδος εκτίμησης πυκνότητας πιθανότητας:** Για να μοντελοποιηθεί η συνάρτηση πυκνότητας πιθανότητας, χρησιμοποιείται το μοντέλο Gaussian blend, το οποίο είναι ένα άθροισμα των σταθμισμένων πυκνοτήτων συστατικών Gaussian μορφής. Αυτή η μέθοδος εξειδικεύει το γενικό μοντέλο επαναπροσδιορίζοντας τις αναλογίες ανάμειξης για κάθε χρήστη δυναμικά μετά από κάθε περίοδο δειγματοληψίας καθώς διατίθενται νέα δεδομένα. Ενώ τα μέσα και οι διακυμάνσεις των ειδικών μοντέλων του χρήστη είναι κοινά, μόνο οι αναλογίες ανάμειξης είναι διαφορετικές μεταξύ των μοντέλων των χρηστών. Προκειμένου να εκτιμηθεί η πυκνότητα της συμπεριφοράς στο παρελθόν, είναι απαραίτητο να ανακτηθούν τα δεδομένα από τις τελευταίες ημέρες και να προσαρμοστούν οι αναλογίες ανάμειξης για να μεγιστοποιηθεί η πιθανότητα προηγούμενης συμπεριφοράς. Αλλά αυτή η προσέγγιση απαιτεί υπερβολική αλληλεπίδραση με το σύστημα χρέωσης που θα χρησιμοποιηθεί στην πράξη. Για να αποφευχθεί αυτή η επαχθείς επεξεργασία δεδομένων, αυτή η μέθοδος διατυπώνει τη διαδικασία μερικής εκτίμησης χρησιμοποιώντας την on-line εκτίμηση. Η on-line έκδοση του αλγόριθμου EM εισήχθη για πρώτη φορά από την Nowlan.

Η JP Morgan Chase (.14) έχει στραφεί στην τεχνολογία καταπολέμησης της τρομοκρατίας για να εντοπίσει τον κίνδυνο απάτης μεταξύ των υπαλλήλων της. Αυτή η τεχνολογία επιτρέπει στην τράπεζα να επεξεργάζεται τεράστιες ποσότητες δεδομένων για τον εντοπισμό ατομικής συμπεριφοράς που θα μπορούσε να αποκαλύψει κινδύνους ή ανοίγματα για να κερδίσει χρήματα. Άλλες τράπεζες έχουν επίσης στραφεί σε Big Data για να εντοπίσουν πιθανούς αθέμιτους εμπόρους, οι οποίοι θα μπορούσαν ενδεχομένως να επιφέρουν τεράστιες απώλειες.

## 2.3 Το Τραπεζικό Σύστημα στην εποχή των Big Data.

Στο κεφάλαιο αυτό θα δούμε ποιο είναι όμως στο συγκριτικό πλεονέκτημα στην χρήση τεχνολογιών Big Data. Ας δούμε μερικά μακροσκοπικά παραδείγματα (case studies) στις τράπεζες σήμερα και μερικά μικροσκοπικά (έρευνα και αποτελέσματα). Αλλά και τις τεχνολογίες και τους οργανισμούς FinTech.

### Χρήση BD εκ των έσω

Πώς όμως η χρήση των Big Data και των τεχνολογιών Data Mining αντιλαμβάνεται από τους χρήστες τους; Στην έρευνα του αυτή ο Vivek Bhambri, (.53) “Data Mining as a Solution for Data Management in Banking Sector” αποτυπώνει την απάντηση αυτή από τη συλλογή δεδομένων από τραπεζικούς υπαλλήλους, διοικητικά στελέχη και στελέχη του τμήματος πληροφορικής, τα αποτελέσματα καταγράφονται στον πίνακα 3..

Πίνακας 3

<i>Πρόβλημα</i>	<i>Τράπεζες που</i>	<i>Τράπεζες που δεν</i>
	<i>χρησιμοποιούν λογισμικό</i>	<i>χρησιμοποιούν κανένα</i>
	<i>εξόρυξης δεδομένων</i>	<i>λογισμικό εξόρυξης δεδομένων</i>
<b>Υπάρχει δυσκολία ανάκτησης δεδομένων από τη βάση δεδομένων κατά απαίτηση.</b>	58% συμφωνούν 37 % Μπερδεύονται σε αυτό το ζήτημα	92% διαφωνούν
<b>Υπάρχει επαρκής πληροφορία σε σχέση με τον όγκο των δεδομένων.</b>	49% συμφωνούν 32% Μπερδεύονται σε αυτό το ζήτημα	88% διαφωνούν
<b>Είναι δύσκολο να γίνει διάκριση μεταξύ πιστών και μη αφοσιωμένων πελατών.</b>	57% συμφωνούν	84% διαφωνούν
<b>Είναι δύσκολη η διαφοροποίηση των συναλλαγών που οδηγούν σε απάτη.</b>	63% συμφωνούν 28% είναι ουδέτερο 9% διαφωνούν	70% διαφωνούν 29% είναι ουδέτεροι
<b>Είναι δύσκολη η παροχή υπηρεσιών προσαρμοσμένη προς τον πελάτη.</b>	57% συμφωνεί 24% είναι ουδέτεροι	84% διαφωνεί 16% είναι ουδέτεροι
<b>το ποσοστό προσέγγισης του τηλεφωνικού κέντρου είναι πολύ χαμηλό.</b>	54% συμφωνεί 37% είναι ουδέτεροι 10% διαφωνούν	74% διαφώνησε 26% είναι ουδέτεροι

Είναι δύσκολο να κρίνουμε τα γούστα και τις προτιμήσεις των πελατών.	52% συμφωνεί 42% είναι ουδέτεροι	67% διαφωνεί 31% είναι ουδέτεροι
Δεν υπάρχει ακριβής μέθοδος για την πρόβλεψη ότι ο πελάτης θα πληρώσει το δάνειό του ή όχι.	48% συμφωνεί 20% διαφωνεί 32% είναι ουδέτεροι	71% διαφωνεί 29% είναι ουδέτεροι
Τα τρέχοντα λογισμικά δεν είναι σε θέση να παρέχουν τις σχετικές πληροφορίες σύμφωνα με τις απαιτήσεις των υπεύθυνων για τη λήψη αποφάσεων.	60% συμφωνεί 31% διαφωνεί 9% είναι ουδέτεροι	77% διαφωνεί 21% είναι ουδέτεροι

Η ανάλυση του πίνακα δείχνει ότι όλα τα προβλήματα που συζητήθηκαν σε αυτή την ενότητα σχετίζονταν με προβλήματα διαχείρισης δεδομένων. Υπάρχει σημαντική διαφορά μεταξύ των επιπέδων απόκρισης των ερωτηθέντων από τις τράπεζες όπου χρησιμοποιείται λογισμικό εξόρυξης δεδομένων και μεταξύ αυτών των τραπεζών όπου αυτό είναι δεν χρησιμοποιείται. Πολλοί από τους ερωτηθέντες των τραπεζών της πρώτης στήλης δεν κατανοούσαν ή βρισκόντουσαν σε σύγχυση σχετικά με τις ερωτήσεις. Από πλευράς των χρηστών των τεχνολογιών που χρησιμοποιούν λογισμικό εξόρυξης δεδομένων το προβάδισμα καθιστά σαφές.

### Μελέτες Περίπτωσης:

#### Αμερικανικές τράπεζες:

Η **JP Morgan Chase** (.54) δημιουργεί τεράστιο όγκο πληροφοριών για τις πιστωτικές κάρτες και άλλα στοιχεία συναλλαγών για τους πελάτες της που εδρεύουν στις Η.Π.Α. Μαζί με τις διαθέσιμες στο κοινό οικονομικές στατιστικές από την αμερικανική κυβέρνηση, η JPMorgan Chase χρησιμοποιεί νέες αναλυτικές δυνατότητες για την ανάπτυξη ιδιόκτητων γνώσεων σχετικά με τις καταναλωτικές τάσεις και με τη σειρά της προσφέρει αυτές τις αναφορές στους πελάτες της τράπεζας. Η αναλυτική τεχνολογία Big Data επέτρεψε στην τράπεζα να καταλείψει την αγορά των καταναλωτών σε μικρότερα τμήματα, ακόμη και σε μεμονωμένα άτομα, και να παράγει αναφορές σε δευτερόλεπτα.

Η **Τράπεζα της Αμερικής (BoA)** (.55) χρησιμοποίησε τις αναλυτικές δυνατότητες του Big Data για να καταλάβει γιατί πολλοί από τους εμπορικούς πελάτες της είχαν αποτύχει σε μικρότερες τράπεζες. Χρησιμοποίησε μια διαδικτυακή πύλη διαχείρισης μετρητών που αποδείχθηκε υπερβολικά άκαμπτη για τους πελάτες που επιθυμούν την ελευθερία πρόσβασης σε βοηθητικές υπηρεσίες

διαχείρισης χρηματικών διαθεσίμων από άλλες επιχειρήσεις παροχής χρηματοπιστωτικών υπηρεσιών. Ανακάλυψαν ότι οι μικρότερες τράπεζες θα μπορούσαν να προσφέρουν πιο αρθρωτές λύσεις. Η BoA χρησιμοποίησε δεδομένα που προέκυψαν από τη συμπεριφορά των πελατών της στον δικό της ιστότοπο, καθώς και από ημερολόγια τηλεφωνικών κέντρων και αντίγραφα προσωπικών συνεντεύξεων πελατών για να διαπιστώσουν γιατί χάθηκαν πελάτες. Στο τέλος, έριξε την προσφορά all-in-one και ξεκίνησε το 2009 ένα πιο ευέλικτο ηλεκτρονικό προϊόν: το Cash Pro Online και μια κινητή έκδοση Cash Pro Mobile το 2010.

Όταν η **Chase Manhattan Bank** (.56) στη Νέα Υόρκη άρχισε να χάνει πελάτες σε ανταγωνιστές, άρχισε να χρησιμοποιεί την εξόρυξη δεδομένων για να αναλύσει τους λογαριασμούς πελατών και να κάνει αλλαγές στις απαιτήσεις του λογαριασμού της, επιτρέποντας έτσι στην τράπεζα να διατηρήσει τους κερδοφόρους πελάτες της. Η εξόρυξη δεδομένων χρησιμοποιείται επίσης από τη Fleet Bank, Boston, για να εντοπίσει τους καλύτερους υποψήφιους για προσφορές αμοιβαίων κεφαλαίων. Οι δημογραφικές πληροφορίες των τραπεζών για τα ανθρακωρυχεία και τα στοιχεία λογαριασμών σε διαφορετικές σειρές προϊόντων για να προσδιορίσουν ποιοι πελάτες ενδέχεται να επενδύσουν σε αμοιβαίο κεφάλαιο και αυτές οι πληροφορίες χρησιμοποιούνται για να στοχεύσουν αυτούς τους πελάτες.

Η **Citi Bank**, (.56) από την πλευρά της, πειραματίζεται επίσης με νέους τρόπους να προσφέρει δεδομένα συναλλαγών για τους εμπορικούς πελάτες, συγκεντρωμένα από την παγκόσμια πελατειακή της βάση, τα οποία μπορούν να χρησιμοποιηθούν για τον εντοπισμό νέων μορφών εμπορίου. Τα στοιχεία θα μπορούσαν, για παράδειγμα, να αποκαλύψουν δείκτες για τις επόμενες μεγάλες πόλεις στις αναδυόμενες αγορές. Σύμφωνα με ένα διευθυντικό στέλεχος, ο οποίος διαχειρίζεται τις εσωτερικές λειτουργίες και την τεχνολογία στη Citi, η τράπεζα μοιράστηκε τις πληροφορίες αυτές με μια μεγάλη ισπανική εταιρεία ειδών ένδυσης, η οποία ήταν σε θέση να προσδιορίσει πού να ανοίξει μια νέα μονάδα παραγωγής και πολλά νέα καταστήματα.

#### Ευρωπαϊκές τράπεζες:

Η **Deutsche Bank** (.14) κατανοεί τις δυνατότητες μετασχηματισμού της Big Data και βρίσκεται πολύ μπροστά από την καμπύλη, πραγματοποιώντας σημαντικές επενδύσεις σε όλες τις περιοχές της Τράπεζας. Η Deutsche Bank διαθέτει σήμερα πλατφόρμες πολλαπλών παραγωγών Hadoop που διατίθενται μέσω του Open Source, επιτρέποντας ένα μειωμένο κόστος όσον αφορά την επεξεργασία δεδομένων. Ιστορικά, οι υιοθετούντες το Big Data ήταν οι τράπεζες της Global Technology & Operations και Finance. Από τότε αναγνωρίζεται ότι η Big Data είναι σε θέση να επωφεληθεί από όλες τις επιχειρήσεις της Τράπεζας και έχει δημιουργηθεί το Data Lab (συνδεδεμένο με το εργαστήριο καινοτομίας), το οποίο κάθεται οριζόντια στις διάφορες επιχειρήσεις της Τράπεζας παρέχοντας εσωτερικά δεδομένα και συναφή αναλυτικά στοιχεία παροχή συμβουλευτικών



υπηρεσιών. Για παράδειγμα, αν κάποια επιχειρηματική διαίρεση ή λειτουργία έχει μια ιδέα που σχετίζεται με τη χρήση δεδομένων που θα ήθελε να εξερευνήσει, το Data Lab θα μπορεί να βοηθήσει με το λειτουργικό μοντέλο και να χρησιμεύσει ως «εκκολαπτήριο» για να δοκιμάσει την επιχειρηματική ιδέα. Παρέχουν τεχνολογικές υπηρεσίες και προσωπικό για να δουλέψουν παράλληλα με την επιχείρηση για να αποδείξουν τη λύση των δεδομένων.

Η **Τράπεζα Πειραιώς** (.51) έχει ήδη επενδύσει στη λύση της SAS Text Analytics, σε συνδυασμό με τη λύση SAS Visual Analytics για αποκτήσει ενόραση στις τεράστιες εισροές feedback από πελάτες (περίπου 100.000 έρευνες πελατών ετησίως με προκαθορισμένες επιλογές είτε με ελεύθερο κείμενο συν το Call Center, Web input και τα Social Media input). Η λύση αυτή προσφέρει στον οργανισμό τη δυνατότητα να αποκαλύψει τις κρυμμένες σχέσεις ανάμεσα σε αδόμητα και δομημένα δεδομένα και την αξία που κρύβεται σε αυτά. Σε μεγάλο όγκο text δεδομένων τα μηνύματα κλειδιά που εμπεριέχουν πολύτιμες πληροφορίες συχνά κρύβονται μέσα στα κείμενα και δεν είναι δυνατή η διάκριση ή η επεξεργασία τους. Ο Οργανισμός ήθελε να είναι σε θέση να συνδυάσει τα insight που προκύπτουν από τα αδόμητα δεδομένα με εσωτερικά δομημένα δεδομένα του οργανισμού και να μπορεί να τα παρακολουθεί μέσα από προηγμένες τεχνικές data visualization, exploration, performance dashboards και KPIs' και να λαμβάνει αποφάσεις που στηρίζονται στην πραγματική «φωνή του πελάτη» και βελτιστοποιούν την εξυπηρέτησή του.

#### Ασιατικές τράπεζες:

Στην Ινδία (.53) οι τράπεζες όπως το ICICI, το IDBI, η τράπεζα Citi, η HDFC και η PNB έχουν αρχίσει να αποκομίζουν τα πλεονεκτήματα των τεχνικών εξόρυξης δεδομένων. Η Citibank είναι ηγέτης στον τομέα αυτό. Έχει δημιουργήσει ένα προφίλ δεδομένων των πελατών της για μια δεκαετία και περισσότερο. Η Citibank χρησιμοποιεί αποτελέσματα εξόρυξης δεδομένων σε συνεργασία με την Hutchison Max στη Βομβάη και την Airtel στο Δελχί. Ομοίως, η HDFC Bank συνεργάζεται με όλους τους παρόχους υπηρεσιών κινητής τηλεφωνίας της - Hutchison Max, BPL Mobile, Tata Cellular, RPG, Airtel, Aircell, Cellphone και Command. Η ICICI Bank ακολούθησε το παράδειγμα με την Hutchison Max, Airtel και Command στην Καλκούτα. «Η εξόρυξη δεδομένων βοηθά στην αύξηση των πωλήσεων με τη στόχευση των σωστών πελατών και την πραγματοποίηση των σωστών προσφορών στους πελάτες. Οι τράπεζες, οι οποίες έχουν τα αυτιά τους στο έδαφος σχετικά με τις προτιμήσεις και τις προτιμήσεις των πελατών τους, συγκεντρώνουν πολλά δεδομένα. Αυτό που κάνει η εξόρυξη δεδομένων είναι ότι μετατοπίζεται σε όλα τα ογκώδη δεδομένα και εξάγει ένα μοτίβο που επιτρέπει στην τράπεζα να προσωποποιήσει την επικοινωνία της τον πελάτη όσο το δυνατόν περισσότερο. Με την εξόρυξη δεδομένων, οι τράπεζες μπορούν να κατανοήσουν καλύτερα το τι οδηγεί τη σχέση πελατών.

## Η μελέτη περίπτωσης της Κίνας

Ένα μεγάλο μέρος των καταναλωτών στις αναδυόμενες χώρες αντιμετωπίζει δυσκολίες στην απόκτηση δανείου. (.50) Στην περίπτωση αυτή, η Κίνα λαμβάνεται ως παράδειγμα, καθώς μόνο το 20% του ενήλικου πληθυσμού έχει πιστωτικό αποτέλεσμα. Στα παραδοσιακά συστήματα λήψης αποφάσεων των τραπεζών, υπάρχει έλλειψη ιστορικών πιστωτικών στοιχείων για την πλειοψηφία των καταναλωτών. Αφού ζητήσουν αρχεία από τους ίδιους τους πελάτες, είναι εμφανές ότι δεν υπάρχει καθόλου αξιόπιστο πιστωτικό ιστορικό αυτών των ατόμων. Αυτή η αναδυόμενη μεσαία τάξη έχει ανάγκη από ένα σύστημα βαθμολόγησης της πιστοληπτικής ικανότητας που θα επιτρέπει την δίκαιη έγκριση δανείων βάσει της πραγματικής πιστοληπτικής τους ικανότητας.

Για το σκοπό αυτό, ο Kshetri (2016) διερεύνησε τη δυνατότητα χρήσης μεγάλων δεδομένων για τον προσδιορισμό της φερεγγυότητας των καταναλωτών και των SMEs. Τα μέτρα πληρεξουσιότητας χρησιμοποιήθηκαν για τον προσδιορισμό της ικανότητας και της προθυμίας του αιτούντος να επιστρέψει. Τα δεδομένα μπορούν να συγκεντρωθούν κυρίως εσωτερικά από τις τράπεζες, αλλά και εξωτερικά μέσω των κοινωνικών μέσων και των μοντέλων χρήσης κινητών τηλεφώνων. Ο συγγραφέας αναφέρει τα οφέλη από τη χρήση πλούσιων αδόμητων δεδομένων δίπλα σε δομημένα δεδομένα.

Στο άρθρο του, το Kshetri (2016) υπογραμμίζει τις πρακτικές περιπτώσεις στις οποίες χρησιμοποιείται το Big Data για τον προσδιορισμό της φερεγγυότητας. Ο κινεζικός διαδικτυακός γίγαντας Alibaba χρησιμοποιεί ήδη το δικό του σύστημα βαθμολόγησης με βάση τα δεδομένα λιανικής και τις συναλλαγές ηλεκτρονικού εμπορίου. Ο Όμιλος Alibaba είχε 300 εκατομμύρια εγγεγραμμένους χρήστες και 37 εκατομμύρια μικρές επιχειρήσεις το 2015. Η Alibaba άρχισε να αξιολογεί την πιστοληπτική ικανότητα των MME το 2007. Το Sesame Credit (Zhima) ιδρύθηκε τον Ιανουάριο του 2015, ένα σύστημα πιστοληπτικής αξιολόγησης που προσφέρει στις επιχειρήσεις και στους καταναλωτές την κατασκευή και την πρόσβαση στα πιστωτικά τους προφίλ. Το σύστημα αντλεί εσωτερικά δεδομένα από το κέντρο επεξεργασίας δεδομένων και το υπόλοιπο οικοσύστημα Alibaba. Τα εξωτερικά δεδομένα παρέχονται από τους συνεργάτες και χρησιμοποιούνται τόσο τα ηλεκτρονικά όσο και τα offline δεδομένα. Παραδείγματα είναι οι δικαστικές εκθέσεις, η έλλειψη χρέους, οι καθυστερημένες επιστροφές των ενοικιαζόμενων αυτοκινήτων και οι συναλλαγές που καταχωρήθηκαν στην Alipay. Το WeBank της Tencent είναι ένα άλλο παράδειγμα ενός κινεζικού διαδικτυακού γίγαντα που διακλαδίζεται με τη χρηματοδότηση μέσω Διαδικτύου. Συνεργάζονται με ένα σύστημα αναδοχής που απαιτεί από τους πελάτες να φορτώνουν μια μετωπική εικόνα του προσώπου τους. Το σύστημα ταιριάζει με την εικόνα με τα δεδομένα που παρέχει το Υπουργείο Δημόσιας Ασφάλειας για την επαλήθευση της ταυτότητας του καταναλωτή. Το σύστημα δίνει στο άτομο μια βαθμολογία πιστοληπτικής ικανότητας βάσει ενός μεγάλου αριθμού πηγών. Παραδείγματα

είναι το ηλεκτρονικό ιστορικό αγορών, οι δραστηριότητες σε κοινωνικά δίκτυα, τα παιχνίδια ιστότοπων και άλλες δραστηριότητες στο διαδίκτυο.

# Κεφάλαιο 3: Νομοθεσία στα Big Data και τον Τραπεζικό Τομέα

## 3.1 Προβλήματα με τα Big Data

Τα μεγάλα δεδομένα δημιουργούν μεγάλους κινδύνους για την προστασία της ιδιωτικής ζωής και συνολικά όπως θα δούμε για τους πολίτες. (.58) Η συγκομιδή μεγάλων συνόλων προσωπικών δεδομένων και η χρήση αναλυτικών στοιχείων σχετικά με την κατάσταση της τέχνης συνεπάγονται σαφώς αυξανόμενες ανησυχίες όσον αφορά την προστασία της ιδιωτικής ζωής. Η προστασία της ιδιωτικής ζωής γίνεται πιο δύσκολη, καθώς οι πληροφορίες πολλαπλασιάζονται και μοιράζονται όλο και περισσότερο σε πολλά μέρη σε όλο τον κόσμο. Δεδομένου ότι υπάρχουν περισσότερες πληροφορίες σχετικά με την υγεία, τα οικονομικά, την τοποθεσία, τη χρήση ηλεκτρικής ενέργειας και τις επιγραμματικές δραστηριότητες, προκύπτουν ανησυχίες σχετικά με τη δημιουργία προφίλ, την παρακολούθηση, τις διακρίσεις, τον αποκλεισμό, την επιτήρηση της κυβέρνησης και την απώλεια ελέγχου. Αυτό το μέρος εξηγεί μερικούς από τους μοναδικούς κινδύνους ιδιωτικής ζωής που παρουσιάζονται από τα μεγάλα δεδομένα και πιο συγκεκριμένα με την Κατασκευή Προφίλ.

### Ρίσκα που ανακύπτουν από την κατασκευή Προφίλ

Ενώ η εξόρυξη δεδομένων και η μορφοποίηση προφίλ είναι ως επί το πλείστον εννοιολογικά πλαισιωμένη ως απειλή για την προστασία της ιδιωτικής ζωής στην πληροφόρηση, (.59) πιστεύουμε ότι η ισχυρή συσχέτιση με την προστασία της ιδιωτικής ζωής υπονομεύει τους πραγματικούς κινδύνους που μπορεί να δημιουργήσει η δημιουργία προφίλ και η εξόρυξη δεδομένων σε ομάδες και άτομα. Κατά την άποψή μας, οι σημαντικότεροι κίνδυνοι που σχετίζονται με τη δημιουργία προφίλ και την εξόρυξη δεδομένων είναι ασυμμετρικές διακρίσεων, διαχωρισμού και πληροφόρησης.

### Διάκριση

Η ταξινόμηση και ο διαχωρισμός βρίσκονται στην καρδιά της (προγνωστικής) εξόρυξης δεδομένων. Ως εκ τούτου, η διάκριση αποτελεί αναπόσπαστο κομμάτι της ανάλυσης και της εξόρυξης δεδομένων. Ωστόσο, υπάρχουν καταστάσεις όπου η διάκριση θεωρείται ανήθικη και ακόμη και παράνομη. Αυτό μπορεί να συμβεί, για παράδειγμα, όταν μια άσκηση εξόρυξης δεδομένων επικεντρώνεται σε χαρακτηριστικά όπως η εθνικότητα, το φύλο, η θρησκεία ή η σεξουαλική προτίμηση. Αλλά ακόμη και χωρίς μια προηγούμενη επιθυμία να κρίνουμε τους ανθρώπους με βάση συγκεκριμένα χαρακτηριστικά, υπάρχει ο κίνδυνος να γίνονται αντιληπτές από ακούσια συγκεκριμένες ομάδες ή άτομα. Ο λόγος για αυτό είναι ότι οι αλγόριθμοι πρόγνωσης εξόρυξης

δεδομένων μπορεί να "μαθαίνουν" να κάνουν διακρίσεις με βάση τα προκατειλημμένα δεδομένα που χρησιμοποιούνται για την εκπαίδευση του αλγορίθμου.

Ένας ταξινομητής πρέπει να εκπαιδευτεί για να ταξινομήσει τα δεδομένα. Όταν τα δεδομένα εκπαίδευσης είναι μολυσμένα (για παράδειγμα επειδή διακρίνονται σε μια συγκεκριμένη ομάδα), ο ταξινομητής θα μάθει να ταξινομεί με προκατειλημμένο τρόπο, ενισχύοντας τις διακριτικές επιπτώσεις. Τέτοιες περιπτώσεις συμβαίνουν φυσικά όταν η διαδικασία λήψης αποφάσεων που οδηγεί στις ετικέτες είναι προκατειλημμένη. Ένα παράδειγμα στον τομέα της επιβολής του νόμου μπορεί να βοηθήσει να το εξηγήσουμε αυτό. Όταν οι αστυνομικοί έχουν στοχεύσει σε μια εθνική μειονότητα δυσανάλογα στο παρελθόν με βάση τη δική τους μεροληψία, είναι πιθανό ότι αυτές οι μειονότητες θα εμφανιστούν πιο εμφανώς στις στατιστικές για το έγκλημα. Αν αυτές οι στατιστικές για το έγκλημα χρησιμοποιούνται στη συνέχεια ως στοιχεία κατάρτισης για έναν ταξινομητή, είναι πιθανό ότι ο ταξινομητής θα μάθει ότι υπάρχει ισχυρή συσχέτιση μεταξύ εθνότητας και εγκληματικότητας. Αυτό με τη σειρά του θα οδηγήσει σε αποτελέσματα διακρίσεων που μπορούν να αποτελέσουν τη βάση για μελλοντικές διακρίσεις. Αυτό το αποτέλεσμα ενισχύεται περαιτέρω από το γεγονός ότι ένας ταξινομητής πιθανότατα δεν θα έχει πρόσβαση σε όλους τους σημαντικούς παράγοντες επί των οποίων θα βασιστεί μια πρόβλεψη επειδή, π.χ., λείπουν στα δεδομένα. Επομένως, η σημασία αυτών των παραγόντων που υπάρχουν στα δεδομένα αυξάνεται και θα είναι ακόμη πιο σημαντική στην πρόβλεψη από ότι ήταν στα δεδομένα εισόδου.

### Περιορισμένη Πρόσβαση και Αποκλεισμός

Μια πρόσθετη ανησυχία που προκαλείται από τα μεγάλα δεδομένα είναι ότι κλίνει μια ήδη ανομοιογενές κλίμακα υπέρ των οργανώσεων και κατά των ατόμων. (58) Τα μεγάλα οφέλη από τα μεγάλα δεδομένα, το επιχείρημα πηγαίνει, συγκεντρώνονται στην κυβέρνηση και στις μεγάλες επιχειρήσεις, όχι σε ιδιώτες - και συχνά έρχονται σε βάρος των ατόμων. Σύμφωνα με την παροιμία, «αν δεν πληρώνεις για αυτό, δεν είσαι ο πελάτης. είσαι το προϊόν ».

Ο αποκλεισμός των ατόμων από τα οφέλη από τη χρήση των δεδομένων τους εκδηλώνεται με δύο βασικούς τρόπους. Πρώτον, οι διαδικτυακές αλληλεπιδράσεις είναι συναλλαγές τύπου barter όπου τα άτομα ανταλλάσσουν προσωπικά δεδομένα για δωρεάν υπηρεσίες. Ωστόσο, οι συναλλαγές αυτές φαίνεται να συμβαίνουν σε μια αναποτελεσματική αγορά που παρεμποδίζεται από απότομες ασυμμετρίες πληροφόρησης, οι οποίες επιδεινώνονται περαιτέρω από μεγάλα δεδομένα. Η συναλλαγή με μια μεγάλη πλατφόρμα δεδομένων είναι σαν ένα παιχνίδι πόκερ όπου ένας από τους παίκτες έχει το χέρι του ανοιχτό και ο άλλος κρατά τα χαρτιά του κοντά. Η σε απευθείας σύνδεση εταιρεία γνωρίζει τις προτιμήσεις του συναλασσόμενου ατόμου μέσα έξω, ίσως καλύτερα από ό, τι το άτομο ξέρει τον εαυτό του. Συνεπώς, μπορεί να σφετεριστεί ολόκληρο το πλεόνασμα της αξίας που είναι διαθέσιμο

στη συναλλαγή με την τιμολόγηση αγαθών ή υπηρεσιών όσο το δυνατόν πλησιέστερα στην τιμή κράτησης του ατόμου.

Δεύτερον, οι οργανώσεις είναι σπάνια προετοιμασμένες να μοιράζονται τον πλούτο που δημιουργούν τα προσωπικά δεδομένα των ατόμων με τα άτομα αυτά. Ο Sir Tim Berners-Lee παρατήρησε πρόσφατα: "Ο υπολογιστής μου έχει μια μεγάλη κατανόηση της κατάστασης της κατάστασής μου, των πραγμάτων που τρώω, των τόπων στους οποίους είμαι. Το τηλέφωνό μου καταλαβαίνει από την τσέπη μου πόση άσκηση έχω πάρει και πόσες σκάλες έχω περπατήσει επάνω και ούτω καθεξής. Η εκμετάλλευση τέτοιων δεδομένων θα μπορούσε να προσφέρει εξαιρετικά χρήσιμες υπηρεσίες σε ιδιώτες, δήλωσε, αλλά μόνο εάν οι υπολογιστές τους είχαν πρόσβαση σε προσωπικά δεδομένα που κρατούσαν για 'αυτά από εταιρίες ιστού. Ένα από τα ζητήματα των σιλό κοινωνικής δικτύωσης είναι ότι έχουν τα δεδομένα και δεν το κάνουν. "

Το δικαίωμα πρόσβασης που παρέχεται στους πολίτες βάσει της ευρωπαϊκής οδηγίας για την προστασία των δεδομένων και οι πρόσθετες αρχές δίκαιης πληροφόρησης εφαρμόστηκαν στενά. Ακόμη και όταν συμμορφώνονται με το νόμο, οι οργανώσεις παρέχουν στους χρήστες λίγες χρήσιμες πληροφορίες.

#### *Από-εξατομίκευση*

Σε πολλές περιπτώσεις η εξόρυξη δεδομένων σχετίζεται σε μεγάλο βαθμό με την ταξινόμηση και συνεπώς υπάρχει ο κίνδυνος τα άτομα να κριθούν βάσει των χαρακτηριστικών της ομάδας και όχι των δικών τους μεμονωμένων χαρακτηριστικών και προσόντων.(.60) Τα προφίλ των ομάδων συνήθως περιέχουν στατιστικά στοιχεία και επομένως τα χαρακτηριστικά της ομάδας τα προφίλ μπορεί να ισχύουν για την ομάδα και τα άτομα ως μέλη αυτής της ομάδας, αν και όχι για τα άτομα ως τέτοια. Ένα παράδειγμα μπορεί να το δείξει αυτό. Για παράδειγμα, όταν οι άνθρωποι στο Άμστερνταμ είναι κατά 3% πιο εγκληματικοί από όσους βρίσκονται στην υπόλοιπη Ολλανδία, αυτό το χαρακτηριστικό ισχύει για την ομάδα (δηλ. Τους ανθρώπους στο Άμστερνταμ), για τα άτομα ως μέλη αυτής της ομάδας (δηλαδή, τυχαία επιλεγμένα άτομα που ζουν σε Άμστερνταμ), αλλά όχι για τα άτομα ως τέτοια (δηλαδή, για τον Ιωάννη, τη Μαρία και τον William που ζουν όλοι στο Άμστερνταμ). Όταν τα άτομα κρίνεται από τα χαρακτηριστικά ομάδας που δεν διαθέτουν ως άτομα, αυτό μπορεί να επηρεάσει έντονα τα πλεονεκτήματα και τα μειονεκτήματα της χρήσης προφίλ ομάδων.(.8) Εκτός από τις αρνητικές επιδράσεις της ομαδοποίησης των ομάδων σε άτομα, ο ομαδοποιητικός χαρακτηρισμός μπορεί επίσης να οδηγήσει στον στιγματισμό των μελών της ομάδας . Επιπλέον, οι διαιρέσεις σε ομάδες μπορούν να βλάψουν την κοινωνική συνοχή. Όταν τα προφίλ ομάδων, είτε είναι σωστά είτε όχι, γίνονται γνωστά στο κοινό, οι άνθρωποι μπορούν να αρχίσουν να αλληλοεπιλογούν ανάλογα. Για παράδειγμα, όταν οι άνθρωποι αρχίζουν να πιστεύουν ότι οι πολίτες του Άμστερνταμ είναι πιο

εγκληματικοί, οι άνθρωποι μπορεί να αρχίσουν να αντιδρούν και να επικοινωνούν με περισσότερη καχυποψία προς τους πολίτες του Άμστερνταμ, ανεξάρτητα από την ορθότητα ενός τέτοιου προφίλ.

### Ασυμμετρίες πληροφοριών

Η εξόρυξη δεδομένων μπορεί να οδηγήσει σε πολύτιμες πληροφορίες για τα μέρη που την χρησιμοποιούν. Όταν η εξόρυξη δεδομένων αποσκοπεί στην καλύτερη κατανόηση ατόμων ή ομάδων, συναντάμε το πρόβλημα της ασυμμετρίας των πληροφοριών. (.58) Οι ασυμμετρίες πληροφόρησης ενδέχεται να επηρεάσουν τους ισότιμους όρους ανταγωνισμού μεταξύ κυβέρνησης και πολιτών και μεταξύ επιχειρήσεων και καταναλωτών, ανατρέποντας την τρέχουσα ισορροπία δυνάμεων μεταξύ των διαφόρων μερών.

Στο πλαίσιο της σχέσης μεταξύ κυβέρνησης και πολιτών, οι ασυμμετρίες μπορούν να επηρεάσουν την αυτονομία των ατόμων. Εάν η εξόρυξη δεδομένων παράγει πράγματι αποδεκτή γνώση, η κυβέρνηση θα έχει περισσότερη δύναμη. Επιπλέον, ο φόβος για ισχυρές δυνατότητες εξόρυξης δεδομένων από την πλευρά της κυβέρνησης μπορεί να "χαλαρώσει" τη βούληση των ανθρώπων να συμμετάσχουν, για παράδειγμα, σε πολιτικές δραστηριότητες, από φόβο να παρακολουθήσουν. Για να υλοποιηθεί αυτός ο πανοπικός φόβος, μια εφαρμογή εξόρυξης δεδομένων δεν χρειάζεται καν να είναι αποτελεσματική.

Οι ασυμμετρίες πληροφόρησης ενδέχεται να διαταράξουν τους οικονομικούς όρους ανταγωνισμού μεταξύ των καταναλωτών και των επιχειρήσεων. Επιπλέον, υπάρχουν περιπτώσεις όπου η εξόρυξη δεδομένων μπορεί να βοηθήσει στη λήψη αποφάσεων σχετικά με τους καταναλωτές που θεωρούνται ανεπιθύμητοι, ανήθικοι ή παράνομοι. Ένα παράδειγμα αυτού θα ήταν να αποκλειστούν συγκεκριμένα άτομα ή ομάδες από αγαθά και υπηρεσίες που βασίζονται στην εθνικότητά τους.

Οι ασυμμετρίες πληροφόρησης ενδέχεται να διαταράξουν τους οικονομικούς όρους ανταγωνισμού μεταξύ των καταναλωτών και των επιχειρήσεων. Επιπλέον, υπάρχουν περιπτώσεις όπου η εξόρυξη δεδομένων μπορεί να βοηθήσει στη λήψη αποφάσεων σχετικά με τους καταναλωτές που θεωρούνται ανεπιθύμητοι, ανήθικοι ή παράνομοι. Ένα παράδειγμα αυτού θα ήταν να αποκλειστούν συγκεκριμένα άτομα ή ομάδες από αγαθά και υπηρεσίες που βασίζονται στην εθνικότητά τους.

### Προβλήματα εφαρμογής

Οι κίνδυνοι που περιγράφηκαν παραπάνω μπορούν να εκδηλωθούν ανεξάρτητα από το γεγονός ότι η εξόρυξη δεδομένων εφαρμόστηκε ορθά με τεχνική έννοια. Αλλά όπως είδαμε στο παράδειγμα των διακρίσεων, όταν η εξόρυξη δεδομένων εφαρμόζεται λανθασμένα, οι κίνδυνοι που

σχετίζονται με τη δημιουργία προφίλ και την εξόρυξη δεδομένων μπορεί να ενισχυθούν. Ως εκ τούτου, θα συζητήσουμε πολλά κοινά θέματα εφαρμογής που σχετίζονται με την εξόρυξη δεδομένων, τα οποία ενδέχεται να δημιουργούν πρόσθετους κινδύνους για τις ομάδες και τα άτομα.

### Ακρίβεια και αξιοπιστία

Η επιτυχία μιας άσκησης εξόρυξης δεδομένων εξαρτάται από την ποιότητα των πρώτων δεδομένων που εξορύσσονται. (.59) Αν τα δεδομένα είναι ανακριβή, τα αποτελέσματα θα είναι επίσης ανακριβή. Αυτό ισχύει τόσο για περιγραφική όσο και για προγνωστική εξόρυξη δεδομένων. Στον τομέα των προβλέψεων, τα ζητήματα εξόρυξης δεδομένων με ακρίβεια και αξιοπιστία είναι ιδιαίτερα προβληματικά, δεδομένου ότι τα αποτελέσματα μιας άσκησης εξόρυξης δεδομένων είναι συχνά χρησιμοποιούμενα για να λαμβάνουν (αυτοματοποιημένες) αποφάσεις σχετικά με άτομα ή / και ομάδες. Αλλά ακόμα και αν τα ανεπεξέργαστα δεδομένα είναι απαλλαγμένα από λάθη, η ακρίβεια και η αξιοπιστία παραμένουν ένα ζήτημα.

Συγκεκριμένα, υπάρχει το πρόβλημα των «ψευδών θετικών» και των «ψευδών αρνητικών». Αυτό σημαίνει ότι οι άνθρωποι που στην πραγματικότητα δεν ταιριάζουν με την τάξη είναι τοποθετημένοι στην τάξη (ένα ψευδώς θετικό) ή οι άνθρωποι που ταιριάζουν στην τάξη έχουν απομείνει ( ψευδές αρνητικό), τα ψευδώς θετικά και τα ψευδώς αρνητικά συμβαίνουν για διάφορους λόγους, ένα από τα οποία είναι ότι δεν υπάρχουν όλες οι πληροφορίες, για παράδειγμα, στο παράδειγμα της ταξινόμησης των καναρινιών, η παρουσία χαρακτηριστικών όπως {κίτρινο, ράμφος, φτερά, ουρά } ήταν ισχυρές ενδείξεις ότι ένα αντικείμενο μπορεί να ταξινομηθεί ως καναρίνι. Ωστόσο, ένα παπαγάλος είναι επίσης κίτρινο και έχει ράμφος, φτερά και ουρά, καθιστώντας ένα ψευδώς θετικό στο σενάριο μας.

### Αιτιότητα έναντι συσχετισμού

Ο στόχος της εξόρυξης δεδομένων είναι να βρεθούν σιωπηρές και προηγουμένως άγνωστες σχέσεις μεταξύ των δεδομένων. Ως εκ τούτου, η εξόρυξη δεδομένων δίνει νέες γνώσεις σχετικά με ένα συγκεκριμένο χώρο προβλημάτων. Στην περιγραφική εξόρυξη δεδομένων, αυτή η γνώση βασίζεται στον συσχετισμό μεταξύ ορισμένων αντικειμένων και χαρακτηριστικών. Ωστόσο, ενώ η εξόρυξη δεδομένων μπορεί να αποδείξει ότι υπάρχει σχέση μεταξύ ορισμένων αντικειμένων και χαρακτηριστικών, δεν εξηγεί γιατί υπάρχει αυτή η σχέση. Ως εκ τούτου, είναι σημαντικό να μην σφάλουμε τον συσχετισμό για την αιτιώδη συνάφεια. Για παράδειγμα, η εξόρυξη δεδομένων μπορεί να αποκαλύψει ότι οι διαρρήκτες χρησιμοποιούν κάνναβη συχνότερα από άλλους ανθρώπους. Παρόλο που είναι δελεαστικό να αναφέρεται η κάνναβη ως αιτία διάρρηξης, τα δεδομένα δεν υποστηρίζουν ένα τέτοιο συμπέρασμα.



Οι εμπειρογνώμονες εξόρυξης δεδομένων προειδοποιούν για το γεγονός ότι η συσχέτιση μεταξύ ορισμένων γεγονότων δεν συνεπάγεται αιτιώδη συνάφεια ούτε εξηγεί γιατί υπάρχει συσχέτιση μεταξύ πραγματικών περιστατικών. Μια τέτοια προειδοποίηση πρέπει να ληφθεί υπόψη, δεδομένου ότι οι προσπάθειες εξόρυξης δεδομένων και τα στατιστικά στοιχεία θα μπορούσαν να συμβάλουν στη χάραξη πολιτικής. Δεδομένου ότι ο στόχος της χάραξης πολιτικής είναι η αντιμετώπιση των αιτιών ενός προβλήματος και όχι η καταπολέμηση των συμπτωμάτων, είναι σημαντικό να γνωρίζουμε περισσότερα σχετικά με το ιστορικό, την εμφάνιση και την αιτιώδη συνάφεια ορισμένων γεγονότων. Ειδικότερα, η άυλη περιγραφική εξόρυξη δεδομένων είναι λιγότερο κατάλληλη, εάν δεν είναι ακατάλληλη για την ανακάλυψη αυτών των πληροφοριών.

### Βυθοκόρησης δεδομένων

Σε άυλη περιγραφική εξόρυξη δεδομένων ψάχνουμε για συσχετισμούς στα δεδομένα χωρίς να χρησιμοποιήσουμε μια προκαθορισμένη υπόθεση εργασίας. Ανάλογα με το μέγεθος του συνόλου δεδομένων και το "διάστημα εμπιστοσύνης" που χρησιμοποιείται για τον προσδιορισμό των συσχετισμών, η άσκηση εξόρυξης δεδομένων θα αποδώσει ορισμένα αποτελέσματα. Αν και αυτά τα αποτελέσματα μπορεί να είναι πραγματικά σημαντικά, υπάρχει επίσης η πιθανότητα να είναι εντελώς τυχαία. Επομένως, τα αποτελέσματα που βρίσκουμε (και η υπόθεση που διατυπώνουμε βάσει αυτών των αποτελεσμάτων) πρέπει να επικυρωθούν για να αποκλειστεί η πιθανότητα ότι ο συσχετισμός είναι στην πραγματικότητα εντελώς τυχαίος. Είναι σημαντικό να σημειωθεί ότι το σύνολο δεδομένων που χρησιμοποιήθηκε για την κατασκευή της υπόθεσης δεν μπορεί να χρησιμοποιηθεί για την επικύρωση αυτής της ίδιας υπόθεσης, δεδομένου ότι τα δεδομένα αυτά είναι "προκατειλημμένα" προς υποστήριξη αυτής της συγκεκριμένης υπόθεσης.

Η βυθοκόρηση δεδομένων των κινδύνων είναι ότι παρουσιάζουμε τα αποτελέσματα της αρχικής άσκησης εξόρυξης δεδομένων ως γεγονότα και όχι ως υπόθεση που πρέπει να εξεταστεί περαιτέρω. Ένα παράδειγμα που επεξηγεί αυτό το ζήτημα είναι το εξής: ας υποθέσουμε ότι θέλουμε να δοκιμάσουμε εάν κάποιοι άνθρωποι έχουν ιδιαίτερες νοητικές ικανότητες και είναι σε θέση να προβλέψουν το μέλλον. Δημιουργήσαμε ένα πείραμα στο οποίο 1000 εθελοντές πρέπει να προβλέψουν ποια ακολουθία κεφαλών και ουρών θα προκύψει από την ανατροπή ενός αμερόληπτου νομίσματος 10 φορές. Οι στατιστικές μας διδάσκουν ότι εάν ένας συμμετέχων κάνει τυχαίες εικασίες, αυτός ή αυτή εξακολουθεί να έχει πιθανότητα 1 στα 1024 για να πάρει την αλληλουχία πλήρως σωστή. Ως εκ τούτου, μπορούμε να αναμένουμε ότι κατά μέσο όρο ένας συμμετέχων θα προβλέψει σωστά το αποτέλεσμα και των 10 υποψηφίων. Φυσικά, αυτό το αποτέλεσμα δεν υποστηρίζει καθόλου τον ισχυρισμό ότι αυτός ο συμμετέχων έχει ειδικές νοητικές ικανότητες. Πρώτα ο συμμετέχων πρέπει να επιβεβαιώσει τα τέλεια αποτελέσματα σε νέα ελεγχόμενα πειράματα. Για να συνδέσετε αυτό το παράδειγμα με την εξόρυξη δεδομένων: θα μπορούσαμε να θεωρήσουμε τα αποτελέσματα της

πρόβλεψης ως δεδομένα εισόδου. Υπάρχουν δοκιμασμένες 1000 υποθέσεις: "ο συμμετέχων 1 μπορεί να προβλέψει τα πάντα σωστά" έως ότου "ο συμμετέχων 1000 μπορεί να προβλέψει τα πάντα σωστά". Ως εκ τούτου, το αποτέλεσμα του πειράματος χρησιμοποιείται για τη δημιουργία του "σχεδίου". π.χ., ο "συμμετέχων 174 μπορεί να προβλέψει τα πάντα σωστά" που θα παρουσιαστούν στον χρήστη. Είναι χρήσιμο να δούμε ότι ακόμη και μια στατιστική δοκιμασία για τη σημασία θα επιβεβαιώσει την εγκυρότητα αυτού του προτύπου.

### Αυτοματοποιημένη λήψη αποφάσεων

Η υποβιβασμός των αποφάσεων σχετικά με τη ζωή ενός ατόμου σε αυτοματοποιημένες διαδικασίες που βασίζονται σε αλγόριθμους και τεχνητή νοημοσύνη εγείρει ανησυχίες σχετικά με τις διακρίσεις, την αυτοδιάθεση και τη μείωση των επιλογών. (.58) Αυτό ισχύει όχι μόνο για τις αποφάσεις σχετικά με την πίστωση, την ασφάλιση ή τις προοπτικές απασχόλησης ενός ατόμου, οι οποίες εδώ και πολλά χρόνια έχουν ρυθμιστεί από νόμους όπως ο νόμος για τη δίκαιη πιστοληπτική ικανότητα, αλλά επίσης και για τις εξαιρετικά προσαρμοσμένες επιλογές σχετικά με τις διαφημίσεις ή το περιεχόμενο που ένας χρήστης θα δει. Στο βιβλίο του "The Daily You: How The New Advertising Industry Is Defining Your Identity And Your Worth", ο Joseph Turow υποστηρίζει ότι η αυξημένη εξατομίκευση που βασίζεται σε αδιαφανείς αλγόριθμους εταιρικού προφίλ δημιουργεί τον κίνδυνο να ανοίξει η κοινωνία και ο δημοκρατικός λόγος. "τα άτομα σε προκαθορισμένες κατηγορίες, η αυτοματοποιημένη λήψη αποφάσεων κατακερματίζει την κοινωνία σε θύλακες παρόμοιων ατόμων. Ο Turow πιστεύει ότι η κυβέρνηση θα πρέπει να ρυθμίζει τους μεσάζοντες πληροφόρησης ώστε να εξασφαλίζεται ότι οι χρήστες έχουν τον πλήρη έλεγχο της κατανάλωσης δεδομένων και περιεχομένου τους.

Σε αυτές τις περιπτώσεις, (.3) οι προβλέψεις που βασίζονται σε συσχετισμούς δεν επηρεάζουν μόνο τα άτομα, τα οποία μπορούν να ενεργούν διαφορετικά από την υπόλοιπη ομάδα στην οποία έχουν ανατεθεί, αλλά και να επηρεάζουν ολόκληρη την ομάδα και να την ξεχωρίζουν από την υπόλοιπη κοινωνία. Ένα παράδειγμα υπό αυτή την έννοια παρέχεται από το γενικό πιστωτικό αποτέλεσμα της γειτονιάς που υιοθετούν οι πιστωτικές εταιρείες, το οποίο ωθεί τις επιχειρήσεις να παρέχουν ευκαιρίες σε άτομα που ζουν σε μια δεδομένη γειτονιά με τρόπο που δεν έχει καμία σχέση με τις ατομικές τους συνθήκες, αλλά βασίζεται στην συνολική βαθμολογία της περιοχής.

Επιπλέον, (.3) οι χρήστες συχνά δεν γνωρίζουν αυτές τις μορφές ανάλυσης δεδομένων και τις επιπτώσεις που ενδέχεται να έχουν κάποιες πληροφορίες στην ιδιότητά τους ως μέλους μιας ή άλλης ομάδας που δημιουργείται από τα αναλυτικά στοιχεία. Τέλος, οι υπεύθυνοι λήψης αποφάσεων χρησιμοποιούν τα αποτελέσματα που παράγονται από τα μεγάλα δεδομένα ανάλυσης για να λαμβάνουν αποφάσεις που επηρεάζουν άτομα και ομάδες χωρίς να τους επιτρέπουν οποιαδήποτε

συμμετοχή στη διαδικασία, η οποία παραμένει βασικά βασισμένη σε συγκεχυμένη διαχείριση δεδομένων και συχνά λαμβάνει χώρα σε καταστάσεις ανισορροπίας υποκείμενα δεδομένων.

Ο Stephen Baker της Business Week προειδοποίησε ότι οι αλγόριθμοι των υπολογιστών (.62) δεν είναι αλάνθαστοι μόνο επειδή είναι αυτοματοποιημένοι: «Οι προκαταλήψεις μιας κοινωνίας αντικατοπτρίζονται στους αλγόριθμους που αναζητούνται». Όταν ερευνούσε το βιβλίο του The Numerati, συνάντησε μια ιστορία για έναν πράκτορα του FBI ο οποίος υποτίθεται ότι συσχετίζει την κατανάλωση hummus από τη γειτονιά, με την πιθανότητα να αποτελέσει καταφύγιο τρομοκρατών. Ο Μπέικερ δεν μπόρεσε να επιβεβαιώσει την ιστορία και ίσως να είναι αποκριμένος, αλλά είπε ότι η ιστορία δείχνει πόσο ανόητες υποθέσεις, όταν συνδέονται σε μια βάση δεδομένων, μπορούν να ληφθούν σοβαρά υπόψη.

### Προγνωστική Ανάλυση

Τα μεγάλα δεδομένα μπορούν να διευκολύνουν την προβλεπτική ανάλυση με έντονες επιπτώσεις για άτομα ευάλωτα σε ασθένειες, (.58) έγκλημα ή άλλα κοινωνικά στιγματιστικά χαρακτηριστικά ή συμπεριφορές. Βεβαίως, η ανάλυση πρόβλεψης μπορεί να χρησιμοποιηθεί για κοινωνικά ευεργετικούς στόχους, όπως ο προγραμματισμός της ανάκαμψης σε καταστροφική περιοχή με βάση τις διαδρομές εκκένωσης των ατόμων και τις ανάγκες αγοράς. Ωστόσο, μπορεί εύκολα να διασχίσει το κατώφλι "σκιά".

Εξετάστε μια πρόσφατη ιστορία στους New York Times, η οποία αποκάλυψε ότι ο giant Target για το λιανικό εμπόριο αποδίδει στους πελάτες τους μια «βαθμολογία πρόβλεψης εγκυμοσύνης» με βάση τις συνήθειες αγοράς τους. Σύμφωνα με την Times, η Target απασχολούσε τους στατιστικολόγους να απομακρύνουν ιστορικά στοιχεία αγοράς των γυναικών που είχαν Οι στατιστικολόγοι ανακάλυψαν λανθάνοντα πρότυπα, όπως η προτίμηση των γυναικών για μη καλλωπισμένες λοσιόν κατά την έναρξη του δεύτερου τριμήνου τους ή η τάση να αγοράζουν συμπληρώματα όπως το ασβέστιο, το μαγνήσιο και ο ψευδάργυρος κατά τις πρώτες 20 εβδομάδες της εγκυμοσύνης. για να προσδιορίσουν ένα σύνολο προϊόντων τα οποία, όταν ομαδοποιηθούν, επέτρεψαν σε αυτά να προβλέψουν με μεγάλη ακρίβεια την εγκυμοσύνη ενός πελάτη και την ημερομηνία λήξης του χρέους. Σε μία περίπτωση, ανέφερε η Times, ένας πατέρας ενός εφηβικού κοριτσιού που εισέβαλε σε ένα κατάστημα Target για να διαμαρτυρηθεί ότι η κόρη του κουπόνια και διαφημίσεις για προϊόντα μωρών. Λίγα μέρες αργότερα, κάλεσε τον διευθυντή του καταστήματος να ζητήσει συγγνώμη, αναγνωρίζοντας ότι: «υπήρξαν κάποιες δραστηριότητες στο σπίτι μου δεν είχα πλήρη επίγνωση του.»

Η προγνωστική ανάλυση είναι χρήσιμη για την επιβολή του νόμου, την εθνική ασφάλεια, τον έλεγχο των πιστώσεων, την ασφάλιση και την απασχόληση. Δημιουργεί ηθικά διλήμματα που έχουν καταγραφεί, για παράδειγμα, στην ταινία Minority Report, όπου ένα αστυνομικό τμήμα «PreCrime»

συλλαμβάνει «εγκληματίες» με βάση την προγνωστική των μελλοντικών τους παραπτωμάτων. Θα μπορούσε να διευκολύνει την παράνομη δραστηριότητα, όπως είναι η «αναπροσαρμογή» - η άρνηση ή η αύξηση του κόστους των υπηρεσιών, όπως τα δάνεια, η ασφάλιση ή η υγειονομική περίθαλψη, στους κατοίκους των συνοικιών που αποτελούνται κυρίως από μειονότητες. Παρόλο που οι πρακτικές αυτές είναι παράνομες σύμφωνα με τους ισχύοντες νόμους, οι επικριτές εξέφρασαν ανησυχίες ότι τα δεδομένα χρησιμοποιούνται κρυφά με τέτοιο τρόπο.

Ένα παράδειγμα αυτού παρέχεται από προβλέψιμες αστυνομικές αποφάσεις (.3) όπως το «PredPol», ένα λογισμικό που χρησιμοποιείται από τις τοπικές αστυνομικές δυνάμεις των ΗΠΑ για την πρόληψη, την πρόληψη και την αποτελεσματικότερη αντιμετώπιση του εγκλήματος, με βάση τα στοιχεία σταυροειδούς ελέγχου, τους τόπους και τις τεχνικές των πρόσφατων εγκλημάτων. Το PredPol και παρόμοιο λογισμικό είναι σε θέση να προβλέψουν τα μελλοντικά εγκλήματα και την τοποθεσία τους, αλλά και να προκαλέσουν έναν "αυτοπεποίθηση κύκλους προκατάληψης". Αυτό οφείλεται στο γεγονός ότι τα αστυνομικά τμήματα διαθέτουν περισσότερους πόρους στους τομείς που προτείνονται από τα αναλυτικά στοιχεία και αυτό αυξάνει την ανίχνευση εγκλημάτων σε τοπικό επίπεδο, με αποτέλεσμα την ενίσχυση της αρχικής πρόβλεψης. Ταυτόχρονα, η μειωμένη παρουσία της αστυνομίας σε άλλους τομείς μειώνει την ανίχνευση εγκλημάτων και προκαλεί δυσμενή πρόβλεψη για αυτές τις περιοχές. Η συνέπεια αυτών των λύσεων λογισμικού είναι μια ενδεχόμενη γεωγραφική διάκριση, η οποία μπορεί να μην επηρεάζει άμεσα τα άτομα, αλλά έχει αντίκτυπο στις τοπικές κοινότητες όσον αφορά το κοινωνικό στίγμα ή την ανεπαρκή παροχή αστυνομικών υπηρεσιών. Υπό αυτή την έννοια, υπάρχει συλλογικό ενδιαφέρον για τη σωστή και σωστή χρήση των δεδομένων.

Η προγνωστική ανάλυση είναι ιδιαίτερα προβληματική όταν βασίζεται σε ευαίσθητες κατηγορίες δεδομένων, όπως η υγεία, η φυλή ή η σεξουαλικότητα. Είναι ένα πράγμα να προτείνουμε σε πελάτες, βιβλία, μουσική ή ταινίες που θα μπορούσε να ενδιαφέρει με βάση τις προηγούμενες αγορές της. Είναι πολύ σημαντικό να εντοπίσουμε πότε είναι έγκυος πριν γνωρίσει η κοντινότερη οικογένειά της. Στον τομέα της επιβολής του νόμου, εγείρει το φάσμα της τοποθέτησης ή ακόμη και της φυλάκισης των ατόμων με βάση τις σκέψεις σε αντίθεση με τις πράξεις. Αυτός ο τύπος δραστηριότητας, αν και είναι σαφώς αντισυνταγματικός σύμφωνα με την ισχύουσα νομοθεσία των Ηνωμένων Πολιτειών, δεν προσελκύεται τόσο σε άλλα μέρη του κόσμου και θα μπορούσε ενδεχομένως να διασχίσει τη γραμμή από τη μυθοπλασία στην πραγματικότητα, λαμβάνοντας υπόψη τις σωστές συνθήκες στις Ηνωμένες Πολιτείες.

Ακόμη και με μη ευαίσθητες κατηγορίες δεδομένων, η προγνωστική ανάλυση μπορεί να έχει επιβραδυντικό αποτέλεσμα για τα άτομα και την κοινωνία, διαιωνίζοντας παλιές προκαταλήψεις. Οι πλούσιοι και μορφωμένοι θα βρουν το γρήγορο δρόμο, ενώ οι φτωχοί και μειονεκτούντες θα στοιβάζονται στο κατάστρωμα ακόμα περισσότερο από, τι πριν. Αν αγνοήσουμε τις αποκλίσεις και υποθέσουμε ότι «αυτό που έχει είναι τι θα είναι», η προγνωστική ανάλυση γίνεται μια

αυτοεκπληρούμενη προφητεία που δίνει έμφαση στην κοινωνική διαστρωμάτωση. Αυτό οδηγεί σε ηθικά αμφισβητούμενα συμπεράσματα, όπως αυτά που συντάχθηκαν από το (το) διάσημο άρθρο του 2001 των John Donohue και Steven Levitt, "Οι επιπτώσεις της νομιμοποιημένης έκτρωσης στην εγκληματικότητα", το οποίο υποστήριξε ότι η νομιμοποίηση των αμβλώσεων στη δεκαετία του 1970 συνέβαλε σημαντικά στις μειώσεις σε ποσοστά εγκληματικότητας που γνώρισαν τη δεκαετία του 1990.

### Προβλήματα από την χρήση στο Τραπεζικό σύστημα

Στην περίπτωση του τραπεζικού τομέα, τα προβλήματα και τα ρίσκα αυτά εμφανίζονται στην πιο βίαιη και άμεση μορφή τους καθώς οι επιπτώσεις καταχρήσεων ,μη ορθής ή σφαλμάτων μπορούν να είναι καθοριστικές ,ακόμα και καταστρεπτικές για τα υποκείμενα των προφίλ. Τα αποτελέσματα αυτών των εφαρμογών καθορίζουν τις προσφορές που θα γίνονται στους καταναλωτές, τα επιτόκια, την έγκριση ή την απόρριψη ενός δανείου (επιχειρηματικό, στεγαστικό, σπουδών, κ.α.). Τα φαινόμενα διάκρισης και αποκλεισμού μπορούν να παρουσιαστούν είτε από τεχνικά αίτια (λανθασμένα στοιχεία, σφάλματα επεξεργασίας, κ.α.) είτε από καταχρηστικές πολιτικές (πολιτικές που προωθούν την διάκριση, είναι κληρονομικά προκατειλημμένες, κ.α.). Ταυτόχρονα δεν υπάρχει μέθοδος για τους πολίτες να έχουν πρόσβαση στο πώς εξήχθησαν τα αποτελέσματα, από το ποια δεδομένα συλλέγονται και επεξεργάζονται έως τις μεθόδους και την πολιτική τους, καθώς οι τεχνολογίες αυτές αποτελούν πνευματική ιδιοκτησία των τραπεζών. Φαινόμενα όπως αυτά των μηχανών κατασκευής προφίλ, των αυτοματοποιημένων αποφάσεων και της προγνωστικής ανάλυσης καθιστούν τους ανθρώπους σε αριθμούς και αναπαράγοντας όλα τα προαναφερθέντα προβλήματα του Profiling. Οι διαχωριστικές γραμμές είναι λεπτές και οι επιπτώσεις μεγάλες.

Ρόλος του να κρατά αυτές τις ισορροπίες, να επιβλέπει την ορθή χρήση και να προστατεύει τους πολίτες από τέτοια φαινόμενα είναι αυτός του έννομου κράτους, δίνοντας πάντα το περιθώριο στις επιχειρήσεις να ασκούν τις δραστηριότητές τους και να επιτυγχάνουν τους στόχους τους. Στα πλαίσια αυτά η προστασία των πολιτών καθίσταται από την ενίοτε κρατική νομοθεσία και τα ανάλογα ρυθμιστικά όργανα.

## 3.2 Προστασία από τα Big Data

### Νομικό πλαίσιο προστασίας Big Data, Data Mining και Profiling

Θα ήταν εσφαλμένο να πει κανείς ότι τα Big Data είναι τομείς για τους οποίους δεν υπάρχουν κανόνες. Ακόμη και αν ενδεχομένως σε κάποιες περιπτώσεις δεν είναι στενευμένοι, κατάλληλα δομημένοι ή προνοητικοί, οι κανόνες για την προστασία προσωπικών δεδομένων αλλά και οι σχετικές

γνώμες των αρμοδίων οργάνων της Ευρωπαϊκής Ένωσης μπορούν να αξιοποιηθούν αναφορικά με τις τεχνολογίες Big Data. Ενδεικτικά μπορεί να αναφερθούν:

- Κανονισμός (ΕΕ) 2016/679 του Ευρωπαϊκού Κοινοβουλίου
- η Οδηγία 2002/58/ΕΚ για την επεξεργασία των δεδομένων προσωπικού χαρακτήρα και την προστασία της ιδιωτικής ζωής στον τομέα των ηλεκτρονικών επικοινωνιών
- ο Ν. 2472/1997 για την προστασία του ατόμου από την επεξεργασία δεδομένων προσωπικού χαρακτήρα. ο Ν. 3471/2006 για την προστασία δεδομένων προσωπικού χαρακτήρα και της ιδιωτικής ζωής στον τομέα των ηλεκτρονικών επικοινωνιών
- η Γνώμη 02/2013 της Ομάδας Εργασίας του άρθρου 291 για τις εφαρμογές των έξυπνων συσκευών
- η Γνώμη 05/2014 της Ομάδας Εργασίας του άρθρου 29 σχετικά με τις τεχνικές ανωνυμοποίησης

Μέσα στο κείμενο της γνωμοδότησης γίνεται αναφορά στους κυριότερους κινδύνους που απειλούν την προστασία των δεδομένων στο εσωτερικό του οικοσυστήματος του ΔτΠ, ενώ εξάγονται κάποια χρήσιμα συμπεράσματα και γίνονται συστάσεις προς τα ενδιαφερόμενα μέρη (κατασκευαστές, σχεδιαστές, πλατφόρμες, οργανισμούς τυποποίησης, κλπ.).

#### *Γενικός κανονισμός Προστασίας προσωπικών Δεδομένων (GDPR)*

Ο νέος Γενικός Κανονισμός για την Προστασία των Προσωπικών Δεδομένων (General Data Protection Regulation - GDPR) είναι το ισχυρότερο μέτρο προστασίας των πολιτών έναντι των καταχρηστικών χρήσεων των Big Data από τους παρόχους υπηρεσιών. Λαμβάνοντας υπόψη τα τέσσερα χρόνια διαπραγματεύσεων και περίπου 4.000 τροποποιήσεις που έγιναν πριν από την έγκρισή του, το GDPR μπορεί σίγουρα να θεωρηθεί ορόσημο όσον αφορά τη ρύθμιση της προστασίας των δεδομένων. αυτό απεικονίζει επίσης τις σημαντικές επιπτώσεις που συνδέονται με τον παρόντα κανονισμό. Ο κανονισμός τέθηκε σε ισχύ στις 25 Μαΐου 2016 και έχει τεθεί σε εφαρμογή από τις 25 Μαΐου 2018.

Το ιστορικό του έχει ως εξής:

- 25 Ιανουαρίου 2012: Η Ευρωπαϊκή Επιτροπή προτείνει τη μεταρρύθμιση των κανόνων προστασίας προσωπικών δεδομένων στην ΕΕ
- 8 Απριλίου 2016: Το Συμβούλιο εγκρίνει τη θέση του σε πρώτη ανάγνωση
- 14 Απριλίου 2016: Ο Κανονισμός και η Οδηγία εγκρίνονται από το Ευρωπαϊκό Κοινοβούλιο
- 27 Απριλίου 2016: Ψηφίζεται ο Γενικός Κανονισμός για την Προστασία Δεδομένων (ΕΕ) 2016/679 που αποτελεί το κύριο νομοθέτημα της νέας δέσμης κανόνων

- 4 Μαΐου 2016: Ο Κανονισμός και η Οδηγία (ΕΕ) 2016/680 δημοσιεύονται στην Επίσημη Εφημερίδα της ΕΕ
- 5 Μαΐου 2016: Η Οδηγία τίθεται σε ισχύ
- 24 Μαΐου 2016: Ο Κανονισμός τίθεται σε ισχύ
- 6 Μαΐου 2018: Τα Κράτη Μέλη πρέπει να έχουν ενσωματώσει την Οδηγία στο εθνικό δίκαιο
- 25 Μαΐου 2018: Ο Κανονισμός τίθεται σε εφαρμογή

Η μεταρρύθμιση της προστασίας των δεδομένων αποτελεί νομοθετική δέσμη με σκοπό την επικαιροποίηση και τον εκσυγχρονισμό των υφιστάμενων κανόνων της προστασίας των δεδομένων. Περιλαμβάνει δύο νομοθετικές πράξεις: τον Γενικό Κανονισμό Προστασίας των Δεδομένων (που αντικαθιστά την οδηγία 95/46/ΕΚ) και την οδηγία προστασίας των δεδομένων στον τομέα της επιβολής του νόμου (που αντικαθιστά την απόφαση-πλαίσιο του 2008 για την προστασία των δεδομένων) (.78).

Δεδομένου ότι οι κύριες ανησυχίες αναφορικά με τις τεχνολογίες Big Data αφορούν στην προστασία των προσωπικών δεδομένων, είναι εκ των ων ουκ άνευ ότι οποιοδήποτε τέτοιο πλάνο προϋποθέτει την έκδοση του πολυαναμενόμενου νέου κανονισμού σχετικά με τα δεδομένα αυτά. Έννοιες όπως Privacy by design, accountability principle αλλά και η ανωνυμοποίηση των δεδομένων είναι έννοιες άρρηκτα συνδεδεμένες με τα Big Data και υπάρχει ειδική μέριμνα στον σχετικό κανονισμό.

Το GDPR έχει έντονο αντίκτυπο στην ψηφιακή βιομηχανία, αλλά η μείωση της συνάφειας του GDPR με τις εταιρείες τεχνολογίας θα οδηγούσε επίσης στην έλλειψη σημαντικών επιπτώσεων σε όλες τις άλλες βιομηχανίες.

Η Ευρωπαϊκή Επιτροπή πρότεινε το GDPR ως μέσο για την ενδυνάμωση των πολιτών και την εξασφάλιση του ελέγχου των προσωπικών τους δεδομένων και την απλούστευση του ρυθμιστικού περιβάλλοντος για τις επιχειρήσεις κατά τη διάρκεια της ψηφιοποίησης. Ο κανονισμός βρίσκεται στο επίκεντρο της υλοποίησης της ψηφιακής ενιαίας αγοράς, προτεραιότητα στην ημερήσια διάταξη της Επιτροπής.

Ως θεματοφύλακας της ενιαίας αγοράς και επιφορτισμένη με τη διασφάλιση ισότιμων όρων ανταγωνισμού, η Επιτροπή είναι πιθανό να παρακολουθεί στενά τη συμμόρφωση με τους κανονισμούς GDPR από τους συμμετέχοντες στην αγορά που εδρεύουν εντός και εκτός της επικράτειάς της. Παρόλο που πρόκειται για κανονισμό, το GDPR παραμένει πολύ αόριστο, απαιτώντας προσεκτική ερμηνεία και συμβουλές από τους φορείς προστασίας δεδομένων και τους νομικούς.

Οι εθνικές αρχές προστασίας δεδομένων (DPA) και το μελλοντικό Ευρωπαϊκό Συμβούλιο Προστασίας Δεδομένων θα δημοσιεύσουν κατευθυντήριες γραμμές για τον καθορισμό βέλτιστων πρακτικών όσον αφορά την εφαρμογή του GDPR, η οποία ενδέχεται να απαιτεί συνεχείς αξιολογήσεις νομικού αντικτύπου. Όσον αφορά τις κατευθυντήριες γραμμές για συγκεκριμένους κλάδους, όπως για παράδειγμα τις χρηματοπιστωτικές υπηρεσίες, η Επιτροπή και οι αρχές προστασίας δεδομένων ενθαρρύνουν την κατάρτιση κωδίκων δεοντολογίας που θα πρέπει να συμβάλλουν στην ορθή εφαρμογή του GDPR. Οι κώδικες αυτοί θα πρέπει να λαμβάνουν υπόψη τα ιδιαίτερα χαρακτηριστικά του τομέα όσον αφορά την επεξεργασία δεδομένων και τις συγκεκριμένες βιομηχανικές ανάγκες (.78).

### *Κύρια καινοτόμα στοιχεία που τέθηκαν από το GDPR*

#### *Συναίνεση*

Για να είναι νόμιμη η επεξεργασία δεδομένων, πρέπει να βασίζεται σε έναν από τους νομικούς λόγους που προβλέπονται στο άρθρο 6 του GDPR. Από τις έξι αυτές δυνατές επιλογές, η συγκατάθεση του υποκειμένου των δεδομένων είναι μία από τις πλέον διαδεδομένες νομικές βάσεις για την επεξεργασία δεδομένων (Άρθρο 6 παρ. 1, GDPR). Επομένως, η ενίσχυση των προϋποθέσεων έγκυρης συναίνεσης από το GDPR έχει μεγάλη σημασία. Σύμφωνα με τον κανονισμό, «συγκατάθεση» νοείται κάθε ελεύθερη, συγκεκριμένη, ενημερωμένη και σαφής ένδειξη των επιθυμιών του υποκειμένου των δεδομένων με την οποία αυτός, με δήλωση ή με σαφή καταφατική ενέργεια, συμφωνεί με την επεξεργασία δεδομένων προσωπικού χαρακτήρα σχετικά με τον εαυτό του (Άρθρο 4 παρ. 11, GDPR). Εάν, βάσει της ισχύουσας οδηγίας 95/46 / ΕΚ, θα μπορούσαν να ληφθούν υπόψη τόσο οι διαδικασίες συμμετοχής όσο και οι μέθοδοι εξαίρεσης για τη λήψη συγκατάθεσης, οι λύσεις opt-out αποκλείονται σαφώς από τον κανονισμό. Το GDPR απαιτεί η συγκατάθεση του υποκειμένου των δεδομένων να εκφράζεται με δήλωση ή σαφή θετική ενέργεια για να επισημανθεί η συμφωνία στην επεξεργασία δεδομένων προσωπικού χαρακτήρα. Τα προ συμπληρωμένα τετραγωνίδια, η σιωπή και η αδράνεια είναι σαφώς ανεπαρκή και δεν μπορούν να θεωρηθούν ότι συνιστούν την αποδοχή της προτεινόμενης επεξεργασίας από το υποκείμενο των δεδομένων (Αιτία 32, GDPR). Οι συνήθεις πρακτικές των τραπεζών που τροποποιούν τους όρους και τις προϋποθέσεις τους είναι απλώς να ενημερώνουν τους χρήστες σχετικά με τέτοιες αλλαγές που ορίζουν ότι η σιωπή ή η αδράνεια των χρηστών εντός συγκεκριμένου χρονικού πλαισίου μπορεί να θεωρηθεί ως συναίνεση. Αυτές οι πρακτικές θα πρέπει να αναθεωρηθούν στο πλαίσιο του GDPR ενόψει της νέας συγκατάθεσης συμμετοχής, η οποία απαιτεί από τα υποκείμενα των δεδομένων να δραστηριοποιηθούν και να εκφράσουν σαφώς τη συγκατάθεσή τους στην επεξεργασία δεδομένων.

Ως αντιστάθμισμα της απαίτησης GDPR ότι η συναίνεση πρέπει να χορηγείται προληπτικά, ο κανονισμός προβλέπει επίσης ρητά το δικαίωμα του υποκειμένου των δεδομένων να αποσύρει τη συγκατάθεσή του. Οι ελεγκτές πρέπει να ενημερώνουν τα υποκείμενα των δεδομένων σχετικά με το



δικαίωμα αυτό και πρέπει να διευκολύνουν την άσκηση αυτού του δικαιώματος ανά πάσα στιγμή (Άρθρο 7 παρ. 3, GDPR) . Η τελευταία αυτή υποχρέωση συνεπάγεται ότι οι ελεγκτές ενδέχεται να υποχρεωθούν να προσαρμόσουν το περιβάλλον πληροφορικής τους για την εκπλήρωση αυτής της απαίτησης του κανονισμού.

Επιπλέον, προκειμένου να αποκτήσουν έγκυρη συγκατάθεση, οι ελεγκτές πρέπει να εξασφαλίσουν ότι τα υποκείμενα των δεδομένων γνωρίζουν καλά ότι συμφωνούν με την επεξεργασία των προσωπικών τους δεδομένων. Επομένως, είναι οι ημέρες που οι ρήτρες προστασίας δεδομένων κρύβονται στους γενικούς όρους και προϋποθέσεις. Σύμφωνα με το GDPR, όταν η συγκατάθεση δίδεται στο πλαίσιο γραπτής δήλωσης που αφορά και άλλα θέματα, η αίτηση συγκατάθεσης θα πρέπει να διακρίνεται σαφώς από άλλα θέματα και να παρουσιάζεται με κατανοητή και εύκολα προσπελάσιμη μορφή, χρησιμοποιώντας σαφή και απλή γλώσσα.

Πρέπει να ενθαρρυνθούν τα μοντέλα πιο συσσωρευμένης συγκατάθεσης. Η ομάδα εργασίας του άρθρου 29 όρισε ήδη το 2011 ότι υπήρχε "απαίτηση για λεπτομερή συγκατάθεση" όσον αφορά τα διάφορα στοιχεία της επεξεργασίας δεδομένων(Άρθρο 29 Data Protection Working Party, Opinion 15/2011 σελ. 17) . Στο πλαίσιο των εφαρμογών για έξυπνες συσκευές, η ομάδα εργασίας υποστήριξε ότι η ευκρίνεια πρέπει να επιτρέπει στους χρήστες να ελέγχουν καλύτερα κάθε τύπο εφαρμογών δεδομένων που ενδέχεται να έχουν πρόσβαση, έτσι ώστε η πρόσβαση να παρέχεται μόνο σε εκείνα τα δεδομένα που είναι απαραίτητα για τη λειτουργία της εφαρμογής (Άρθρο 29 Data Protection Working Party, Opinion 02/2013 σελ. 15-27) . Πολύ πρόσφατα, ο ΕΕΠΔ προέβλεψε, όσον αφορά τα δεδομένα προσωπικού χαρακτήρα που υποβάλλονται σε επεξεργασία από κινητές εφαρμογές, την ανάγκη πραγματικών επιλογών για την επεξεργασία δεδομένων προσωπικού χαρακτήρα, σύμφωνα με τα οποία: "Η εφαρμογή κινητής τηλεφωνίας πρέπει να ζητήσει συγκαταβατική άδεια για κάθε κατηγορία δεδομένων προσωπικού χαρακτήρα που επεξεργάζεται σχετική χρήση " (Ευρωπαϊός Επόπτης Προστασίας Δεδομένων, Κατευθυντήριες γραμμές σχετικά με την προστασία των δεδομένων προσωπικού χαρακτήρα που υποβάλλονται σε επεξεργασία από κινητές εφαρμογές, 11.2016, p. 9). Μια πολυεπίπεδη, κοκκώδης συγκατάθεση φαίνεται να αποτελεί τον καλύτερο τρόπο για να αντιμετωπίσει τις τρέχουσες πραγματικότητες επεξεργασίας δεδομένων, πέραν του πλαισίου των εφαρμογών κινητής τηλεφωνίας. Είναι καλύτερα ευθυγραμμισμένη με τις θεμελιώδεις αρχές προστασίας δεδομένων όσον αφορά την ελαχιστοποίηση των δεδομένων και τον περιορισμό του σκοπού και επιτρέπει στα πρόσωπα στα οποία αναφέρονται τα δεδομένα να διαθέτουν ενισχυμένο έλεγχο.

Συχνά λέγεται ότι τα δεδομένα είναι το νέο νόμισμα, ειδικά στο Διαδίκτυο. Και στην πραγματικότητα, για να επωφεληθούν από μια υπηρεσία, τα υποκείμενα των δεδομένων πρέπει συχνά να συμφωνούν με την επεξεργασία των προσωπικών τους δεδομένων μετά από ένα μοντέλο «πάρτε-ή αφήστε το». Ο Daniel Solove παρατήρησε ότι "οι επιλογές που δίδονται στους ανθρώπους σχετικά με

την πληροφόρησή τους δεν είναι καθόλου επιλογές. Οι άνθρωποι πρέπει να παραιτηθούν από τα προσωπικά δεδομένα για να κερδίσουν εργασία, να προμηθεύσουν ασφάλεια, να αποκτήσουν πιστωτική κάρτα ή να συμμετάσχουν με άλλο τρόπο σαν κανονικός πολίτης στη σημερινή οικονομία

Η συγκατάθεση είναι ουσιαστικά χωρίς νόημα σε πολλά πλαίσια. Όταν οι άνθρωποι δίνουν τη συγκατάθεσή τους, πρέπει συχνά να συναινούν σε μια απόλυτη παραίτηση ελέγχου των πληροφοριών τους (Άρθρο 7 παρ. 4, GDPR) (.79) . Δεδομένου ότι η διαπραγματευτική ισχύς των υποκειμένων των δεδομένων που αντιμετωπίζουν τέτοιες πολιτικές είναι σχεδόν μηδενική, αμφισβητείται η ελευθερία συγκατάθεσης. Ο κανονισμός υπογραμμίζει ότι η συναίνεση πρέπει να παρέχεται ελεύθερα.

Προκειμένου να εκτιμηθεί κατά πόσον παρέχεται ελεύθερα, «λαμβάνεται ιδιαίτερα υπόψη το κατά πόσον, μεταξύ άλλων, η εκτέλεση μιας σύμβασης, συμπεριλαμβανομένης της παροχής υπηρεσίας, εξαρτάται από τη συγκατάθεση για την επεξεργασία δεδομένων προσωπικού χαρακτήρα που δεν είναι απαραίτητη για την εκτέλεση της εν λόγω σύμβασης (Άρθρο 7 παρ. 4, GDPR). Οι ελεγκτές έχουν συνεπώς αυξημένο καθήκον να εξασφαλίσουν ότι τα υποκείμενα των δεδομένων έχουν πραγματική επιλογή. Αυτό μπορεί να είναι μάλλον προβληματικό υπό ορισμένες συνθήκες, όπως στο πλαίσιο της απασχόλησης, όπου οι εργοδότες βασίζονται στη συναίνεση του υποκειμένου των δεδομένων για την επεξεργασία των προσωπικών δεδομένων των εργαζομένων. Στις περιπτώσεις αυτές και γενικά, όταν υπάρχει ανισορροπία ισχύος μεταξύ του υποκειμένου των δεδομένων και του υπεύθυνου επεξεργασίας, πρέπει να αποφεύγεται η συναίνεση και η επεξεργασία πρέπει να βασίζεται σε άλλη νομική βάση.

Όσον αφορά την προστασία δεδομένων στο πλαίσιο της ρύθμισης των χρηματοπιστωτικών υπηρεσιών της ΕΕ, ο Ευρωπαϊός Επόπτης Προστασίας Δεδομένων (ΕΕΠΔ) έχει ήδη επισημάνει ότι μολονότι η συναίνεση είναι προφανής νομικός λόγος, ενδέχεται να μην είναι πάντοτε κατάλληλος λόγος των όρων που πρέπει να πληρούνται να είναι έγκυρη. Όσον αφορά τη διαφάνεια των περιουσιακών στοιχείων των οφειλετών, ο ΕΕΠΔ ισχυρίστηκε ότι η νόμιμη επεξεργασία πληροφοριών σχετικά με τα περιουσιακά στοιχεία των οφειλετών δεν πρέπει να βασίζεται στη συναίνεση, αλλά στην τήρηση μιας νομικής υποχρέωσης ή της εκπλήρωσης δημοσίου συμφέροντος. Όσον αφορά τα μέτρα για την καταπολέμηση της νομιμοποίησης εσόδων από παράνομες δραστηριότητες, ο ΕΕΠΔ πρότεινε ότι η αναγκαιότητα συμμόρφωσης με μια νομική υποχρέωση θα μπορούσε να χρησιμεύσει ως κατάλληλη νομική βάση. Επιπλέον, τα μέτρα που αποσκοπούν στην αύξηση της διαφάνειας στις χρηματοπιστωτικές αγορές πρέπει να ξεετάζουν την εκτέλεση ενός καθήκοντος προς το δημόσιο συμφέρον και όχι τη συναίνεση (ΕΕΠΔ Κατευθυντήριες γραμμές για την προστασία δεδομένων στο πλαίσιο της ρύθμισης των χρηματοπιστωτικών υπηρεσιών της ΕΕ,

26.11.2014, p. 37) . Ο ΕΕΠΔ έκρινε ότι η συγκατάθεση μπορεί να είναι ενδεδειγμένη σε περιπτώσεις έκτακτης ανάγκης ad hoc, όπου δεν υπάρχει πιθανότητα άσκησης υπερβολικής πίεσης στο υποκείμενο των δεδομένων ή ως πρόσθετο επίπεδο προστασίας για ιδιαίτερα εμπιστευτικές πληροφορίες.

### Δικαίωμα στη λήθη (δικαίωμα διαγραφής)

Υποστηριζόμενο ως ένας από τους κύριους κανόνες του κανονισμού που αποσκοπεί στην ενίσχυση των προσώπων στα οποία αναφέρονται τα δεδομένα, το "δικαίωμα να ξεχαστεί" είναι στην πραγματικότητα ένα ενισχυμένο δικαίωμα διαγραφής, όπως το γνωρίζουμε ήδη βάσει της οδηγίας 95/46 / ΕΚ (Άρθρο 12b, Directive 95/46/EC). Αυτό το υφιστάμενο δικαίωμα είναι, ωστόσο, "ακονισμένο" στο πλαίσιο του GDPR, ώστε να επιτρέπεται στα υποκείμενα των δεδομένων καλύτερη διαχείριση των κινδύνων προστασίας δεδομένων σε απευθείας σύνδεση. Σύμφωνα με την οδηγία 95/46 / ΕΚ, τα υποκείμενα των δεδομένων μπορούν να ζητήσουν τη διαγραφή δεδομένων που δεν συμμορφώνονται με τις διατάξεις της παρούσας οδηγίας, ιδίως λόγω της ελλιπούς ή ανακριβούς φύσης των δεδομένων. Ο κανονισμός επεκτείνει τις περιπτώσεις κατά τις οποίες το άτομο μπορεί να ζητήσει διαγραφή σε καταστάσεις κατά τις οποίες τα δεδομένα δεν είναι πλέον απαραίτητα σε σχέση με τους σκοπούς, όταν το πρόσωπο στο οποίο αναφέρονται τα δεδομένα αποσύρει τη συγκατάθεση ή αντιτίθεται στη μεταποίηση, όταν τα δεδομένα προσωπικού χαρακτήρα έχουν υιοθετήσει παράνομη επεξεργασία ή πρέπει διαγράφονται σύμφωνα με μια νομική υποχρέωση καθώς και όταν έχουν συλλεχθεί σε σχέση με την προσφορά υπηρεσιών της κοινωνίας της πληροφορίας σε ένα παιδί (Άρθρο 17 παρ. 1, GDPR).

Το Δικαστήριο της Ευρωπαϊκής Ένωσης επιβεβαίωσε ήδη το δικαίωμα του ατόμου να αποσυρθεί από τη λίστα στην περιφέρεια υπόθεση Google Ισπανίας (Άρθρο 17 παρ. 2, GDPR) (.80) που βασίζεται στην οδηγία 95/46 / ΕΚ. Σύμφωνα με την απόφαση του Δικαστηρίου, ένα άτομο μπορεί, υπό ορισμένες περιστάσεις, να ζητήσει την απομάκρυνση των συνδέσμων στις προσωπικές πληροφορίες που εμφανίζονται από τις μηχανές αναζήτησης μετά από αναζήτηση που βασίζεται στο όνομα του συγκεκριμένου προσώπου. Το GDPR εισάγει τώρα μια εντελώς νέα διάταξη που υποχρεώνει τους ελεγκτές, οι οποίοι υποχρεούνται να διαγράψουν τα προσωπικά δεδομένα που έχουν δημοσιοποιήσει, να ενημερώσουν τους άλλους ελεγκτές που επεξεργάζονται τα προσωπικά δεδομένα "για το αίτημα του υποκειμένου των δεδομένων σχετικά με τη διαγραφή οποιουδήποτε συνδέσμου ή αντιγράφου ή την αναπαραγωγή αυτών των προσωπικών δεδομένων" (Άρθρο 17 παρ. 2, GDPR) . Παρόλο που η διάταξη μπορεί να θεωρηθεί επιβαρυντική για τους υπεύθυνους επεξεργασίας δεδομένων, είναι σημαντικό να υπογραμμιστεί ότι αυτό δεν αποτελεί υποχρέωση αποτελέσματος, αλλά μόνο ένα από τα μέσα «να λάβουμε όλα τα εύλογα μέτρα, συμπεριλαμβανομένων των τεχνικών μέτρων». Επομένως, οι υπεύθυνοι επεξεργασίας δεν χρειάζεται να εξασφαλίσουν την πραγματική διαγραφή των συνδέσμων, των αντιγράφων ή της αναπαραγωγής των δεδομένων, αλλά πρέπει να ενημερώσουν τους άλλους ελεγκτές σχετικά με το αίτημα του υποκειμένου των δεδομένων. Το "δικαίωμα να ξεχαστεί" είναι ως εκ τούτου κάπως αραιωμένο και επομένως πιο εφικτό όσον αφορά την πιθανή εφαρμογή του.

Είναι εξίσου σημαντικό να επισημάνουμε ότι το δικαίωμα διαγραφής σύμφωνα με το GDPR δεν είναι απόλυτο δικαίωμα. Το GDPR απαριθμεί τις περιπτώσεις κατά τις οποίες το δικαίωμα αυτό δεν θα εφαρμόζεται στο βαθμό που η επεξεργασία είναι απαραίτητη για την άσκηση της ελευθερίας έκφρασης και ενημέρωσης, για την τήρηση νομικής υποχρέωσης, για λόγους δημοσίου συμφέροντος στον τομέα της δημόσιας υγείας κλπ(Άρθρο 17 παρ. 3, GDPR).

### Φορητότητα

Όσο περισσότεροι χρήστες δεδομένων τροφοδοτούν ένα σύστημα, μια πλατφόρμα ή μια υπηρεσία, τόσο πιο δύσκολο γίνεται για αυτούς να μεταναστεύσουν σε άλλο. Αυτό το φαινόμενο είναι γενικά γνωστό ως φαινόμενο κλειδώματος. Οι ιστότοποι κοινωνικής δικτύωσης αποτελούν ένα καλό παράδειγμα. Οι χρήστες γενικά ξοδεύουν πολύ χρόνο προσθέτοντας τους φίλους και την οικογένειά τους, μεταφορτώνοντας τις εικόνες, τα αγαπημένα τους βιβλία και μουσική κλπ. Η χρονοβόρα διαδικασία προσαρμογής του προφίλ σας καθιστά πιθανή τη μετάβαση σε άλλη υπηρεσία που προσφέρει καλύτερη εξυπηρέτηση πελατών ή πιο ελκυστικά χαρακτηριστικά. Η σύγχρονη ηλεκτρονική τραπεζική συγκρίθηκε με ιστότοπους κοινωνικής δικτύωσης, καθώς περιλαμβάνει την επιλογή προσαρμογής και εξατομίκευσης της διασύνδεσης, με τρόπο παρόμοιο με αυτόν του προφίλ κοινωνικής δικτύωσης. Οι τράπεζες, όπως και άλλοι φορείς εκμετάλλευσης πλατφορμών, έχουν ήδη ξεκινήσει στρατηγικές υλοποίησης για κερδοσκοπικά κέρδη για το όφελος αυτής της δέσμευσης.

Με την εισαγωγή του δικαιώματος στη φορητότητα δεδομένων, το GDPR θα είναι σε θέση να αντισταθμίσει ακριβώς αυτά τα φαινόμενα επιτρέποντας μια αυξημένη «κινητικότητα» των δεδομένων και κατά συνέπεια μεγαλύτερη ευελιξία για τα υποκείμενα των δεδομένων. Ο κανονισμός προβλέπει ρητώς ότι το υποκείμενο των δεδομένων έχει το δικαίωμα να λαμβάνει τα προσωπικά δεδομένα που τον αφορούν σε δομημένη, ευρέως χρησιμοποιούμενη και μηχανικά αναγνώσιμη μορφή και έχει το δικαίωμα να διαβιβάσει τα δεδομένα αυτά σε άλλο υπεύθυνο επεξεργασίας (Άρθρο 20, GDPR). Για το δικαίωμα υποβολής αίτησης, το υποκείμενο των δεδομένων πρέπει να έχει παράσχει τα δεδομένα στον εαυτό του υπεύθυνο επεξεργασίας, η επεξεργασία πρέπει να βασίζεται σε συγκατάθεση ή σύμβαση και να πραγματοποιείται με αυτοματοποιημένα μέσα.

Το δικαίωμα μεταφοράς δεδομένων εξακολουθεί να προκαλεί προβλήματα. Έτσι, για παράδειγμα, δεν είναι σαφές τι σημαίνει ότι τα δεδομένα πρέπει να έχουν παρασχεθεί από το υποκείμενο των δεδομένων στον υπεύθυνο επεξεργασίας. Επιπλέον, το GDPR δηλώνει ότι το υποκείμενο των δεδομένων έχει το δικαίωμα να διαβιβάσει τα δεδομένα απευθείας από έναν ελεγκτή σε άλλο, όταν αυτό είναι τεχνικά εφικτό. Αυτό δημιουργεί φυσικά το ζήτημα της αναγκαιότητας μιας διαλειτουργικότητας των συστημάτων (Αιτία 68, GDPR), η οποία, όπως αποδεικνύεται και από τη σύσταση στις αιτιολογικές σκέψεις για τους υπεύθυνους επεξεργασίας δεδομένων για την ανάπτυξη διαλειτουργικών μορφών, δεν υπάρχει προς το παρόν.

Σε κάθε περίπτωση, είναι φανερό ότι στο εγγύς μέλλον, στο πλαίσιο του GDPR, κάποιος θα μπορούσε να ζητήσει τη φορητότητα ενός αριθμού IBAN λίγο πολύ με τον ίδιο τρόπο που κάποιος είναι πλέον σε θέση να ζητήσει τη φορητότητα ενός αριθμού κινητού τηλεφώνου από έναν πάροχο αλλο. Θα μπορούσε επίσης να προβλεφθεί η δυνατότητα να γίνει χρήση του νέου δικαιώματος ζητώντας να μεταφερθεί ο κατάλογος των αποταμιευθέντων δικαιούχων προσωπικών πληρωμών από μια τράπεζα σε άλλη. Ωστόσο, πρέπει να δούμε πώς θα ξετυλίγεται αυτό το νέο δικαίωμα στο πλαίσιο του GDPR.

*Αυτοματοποιημένη λήψη αποφάσεων, δημιουργία προφίλ και δικαίωμα ένστασης* Σήμερα, αναμφισβήτητα αντιμετωπίζουμε μια αυξανόμενη εξατομίκευση προϊόντων και υπηρεσιών που βασίζονται σε αναλύσεις δεδομένων. Είναι γνωστό ότι, στον χρηματοπιστωτικό τομέα, οι τράπεζες και οι προμηθευτές πληροφοριών για τις πιστώσεις χρησιμοποιούν προσωπικές πληροφορίες για ανάλυση κινδύνου, βάσει των οποίων ταξινομούν τους δυνητικούς τους πελάτες. Μια τέτοια ταξινόμηση θα επηρεάσει τα ασφάλιστρα ή τα επιτόκια των σημερινών ή δυνητικών πελατών ανάλογα με το προφίλ κινδύνου τους ή μπορεί ακόμη και να οδηγήσει στην άρνηση σύναψης σύμβασης βάσει του προφίλ υψηλού κινδύνου ενός ατόμου (.81).

Το GDPR προβλέπει, διατηρώντας την προσέγγιση της οδηγίας 95/46 / ΕΚ, ότι τα υποκείμενα των δεδομένων έχουν το δικαίωμα να μην αξιολογούνται μόνο με βάση την αυτοματοποιημένη επεξεργασία των προσωπικών τους δεδομένων (Άρθρο 22 παρ. 1, GDPR). Οι αποφάσεις που παράγουν έννομες συνέπειες για το πρόσωπο στο οποίο αναφέρονται τα δεδομένα ή που τον επηρεάζουν σημαντικά δεν πρέπει να βασίζονται σε αυτοματοποιημένη επεξεργασία που αποσκοπεί στην αξιολόγηση προσωπικών στοιχείων όπως η απόδοση στην εργασία, η φερεγγυότητα, η αξιοπιστία, η συμπεριφορά κ.λπ. ο χαρακτηρισμός ακολουθεί το ίδιο νομικό καθεστώς.

Ο κανονισμός διατηρεί επίσης τις εξαιρέσεις της οδηγίας, σύμφωνα με τις οποίες οι αποφάσεις αυτές επιτρέπονται όταν λαμβάνονται κατά τη σύναψη ή την εκτέλεση μιας σύμβασης ή όταν επιτρέπονται από το νόμο. Επιπλέον, το GDPR δηλώνει ότι οι αυτοματοποιημένες αποφάσεις επιτρέπονται επίσης εάν το υποκείμενο των δεδομένων έχει συναινέσει ρητά. Σε μια τέτοια περίπτωση, ωστόσο, ο υπεύθυνος επεξεργασίας πρέπει να εφαρμόσει τις κατάλληλες διασφαλίσεις που επιτρέπουν στα υποκείμενα των δεδομένων να είναι τουλάχιστον σε θέση να αποκτήσουν ανθρώπινη παρέμβαση, να εκφράσουν την άποψή τους και να αμφισβητήσουν την απόφαση (Άρθρο 22 παρ. 3, GDPR). Ωστόσο, οι αυτοματοποιημένες αποφάσεις δεν πρέπει να βασίζονται σε ειδικές κατηγορίες δεδομένων.

Όσον αφορά το δικαίωμα αντίρρησης, το GDPR διατηρεί το γενικό δικαίωμα υποβολής αντιρρήσεων του υποκειμένου των δεδομένων βάσει της ιδιαίτερης κατάστασής του. Εντούτοις, αντιστρέφει το βάρος της απόδειξης, προβλέποντας ότι - προκειμένου να συνεχίσει η επεξεργασία, ο

υπεύθυνος της επεξεργασίας πρέπει να αποδείξει - διαφορετικά από την κατάσταση που προβλέπει η οδηγία 95/46 / ΕΚ - επιβλητικά θεμιτά δικαιώματα που υπερισχύουν των δικαιωμάτων του υποκειμένου των δεδομένων, ή ότι η επεξεργασία είναι απαραίτητη σε σχέση με την άσκηση ή την υπεράσπιση νομικών αξιώσεων (Άρθρο 21 παρ. 1, GDPR). Επιπρόσθετα, το GDPR εισάγει δικαίωμα επίκλησης με βάση την ιδιαίτερη κατάσταση του υποκειμένου των δεδομένων, όταν τα δεδομένα υποβάλλονται σε επεξεργασία για επιστημονικούς ή ιστορικούς ερευνητικούς σκοπούς ή για στατιστικούς σκοπούς σύμφωνα με το άρθρο 89 παράγραφος 1 του GDPR (Άρθρο 21 παρ. 6, GDPR). Το υποκείμενο των δεδομένων διατηρεί το δικαίωμά του να αντιτίθεται ανά πάσα στιγμή στην επεξεργασία για λόγους άμεσης εμπορίας (Άρθρο 21 παρ. 2, GDPR). Οι υπεύθυνοι επεξεργασίας υποχρεούνται να ενημερώνουν τα υποκείμενα των δεδομένων σχετικά με το δικαίωμα να διατυπώσουν αντιρρήσεις "το αργότερο κατά τη στιγμή της πρώτης επικοινωνίας με το υποκείμενο των δεδομένων", με τρόπο σαφή και χωριστά από κάθε άλλη πληροφορία (Άρθρο 21 παρ. 4, GDPR).

### The GDPR και ο Τραπεζικός τομέας

Το GDPR προσθέτει ένα επιπρόσθετο επίπεδο υποχρεώσεων στο ήδη πολύπλοκο και εξαιρετικά ρυθμισμένο τραπεζικό περιβάλλον. Εκτός από τη νομοθεσία για την προστασία των δεδομένων, οι ελβετικές τράπεζες υπόκεινται σε άλλους κανονισμούς με ισχυρές επιπτώσεις στις προσωπικές τους πληροφορίες, όπως οι κανονισμοί της εποπτικής αρχής της χρηματοπιστωτικής αγοράς σχετικά με τα δεδομένα πελατών (δηλ. Εγκύκλιος FINMA 2008/21 παράρτημα 3), καθήκοντα Know-Your-Customer (KYC) (MiFID II) ή την αυτόματη ανταλλαγή πληροφοριών (ΑΕΟΙ) σε φορολογικά θέματα.

Οι υποχρεώσεις τεκμηρίωσης και αρχειοθέτησης αποτελούν άλλα καθήκοντα με διάσταση προσωπικού χαρακτήρα για τις τράπεζες. Ως εκ τούτου, η ευθυγράμμιση των ρυθμιστικών πρωτοβουλιών είναι το κλειδί για την εξοικονόμηση κόστους και προσπάθειας κατά την προετοιμασία με μια στέρεη βάση για τους προσεχείς κανονισμούς στον τομέα της προστασίας δεδομένων και πέραν αυτών.

Η συμμόρφωση με το GDPR, τηρώντας όλες τις άλλες υποχρεώσεις, αποτελεί σημαντική πρόκληση για τις τράπεζες εντός και εκτός της ΕΕ. Ωστόσο, η συμμόρφωση θα αποτελέσει προϋπόθεση για την εξυπηρέτηση των πελατών στην ΕΕ. Η μη συμμόρφωση θα εκθέσει τις τράπεζες που επηρεάζονται από το GDPR σε σημαντικούς νομικούς κινδύνους και ενδέχεται τελικά να προκαλέσει βαριά διοικητικά πρόστιμα.

### *Συνέπειες της μη συμμόρφωσης με το GDPR*

Όπως έχει ήδη αναφερθεί, το GDPR παρέχει στα DPA νέα δικαιώματα ελέγχου και επιβολής διοικητικών προστίμων που μπορούν να ανέλθουν σε 20 εκατ. Ευρώ κατ' ανώτατο όριο ή 4% του συνολικού ετήσιου κύκλου εργασιών μιας επιχείρησης - όποια από τις δύο είναι υψηλότερη (.78).

Σε κάθε περίπτωση, η μη συμμόρφωση με το GDPR από τις τράπεζες είναι πιθανό να προκαλέσει σημαντική ζημιά στη φήμη, ιδίως σε πελάτες στην ΕΕ. Η άμεση συμμόρφωση της Google με την Απόφαση της Ισπανίας - η επέκταση της εμβέλειας της Οδηγίας της ΕΕ για την Προστασία των Δεδομένων το 2014 - υποδεικνύει ότι ακόμη και οι μεγάλες διεθνείς εταιρείες προτιμούν γενικά να συμμορφώνονται με τις προειδοποιήσεις για την επιβολή της προστασίας δεδομένων, καθιστώντας έτσι άσχετο το θέμα της πρακτικής επιβολής.

Τέλος, ενώ η διαδικασία επιβολής διοικητικών προστίμων για μη συμμόρφωση ενδέχεται να παραμείνει αβέβαιη και θα πρέπει να αξιολογηθεί κατά περίπτωση, οι απαιτήσεις των ιδιωτών πιθανόν να αντιμετωπίζονται διαφορετικά. Βάσει του ιδιωτικού διεθνούς δικαίου, η εφαρμογή του GDPR κατά τη διαδικασία κατά μιας τράπεζας είναι σχεδόν βέβαιο αν ο κάτοικος της ΕΕ επηρεάζει το GDPR ως το εφαρμοστέο δίκαιο. Αυτή η επιλογή είναι επίσης διαθέσιμη στους κατοίκους της ΕΕ που αναλαμβάνουν δράση στα κρατικά δικαστήρια. Ενώ η εφαρμογή του νόμου για την προστασία των δεδομένων στο εξωτερικό δεν είναι ιδιαίτερα νέα, δεδομένου του πολύ ευρύτερου πεδίου εφαρμογής και των νέων υποχρεώσεων που απορρέουν από το GDPR σε σύγκριση με τη νομοθεσία του κράτους, η μη συμμόρφωση με το GDPR συνεπάγεται υψηλό νομικό και χρηματοοικονομικό κίνδυνο (.78).

### *Προσδιορισμός των επιπτώσεων της GDPR στην τραπεζική*

Όσον αφορά τη συμμόρφωση και την εφαρμογή, το GDPR επηρεάζει την τραπεζική με διαφορετικό τρόπο ανάλογα με τον επιχειρηματικό τομέα. Όσον αφορά τα δεδομένα πελατών, εξαρτάται σε μεγάλο βαθμό από τον τομέα πελατών, τη βάση αλληλεπίδρασης και τις προσφερόμενες υπηρεσίες. Όλα αυτά καθορίζουν τη δομή των προσωπικών δεδομένων πελατών που υποβάλλονται σε επεξεργασία από μια τράπεζα. Η συμμόρφωση με το GDPR είναι γενικά ευκολότερη όταν η επεξεργασία δεδομένων έχει ως αποτέλεσμα καλά οργανωμένα προσωπικά δεδομένα. Επιπλέον, όσο μεγαλύτερη είναι η τράπεζα, τόσο πιο σύνθετη θα είναι η δομή των δεδομένων ως συνάρτηση της νομικής της οργάνωσης (οντότητες), της ποικιλομορφίας των λειτουργιών και των ροών δεδομένων μεταξύ και μεταξύ των τμημάτων, όπως αναγράφονται στον πίνακα 4 (.78).

Πίνακας 4

	Σχετιζόμενες με Φυσικό Πρόσωπο - Πελάτη		Σχετιζόμενες με Μη φυσικό Πρόσωπο - Πελάτη		
	Ιδιωτική Περιουσία	Λιανική Τραπεζική	Εταιρικοί Πελάτες	Διαχείριση περιουσιακών	Επενδυτική Τραπεζική
<b>Τύπος πελάτη</b>	πελάτες υψηλού προφίλ - VIP	Λιανικοί πελάτες	Εκπρόσωποι των θεσμικών πελατών, επωφελής	Εκπρόσωποι των θεσμικών πελατών, πραγματικοί δικαιούχοι	Εκπρόσωποι των θεσμικών πελατών, πραγματικοί δικαιούχοι
<b>Αλληλεπίδραση πελάτη</b>	Πολλές άμεσες αλληλεπιδράσεις, συμπεριλαμβανομένων ευαίσθητων δεδομένων.	Τυποποιημένος τρόπος άμεσης αλληλεπίδρασης πελάτη, όλο και πιο αυτοματοποιημένος, τυποποιημένος και ψηφιοποιημένος	Πολύ λίγες και τυποποιημένες πληροφορίες σχετικά με τα φυσικά πρόσωπα	Οι περισσότερες τυποποιημένες και περιορισμένες πληροφορίες σχετικά με φυσικούς πελάτες.	Οι περισσότερες τυποποιημένες και περιορισμένες πληροφορίες σχετικά με φυσικούς πελάτες.
<b>Υπηρεσίες και προϊόντα που προσφέρονται</b>	Μεγάλη ποικιλία προσαρμοσμένων υπηρεσιών και προϊόντων	Τυποποιημένες υπηρεσίες και προϊόντα	Καμία υπηρεσία ή προϊόν σε πελάτες φυσικών προσώπων	Καμία υπηρεσία ή προϊόν σε πελάτες φυσικών προσώπων	Καμία υπηρεσία ή προϊόν σε πελάτες φυσικών προσώπων
<b>Τύπος δεδομένων</b>	Προσωπικά δεδομένα και ειδικές κατηγορίες προσωπικών δεδομένων (ευαίσθητες πληροφορίες)	Προσωπικά δεδομένα (δυστηκτικά περιορισμένα ευαίσθητα δεδομένα)	Προσωπικά δεδομένα	Προσωπικά δεδομένα	Προσωπικά δεδομένα
<b>Δομή δεδομένων</b>	Πολλά μη δομημένα δεδομένα	Δομημένα και αδόμητα δεδομένα	Δομημένα δεδομένα για φυσικά πρόσωπα	Δομημένα δεδομένα για φυσικά πρόσωπα	Δομημένα δεδομένα για φυσικά πρόσωπα
<b>Αναμενόμενο αντίκτυπο από GDPR</b>	<b>Υψηλό</b>	<b>Μεσαίο</b>	<b>Χαμηλό</b>	<b>Χαμηλό έως μέτριο</b>	<b>Χαμηλό έως μέτριο</b>

#### *Επίδραση του GDPR σε διάφορους τραπεζικούς τομείς*

Καθώς το GDPR αφορά μόνο τα φυσικά πρόσωπα, οι τομείς δραστηριότητας που συνδέονται άμεσα με τον πελάτη, όπως η διαχείριση του πλούτου και η λιανική κατάταξη, επηρεάζονται περισσότερο από τον κανονισμό από ό, τι άλλα τμήματα. Ενώ το επίκεντρο αυτής της Λευκής Βίβλου είναι η προστασία των δεδομένων των πελατών, είναι σημαντικό να θυμόμαστε ότι το GDPR ισχύει για την επεξεργασία προσωπικών δεδομένων οποιουδήποτε κατοίκου της ΕΕ. Έτσι, μπορεί να αφορά όλα τα είδη προσωπικών δεδομένων που μπορούν να βρεθούν μέσα σε μια τράπεζα (.78) .

Κατά κανόνα, η διαχείριση του πλούτου με το υψηλό επίπεδο των εξατομικευμένων υπηρεσιών και την αλληλεπίδραση των πελατών θα επηρεαστεί περισσότερο από το GDPR. Ένα βασικό χαρακτηριστικό του επιχειρησιακού μοντέλου διαχείρισης περιουσίας και ενός πακέτου



αισθητικής συμπεριφοράς είναι η ευρεία επιλογή πρόσθετων υπηρεσιών, για παράδειγμα υπηρεσίες θυρωρού. Αυτές οι υπηρεσίες παρέχονται από τα τμήματα της συμβουλευτικής και του χαρτοφυλακίου διαχείρισης πελατών. Αναπόφευκτα, συλλέγονται σημαντικά ποσά ευαίσθητων δεδομένων για την παροχή τέτοιων πρόσθετων υπηρεσιών. Αυτό περιλαμβάνει πληροφορίες σχετικά με την οικογένεια ενός πελάτη, τις διαδρομές ταξιδιού ή την κατάσταση υγείας. Επιπλέον, είναι μια αβεβαιότητα στον τραπεζικό τομέα ότι τα περιουσιακά στοιχεία ενός διαχειριστή σχέσεων είναι πληροφορίες σχετικά με τον πελάτη και το περιβάλλον του πελάτη που συλλέγονται προκειμένου να εξυπηρετούνται καλύτερα και να ανταποκρίνονται στις προσδοκίες του πελάτη. Εδώ, οι τράπεζες αντιμετωπίζουν δύσκολα νομικά ζητήματα, ιδίως όσον αφορά ειδικές κατηγορίες προσωπικών δεδομένων, όπως η υγεία, η θρησκεία ή η πολιτική αλληλεγγύη (.78).

Κατά συνέπεια, η τράπεζα αντιμετωπίζει τρία βασικά ζητήματα στον τομέα διαχείρισης περιουσίας:

1. Επί του παρόντος, στις περισσότερες τράπεζες, τα δεδομένα κατανέμονται σε διάφορα συστήματα και διατηρούνται σε διάφορους αποθηκευτικούς χώρους. Για να συμμορφωθούν με το GDPR, οι τράπεζες πρέπει να κατανοήσουν πού και πώς τα δεδομένα συλλαμβάνονται και αποθηκεύονται επί του παρόντος και πρέπει να απαντήσουν στην ερώτηση σχετικά με τον τρόπο δομής των διαδικασιών για το μέλλον, προκειμένου να έχουν μια εύκολα προσιτή ολιστική άποψη των δεδομένων των πελατών.

2. Θεωρείται ότι οι διαχειριστές πλούτου συγκεντρώνουν επί του παρόντος όλα τα είδη πληροφοριών πελατών, ακόμα και ευαίσθητα δεδομένα, συχνά μη δομημένα. Οι τράπεζες πρέπει να αποφασίσουν πώς να αντιμετωπίσουν τη συλλογή και επεξεργασία δεδομένων προχωρώντας, διατηρώντας παράλληλα μια συνολική εικόνα των δεδομένων. Οι τράπεζες μπορούν να εφαρμόσουν ένα σαφές πλαίσιο πολιτικής και περαιτέρω καθοδήγηση στους κανόνες συμπεριφοράς για τους διαχειριστές σχέσεων με πελάτες για να μετριάσουν τον κίνδυνο παραβίασης του GDPR.

3. Πώς να χειριστείτε τα συλλεχθέντα προσωπικά δεδομένα που σχετίζονται με το περιβάλλον του πελάτη, όπως πληροφορίες σχετικά με τους φίλους και την οικογένεια του πελάτη;

Λόγω της σημασίας της διασυνοριακής διαχείρισης πλούτου για τις τράπεζες, είναι επίσης σημαντικό να αντιμετωπιστεί η διαβίβαση δεδομένων προσωπικού χαρακτήρα. Αυτό θα συνεπάγεται αλληλεπίδραση με άλλα τμήματα της τράπεζας, κυρίως την επενδυτική τραπεζική στην πλευρά των προϊόντων.

Η λιανική τραπεζική αντιμετωπίζει παρόμοιες προκλήσεις με τη διαχείριση του πλούτου σε σχέση με την τακτική αλληλεπίδραση με φυσικά πρόσωπα. Ωστόσο, σε αντίθεση με την τελευταία, η λιανική τραπεζική - χάρη στα κανάλια εξυπηρέτησης της - ωφελείται από τη συλλογή προσωπικών δεδομένων, η οποία είναι πολύ πιο τυποποιημένη και λιγότερο επαναλαμβανόμενη ή εξατομικευμένη. Η ψηφιοποίηση αναμένεται να συμβάλει στην τυποποίηση, ιδίως όσον αφορά τον τρόπο επεξεργασίας

των προσωπικών δεδομένων στη λιανική τραπεζική. Τέλος, οι λιανικές πελάτες των τραπεζών στην πλειονότητα των περιπτώσεων έχουν την κατοικία τους στην ενίοτε χώρα και, κατά συνέπεια, υπερβαίνουν το πεδίο εφαρμογής του GDPR (.78).

Για να συμμορφωθούν με τις GDPR, οι τράπεζες ίσως χρειαστεί να επανασχεδιάσουν ουσιαστικά τις επιχειρηματικές διαδικασίες, ιδίως στους τομείς της λιανικής τραπεζικής και της διαχείρισης περιουσίας. Αυτό είναι πιθανόν να έχει σημαντικό αντίκτυπο στη διαχείριση του κύκλου ζωής ενός πελάτη για μια τράπεζα, δηλαδή στις διάφορες φάσεις μιας σχέσης πελάτη: αναζήτηση, επιβίβαση, εξυπηρέτηση / σταυροειδής πώληση και πώληση και τερματισμός.

Ως εκ τούτου, κάθε μέτρο εφαρμογής GDPR στο πλαίσιο των προσωπικών δεδομένων πελατών πρέπει να σχεδιαστεί προσεκτικά για να μειώσει στο ελάχιστο τις αρνητικές επιπτώσεις στις επιχειρήσεις μιας τράπεζας. Η εικόνα είναι διαφορετική στους τομείς εταιρικής τραπεζικής, διαχείρισης περιουσιακών στοιχείων και επενδυτικής τραπεζικής. Αυτές οι υποδιαίρεσεις μόνο ή κυρίως ασχολούνται με θεσμικούς πελάτες και επομένως νομικά πρόσωπα. Ως εκ τούτου, για τις περισσότερες τράπεζες, η αλληλεπίδραση με τα φυσικά πρόσωπα περιορίζεται στους εκπροσώπους των θεσμικών πελατών και των τελικών πραγματικών δικαιούχων και οι συναλλαγές αυτές βασίζονται σε τυποποιημένες ή και αυτοματοποιημένες διαδικασίες. Αυτά τα δεδομένα περιορίζονται γενικά σε στοιχεία επικοινωνίας και ταυτοποίησης και είναι σε δομημένη μορφή. Η εφαρμογή του GDPR σε αυτά τα τραπεζικά τμήματα φαίνεται να είναι λιγότερο περίπλοκη (.78).

Ανεξάρτητα από τον τραπεζικό τομέα, τα δεδομένα συναλλαγών - το μεγαλύτερο σύνολο δεδομένων που τηρούν οι τράπεζες - είναι ένα σημαντικό θέμα. Περιλαμβάνει διατραπεζικές συναλλαγές, συναλλαγές με τρίτους και διασυννοριακές συναλλαγές. Δεδομένου ότι τα δεδομένα συναλλαγών μπορούν να οδηγήσουν στην αναγνώριση ενός πελάτη, θεωρούνται δεδομένα προσωπικού χαρακτήρα σύμφωνα με το GDPR.

#### *Καταφέρνει το Νομικό πλαίσιο το στόχο του;*

Η έλευση του Big Data προκαλεί μια σειρά σοβαρών προβλημάτων. Σε πολλές δικαιοδοσίες, αυτές αντιμετωπίζονται άμεσα μέσω κανονισμών. Το ερώτημα είναι εάν ο κανονισμός για την προστασία της ιδιωτικής ζωής παρέχει επαρκή προστασία για τα άτομα, ενώ παράλληλα αναγνωρίζει τα πολλά οφέλη του Big Data για την ευρύτερη κοινωνία και τις επιβαρύνσεις που βαρύνει τις επιχειρήσεις.

#### *Ενίσχυση της διαφάνειας και της συναίνεσης*

Τα άρθρα 11 και 14 προτείνουν αυστηρότερες υποχρεώσεις διαφάνειας, (.63) απαιτώντας ότι οι πληροφορίες που απευθύνονται στα πρόσωπα στα οποία αναφέρονται τα δεδομένα πρέπει να είναι «εύκολα προσβάσιμες» και «κατανοητές» και να απαριθμούνται λεπτομερώς οι τύποι πληροφοριών

που πρέπει να παρέχουν οι υπεύθυνοι επεξεργασίας κατά τη συλλογή προσωπικών δεδομένων. Το άρθρο 4 παράγραφος 8 και το άρθρο 7 παράγραφος 1 προτείνουν αυστηρότερους ορισμούς της συγκατάθεσης διευκρινίζοντας ότι δεν πρέπει να παρέχεται μόνο ελεύθερη, συγκεκριμένη και ενημερωμένη αλλά «ρητή» - επομένως ούτε η σιωπή ούτε η αδράνεια δεν μπορούν να συνιστούν έγκυρη συναίνεση. Επιπλέον, οι ελεγκτές φέρουν το βάρος της απόδειξης ότι τα υποκείμενα των δεδομένων συμφώνησαν στην επεξεργασία των προσωπικών τους δεδομένων. Παρόλο που αυτές οι αλλαγές μπορούν να γίνουν καλά, είναι δύσκολο να φανταστούμε ότι θα ξεπεράσουν τις μακροχρόνιες ελλείψεις του ενημερωμένου μοντέλου επιλογής ή θα δημιουργήσουν μια νέα εποχή στην οποία οι καταναλωτές κατανοούν τα δικαιώματά τους και ενεργούν πάνω τους. Ο ισχυρισμός μου είναι απλός εάν είναι ριζοσπαστικός: το ενημερωμένο μοντέλο επιλογής σπάει πέρα από οποιαδήποτε ρυθμιστική επισκευή και ο μόνος τρόπος για να τον αναζωογονήσουμε είναι να αλλάξουμε τις σχετικές αγορές πληροφοριών.

### *Profiling*

Το άρθρο 20 του κανονισμού αντικαθιστά το άρθρο 15 της DPD, τις διατάξεις που έχουν άμεση σχέση με τον προσδιορισμό του προφίλ. (.63) Ωστόσο, υπάρχουν ελλείψεις στην αρχική έκδοση και παραμένουν ως επί το πλείστον στη νέα. Το άρθρο 15, παράγραφος 1, της DPD πραγματεύεται «αυτοματοποιημένες ατομικές αποφάσεις» και παρέχει στο κάθε πρόσωπο «το δικαίωμα να μην υπόκειται σε απόφαση η οποία παράγει έννομα αποτελέσματα για τον ίδιο ή τον επηρεάζει σημαντικά και η οποία βασίζεται αποκλειστικά στην αυτοματοποιημένη επεξεργασία δεδομένων να αξιολογήσει ορισμένες προσωπικές πτυχές που σχετίζονται με αυτόν, όπως η απόδοσή του στην εργασία, η φερεγγυότητα, η αξιοπιστία, η συμπεριφορά κ.λπ. " Το άρθρο 15 παράγραφος 2 προβλέπει εξαιρέσεις όταν λαμβάνεται απόφαση κατά τη σύναψη ή την εκτέλεση σύμβασης και πληρούνται ορισμένες προϋποθέσεις ή «επιτρέπονται από νόμο ο οποίος προβλέπει επίσης μέτρα για την προστασία των νόμιμων συμφερόντων του υποκειμένου των δεδομένων». Αυτή η διάταξη δεν απαγορεύει τη δημιουργία προφίλ αλλά μόνο τον τρόπο με τον οποίο εφαρμόζονται. Το άρθρο 15 φαίνεται να έχει κίνητρο από τις δίδυμες ανησυχίες ότι η αυτοματοποιημένη λήψη αποφάσεων θα μειώσει τον ρόλο των προσώπων επηρεάζοντας τις αποφάσεις που τις επηρεάζουν και ότι αυτές οι αποφάσεις θα έχουν υπερβολικό σεβασμό (όπως και αν το απλό γεγονός ότι προκύπτουν από εξελιγμένες διαδικασίες πληροφορικής τις καθιστά πιο αντικειμενικό). Τα τρία βασικά προβλήματα με το άρθρο 15 είναι το περιορισμένο πεδίο εφαρμογής του και παρέχει μόνο περιορισμένο δικαίωμα και περιορισμένη θεραπεία.

Όσον αφορά το πεδίο εφαρμογής, το άρθρο 15 παράγραφος 1 εφαρμόζεται μόνον εφόσον πληρούνται όλες οι σχετικές προϋποθέσεις. Αυτό καθιστά ανεφάρμοστο κάθε φορά που ένα άτομο ασκεί κάποιο επίπεδο επιρροής στη διαδικασία λήψης αποφάσεων. Οι ασάφειες κατά την έννοια

πολλών βασικών όρων (π.χ. «απόφαση», «σημαντικά», μόνο, «ορισμένες προσωπικές πτυχές») μπορούν περαιτέρω να περιορίσουν το πεδίο εφαρμογής της. Αυτό περιλαμβάνει τις παρεκκλίσεις του άρθρου 15, παράγραφος 2, οι οποίες επιτρέπουν την αυτοματοποιημένη λήψη αποφάσεων σχετικά με ένα πρόσωπο εάν «συνάπτονται τα κατάλληλα μέτρα (με σύμβαση ή με νόμο) για την προστασία των» νόμιμων συμφερόντων »του. Το περιορισμένο δικαίωμα που παρέχει το άρθρο 15 είναι να αντισταθεί στις αυτοματοποιημένες αποφάσεις και να επιδιώξει την ανθρώπινη παρέμβαση. Ωστόσο, αυτό δεν είναι το δικαίωμα ακύρωσης τέτοιων αποφάσεων (για παράδειγμα, εκτός εάν το πρόσωπο στο οποίο αναφέρονται τα δεδομένα συμφωνεί με αυτά). Αντίθετα, το εν λόγω δικαίωμα διατύπωσης αντιρρήσεων τίθεται σε εφαρμογή μόνον εάν το υποκείμενο των δεδομένων γνωρίζει ότι υπόκειται σε τέτοιες αποφάσεις και υποβάλλει ένσταση. Όπως διαπιστώθηκε από την προηγούμενη ανάλυση, οι βλάβες στην ιδιωτική ζωή που συνδέονται με το Big Data προκύπτουν χωρίς τη γνώση ή την ευαισθητοποίηση του υποκειμένου των δεδομένων, οδηγώντας στην κριτική ότι το αόρατο της μορφοποίησης καθιστά το άρθρο 15 «Άσσο στο μανίκι». Ούτε το δικαίωμα του άρθρου 12 (α) να ανακαλύψει τη «γνώση της λογικής» των αυτοματοποιημένων διαδικασιών θεραπεύει αυτό το πρόβλημα, αφού και αυτό ωφελεί μόνο κάποιον που γνωρίζει ότι είναι προφίλ. Τέλος, για να αποκατασταθεί το ενδεχόμενο, εάν ένα πρόσωπο στο οποίο αναφέρονται τα δεδομένα παρουσιάζει διακρίσεις ή υπόκειται σε διακρίσεις, το άρθρο 15 παρέχει περιορισμένη άνεση. Το ανώτατο όριο απαιτεί από τον υπεύθυνο επεξεργασίας δεδομένων να επιβάλει κάποια ανθρώπινη κρίση σε μια απόφαση, εξετάζοντας τους παράγοντες που αποτελούν τη βάση της αυτοματοποιημένης απόφασης. Όπως σημειώνει ο Bygrave: «Ο υπεύθυνος επεξεργασίας δεν υποχρεούται ούτε να αλλάζει αυτά τα κριτήρια ούτε να τα συμπληρώνει με άλλα κριτήρια / παράγοντες» 44

Το άρθρο 20 του κανονισμού τροποποιεί το άρθρο 15 με διάφορους τρόπους. Για παράδειγμα, χαρακτηρίζει την αυτοματοποιημένη διαδικασία λήψης αποφάσεων ευρύτερα, προσφέρει μια νέα εξαίρεση βάσει συναίνεσης, απαγορεύει την αυτοματοποιημένη επεξεργασία βασιζόμενη αποκλειστικά σε ευαίσθητα δεδομένα και εξουσιοδοτεί την Επιτροπή να εγκρίνει κατ'εξουσιοδότηση πράξεις για να διευκρινίσει περαιτέρω τη φύση των «κατάλληλων μέτρων για τη διασφάλιση του υποκειμένου των δεδομένων έννομα συμφέροντα ». Εκτός από αυτές τις αλλαγές, ωστόσο, η κύρια ώθηση της είναι σε μεγάλο βαθμό η ίδια με την προηγούμενη έκδοση.

Παρόλο που η Hildebrandt επικρίνει το άρθρο 20, προτείνει επίσης ότι αν η υποχρέωση προειδοποίησης που αναφέρεται στο άρθρο 20 παράγραφος 4 «δεν εξαρτάται από πρόσωπο που το ζητεί, αυτό θα κατέληγε σε ριζική διαφορά με το σημερινό επίπεδο προστασίας». Επιπλέον, περιγράφει την υποχρέωση γνωστοποίησης των «προβλεπόμενων αποτελεσμάτων» ως «επαναστατικής νέας νομικής απαίτησης». Αυτό φαίνεται υπερβολικό. Πρώτον, όπως ομολογεί η Hildebrandt, η διατύπωση του άρθρου 20, παράγραφος 4, είναι διφορούμενη και μπορεί να στηρίξει την αντίθετη ερμηνεία (η εν λόγω προειδοποίηση οφείλεται μόνο εφόσον ζητηθεί). Δεύτερον,

ανεξάρτητα από το κατά πόσον οι υπάλληλοι που το ζητούν, οι υποχρεώσεις του υπεύθυνου της ελεγκτικής υπηρεσίας βάσει του άρθρου 20 παράγραφος 4 ισχύουν μόνο «στις περιπτώσεις που αναφέρονται στην παράγραφο 2», δηλαδή σε παρεκκλίσεις που συνεπάγονται τη σύναψη ή την εκτέλεση ορισμένων συμβάσεων, ή συγκατάθεση. Έτσι, οι υποχρεώσεις προειδοποίησης δεν είναι καθόλου καθολικές. Τέλος, παρόλο που η Hildebrandt ερμηνεύει το άρθρο 20 ως «αναγκάζοντας τους υπεύθυνους επεξεργασίας δεδομένων να μας ενημερώσουν για τον κίνδυνο που αντιμετωπίζουμε διαρροής των δεδομένων μας», είναι δύσκολο να καταλάβουμε γιατί τα άτομα είναι πιθανότερο να διαβάζουν ή να κατανοούν ανακοινώσεις σχετικά με την ύπαρξη και την προβλεπόμενη επίδραση του προφίλ από ό, τι έχουν σχέση με άλλες ειδοποιήσεις απορρήτου, ειδικά δεδομένης της καινοτομίας, της πολυπλοκότητας και της ασαφούς διαμόρφωσης του προφίλ στον μέσο χρήστη του Διαδικτύου.

### *Νέα ατομικά δικαιώματα*

Το άρθρο 17, το δικαίωμα να ξεχαστεί και να διαγραφεί, αποτελεί μια εξαιρετικά αμφισβητούμενη διάταξη που βασίζεται στο υφιστάμενο δικαίωμα διαγραφής δεδομένων (άρθρο 12 της DPD) και επιδιώκει να αντιμετωπίσει αποτελεσματικότερα (.63) την ιδιωτική ζωή και τις αξιόποινες βλάβες (συμπεριλαμβανομένης της βλάβης της φήμης) που σχετίζονται με τη διάδοση και, επομένως, την εμμονή των δεδομένων που έχουν ανταλλαγή οικειοθελώς στην κοινωνική δικτύωση και άλλες υπηρεσίες Web 2.0. Το άρθρο 17, παράγραφος 2, θα απαιτούσε από τους ελεγκτές να λαμβάνουν «εύλογα μέτρα, συμπεριλαμβανομένων τεχνικών μέτρων», για να ενημερώνουν τρίτους όταν ένα υποκείμενο των δεδομένων ζήτησε τη διαγραφή των προσωπικών δεδομένων που είχαν προηγουμένως δημοσιευθεί σχετικά με αυτά, αλλά αυτό θα μπορούσε να αποδειχθεί επαχθές ή και αδύνατο αριθμό σεναρίων. Επιπλέον, το δικαίωμα να ξεχαστεί δεν είναι μόνο κάπως αόριστο και πρακτικό, όπως συντάχθηκε, αλλά δημιουργεί σοβαρές και ενδεχομένως ανεπίλυτες συγκρούσεις με τα δικαιώματα της ελεύθερης έκφρασης. Για τους παρόντες σκοπούς, το βασικό σημείο είναι ότι το δικαίωμα να ξεχαστεί περιορίζεται από τους όρους του στα προσωπικά δεδομένα. Επομένως, δεν είναι καν σαφές εάν το άρθρο 17 θα εφαρμοζόταν σε προγνωστικά συμπεράσματα βασισμένα σε δεδομένα προσωπικού χαρακτήρα που μπορεί να έχουν ανωνυμοποιηθεί ή να γενικευθούν ως αποτέλεσμα αναλυτικών τεχνικών στην καρδιά του Big Data.

Το άρθρο 18, το οποίο δημιουργεί ένα νέο δικαίωμα στη φορητότητα δεδομένων, εξυπηρετεί τον ιδιαίτερα αξιότιμο στόχο που είναι να επιτρέπει στα άτομα να αποσπάσουν τα προσωπικά τους δεδομένα (προσωπικά προφίλ, φωτογραφίες, δημοσιεύσεις, λίστες επαφών κ.λπ.) από μια εφαρμογή ή υπηρεσία και να τα μεταφέρουν σε άλλο δεδομένου ότι τα δεδομένα «υποβάλλονται σε επεξεργασία με ηλεκτρονικά μέσα και σε δομημένη και συνηθισμένη μορφή». Ενώ μπορεί να είναι ανεπιθύμητο για την Επιτροπή να καθορίσει αυτές τις μορφές ή τυχόν συναφή τεχνικά πρότυπα, όπως επιτρέπονται από το άρθρο 18 παράγραφος 3, η διασφάλιση της φορητότητας των δεδομένων παραμένει ένα

κρίσιμο βήμα. Η φορητότητα των δεδομένων αποτελεί βασικό παράγοντα για την προώθηση του ανταγωνισμού μεταξύ των υφιστάμενων υπηρεσιών και για την αποτροπή του «κλειδώματος». Ταυτόχρονα, διευκολύνει σημαντικά τη δημιουργία νέων υπηρεσιών όπως οι υπηρεσίες προσωπικών δεδομένων που περιγράφονται παρακάτω.

#### *Νέες ευθύνες των διαχειριστών δεδομένων*

Το άρθρο 22 συνοψίζει τις ευθύνες των ελεγκτών, συμπεριλαμβανομένων των γενικών υποχρεώσεων, όπως η τεκμηρίωση (άρθρο 28), η ασφάλεια των δεδομένων (άρθρο 30), οι εκτιμήσεις επιπτώσεων (άρθρο 33), η προηγούμενη έγκριση ή διαβούλευση (άρθρο 34) και ο ορισμός υπευθύνου προστασίας δεδομένων 35). Επιπλέον, το άρθρο 23 παράγραφος 2 δημιουργεί μια πιο συγκεκριμένη υποχρέωση για τους ελεγκτές να εφαρμόζουν μηχανισμούς για την εξασφάλιση, εξ ορισμού, ότι πληρούνται οι απαιτήσεις ελαχιστοποίησης των δεδομένων. (Αυτή η νέα απαίτηση της προστασίας δεδομένων «από το σχεδιασμό και την προεπιλογή» είναι πολύ ελπιδοφόρα αλλά εξαρτάται σε μεγάλο βαθμό από τον τρόπο εφαρμογής της). Όλες αυτές οι νέες διατάξεις αποσκοπούν στο να διασφαλίσουν ότι οι ελεγκτές επεξεργάζονται τα δεδομένα σύμφωνα με τις βασικές αρχές της προστασίας της ιδιωτικής ζωής του κανονισμού. Αν όμως αυτές οι βασικές διατάξεις δεν ικανοποιήσουν ικανοποιητικά την Big Data, δεν είναι καθόλου σαφές ότι αυτές οι πρόσθετες υποχρεώσεις θα διορθώσουν αυτό το κενό.

Εν ολίγοις, η τάση των μεγάλων δεδομένων - περισσότερα δεδομένα, ταχύτεροι υπολογιστές και νέες αναλυτικές τεχνικές - θέτει σοβαρές προκλήσεις στον νόμο περί προστασίας δεδομένων, ο οποίος όχι μόνο απέτυχε να συμβαδίσει με τις τεχνολογικές αλλαγές αλλά είναι ακόμη πιο πιθανό να μείνει πίσω όταν αντιμετωπίζει με το Big Data. Ακόμη και ο κανονισμός, παρά τις αξιέπαινες προσπάθειες για την κάλυψη ορισμένων ελλείψεων της DPD, δεν αλλάζει αυτή την ετυμολογία. Ποιες είναι οι εναλλακτικές λύσεις;

#### *Ευέλικτες και κλιμακούμενες νομικές υποχρεώσεις*

Μία από τις σημαντικότερες καινοτομίες του νέου Κανονισμού (.64) για την προστασία των γενικών δεδομένων είναι η ενσωμάτωση της αρχής της λογοδοσίας, κάτι που έχει καθιερωθεί στο δίκαιο του ανταγωνισμού, αλλά είναι σχετικά νέο για τη νομοθεσία περί προστασίας δεδομένων. Σύμφωνα με την αρχή αυτή, οι οργανισμοί που υπόκεινται σε υποχρεώσεις προστασίας δεδομένων αναμένεται να είναι σε θέση να αποδείξουν ότι έχουν ληφθεί τα απαραίτητα μέτρα για να εξασφαλιστεί η συμμόρφωση με τους κανόνες, ενώ οι αρχές προστασίας δεδομένων παρεμβαίνουν μόνο για τον έλεγχο, ιδίως όταν υπάρχει ένδειξη αθέτησης. Οι υποχρεώσεις σε κάθε έναν από αυτούς τους τομείς είναι κλιμακωτές: όταν μια επιχείρηση, για παράδειγμα, έχει μεγαλύτερη ισχύ στην αγορά (ανησυχία για τις αρχές ανταγωνισμού) ή ισχυρότερη συμβατική θέση (προστασία του καταναλωτή) ή

είναι υπεύθυνη για πιο επικίνδυνες διαδικασίες επεξεργασίας δεδομένων , πρέπει να είναι πιο επιμελείς στα βήματα που παίρνουν για να συμμορφωθούν.

### Μελέτη περίπτωσης: Εξατομικευμένη τιμολόγηση υπό GDPR

Στο πλαίσιο Στην Παράγραφο αυτή θα μελετήσουμε την περίπτωση της εξατομικευμένη τιμολόγησης, μια μέθοδο που χρησιμοποιείτε τακτικά από πάσης φύσεως πωλητές, εταιρίες promotion και marketing, και τα αντίστοιχα τμήματα των τραπεζών. Ποιες είναι οι ιδιαιτερότητες και ιδιομορφίες της εξατομικευμένης τιμολόγησης, πώς και αν πάσχει υπό το νέο κανονισμό.

#### *Διάκριση τιμών*

Στην οικονομία, οι διακρίσεις ως προς τις τιμές περιγράφουν την πρακτική της πώλησης των ίδιων υπηρεσιών ή αγαθών σε διαφορετικές τιμές σε διαφορετικούς αγοραστές(.84). Ο βασικός λόγος για τη διάκριση των τιμών είναι η τιμολόγηση της τιμής προς τη ζήτηση των αγοραστών και η προθυμία τους να πληρώσουν, σε αντίθεση με τη σύνδεση της τιμής με το κόστος παραγωγής (.85). Η χρήση διαφορετικών σημείων τιμών για διάφορους αγοραστές απαιτεί γνώση αυτών των αγοραστών και την ικανότητά τους να πληρώνουν καθώς και της αγοράς. Αυτό απαιτεί εκτίμηση της κλίσης των αγοραστών και δυνατότητα πληρωμής για ένα προϊόν ή μια υπηρεσία. Η διάκριση ως προς τις τιμές διαιρείται γενικά σε τρεις τακτικές (.86).

**Διακρίσεις τρίτου βαθμού:** επίσης γνωστό «τιμολόγηση της ομάδας», αναφέρεται στη δραστηριότητα της ανάθεσης μια ορισμένη τιμή σε μια συγκεκριμένη ομάδα ανθρώπων (.87). Η τιμή μπορεί να αποφασιστεί με ιστορία αγοράς, γεωγραφικές τοποθεσίες, μοντέλα συμπεριφοράς, ηλικία ή ό, τι κάνει διαφορετική την ομάδα των ανθρώπων. Τυπικές καταστάσεις για τη χρήση της τιμολόγησης του ομίλου είναι οι εκπτώσεις φοιτητών και οι εκπτώσεις για ηλικιωμένους πολίτες. Ένας κοινός λόγος για την τιμολόγηση της ομάδας είναι η ευαισθησία των τιμών (.88). Επομένως, η τιμολόγηση της ομάδας χρησιμοποιεί γενικά χαρακτηριστικά των ομάδων ανθρώπων για να καθορίσει τιμές συγκεκριμένες τιμές για ορισμένες ομάδες ανθρώπων.

**Διακρίσεις δευτέρου βαθμού,** γνωστή και ως "τιμολόγηση μενού" ή " τιμολόγηση έκδοσης", αναφέρεται σε περιπτώσεις όπου οι πωλητές αποφασίζουν την τιμή σε σχέση με τον τρόπο πώλησης του προϊόντος (.89). Ο τρόπος με τον οποίο πωλείται ένα προϊόν ή μια υπηρεσία μπορεί να αναφέρεται σε ειδικά διαμορφωμένα πακέτα για την αγορά όγκων ή τιμών εισιτηρίων που αλλάζουν ανάλογα με το χρόνο που θα χρειαστεί να χρησιμοποιηθεί (.90).

**Διακρίσεις πρώτου βαθμού:** ή "εξατομικευμένη τιμολόγηση" (.91). Οι εξατομικευμένες τιμές είναι οι τιμές που έχουν καθοριστεί για ένα συγκεκριμένο άτομο για ένα συγκεκριμένο αντικείμενο ή υπηρεσία. Οι υπολογιστές αναλύουν και κατηγοριοποιούν τη συμπεριφορά των

αγοραστών, καταλήγοντας σε συστήματα με δεδομένα πωλήσεων, λίστες πελατών και προτάσεις τιμών (.92).

### *Εξατομικευμένη τιμολόγηση σύμφωνα με το άρθρο 22 GDPR*

Ενώ η επεξεργασία δεδομένων για τη δημιουργία προφίλ πρέπει να ανταποκρίνεται στις απαιτήσεις που ορίζονται στο GDPR σχετικά με την επεξεργασία δεδομένων γενικά (Αιτία 22, GDPR), για την προστασία του άρθρου 22 (1) που πρέπει να ενεργοποιηθεί, πρέπει να ληφθεί απόφαση βασισμένη αποκλειστικά στην αυτόματη επεξεργασία, συμπεριλαμβανομένου του προφίλ που παράγει νομικές συνέπειες για το υποκείμενο των δεδομένων ή τον επηρεάζει ομοίως. Δεδομένου ότι οι εξατομικευμένες τιμές καθορίζονται αυτόματα με αλγορίθμους για την αξιοποίηση σειρών δεδομένων προσωπικού χαρακτήρα, η πρακτική θα ανταποκριθεί στην πρώτη απαίτηση της πρόβλεψης.

Όταν χρησιμοποιείται προφίλ για εξατομικευμένη τιμολόγηση, ένα προϊόν μπορεί να προσφέρεται σε εξατομικευμένη τιμή, η οποία μπορεί να είναι υψηλότερη ή χαμηλότερη από την αγοραία τιμή. Εάν το υποκείμενο των δεδομένων και ο επεξεργαστής δεδομένων (ή ο υπεύθυνος επεξεργασίας δεδομένων) συνάψουν σύμβαση, η τιμολόγηση θα έχει οδηγήσει σε νομική ισχύ. Ωστόσο, η σύμβαση δεν αποτελεί άμεσο αποτέλεσμα της επεξεργασίας δεδομένων. Συνεπώς, η επεξεργασία δεν έχει νομική ισχύ για το υποκείμενο των δεδομένων. Επομένως, το ζήτημα είναι εάν η εξατομικευμένη τιμολόγηση επηρεάζει σημαντικά το υποκείμενο των δεδομένων.

Παρόλο που δεν είναι πανομοιότυπο, είναι δυνατό να συγκριθεί η εξατομικευμένη τιμολόγηση και η κατευθυνόμενη διαφήμιση. Η διαπίστωση του εάν το προφίλ που χρησιμοποιείται για εξατομικευμένη τιμολόγηση έχει παρόμοια σημαντική επίδραση στο υποκείμενο των δεδομένων, επομένως, ποικίλλει σε μεγάλο βαθμό ανάλογα με την κατάσταση. Ένα προφίλ μπορεί να αποκαλύψει ότι ένα άτομο έχει μικρότερο προϋπολογισμό. Αυτό θα μπορούσε τότε να οδηγήσει στο να προσφερθεί στο πρόσωπο αυτό αντικείμενα σε τιμές χαμηλότερες από τις τιμές αγοράς, να τονώσει τις αγοραστικές του συνήθειες και να αναπτύξει τις επιχειρηματικές σχέσεις. Εάν οι τιμές ορισμένων ειδών αυξήθηκαν αφού ενθάρρυναν την αγορά και δημιούργησαν μια σχέση, η πρακτική θα ήταν ενοχλητική, κάτι που το άτομο δεν περίμενε και λόγω της οικοδόμησης εμπιστοσύνης, το άτομο θα ήταν ευάλωτο. Επομένως, ο καθορισμός του κατά πόσον κάτι έχει παρόμοια σημαντική επίδραση εξαρτάται από την εν λόγω εικόνα και την τιμολόγηση. Αυτό θα μπορούσε ενδεχομένως να έχει παρόμοια σημαντική επίδραση.

Οι εξαιρέσεις του άρθρου 22 (2): Εάν πληρούνται οι απαιτήσεις του άρθρου 22 (1), υπάρχουν εξαιρέσεις στο άρθρο 22 (2) του GDPR. Το άρθρο 22, παράγραφος 2, επιτρέπει την αυτοματοποιημένη λήψη αποφάσεων, συμπεριλαμβανομένης της διαμόρφωσης προφίλ, α) είναι



αναγκαία για τη σύναψη σύμβασης, β) έχει εγκριθεί από το δίκαιο της ΕΕ ή γ) υπάρχει ρητή συγκατάθεση από το υποκείμενο των δεδομένων.

Εάν η εξατομικευμένη τιμολόγηση έχει νόμιμη ή παρόμοια σημαντική επίπτωση, θα ήταν δύσκολο να δικαιολογηθεί βάσει των εξαιρέσεων του άρθρου 22 παράγραφος 2. Για να είναι αναγκαία για την εκτέλεση μιας σύμβασης, η χρήση του προφίλ για εξατομικευμένη τιμολόγηση θα πρέπει να κάνει τη διαδικασία λήψης αποφάσεων (εξατομικευμένη ρύθμιση τιμών σε αυτή την περίπτωση) πιο δίκαιη ή συνεπή (Άρθρο 29 Working Party on automated decision making, p. 12.). Παρόλο που έχει θετικά αποτελέσματα, δεν είναι απαραίτητο να εξατομικεύεται η τιμολόγηση για την πώληση αντικειμένων μέσω του Διαδικτύου. Υποστηρίζεται ότι ο χαρακτηρισμός χρησιμοποιείται για τη μεγιστοποίηση των κερδών για τους πωλητές και ότι δεν καθιστά τη διαδικασία λήψης αποφάσεων πιο δίκαιη ή συνεπή. Αν στόχος του προφίλ, ωστόσο, ήταν να εξισορροπηθεί η αγοραστική δύναμη των οικογενειών χαμηλού εισοδήματος, θα ήταν δυνατόν να υποστηριχθεί ότι θα καθιστούσε τη διαδικασία λήψης αποφάσεων πιο δίκαιη ή συνεπή.

Η εξατομικευμένη τιμολόγηση δεν επιτρέπεται ούτε από τη νομοθεσία της Ένωσης ούτε από τα κράτη μέλη. Σύμφωνα με την αιτιολογική σκέψη 71, ο σκοπός θα μπορούσε να είναι η παρακολούθηση και πρόληψη της απάτης και της φοροδιαφυγής ή η διασφάλιση της ασφάλειας και της αξιοπιστίας μιας υπηρεσίας που παρέχεται από τον ελεγκτή. Δεδομένου ότι η εξατομικευμένη τιμολόγηση είναι ένας τρόπος για μια επιχείρηση να βελτιώσει τις πωλήσεις, δεν θα υπαγόταν ούτε στην εξαίρεση αυτή. Θα ήταν δυνατό να συγκεντρωθεί ρητή συγκατάθεση.

Η έλλειψη σαφών κατευθυντήριων γραμμών σχετικά με τη ρητή συγκατάθεση καθιστά την εξέταση του κατά πόσον είναι δυνατόν να γίνει πιο δύσκολη. Οι απαιτήσεις για μια "συνηθισμένη" συγκατάθεση είναι ήδη υψηλές με το GDPR. Η προσθήκη της λέξης "ρητή" συνεπώς υποδηλώνει ότι υπάρχει ακόμη μεγαλύτερη απαίτηση για να επιτραπεί. Λαμβάνοντας υπόψη τη φύση της αυτοματοποιημένης λήψης αποφάσεων και την ασυμμετρία πληροφόρησης που συνδέεται με αυτήν, θα ήταν φυσικό για το υποκείμενο των δεδομένων να γνωρίζει ακόμη περισσότερο τους κινδύνους που συνδέονται με τη διαδικασία.

#### *Η εξατομικευμένη τιμολόγηση πάσχει υπό το GDPR*

Η αντιστοίχιση δεδομένων επιτρέπει την αξιοποίηση δεδομένων σε τεράστια κλίμακα. Για ένα επιτυχημένο εξατομικευμένο σχήμα τιμολόγησης είναι απαραίτητες πληροφορίες σε μεγάλη κλίμακα. Ο συνδυασμός της μορφοποίησης των δεδομένων και της εξατομικευμένης τιμολόγησης διακυβεύει επομένως τη δημιουργία ασυμμετρίας πληροφόρησης για τους αγοραστές. Ενώ οι αγοραστές είναι συχνά σε χειρότερη θέση από τους πωλητές, αυτή η ασυμμετρία πληροφορίας είναι πιο εμφανής. Δεδομένου ότι οι αγοραστές κινδυνεύουν να μην γνωρίζουν ότι είναι προφίλ, δημιουργεί ένα πρόβλημα (.93).

Ένα γενικό καθεστώς δυναμικής τιμολόγησης που ισχύει για όλους τους αγοραστές διαφέρει με την έννοια ότι όλοι αντιμετωπίζονται εξίσου. Η ευρεία χρήση έχει επιτρέψει στους ανθρώπους να συνηθίσουν στο γεγονός ότι οι τιμές αλλάζουν και δεν αποτελεί έκπληξη το γεγονός ότι μια πτήση μπορεί να είναι φθηνότερη εάν αγοραστεί δύο μήνες νωρίτερα σε αντίθεση με μία ημέρα πριν από την αναχώρηση. Ωστόσο, η εξατομικευμένη τιμολόγηση δεν είναι τόσο επιφανής, είναι πιο δύσκολο να παρατηρηθεί και πιο δύσκολο να ληφθεί τεκμηριωμένη απόφαση. Δεδομένων των πληροφοριακών απαιτήσεων που ορίζονται στο GDPR, όπου το υποκείμενο των δεδομένων πρέπει να ενημερώνεται για την πρακτική και το σκεπτικό πίσω από αυτό, πολλοί άνθρωποι δεν θα αφιερώσουν χρόνο και θα ενημερωθούν επαρκώς, οδηγώντας σε αυτούς τους ανθρώπους που εμπιστεύονται το σύστημα(.93).

Αυτό οδηγεί στο ερώτημα εάν υπάρχει πραγματική ζημιά - εάν οι τιμές καθορίζονται σύμφωνα με το μέγιστο που θέλουν να πληρώσουν οι άνθρωποι και αν οι άνθρωποι βρουν τα προϊόντα ή τις υπηρεσίες που αξίζει τον κόπο. Παρόλο που ένα άτομο μπορεί να πληρώσει υψηλότερη τιμή για ένα συγκεκριμένο προϊόν ή υπηρεσία, μπορεί να μην έχει υπερβολικά αρνητικό αντίκτυπο στην οικονομία του. Ωστόσο, η διαφορά τιμών βασίζεται σε μια ασυμμετρία πληροφόρησης. Όταν οι τιμές αλλάζουν για όλους, όλοι επηρεάζονται με τον ίδιο τρόπο και είναι πιο εύκολο να παρακολουθείτε και να καταλαβαίνετε. Οι περιπλοκές της μορφοποίησης των δεδομένων και της εξατομικευμένης τιμολόγησης δεν είναι, ωστόσο, εύκολο να καταλάβουν το κοινό. Αν οι αγοραστές χαρτογραφούσαν όλες τις εξατομικευμένες τιμές τους και τις συνέκριναν, αυτή η ασυμμετρία πληροφορίας θα μπορούσε να θεραπευτεί. Υπάρχουν διαθέσιμες ιστοσελίδες σύγκρισης τιμών που βοηθούν τους αγοραστές να καταπολεμήσουν την ασυμμετρία των πληροφοριών(.93).

Το πρόβλημα όμως είναι ότι οι πωλητές σε αυτό το σενάριο έχουν πολύ μεγαλύτερη ισχύ από τους αντίστοιχους πελάτες τους. Εκτός εάν ένα υποκείμενο δεδομένων ζητά τα επεξεργασμένα προσωπικά δεδομένα, δεν θα γνωρίζει πόσα δεδομένα επεξεργάζεται ο υπεύθυνος επεξεργασίας δεδομένων. Αυτό κινδυνεύει να οδηγήσει στον πωλητή να ζωγραφίσει την εικόνα των αντικειμένων που πωλούνται στην τιμή της αγοράς, να αναρωτιέται ποιος θα ήταν συνεχώς ύποπτος για την τιμολόγηση και ότι οι αγοραστές πιστεύουν ότι προσφέρονται αγαθά και υπηρεσίες σε τιμή αγοράς(.93).

## Πως θα πρέπει να αντιμετωπίζεται οι Παραβιάσεις

### Ο ρόλος των αρχών προστασίας

Η μεταφορά της οδηγίας 95/46 / ΕΚ από τα διάφορα κράτη μέλη δημιούργησε διαφορές μεταξύ των εθνικών νομοθεσιών εφαρμογής. Οι διαφορές αυτές ήταν επίσης εμφανείς στο επίπεδο των DPA, των οποίων οι αρμοδιότητες διέφεραν από το ένα κράτος μέλος στο άλλο. Μία από τις βασικές προόδους της GDPR από την άποψη αυτή είναι ότι, με την ενοποίηση του κοινοτικού δικαίου, παρείχε επίσης τις εθνικές DPA σε ολόκληρη την Ένωση με την ίδια εξουσία. Επιπλέον, η

καταναγκαστική τους λειτουργία ενισχύθηκε σημαντικά από το γεγονός ότι θα μπορούν να επιβάλλουν πρόστιμα ύψους 20 εκατομμυρίων ευρώ ή το 4% του συνολικού ετήσιου κύκλου εργασιών παγκοσμίως (Άρθρο 83 παρ. 6, GDPR). Συνεπώς, οι οργανισμοί αναγκάζονται να εξετάσουν πολύ πιο σοβαρά την προστασία των δεδομένων, προκειμένου να αποφύγουν τις ενέργειες επιβολής από τις εθνικές αρχές προστασίας δεδομένων.

Ο τρόπος με τον οποίο τα συλλογικά συμφέροντα πρέπει να προστατεύονται από τις διακρίσεις και την κοινωνική εποπτεία στη χρήση μεγάλων αναλυτικών στοιχείων δεδομένων είναι σε μεγάλο βαθμό θέμα των υπευθύνων χάραξης πολιτικής. (.65) Τα διαφορετικά νομικά συστήματα και οι διαφορετικές ισορροπίες μεταξύ των συνιστωσών της κοινωνίας προτείνουν διαφορετικές λύσεις. Επομένως, ο προσδιορισμός της ανεξάρτητης αρχής που είναι αρμόδιος για την προστασία των συλλογικών συμφερόντων μπορεί να είναι δύσκολος. Πολλές χώρες διαθέτουν ανεξάρτητα όργανα υπεύθυνα για την εποπτεία συγκεκριμένων δραστηριοτήτων κοινωνικής εποπτείας και άλλοι φορείς επικεντρώνονται σε ενέργειες κατά των διακρίσεων. Σε άλλες χώρες, αυτή η ευθύνη κατανέμεται σε διάφορες αρχές, οι οποίες υιοθετούν διαφορετικές προσεγγίσεις, χρησιμοποιούν διαφορετικά διορθωτικά μέτρα και δεν συνεργάζονται απαραίτητα επίλυση περιπτώσεων με πολλαπλές επιπτώσεις.

Ένα κεντρικό στοιχείο στην αξιολόγηση κινδύνου των μεγάλων αναλυτικών στοιχείων είναι η προσέγγιση που λαμβάνει υπόψη τις διάφορες αρνητικές **επιπτώσεις που προκύπτουν από τη χρήση μεγάλων δεδομένων** (διακρίσεις, αθέμιτες καταναλωτικές πρακτικές, κοινωνικός έλεγχος κλπ.). Οι αρχές προστασίας δεδομένων έχουν ανατεθεί με την αντιμετώπιση συλλογικών ζητημάτων και έχουν ήδη αποδείξει ότι εξετάζουν τόσο την ατομική όσο και την ευρύτερη συλλογική διάσταση της επεξεργασίας δεδομένων. Εστιάζοντας στην προστασία των δεδομένων και στα θεμελιώδη δικαιώματα, για την εξισορρόπηση των συγκρουόμενων συμφερόντων γύρω από τη χρήση των δεδομένων.

Στο πλαίσιο αυτό, οι αρχές προστασίας δεδομένων μπορούν να εμπλέκουν στη διαδικασία αξιολόγησης τους διάφορους εμπλεκόμενους φορείς, οι οποίοι αντιπροσωπεύουν τα συλλογικά συμφέροντα που επηρεάζονται από συγκεκριμένα έργα επεξεργασίας δεδομένων. Αυτό θα οδηγούσε στον ορισμό ενός νέου μοντέλου στον οποίο οι επιχειρήσεις που σκοπεύουν να χρησιμοποιήσουν τα μεγάλα δεδομένα ανάλυσης θα υποβάλλονταν σε αξιολόγηση πριν από τη συλλογή και την επεξεργασία δεδομένων.

Αλλά ίσως μια από τις κύριες καινοτομίες του κανονισμού είναι η εισαγωγή ενός λεγόμενου μηχανισμού «ενιαίας θυρίδας». Σύμφωνα με την οδηγία 95/46 / ΕΚ, μια εταιρεία που ασκεί επιχειρηματικές δραστηριότητες σε πολλά κράτη μέλη οφείλει να χειρίζεται και να συμμορφώνεται με τους εθνικούς κανόνες που εφαρμόζονται και ερμηνεύονται από διάφορες αρχές προστασίας δεδομένων, οι οποίες είναι αρμόδιες για την εποπτεία των δραστηριοτήτων επεξεργασίας της ίδιας εταιρείας. Η κατάσταση αυτή δεν είναι μόνο επαχθής για τις επιχειρήσεις στην ΕΕ αλλά και επιζήμια

για τα δικαιώματα των υποκειμένων των δεδομένων. Ωστόσο, σύμφωνα με τον νέο κανονισμό, αυτές οι εταιρείες θα έχουν ένα ενιαίο DPA ως σημείο επαφής. Εάν μια εταιρεία ασκεί τις δραστηριότητές της σε περισσότερα του ενός κράτη μέλη και έχει τουλάχιστον μία εγκατάσταση (.83) στην ΕΕ, θα πρέπει να αντιμετωπίσει μια ενιαία DPA (επικεφαλής DPA), η οποία θα συντονίζει την εποπτεία όλων των δραστηριοτήτων επεξεργασίας αυτής της επιχείρησης την ΕΕ, σε συνεργασία με άλλες «ενδιαφερόμενες» DPA. Εάν οι DPA δεν καταλήξουν σε συμφωνία, αλλά και σε ορισμένες περιπτώσεις που ορίζονται σαφώς στον κανονισμό, το ζήτημα θα εξεταστεί σε επίπεδο ΕΕ από το νέο Ευρωπαϊκό Συμβούλιο Προστασίας Δεδομένων (Binding Corporate Rules, Άρθρο 64 GDPR), στο πλαίσιο του μηχανισμού συνέπειας (Άρθρο 63 - 76, GDPR).

Στην περίπτωση τραπεζών με πολλαπλές εγκαταστάσεις στην ΕΕ, θα είναι απαραίτητο να προσδιοριστεί το κύριο ίδρυμα και η σχετική επικεφαλής DPA, αρμόδια για την εποπτεία των δραστηριοτήτων επεξεργασίας της τράπεζας σε ολόκληρη την Ένωση. Εκτός εάν οι αποφάσεις σχετικά με τους σκοπούς και τα μέσα επεξεργασίας λαμβάνονται αλλού, ο τόπος της κεντρικής διοίκησης του υπεύθυνου της επεξεργασίας στην Ένωση θα θεωρείται ως κύρια εγκατάσταση (Άρτο 4, παρ 16 GDPR). Είναι, ωστόσο, η πραγματική και πραγματική άσκηση των διαχειριστικών δραστηριοτήτων όσον αφορά τους σκοπούς και τα μέσα επεξεργασίας που θα είναι καθοριστικοί για τον προσδιορισμό της κύριας εγκατάστασης (Αίτια 36, GDPR). Τα υψηλά πρόστιμα που μπορούν να επιβάλουν τα DPA αναμένεται επίσης να έχουν σημαντικό αντίκτυπο, ιδίως επειδή στο μέλλον θα υπολογιστούν με βάση τον ετήσιο παγκόσμιο κύκλο εργασιών της τράπεζας, αντί να βασίζονται μόνο στα τοπικά έσοδα. Επομένως, τα DPA θα διαδραματίσουν ενισχυμένο ρόλο στο πλαίσιο του GDPR και θα βρεθούν σε ισχυρότερη θέση για να παρακολουθήσουν και να επιβάλουν την τήρηση των νέων κανόνων προστασίας δεδομένων.

#### Ορισμός υπεύθυνου προστασίας δεδομένων (ΥΠΔ)

Οι υπεύθυνοι προστασίας δεδομένων (DPO) θα πρέπει να βοηθούν τους υπεύθυνους επεξεργασίας και τους μεταποιητές στην εκπλήρωση των υποχρεώσεών τους βάσει του GDPR. Οι ΥΠΔ είναι υποχρεωτικές σε τρεις περιπτώσεις που προβλέπονται ρητά στον κανονισμό. Η πρώτη περίπτωση αφορά την περίπτωση κατά την οποία η επεξεργασία πραγματοποιείται από δημόσια αρχή ή οργανισμό. Η δεύτερη και η τρίτη περίπτωση είναι ίσως λιγότερο σαφείς στον κανονισμό. Έτσι, ένας ΥΠΔ απαιτείται όταν οι βασικές δραστηριότητες του υπεύθυνου επεξεργασίας ή του μεταποιητή συνίστανται είτε σε πράξεις επεξεργασίας που απαιτούν τακτική και συστηματική παρακολούθηση των υποκειμένων των δεδομένων σε μεγάλη κλίμακα είτε σε επεξεργασία ειδικών κατηγοριών δεδομένων ή δεδομένων που αφορούν ποινικές καταδίκες και αδικήματα σε μεγάλη κλίμακα. Παρόλο που ο κανονισμός δεν παρέχει περαιτέρω λεπτομέρειες ή παραδείγματα, μπορούμε να προσδιορίσουμε δύο προϋποθέσεις για την εφαρμογή αυτής της υποχρέωσης: η επεξεργασία αυτή αποτελεί τις βασικές

δραστηριότητες του υπεύθυνου επεξεργασίας ή του μεταποιητή και εκτελείται σε μεγάλη κλίμακα. Επιπλέον, ο ευαίσθητος χαρακτήρας των επεξεργασμένων δεδομένων προσωπικού χαρακτήρα ή το γεγονός ότι τα υποκείμενα των δεδομένων παρακολουθούνται τακτικά και συστηματικά μπορεί να καθορίσει την αναγκαιότητα ενός ΥΠΔ (.82).

Το GDPR εξακολουθεί να περιέχει πολλά σημεία και ιδέες, οι οποίες μπορούν να αποτελέσουν αντικείμενο ερμηνείας και θα πρέπει να διευκρινιστούν στο μέλλον. Κάποιες διευκρινίσεις προσφέρονται, ωστόσο, από τις αιτιολογικές σκέψεις. Ως εκ τούτου, όσον αφορά την έννοια της «παρακολούθησης» του προσώπου στο οποίο αναφέρονται τα δεδομένα, οι αιτιολογικές σκέψεις αναφέρουν ότι «προκειμένου να προσδιοριστεί κατά πόσον μια δραστηριότητα επεξεργασίας μπορεί να θεωρηθεί ότι παρακολουθεί τη συμπεριφορά των προσώπων στα οποία αναφέρονται τα δεδομένα, που συμπεριλαμβάνεται στην πιθανή επακόλουθη χρήση τεχνικών επεξεργασίας δεδομένων προσωπικού χαρακτήρα, οι οποίες συνίστανται στο χαρακτηρισμό ενός φυσικού προσώπου, ιδίως προκειμένου να ληφθούν αποφάσεις που αφορούν την ίδια ή για την ανάλυση ή την πρόβλεψη των προσωπικών του προτιμήσεων, συμπεριφορών και συμπεριφορών (Αιτία 36, GDPR). Η GDPR προβλέπει ότι «στον ιδιωτικό τομέα, οι βασικές δραστηριότητες ενός ελεγκτή σχετίζονται με τις κύριες δραστηριότητές του και δεν σχετίζονται με την επεξεργασία δεδομένων προσωπικού χαρακτήρα ως βοηθητικές δραστηριότητες (Αιτία 97, GDPR).

Θα μπορούσε να προβλεφθεί ότι, σύμφωνα με τον κανονισμό, οι τράπεζες και άλλοι παράγοντες του χρηματοπιστωτικού τομέα θα πρέπει να ορίσουν έναν ΥΠΔ, όχι μόνο λόγω της φύσης των δεδομένων που επεξεργάζονται, όπως είναι τα δεδομένα σχετικά με ποινικές καταδίκες και αδικήματα, αλλά και λόγω το γεγονός ότι παρακολουθούν τακτικά και συστηματικά τους πελάτες τους μέσω δραστηριοτήτων όπως η δημιουργία προφίλ. Παρόλα αυτά, θα μπορούσε κανείς να υποστηρίξει ότι αυτή η επεξεργασία δεν αποτελεί τις «βασικές δραστηριότητές» του (.82).

Το GDPR επιτρέπει τον διορισμό ενός ενιαίου ΥΠΔ από ομάδα επιχειρήσεων. Ωστόσο, είναι αμφίβολο το γεγονός ότι σε μεγάλες πολυεθνικές εταιρείες, ένας ενιαίος ΥΠΔ θα είναι σε θέση να συμβουλευεί και να ενημερώνει αποτελεσματικά τον υπεύθυνο για όλες τις υποχρεώσεις του, να παρακολουθεί τη συμμόρφωση με τους νόμους περί προστασίας των δεδομένων της ΕΕ και των κρατών μελών και να εκπληρώνει όλα τα καθήκοντα και τις αποστολές του. Στην πραγματικότητα, ένα δομημένο δίκτυο ΥΠΔ είναι πολύ πιθανότερο να καταστήσει δυνατή την καλή κατανόηση και επισκόπηση των ροών δεδομένων που πραγματοποιούνται και των μέσων που είναι απαραίτητα για την προστασία των προσωπικών δεδομένων και τη συμμόρφωση με τον κανονισμό.

Το GDPR καθιστά επίσης δυνατή την ύπαρξη εσωτερικού ΥΠΔ, μέρους του προσωπικού του υπεύθυνου επεξεργασίας ή του μεταποιητή, ή ενός εξωτερικού υπάλληλου που προσλαμβάνεται βάσει σύμβασης παροχής υπηρεσιών. Και οι δύο επιλογές έχουν μειονεκτήματα και οφέλη. Ενώ η πρώτη επιλογή έχει το πλεονέκτημα ότι ο ΥΠΔ θα έχει πιθανώς καλύτερη κατανόηση των πράξεων

επεξεργασίας δεδομένων, της ροής εργασιών και της γενικής δομής του οργανισμού, έχει επίσης το μειονέκτημα ότι μπορεί να έχει αρνητικό αντίκτυπο στην ανεξαρτησία του ΥΠΔ. Ο κανονισμός προβλέπει την ανάγκη διασφάλισης της ανεξαρτησίας και της σταθερότητας αυτής της θέσης. Ο ΥΠΔ δεν λαμβάνει εντολές όσον αφορά την εκπλήρωση των καθηκόντων αυτών ούτε απολύεται ή τιμωρείται για την εκτέλεση των καθηκόντων του και αναφέρεται απευθείας στην ανώτατη διοίκηση (Άρθρο 39, GDPR). Η τελευταία αυτή επιλογή έχει το πλεονέκτημα ότι εξασφαλίζει μεγαλύτερη ανεξαρτησία για τους ΥΠΔ, εάν προσλαμβάνονται βάσει σύμβασης παροχής υπηρεσιών. Παρόλα αυτά, η ύπαρξη εξωτερικών στοιχείων μπορεί να θέσει υπό αμφισβήτηση την επαρκή γνώση τους σχετικά με την οργάνωση, τις εργασίες και τις ροές δεδομένων του υπεύθυνου επεξεργασίας. Παρά την υποχρέωση του ΥΠΔ να τηρεί το απόρρητο ή την εμπιστευτικότητα (Άρθρο 38 παρ. 5, GDPR), οι οργανισμοί μπορεί να αισθάνονται λιγότερο άνετοι να μοιράζονται τα επιχειρηματικά μυστικά με έναν «ξένο». Ωστόσο, οι πληροφορίες αυτές, όσον αφορά, για παράδειγμα, τα ισχύοντα συστήματα ασφαλείας, είναι απαραίτητα για την εκπλήρωση των αποστολών του ΥΠΔ.

Μια άλλη σημαντική πτυχή είναι το προφίλ του ΥΠΔ, ο οποίος ορίζεται ως "με βάση τα επαγγελματικά προσόντα και, ειδικότερα, την εξειδικευμένη γνώση του νόμου περί προστασίας των δεδομένων" (Άρθρο 37 παρ. 5, GDPR). Ενώ ο κανονισμός δεν προβλέπει ρητά την προϋπόθεση ότι ο ΥΠΔ έχει νομικό υπόβαθρο, η προϋπόθεση αυτή προϋποθέτει ότι τα καθήκοντα του ΥΠΔ περιλαμβάνουν την παροχή πληροφοριών και συμβουλών σχετικά με τις υποχρεώσεις των υπευθύνων επεξεργασίας δυνάμει του GDPR και της εθνικής προστασίας δεδομένων να παρακολουθεί τη συμμόρφωση με τους νόμους αυτούς, να συνεργάζεται και να ενεργεί ως σημείο επαφής της εθνικής ΑΠΔ (Άρθρο 39 GDPR). Δεδομένης της μερικής αλληλεπικάλυψης μεταξύ της προστασίας δεδομένων και της τεχνολογίας των πληροφοριών, μπορεί κανείς να προβλέψει την ανάγκη να έχει ο ΥΠΔ τουλάχιστον μια βασική γνώση και κατανόηση των νέων τεχνολογιών πληροφορικής και των διαδικασιών και υπηρεσιών πληροφορικής, ώστε να είναι σε θέση να εκπληρώσει επαρκώς τα καθήκοντά του.

### Κοινές Λύσεις

Στην Γνώμη του EDPS (European Data Protection Director) 8/2016 αναφέρεται ότι: (.64) Ο συνεχιζόμενος διάλογος κατά τη διάρκεια των δύο τελευταίων ετών υπογράμμισε ειδικότερα την έννοια της δικαιοσύνης που διαπνέει καθένα από αυτά τα πεδία και κατοχυρώνεται στα σχετικά άρθρα του Χάρτη της ΕΕ και της ΤFEU (Treaty on the Functioning of the European Union):

- η δικαιοσύνη είναι ίσως το πιο θεμελιώδες κριτήριο για τις νόμιμες εμπορικές πρακτικές στο δίκαιο των καταναλωτών.
- η δικαιοσύνη της επεξεργασίας των προσωπικών δεδομένων αποτελεί βασική αρχή παράλληλα με τη νομιμότητα και τη διαφάνεια.

- το δίκαιο του ανταγωνισμού κάνει παραχωρήσεις σε συμφωνίες που αντιβαίνουν στον ανταγωνισμό, επιτρέποντας στους καταναλωτές δίκαιο μερίδιο "του οφέλους και περιλαμβάνει στον ορισμό του την καταχρηστική εκμετάλλευση δεσπόζουσας θέσης" που επιβάλλει αθέμιτες τιμές αγοράς ή πώλησης "36.

Παρά ταύτα, η συνεργασία μεταξύ των αρχών σε ευρωπαϊκό επίπεδο εξακολουθεί να είναι περιορισμένη. Βασικό ζήτημα για τις αρμόδιες αρχές είναι η κοινή ανησυχία για τον καταναλωτή. Ωστόσο, η έννοια της «ευημερίας των καταναλωτών» στο δίκαιο του ανταγωνισμού δεν έχει προσδιοριστεί με σαφήνεια και έχει τείνει να χρησιμοποιείται για την αντιμετώπιση της διάρθρωσης της αγοράς και της οικονομικής αποτελεσματικότητας και μόνο έμμεσα αντιμετωπίζει μεμονωμένες ανησυχίες των καταναλωτών, όπως το ιδιωτικό απόρρητο. Η TFEU (Treaty on the Functioning of the European Union- Συνθήκη για τη λειτουργία της Ευρωπαϊκής Ένωσης) στο άρθρο 102 απαγορεύει την κατάχρηση δεσπόζουσας θέσης υπό τη μορφή «αθέμιτων αγορών ή τιμών πώλησης ή άλλων αθέμιτων όρων συναλλαγών»: ωστόσο, οι αρχές ανταγωνισμού τείνουν να αφήνουν ενέργειες κατά της εκμετάλλευσης αυτών των εκμεταλλευτικών συμπεριφορών, ενώ μερικές φορές οι φορείς επιβολής του νόμου αφήνουν με τη σειρά τους άδικες υποθέσεις καταναλωτικών όρων στις αρχές προστασίας δεδομένων.

Ορισμένες αλληλεπιδράσεις μεταξύ αρχών σε εθνικό επίπεδο έχουν επιφέρει αποτελέσματα, για παράδειγμα:

- σύμφωνα με μια προσωρινή απόφαση της γαλλικής αρχής ανταγωνισμού τον Σεπτέμβριο του 2014, η GDF Suez είχε καταχραστεί τη δεσπόζουσα θέση της χρησιμοποιώντας προσωπικά δεδομένα που συλλέχθηκαν όταν ήταν κρατικό μονοπώλιο για την προώθηση ενός πακέτου φυσικού αερίου και ηλεκτρικής ενέργειας σε μια ανοιχτή μη ρυθμιζόμενη αγορά. Η αρχή ανέθεσε στην GDF να αποκαλύψει μέρος της βάσης δεδομένων των πελατών της σε ανταγωνιστές, αφού έδωσε στα άτομα την ευκαιρία να αποχωρήσουν από την αποκάλυψη.

- η βρετανική αρχή προστασίας δεδομένων ενημέρωσε την αρμόδια για τον ανταγωνισμό αρχή του Ηνωμένου Βασιλείου τον Αύγουστο του 2015 σχετικά με την πρότασή της να καλέσει τα νοικοκυριά που δεν είχαν μεταβιβάσει προμηθευτές ενέργειας για τρία ή περισσότερα έτη για να μην συμμετάσχουν στα ανταγωνιστικά τους προμηθευτές.

- το Σεπτέμβριο του 2015 η βελγική αρχή ανταγωνισμού επέβαλε πρόστιμο ύψους 1,9 εκατομμυρίων ευρώ στη βελγική εθνική λαχειοφόρο αγορά για τη χρήση των προσωπικών δεδομένων που αποκτήθηκαν επίσης ως δημόσιο μονοπώλιο για τον ασυμβίβαστο σκοπό της εμπορικής προώθησης μιας εμπορικής υπηρεσίας στοιχημάτων "Score!" στην παρακείμενη αγορά αθλητικών στοιχημάτων. Η αρχή θεώρησε ότι πρόκειται

για κατάχρηση δεσπόζουσας θέσης κατά τη χρήση πληροφοριών που δεν μπορούσαν να αναπαραχθούν από τους ανταγωνιστές της.

- το 2016 ξεκίνησε μια έρευνα από το Bundeskartellamt σχετικά με τις πολιτικές απορρήτου που εφαρμόζει η φερόμενη ως κυρίαρχη εταιρία κοινωνικών μέσων Facebook, έχοντας στενή επαφή με τις αρχές προστασίας δεδομένων, τις ενώσεις προστασίας των καταναλωτών και άλλες εθνικές αρχές ανταγωνισμού<sup>40</sup>.

Παρ' όλα αυτά, συνολικά υπάρχει ένα αρκετά κατακερματισμένο σενάριο για την επιβολή των κανόνων της ΕΕ, ενώ οι αρμόδιες αρχές δεν μιλούν ο ένας τον άλλο, ενώ αντιμετωπίζουν περιπτώσεις που χαρακτηρίζονται από σημαντικές επικαλύψεις από πλευράς ουσίας. Για παράδειγμα, θα ήταν χρήσιμες οι κοινές συνεδριάσεις της ομάδας εργασίας του άρθρου 29, του ευρωπαϊκού δικτύου ανταγωνισμού και του δικτύου συνεργασίας για την προστασία των καταναλωτών, των αντίστοιχων οργάνων συντονισμού σε επίπεδο ΕΕ.

#### Διαφορετικές αλλά συσχετιζόμενες δικαιοδοσίες

Οι ρυθμιστικές αρχές συχνά υπόκεινται σε μεγάλη πίεση για να ανταποκριθούν στις προσδοκίες των πολιτών με περιορισμένους πόρους και αυξανόμενο φόρτο εργασίας και είναι φυσικό να επικεντρωθεί στις δικές τους αρμοδιότητες. (.64) Τα όρια των αντίστοιχων αρμοδιοτήτων και αρμοδιοτήτων των οργάνων πρέπει να γίνονται σεβαστά: σαφώς οι αρχές δεν θα πρέπει, και πιθανώς δεν μπορούν, να επιβάλλουν την εφαρμογή νόμων σε άλλους νομικούς τομείς<sup>41</sup>. Κανένα ενιαίο δίκαιο δεν αποτελεί πανάκεια για όλα τα προβλήματα και θα ήταν ακατάλληλο για έναν τομέα ρύθμισης να κοιτάξει σε μια άλλη περιοχή για να αντισταθμίσει τις δικές του αδυναμίες. Οι αρχές σε κάθε περιοχή έχουν περιορισμένα εργαλεία στη διάθεσή τους, όπως για παράδειγμα η επιβολή του ανταγωνισμού μπορεί να αντιμετωπίσει μόνο την κατάχρηση δεσπόζουσας θέσης, τη συμπεριφορά του καρτέλ και τις συγχωνεύσεις που δεν είναι προς το συμφέρον του καταναλωτή. οι καταχρηστικές συνθήκες υπηρεσίας δεν είναι κατ' ανάγκη αντιμονοπωλιακό ζήτημα.

Π.Χ. Μια σημαντική υπόθεση συγχώνευσης μετά την προκαταρκτική γνωμοδότησή μας αφορούσε την απόκτηση της εφαρμογής WhatsApp - μια δημοφιλής εφαρμογή ανταλλαγής μηνυμάτων που σαρώνει ολόκληρα βιβλία διευθύνσεων αλλά δεν διαθέτει στο εμπόριο τις πληροφορίες χρήστη - από το Facebook - το οποίο έχει μια πολύ διαφορετική προσέγγιση χρήσης δεδομένων. Η Ομοσπονδιακή Επιτροπή Εμπορίου των ΗΠΑ απαίτησε από τα συμβαλλόμενα μέρη να δώσουν ειδοποίηση στους πελάτες και να επιλέξουν εάν απομακρυνθούν από τους όρους. Η Ευρωπαϊκή Επιτροπή, ενεργώντας ως αρχή ανταγωνισμού της ΕΕ, διαπίστωσε ότι ο κανονισμός περί συγκεντρώσεων δεν είχε ως βάση να απαιτήσει από την αποκτώσα εταιρεία να τηρήσει τη συμφωνία περί ιδιωτικού απορρήτου που είχε συνάψει με τους πελάτες της WhatsApp. Ωστόσο, κάθε προσέγγιση υποδήλωνε ότι οι χρήστες των υπηρεσιών ανταλλαγής μηνυμάτων έπρεπε να αποδεχθούν



τις νέες συνθήκες ή να τους απαγορευθεί η χρήση των υπηρεσιών. Πρόσφατα μια αλλαγή στους όρους προστασίας της ιδιωτικής ζωής της υπηρεσίας WhatsApp Messaging οδήγησε τον Επίτροπο Ανταγωνισμού να θέσει ερωτήσεις της νέας οντότητας. Σε περίπτωση μελλοντικών συγχωνεύσεων παρόμοιου χαρακτήρα, τα άτομα θα μπορούσαν να επωφεληθούν από μια πιο συνεκτική απάντηση από τον ανταγωνισμό, τις αρχές προστασίας των καταναλωτών και των δεδομένων. Οι εποπτικές αρχές πρέπει να είναι πλήρως εξοπλισμένες ώστε να προβλέπουν και να αποτρέπουν τόσο τη συμπεριφορά όσο και τις συγκεντρώσεις που μπορεί να βλάψουν το άτομο.

Καμία από αυτές τις ρυθμιστικές δικαιοδοσίες δεν είναι σφραγισμένη ερμητικά από τους άλλους. Η υψηλή συγκέντρωση στις αγορές θα μπορούσε να υπονομεύσει την προστασία αυτών των θεμελιωδών δικαιωμάτων, ακόμη και στις περιπτώσεις όπου δεν επιβάλλονται αντίθετες προς τον ανταγωνισμό συμπεριφορές που αξιολογούνται από αντιμονοπωλιακούς φορείς. Οι αρχές αναμένουν ήδη, σύμφωνα με τη νομολογία, να εξετάσουν τα πιθανά κίνητρα για κατάχρηση δεσπόζουσας θέσης μετά τη συγκέντρωση (.66). Η εφαρμογή του κοινοτικού δικαίου στον τομέα του ανταγωνισμού χρησιμοποιήθηκε στο παρελθόν για πιο συγκεκριμένους στόχους πολιτικής, όπως για την απελευθέρωση της αγοράς τηλεπικοινωνιών. Συγκεκριμένα, το άρθρο 21 παράγραφος 4 του κανονισμού για τις συγκεντρώσεις προβλέπει ότι τα κράτη μέλη εφαρμόζουν πρόσθετους ελέγχους προκειμένου να προστατεύσουν την πολυφωνία των μέσων επικοινωνίας, γεγονός που ανταποκρίνεται στις ανησυχίες ότι η συγκέντρωση στη βιομηχανία των μέσων ενημέρωσης θα μπορούσε να υπονομεύσει την ανεξαρτησία και την ελευθερία έκφρασης που προβλέπονται στο άρθρο 11 της το Χάρτη (.67).

«Ακόμη και αν εξυπηρετούν διαφορετικούς στόχους», σύμφωνα με μια κοινή έκθεση για το δίκαιο του ανταγωνισμού και τα δεδομένα που δημοσιεύτηκαν τον Μάιο του 2016 από τις γαλλικές και τις γερμανικές αρχές ανταγωνισμού, «τα ζητήματα ιδιωτικού απορρήτου δεν μπορούν να αποκλειστούν από την εξέταση του δικαίου ανταγωνισμού απλώς και μόνο λόγω της φύσης τους». Οι αρχές προστασίας δεδομένων μπορούν να βοηθήσουν στη διάχυση του τρόπου και του βαθμού στον οποίο ο έλεγχος των προσωπικών δεδομένων είναι τόσο σημαντικός για τις επιχειρήσεις στις αγορές. Οι συνέργιες μεταξύ των νομικών τομέων, οι οποίες συζητήθηκαν εντατικά τα τελευταία χρόνια, θα μπορούσαν να προωθήσουν τη στενότερη συνεργασία μεταξύ των αρχών, ιδίως όταν δεν υπάρχει ούτε καθοδήγηση ούτε νομολογία. Δεν πρόκειται για «οργάνωση» ενός άλλου τομέα δικαίου αλλά για συγχρονισμό των πολιτικών και των δραστηριοτήτων επιβολής της ΕΕ, προσθέτοντας αξία όταν μια εποπτική αρχή στερείται εμπειρογνομosύνης ή νομικής ικανότητας στην ανάλυση.

### Ευκαιρίες συνεργασίας

Η στρατηγική ψηφιακής ενιαίας αγοράς (Digital Single Market Strategy) περιέχει πολλές ελπιδοφόρες προτάσεις για τη βελτίωση των κανονιστικών πλαισίων για την προστασία των

καταναλωτών και των δεδομένων. Ωστόσο, η στρατηγική θα μπορούσε να βελτιωθεί με ένα μηχανισμό για τη συνεκτική επιβολή των υποχρεώσεων σε σχέση με τα δικαιώματα και τα συμφέροντα των ατόμων<sup>48</sup> στους διάφορους τομείς του κοινοτικού δικαίου. Από τη σκοπιά των θεμελιωδών δικαιωμάτων, η στρατηγική έπρεπε επίσης να εξετάσει τον τρόπο με τον οποίο οι περισσότεροι άνθρωποι αλληλεπιδρούν σήμερα με το Ίντερνετ, με τις καθημερινές υπηρεσίες που βασίζονται στο διαδίκτυο, οι οποίες βασίζονται στην αυξανόμενη κοκκώδη εποπτεία των χρηστών από τους παρόχους υπηρεσιών, η οποία συχνά έρχεται σε αντίθεση με την αδιαφάνεια αυτών οι ίδιοι πάροχοι υπηρεσιών επεξεργάζονται προσωπικές πληροφορίες (γνωστές ως φαινόμενο «μαύρου κουτιού»).

Η πρόσφατη ανακοίνωση της Επιτροπής σχετικά με τις ηλεκτρονικές πλατφόρμες αναγνώρισε ότι ο διασυνοριακός χαρακτήρας της επιχείρησης απαιτούσε «καλή συνεργασία μεταξύ των αρμόδιων αρχών» (.68). Το ψήφισμα του 2016 των εθνικών αρχών για την προστασία των δεδομένων στην Ευρώπη προχώρησε περισσότερο στην προώθηση "ενός μεγαλύτερου διαλόγου και ανταλλαγής πληροφοριών με άλλα ρυθμιστικά όργανα αρμόδια για τη διασφάλιση των δικαιωμάτων και των συμφερόντων των ατόμων στην ψηφιακή κοινωνία και την οικονομία", αναγνωρίζοντας τις προσπάθειες ενίσχυσης τις συνέργειες μεταξύ ρυθμιστικών πλαισίων για τους καταναλωτές, την αντιμονοπωλιακή νομοθεσία και την προστασία των δεδομένων (.69). Στο πλαίσιο των αρχών χρηστής διακυβέρνησης και της αρχής της ειλικρινούς συνεργασίας, οι αρχές προστασίας δεδομένων θα πρέπει σε κάθε περίπτωση να συνεργάζονται με τους οργανισμούς της ΕΕ και τις εθνικές ρυθμιστικές αρχές με αρμοδιότητες σε άλλους τομείς πολιτικής (.70). Μια ομάδα εργασίας του Κοινού Κέντρου Ερευνών της Επιτροπής υποστήριξε ότι μια εξειδικευμένη υπηρεσία θα παρέχει τεχνική υποστήριξη σε εποπτικά όργανα που ερευνούν υποθέσεις στην ψηφιακή αγορά και θα παρακολουθεί τη συμμόρφωση των διαδικτυακών πλατφορμών για τη διευκόλυνση της «συνοχής μεταξύ των ρυθμιστικών αρχών στους αντίστοιχους τομείς»(71).

Θεωρούμε ότι το άρθρο 80 του Κανονισμού περί Γενικής Προστασίας Δεδομένων παρέχει μια σημαντική ευκαιρία συλλογικής επιβολής. Τα κράτη μέλη πρέπει να εφαρμόζουν τη διάταξη αυτή για τη συλλογική έννομη προστασία χωρίς να απαιτούν συγκεκριμένη εντολή από ένα πρόσωπο στο οποίο αναφέρονται τα δεδομένα. Οι ομάδες υποστήριξης έχουν ήδη αρχίσει να ασκούν δράσεις σύμφωνα τόσο με τους κανόνες προστασίας των καταναλωτών όσο και με τους κανόνες προστασίας δεδομένων, για παράδειγμα:

- Η UFC-Que Choisir και η Ομοσπονδία Γερμανών Καταναλωτικών Οργανισμών (VZBV) εισήγαγαν μέτρα κατά των κοινωνικών μέσων και των παρόχων υπηρεσιών online για άδικοι συμβατικούς όρους, αθέμιτες εμπορικές πρακτικές και παραβιάσεις της νομοθεσίας για την προστασία των δεδομένων(.72).

- το Νορβηγικό Συμβούλιο Προστασίας των Καταναλωτών εξέδωσε μελέτη σχετικά με τους τυποποιημένους όρους που χρησιμοποιούνται από επτά παρόχους υπηρεσιών cloud, παρέχοντας συγκριτικές επισκοπήσεις πολλών όρων, συμπεριλαμβανομένων των πολιτικών απορρήτου. Η μελέτη οδήγησε σε καταγγελία σχετικά με τους όρους της Apple με τον Νορβηγό Διαμεσολαβητή Καταναλωτών για παραβίαση του νορβηγικού και ευρωπαϊκού νόμου περί καταναλωτών. Η Apple συμφώνησε να τροποποιήσει τους όρους της, και ειδικότερα το μονομερές της δικαίωμα να αλλάξει τη συμφωνία ανά πάσα στιγμή, κατά τη διακριτική της ευχέρεια και χωρίς να δίνει στους χρήστες καμία ειδοποίηση(.73).

- ένας αυστριακός οργανισμός προστασίας των καταναλωτών αμφισβήτησε τους όρους και τις προϋποθέσεις που επέβαλε μονομερώς η Amazon για λόγους της οδηγίας περί αθέμιτων συμβατικών όρων και της οδηγίας για την προστασία των δεδομένων. (Το Σεπτέμβριο το Δικαστήριο ΔΕΕ αποφάνθηκε επί των αρμοδιοτήτων που του υποβλήθηκαν (.74)).

Η συνεργαζόμενη εφαρμογή και η υπέρβαση του «ρυθμιστικού κατακερματισμού» έχει πλέον γίνει επείγουσα ανάγκη, αναγνωρισμένη από την Ευρωπαϊκή Επιτροπή, με τον Πρόεδρο Juncker στην αρχή της θητείας του να καλεί την Επιτροπή να ξεπεράσει τις σιλονοθήκες, και από τον BEUC, τον Ευρωπαϊκό Οργανισμό Καταναλωτή (.75). Τώρα είναι μια ιδανική στιγμή να μετατρέψουμε αυτές τις θεωρητικές συνέργιες σε θετική δράση.

### Προτεινόμενες Λύσεις

Επί του παρόντος, υπάρχει ελάχιστη υποχρέωση λογοδοσίας και εποπτείας όσον αφορά την εξόρυξη δεδομένων. (.24) Το μεγαλύτερο μέρος αυτό οφείλεται στο γεγονός ότι η διαδικασία εξόρυξης δεδομένων είναι ένα μαύρο κουτί για τους ξένους. Ο χαρακτήρας της εξόρυξης δεδομένων στο μαύρο πλαίσιο ενισχύεται από το γεγονός ότι τα άτομα που ασχολούνται παραδοσιακά με την αξιολόγηση της νομιμότητας των επιχειρηματικών διαδικασιών (πολιτικοί και νομικοί) δεν είναι σε επαρκή βαθμό ικανοί να αξιολογήσουν αποτελεσματικά τη νομιμότητα των αλγορίθμων εξόρυξης δεδομένων. Εν μέρει, το γεγονός ότι οι υπεύθυνοι για τη χάραξη πολιτικής και οι νομικοί επαγγελματίες συχνά δεν διαθέτουν τις απαιτούμενες γνώσεις στον τομέα της πληροφορικής και των μαθηματικών διαδραματίζει επίσης σημαντικό ρόλο.

### Περιγραφική εξόρυξη δεδομένων

Δεδομένου ότι η περιγραφική εξόρυξη δεδομένων χρησιμοποιείται κυρίως για καθήκοντα υποστήριξης λήψης αποφάσεων και αποδίδει στοιχεία για τη χάραξη πολιτικής, μπορεί (έμμεσα) να επηρεάσει τη (νομική) θέση των ομάδων και των ατόμων. Ενώ οι επαγγελματίες εξόρυξης δεδομένων προειδοποιούν ότι τα αποτελέσματα εξόρυξης δεδομένων δεν είναι πάντα άμεσα προσαρμοσμένα ως

εισροές για τη χάραξη πολιτικής, τα στατιστικά στοιχεία και τα αυτοματοποιημένα προφίλ που δημιουργούνται με την περιγραφική εξόρυξη δεδομένων διαδραματίζουν σημαντικό ρόλο στις πολιτικές αποφάσεις και τις επιχειρηματικές αποφάσεις. Επομένως, θα πρέπει να ληφθούν μέτρα για τον εντοπισμό και την αποφυγή σφαλμάτων που προκύπτουν από εσφαλμένη εφαρμογή ή ερμηνεία περιγραφικής εξόρυξης δεδομένων. Τα προβλήματα με την ακρίβεια και την αξιοπιστία, η συσχέτιση της αιτιότητας ως προς την αιτιώδη συνάφεια και η βυθοκόρηση δεδομένων αποτελούν παραδείγματα ζητημάτων εφαρμογής που πρέπει να ληφθούν υπόψη όταν πρόκειται για περιγραφική εξόρυξη δεδομένων.

Πάνω απ' όλα, είναι απαραίτητη η αυξημένη ευαισθητοποίηση των υποκειμένων των δεδομένων και των υπευθύνων επεξεργασίας δεδομένων σχετικά με τις δυνατότητες, τα όρια και τους κινδύνους της εξόρυξης δεδομένων. Ένας από τους λόγους για την εισαγωγή του άρθρου 15 στην οδηγία για την προστασία των δεδομένων είναι ότι οι υπεύθυνοι λήψης αποφάσεων στον άνθρωπο τείνουν να αποδίδουν υπερβολική βαρύτητα στα αποτελέσματα των αυτοματοποιημένων εργαλείων λήψης αποφάσεων, βλέποντας τα αποτελέσματά τους προφανώς αντικειμενικά και αναμφισβήτητα. Το ίδιο μοτίβο εμφανίζεται στην πολιτική. πολύ, αν όχι πάρα πολύ, συνδέεται το βάρος με τα αποτελέσματα εξόρυξης δεδομένων και τα στατιστικά στοιχεία. Ως εκ τούτου, η ανάγκη ευαισθητοποίησης επεκτείνεται και στους υπεύθυνους χάραξης πολιτικής.

Αν και θα ήταν ευεργετικό να υπάρχουν σαφείς διαδικαστικοί μηχανισμοί για την επαλήθευση και την επικύρωση των αποτελεσμάτων των ασκήσεων εξόρυξης δεδομένων που χρησιμοποιούνται για τη στήριξη της διαδικασίας λήψης πολιτικών αποφάσεων, οι μηχανισμοί αυτοί είναι δύσκολοι αν όχι αδύνατο να σχεδιαστούν. Θα πρέπει μάλλον να στραφούμε προς την εκπαίδευση των υπευθύνων χάραξης πολιτικής σχετικά με τα όρια της εξόρυξης δεδομένων και την ερμηνεία των αποτελεσμάτων. Μια άλλη προσέγγιση μπορεί να είναι η λήψη μιας "δεύτερης γνώμης" από εμπειρογνώμονες για την εξόρυξη δεδομένων σε συνεργασία με νομικούς εμπειρογνώμονες για αποτελέσματα εξόρυξης δεδομένων που διαδραματίζουν σημαντικό ρόλο στη λήψη αποφάσεων ή τη χάραξη πολιτικής.

### Προγνωστική εξόρυξη δεδομένων

Η προληπτική εξόρυξη δεδομένων είναι ιδιαίτερα κατάλληλη για τη λήψη αποφάσεων σε συγκεκριμένες περιπτώσεις. Συγκεκριμένα, η πρακτική της ταξινόμησης χρησιμοποιείται για τη λήψη αποφάσεων σχετικά με ομάδες και ιδιώτες (για παράδειγμα, σηματοδύνοντας κάποιον ως δυνητικό τρομοκράτη ή αρνούμενος σε κάποιον δάνειο). Η προστασία από τους κινδύνους της πρόβλεψης εξόρυξης δεδομένων θα πρέπει πρωτίστως να στοχεύει στη μεγαλύτερη διαφάνεια της διαδικασίας λήψης αποφάσεων, αφαιρώντας κάθε δυνατότητα για την ανακριβή ή εσφαλμένη εφαρμογή

αλγορίθμων πρόβλεψης εξόρυξης δεδομένων και εντοπίζοντας τυχόν παράνομες ή ανήθικες αποφάσεις (πιθανότατα εκ των υστέρων).

Εξετάστε, π.χ., μια τράπεζα που θέλει να χρησιμοποιήσει ιστορικά στοιχεία σχετικά με τα προσωπικά δάνεια για να μάθει μοντέλα για να προβλέψει για τους νέους υποψηφίους δανείου την πιθανότητα να χρεοκοπήσουν το δάνειό τους. Θα μπορούσε να είναι πολύ καλά ότι τα στοιχεία αυτά δείχνουν ότι τα μέλη ορισμένων εθνοτικών ομάδων είναι πιο πιθανό να χρεοκοπήσουν το δάνειό τους. Εντούτοις, από ηθική και νομική άποψη, είναι απαράδεκτο να χρησιμοποιεί κανείς την εθνικότητα ενός ατόμου για να αρνηθεί το δάνειο σε αυτόν, καθώς αυτό θα αποτελούσε παραβίαση των νόμων περί διακρίσεων. Σε τέτοιες περιπτώσεις, η εθνικότητα ενός προσώπου είναι πιθανό να είναι ένας φορέας πληροφοριών παρά ένας διακριτικός παράγοντας. οι άνθρωποι από μια συγκεκριμένη εθνοτική ομάδα είναι πιθανότερο να χρεοκοπούν το δάνειό τους, επειδή π.χ. το μέσο επίπεδο εκπαίδευσης σε αυτή την ομάδα είναι χαμηλότερο. Σε μια τέτοια κατάσταση είναι γενικά απολύτως αποδεκτό να χρησιμοποιείται επίπεδο εκπαίδευσης για την επιλογή υποψηφίων δανείων, παρόλο που αυτό θα οδηγούσε στην ευνοϊκή μεταχείριση μιας εθνικής ομάδας έναντι μιας άλλης. Η τράπεζα θα μπορούσε νόμιμα να αποφασίσει να χωρίσει την ομάδα των υποψηφίων δανείων ανάλογα με το επίπεδο εκπαίδευσης και να μάθει περισσότερα μοντέλα για κάθε μία από αυτές τις ομάδες ξεχωριστά. Προϋπόθεση για αυτήν την προσέγγιση ομαδοποίησης ή στρωματοποίησης είναι φυσικά ότι το επίπεδο εκπαίδευσης χαρακτηριστικών υπάρχει στο σύνολο δεδομένων. Όταν το χαρακτηριστικό αυτό δεν υπάρχει στο σύνολο δεδομένων, επειδή, π.χ., δεν καταγράφηκε, βρισκόμαστε σε μια κατάσταση που με τη χρήση της εθνικότητας, άμεσα ή έμμεσα, θα βελτιώσει την προβλεπτική ακρίβεια. Επί του παρόντος, οι αλγόριθμοι εξόρυξης δεδομένων για προγνωστικά μοντέλα εξόρυξης χρησιμοποιούν την ακρίβεια ως το πιο σημαντικό μέτρο ποιότητας και κριτήριο που πρέπει να βελτιστοποιηθεί. Στις περιπτώσεις που περιγράφονται παραπάνω, η συμπεριφορά αυτή μπορεί να οδηγήσει σε διακριτικά μοντέλα.

### Λογοδοσία

Πρέπει να σημειωθεί ότι σε μεμονωμένες περιπτώσεις είναι σχεδόν αδύνατο να αποδειχθεί η διάκριση. δεδομένης της πληθώρας των διαθέσιμων δεδομένων, σχεδόν πάντα είναι δυνατόν να βρεθεί εκ των υστέρων μια ουδέτερη εξήγηση για το λόγο για τον οποίο το άτομο δεν δέχτηκε δάνειο, χαρακτηριζόμενο ως τρομοκράτης κ.λπ. Εκτός από τα εντυπωσιακά στοιχεία όπως τα εσωτερικά σημειώματα που εξηγούν τη διακριτική διαδικασία που ακολουθήθηκε, τα δεδομένα είναι ανεπαρκή. Συνεπώς, η ίδια η διαδικασία λήψης αποφάσεων πρέπει να αποτελεί μέρος της αξιολόγησης των διακρίσεων.

Υπάρχουν δύο (συμπληρωματικές) προσεγγίσεις για την ενίσχυση της λογοδοσίας στην εξόρυξη δεδομένων πρόβλεψης. Το πρώτο είναι να βελτιωθεί η ανίχνευση της ιδιωτικής ζωής και των

ηθικών ευαισθησιών στα αποτελέσματα εξόρυξης δεδομένων ex ante. Οι Fule και Roddick για παράδειγμα έχουν προτείνει ένα σύστημα που σταθμίζει την ευαισθησία της επεξεργασίας δεδομένων και επισημαίνει αυτούς τους αλγορίθμους που είναι πιο πιθανό να περιέχουν ηθικές και ευαισθησίες στην ιδιωτικότητα. Ο Schreuders πρότεινε μια πιο γενική προσέγγιση που επικεντρώνεται στην ανασκόπηση της εξόρυξης δεδομένων και των διαδικασιών λήψης αποφάσεων χρησιμοποιώντας μια σειρά στοχοθετημένων ερωτήσεων.

Η δεύτερη προσέγγιση είναι να καθοριστεί αν ένας αλγόριθμος έχει κάνει διακρίσεις εκ των υστέρων. Αυτή είναι μια περιοχή που εξακολουθεί να είναι σε μεγάλα μέρη πειραματική. Ruggieri et al. προτείνουν μια προσέγγιση ειδικά προσαρμοσμένη στην ανίχνευση πιθανών διακρίσεων στην εξόρυξη δεδομένων. Στην προσέγγισή τους, οι Ruggieri et al. χρησιμοποιώντας δεδομένα ατόμων που εμπλουτίστηκαν με την απόφαση που ελήφθη για αυτά τα άτομα να ανιχνεύσουν κανόνες που καθορίζουν υποομάδες στα δεδομένα για τα οποία οι αποφάσεις διαφέρουν σημαντικά από τις συνολικές αποφάσεις.

### Διαφάνεια

Σχετικά με το θέμα της λογοδοσίας είναι η διαφάνεια. Μία από τις βασικές ανησυχίες σχετικά με τη δημιουργία προφίλ και την εξόρυξη δεδομένων είναι η έλλειψη διαφάνειας και τα επακόλουθα όρια αποκατάστασης από την αυτοματοποιημένη λήψη αποφάσεων. Η μεγαλύτερη διαφάνεια είναι επομένως σημαντική όταν πρόκειται για τη νόμιμη χρήση του προφίλ και της εφαρμογής της εξόρυξης δεδομένων. Η ευρωπαϊκή οδηγία για την προστασία των δεδομένων περιέχει ήδη μια διάταξη (άρθρο 15) που αναφέρει ότι τα υποκείμενα των δεδομένων δεν μπορούν να υποβάλλονται σε αυτοματοποιημένες διαδικασίες λήψης αποφάσεων που έχουν νομική ισχύ ή επηρεάζουν σημαντικά τη θέση τους. Στο σημερινό λόγο για την προστασία της ιδιωτικής ζωής, ωστόσο, το άρθρο 15 διαδραματίζει έναν δευτερεύοντα ρόλο. Ο νορβηγός εμπειρογνώμονας προστασίας δεδομένων Bygrave υποστήριξε ότι το άρθρο 15 πρέπει να θεωρηθεί ως μία από τις βασικές αρχές προστασίας δεδομένων. Συμφωνούμε με την Bygrave επί του σημείου αυτού και υποστηρίζουμε περαιτέρω ότι πρέπει να δοθεί μεγαλύτερη προσοχή στις απαιτήσεις των άρθρων 10, 11 και 12 της οδηγίας σχετικά με την ανακοίνωση πληροφοριών. Εκτός από την ενημέρωση του υποκειμένου των δεδομένων σχετικά με το σκοπό της επεξεργασίας των πληροφοριών, τα υποκείμενα των δεδομένων πρέπει να έχουν περισσότερες γνώσεις σχετικά με την πραγματική διαδικασία λήψης αποφάσεων που βασίζεται σε αυτοματοποιημένη απόφαση. Το άρθρο 12, παράγραφος α, τρίτη παύλα, της οδηγίας αναφέρει ήδη ότι τα υποκείμενα των δεδομένων πρέπει να ενημερώνονται για τη λογική και τη λογική της διαδικασίας λήψης αποφάσεων, αλλά από την στιγμή που η εφαρμογή του άρθρου 12 στην πράξη φαίνεται πολύ περιορισμένη.

Το νόμιμο καθήκον να είναι διαφανές θα πρέπει να μεταφερθεί στο σχεδιασμό των αλγορίθμων εξόρυξης δεδομένων και των συνοδευτικών υποδομών πληροφορικής, δηλαδή "διαφάνεια από το σχεδιασμό." Σε αυτό τον τομέα οι Weitzner et al., Προτείνουν ένα σύστημα στο οποίο ο χρήστης μπορεί να επαληθεύσει ορισμένες στοιχία για τον εαυτό του και να ελέγξει εάν οι διαδικασίες λήψης αποφάσεων με αυτόματη απάντηση είναι πράγματι σωστές.<sup>19</sup> Ένα τέτοιο σύστημα διαφάνειας που συμπληρώνεται από τον έλεγχο των χρηστών θα μπορούσε, σε μεγάλο βαθμό, να αφαιρέσει τους φόβους προστασίας της ιδιωτικής ζωής που προέρχονται από τη "μεταφορά Kafka". Θα πρέπει να σημειωθεί ότι δεν αφαιρεί τα αποτελέσματα του Big Brother (μάλλον θα μπορούσε να ενισχύσει και να επιβάλει αυτά τα αποτελέσματα, καθώς οι χρήστες έχουν μεγαλύτερη επίγνωση - μέσω αντιπαράθεσης - σχετικά με την ισχύ των δεδομένων εξόρυξης), ούτε η ίδια η διαφάνεια νομιμοποιεί τη χρήση προσωπικών δεδομένων.

### Αλγόριθμοι εξόρυξης δεδομένων με επίγνωση κινδύνου

Εκτός από την ανίχνευση αλγορίθμων εξόρυξης δεδομένων που ενδέχεται να επηρεάσουν αρνητικά τις ομάδες και τα άτομα, πρέπει να διασφαλίσουμε ότι αυτοί οι κίνδυνοι μετριάζονται μέσω του ίδιου του κώδικα προγραμματισμού. Αυτή η εφαρμογή της ιδέας του Lessig για τον "κώδικα ως νόμο" μπορεί να λάβει διάφορες μορφές.

Η ανάπτυξη αλγορίθμων εξορυκτικών δεδομένων για τη διατήρηση της ιδιωτικής ζωής υπήρξε εδώ και αρκετό καιρό πεδίο έρευνας. (.24) Η έννοια των αλγορίθμων εξόρυξης δεδομένων που είναι γνωστές για την προστασία της ιδιωτικής ζωής, όπως προτείνεται μεταξύ άλλων από τους Agrawal και Srikant, χρησιμοποιεί μεθόδους όπως ο αποκλεισμός ή η κάλυψη ευαίσθητων χαρακτηριστικών για τη διαφύλαξη της ιδιωτικής ζωής. Ωστόσο, όπως περιγράφεται παραπάνω, το πιο πιθανό είναι πιο αποτελεσματικό να κοιτάξουμε πέρα από το ιδιωτικό απόρρητο (το ιδιωτικό απόρρητο είναι μέσο και όχι αυτοσκοπός) και αλγόριθμοι προγραμμάτων για να κατανοήσουμε τους υποκείμενους κινδύνους, όπως είναι οι διακρίσεις. Σε αυτή την προσέγγιση θα χρειαζόταν να καθοδηγήσουμε το σύστημα σε ποιες πληροφορίες θα έπρεπε (ή δεν έπρεπε) να στηρίξει τις αποφάσεις του<sup>30</sup>. Στην περίπτωση διακρίσεων, για παράδειγμα, είναι επιθυμητό να έχουμε ένα μέσο να "πούμε" τον αλγόριθμο ότι οι προβλέψεις του να είναι ανεξάρτητο από ευαίσθητα χαρακτηριστικά όπως φύλο και εθνικότητα. Ονομάζουμε περιορισμούς ως προς την ανεξαρτησία ". (.76)

Χρησιμοποιώντας αυτούς τους περιορισμούς, μπορούμε να κατασκευάσουμε συστήματα που είναι γνωστά από το σχεδιασμό των διακρίσεων.

## Συμπεράσματα

Τα Big Data είναι ο επιστημονικός κλάδος “ομπρέλα” πολλών άλλων επιστημών, όπως το Data Mining και το Profiling. Σε γενικές γραμμές μπορούμε να πούμε ότι τα Big Data αναφέρονται στα δεδομένα αυτούσια (όγκος, ποικιλότητα και ταχύτητα), το Data Mining είναι η μέθοδος απόσπασης χρήσιμων και προηγουμένως αδιανόητων και/ή αδιάκριτων πληροφοριών από αυτά, ενώ το Profiling είναι μια προχωρημένη μέθοδος εξαγωγής συμπερασμάτων για τα υποκείμενα/πηγές των δεδομένων. Πρωταγωνιστής μας εδώ είναι το Profiling γιατί ενώ όπως αναφέραμε προηγουμένως υπάρχει από τότε που υπάρχουν άνθρωποι η δυναμική των Big Data και Data Mining του έχουν δώσει μια ασύλληπτη δυναμική για τον ανθρώπινο νου και οι επιπτώσεις της λάθος και αυθαίρετης χρήσης τους είναι ζωτικής σημασίας για τα υποκείμενα προς εξέταση, τους ανθρώπους.

Οι τράπεζες, είναι φυσικό συμπέρασμα, όχι μόνο έχουν αντιληφθεί την ανάγκη χρήσης των τεχνολογιών Big Data και Data Mining αλλά έχουν αρχίσει την εντατική υιοθέτηση τους, τη σταδιακή χρήση είτε σε αρχάρια και πιλοτικά επίπεδα είτε με τη μορφή στιβαρών επενδύσεων, και τις περιλαμβάνουν στο στρατηγικό τους όραμα. Οι τεχνολογίες αυτές δεν υποθέτετε αλλά έχουν ήδη αρχίσει να παράσχουν λύσεις σε προηγουμένως αν όχι απροσπέλαστα, σίγουρα δύσκολα και δαπανηρά προβλήματα, με πιο περίτρανα την παροχή πίστωσης σε αχαρτογράφητα κοινά και αναδυόμενες αγορές.

Πως λοιπόν μπορούμε να περιφρουρήσουμε την ορθή χρήση, να εμποδίσουμε τις καταχρήσεις και να προστατεύσουμε τον μέσω πολίτη από τις συνέπειές τους. Η Ευρωπαϊκή Ένωση βρίσκεται στο ηγετικό ρόλο της προστασίας των πολιτών από τέτοια φαινόμενα. Με το νέο Κανονισμό Προστασίας Προσωπικών Δεδομένων οι επιχειρεί αν ενισχύσει τη θέση των πολιτών της έναντι σε όλες τις φάσεις των Big Data και των επιπτώσεων του Profiling. Παρά τα ελαττώματα ή προκλήσεις του νέου κανονισμού, παραμένει το πιο ισχυρό μέσω προστασίας των πολιτών της. Ειδικά σε παγκόσμιο επίπεδο η ΕΕ έχει την πιο σαφή θέση πάνω στην προστασία των δεδομένων.



## Βιβλιογραφία

1. "The Big Data Market: 2014 - 2020 - Opportunities, Challenges, Strategies, Industry Verticals and Forecasts." The Big Data Market: 2014 - 2020 - Opportunities, Challenges, Strategies, Industry Verticals and Forecasts. N.p., n.d. Web. 19 Aug. 2014. <<http://www.marketresearchreports.biz/analysis-details/the-big-data-market-2014-2020-opportunities-challenges-strategies-industry-verticals-and-forecasts>>.
2. "Big data in the banking sector -." Big data in the banking sector. N.p., n.d. Web. 19 Aug. 2014. <<http://onetree-solutions.com/big-data-in-the-banking-sector/>>.
3. Six Provocations for Big Data
4. "OF DATA AND MEN" FUNDAMENTAL RIGHTS AND FREEDOMS IN A WORLD OF BIG DATA Gigabytes=10<sup>6</sup> byte, Terabytes=10<sup>9</sup> byte, Exabytes=10<sup>12</sup> byte, Zettabytes=10<sup>15</sup> byte
5. We shall not discuss here the indirect costs for the environment and health of the development of digital technologies.
6. or every 24 months, depending on the version consulted. Moore's law logically comes up against the physical limits of miniaturisation.
7. <http://www.nnaroup.com/articles/law-of-bandwidth/>
8. See, in particular, the report by the OECD's Working Party on the Information Economy and the Working Party on Information Security and Privacy, *Exploring the economics of personal data: a survey of methodologies for measuring monetary value*, 2 April 2013, DSTI/ICCP/IE/REG(2011)2/FINAL, [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP/IE/REG\(2011\)2/FINAL&docLanguage=EN](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DSTI/ICCP/IE/REG(2011)2/FINAL&docLanguage=EN)
9. For a very clear description of the phenomenon, see Delort P., *Le Big Data*, PUF, Collection "Que sais-je?", 2015.
10. Ibid
11. <http://online.wsi.com/news/articles/SB10001424052748704901104575423294099527212>
12. Sadin E., *Surveillance globale. Enquete sur les nouvelles formes de controle. Climats/Flammarion, 2009*
13. "Big Data." : Technologies and players. N.p., n.d. Web. 19 Aug. 2014. <<http://bigdata-tutorial.blogspot.co.uk/2012/01/big-data-technologies-and-players.html>>.
14. Big Data How it became a Differentiator (Deutsche Bank)
15. "Big Data Vendor Revenue And Market Forecast 2012-2017 - Wikibon." Big Data Vendor Revenue And Market Forecast 2012-2017 - Wikibon. N.p., n.d. Web. 19 Aug. 2014. [http://wikibon.org/wiki/v/Big\\_Data\\_Vendor\\_Revenue\\_and\\_Market\\_Forecast\\_2012-2017](http://wikibon.org/wiki/v/Big_Data_Vendor_Revenue_and_Market_Forecast_2012-2017)
16. Bagging και Boosting Μέθοδοι για την Κατασκευή Μοντέλων με Εφαρμογές στα Χρηματοοικονομικά, ΟΛΥΜΠΙΟΣ ΣΩΚΡΑΤΗΣ Ε.Μ.Π 2013
17. Challenges and Opportunities with Big Data, A community white paper developed by leading researchers across the United States. Executive Summary, 2015
18. Data Mining Techniques and its Applications in Banking Sector, Dr. K. Chitra<sup>1</sup>, B. Subashini Assistant Professor, Department of Computer Science, Government Arts College, Melur, Madurai. Assistant Professor, Department of Computer Science, V. V. Vanniaperumal College for Women, Virudhunagar.
19. DATA MINING TECHNIQUES AND APPLICATIONS, Mrs. Bharati M. Ramageri, Lecturer Modern Institute of Information Technology and Research, Department of Computer Application, Yamunanagar, Nigdi Pune, Maharashtra, India.
20. Customer Successes in your industry from [http://www.spss.com/success/?source=homepage&hpzone=nav\\_bar](http://www.spss.com/success/?source=homepage&hpzone=nav_bar).
21. Hildebrandt, Mireille; Gutwirth, Serge (2008). *Profiling the European Citizen. Cross Disciplinary Perspectives*. Springer, Dordrecht. doi:10.1007/978-1-4020-6914-7. ISBN 978-1-4020-6913-0.
22. H.R. Maturana and F. J. Varela, *Autopoiesis and Cognition: The Realization of the Living*, Dordrecht, Reidel, 1991.
23. A. Rouvroy, *Privacy, Data Protection, and the Unprecedented Challenges of Ambient Intelligence*, 2 *Studies in Ethics, Law, and Technology* 1, Article 3 (2008), at: <http://www.bepress.com/selt/vol2/iss1/art3>.
24. The limits of privacy in automated profiling and data mining, Bart W. Schermer
25. McKinsey Global Institute, *Big data: The next frontier for innovation, competition, and productivity*, June 2011 ("MGI Report"), [http://www.mckinsey.com/Insights/MGI/Research/Technology\\_and\\_Innovation/Big\\_data\\_The\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation); Thomas Edsall, *Let the Nanotargeting Begin*, NY Times Campaignstops Blog, April 15, 2012, <http://campaignstops.blogs.nytimes.com/2012/04/15/let-the-nanotargeting-begin>.
26. Adverse Event Reporting System (AERS), <http://www.fda.gov/drugs/guidancecomplianceregulatoryinformation/surveillance/adversedrugeffects/default.htm>.

27. Rita Rubin, How did Vioxx debacle happen?, USA Today, October 12, 2004, [http://www.usatoday.com/news/health/2004-10-12-vioxx-cover\\_x.htm](http://www.usatoday.com/news/health/2004-10-12-vioxx-cover_x.htm).
28. Jeremy Ginsberg, Matthew Mohebbi, Rajan Patel, Lynnette Brammer, Mark Smolinski & Larry Brilliant, Detecting Influenza Epidemics Using Search Engine Query Data, 457 Nature 1012 (2009).
29. Big Data for Social Good Initiative, Engineering Social Systems, <http://www.hsph.harvard.edu/ess/bigdata.html>; see, e.g., Amy Wesolowski & Nathan Eagle, Parameterizing the Dynamics of Slums, AAAI Spring Symposium 2010 on Artificial Intelligence for Development (AI-D), <http://ai-d.org/pdfs/Wesolowski.pdf>.
30. Washington Okori & Joseph Obua, Machine Learning Classification Technique for Famine Prediction, Proceedings of the World Congress on Engineering 2011, [http://www.iaeng.org/publication/WCE2011/WCE2011\\_pp991-996.pdf](http://www.iaeng.org/publication/WCE2011/WCE2011_pp991-996.pdf)
31. Jmeson Toole, Nathan Eagle and Joshua Plotkin, Quantifying Crime Waves, AAAI Spring Symposium 2010 on Artificial Intelligence for Development (AI-D), <http://ai-d.org/pdfs/Toole.pdf>.
32. Massoud Moussavi and Noel McGinn, A Model for Quality of Schooling, American Associations for the Advancement of Artificial Intelligence, 2010 AAAI Spring Symposium Series.
33. Information and Privacy Commissioner of Ontario & Future of Privacy Forum, Smart Privacy for the Smart Grid: Embedding Privacy into the Design of Electricity Conservation, November 2009, <http://www.ipc.on.ca/images/Resources/pbd-smartpriv-smartgrid.pdf>.
34. Katie Fehrenbacher, Introducing the Facebook Social Energy App, October 17, 2011, GigaOm <http://gigaom.com/cleantech/introducing-the-facebook-social-energy-app>.
35. In Europe: Directive 2004/52/EC of the European Parliament and of the Council of 29 April 2004 on the interoperability of electronic road toll systems in the Community; Commission Decision 2009/750/EC of 6 October 2009 on the definition of the European Electronic Toll Service and its technical element.
36. See, e.g., Carlo Ratti, Riccardo Maria Pulselli, Sarah Williams & Dennis Frenchman, Mobile Landscapes: Using Location Data from Cell-Phones for Urban Analysis, 33 Environment and Planning B: Planning and Design 727 (2006).
37. MGI Report, *supra* note 21.
38. Sastry Duri, Jeffrey Elliot, Marco Gruteser, Xuan Liu, Paul Moskowitz, Ronald Perez, Moninder Singh & Jung-Mu Tang, Data Protection and Data Sharing in Telematics, 9(6) Mobile Networks & Applications, 693 (2004).
39. See Article 29 Working Party, Opinion 2/2010 on online behavioral advertising (WP 171), 22 June 2010, [http://ec.europa.eu/justice\\_home/fsi/privacy/docs/wpdocs/2010/wp171\\_en.pdf](http://ec.europa.eu/justice_home/fsi/privacy/docs/wpdocs/2010/wp171_en.pdf); FTC Staff Report, SelfRegulatory Principles for Online Behavioral Advertising, February 2009, <http://www.ftc.gov/os/2009/02/P085400behavadreport.pdf>.
40. Omer Tene & Jules Polonetsky, To Track or 'Do Not Track': Advancing Transparency and Individual Control in Online Behavioral Advertising, 13 Minn. J. L. Sci. & Tech. 282 (2012).
41. MGI Report, *supra* note 21, at p. 68.
42. Duncan Douglass, An Examination of the Fraud Liability Shift in Consumer Card-Based Payment Systems, Federal Reserve Bank of Chicago (2009), [www.chicagofed.org/digital/ep\\_1qtr2009\\_part7\\_douglass.pdf](http://www.chicagofed.org/digital/ep_1qtr2009_part7_douglass.pdf).
43. Melissa Fach, Stats on Facebook 2012, Search Engine Journal, February 17, 2012, <http://www.searchenginejournal.com/stats-on-facebook-2012-infographic/40301>; also see 10 Key Statistics About Facebook, Experian Hitwise Blog, February 2, 2012, <http://www.experian.com/blogs/hitwise/2012/2/2/10-key-statistics-about-facebook>.
44. <http://www.google.com/intl/en/about/products/index.html>.
45. See Statement of the Federal Trade Commission Concerning Google/DoubleClick, FTC File No. 071-0170, December 20, 2007, <http://www.ftc.gov/os/caselist/0710170/071220statement.pdf>; and see Dissenting Statement of Commissioner Pamela Jones Harbour, <http://www.ftc.gov/os/caselist/0710170/071220harbour.pdf>.
46. Balachander Krishnamurthy & Craig Wills, Privacy Diffusion on the Web: A Longitudinal Perspective, October 2009, submitted as public comment to Federal Trade Commission Exploring Privacy Roundtable Series, <http://www.ftc.gov/os/comments/privacvroundtable/544506-00009.pdf>.
47. Nicole Perlroth, Revamping at Yahoo to Focus on Its Media Properties and Customer Data, NY Times, April 11, 2012, <http://bit.ly/HUeMM>.
48. A Historical Appraisal of Information Technology in Commercial Banking, BERNARDO BÁTIZ-LAZO AND DOUGLAS WOOD, 2002
49. Taming the Beast: A Scientific Definition of Fintech, Patrick Schueffel 2016
50. Data Driven Banking: Applying Big Data to accurately determine consumer creditworthiness, Shen Yi Man 2016
51. Η Τράπεζα Πειραιώς εισέρχεται δυναμικά στο χώρο της αξιοποίησης των Big Data \_ Newsbeast <https://www.newsbeast.gr/financial/arthro/1974053/i-trapeza-pireos-iserchete-dinamika-sto-choro-tis-axiopiis-ton-big-data>

52. Data Mining in Banking and Finance A Note for Bankers, Rajanish Dass, Indian Institute of Management Ahmedabad
53. Data Mining as a Solution for Data Management in Banking Sector, Vivek Bhambri 2013 JCSE International Journal of Computer Sciences and Engineering Open Access
54. "Counter-terrorism tools used to spot fraud - FT.com." Financial Times. N.p., n.d. Web. 19 Aug. 2014. <<http://www.ft.com/cms/s/2/796b412a-4513-11e2-838f-00144feabdc0.html#axzz35fP00Qua>>.
55. "Big data lends new Zest to banks' credit judgments - FT.com." Financial Times. N.p., n.d. Web. 19 Aug. 2014. <http://www.ft.com/cms/s/70/df64c0c-fadd-11e3-8959-00144feab7de.html#axzz35f8d6WvY>.
56. "Banks Using Big Data to Discover 'New Silk Roads'." The CIO Report RSS. N.p., n.d. Web. 19 Aug. 2014. <<http://blogs.wsj.com/cio/2013/02/06/banks-using-big-data-to-discover-new-silk-roads/>>.
57. Data Mining: A Competitive Tool in the Banking and Retail Industries, Dr. Madan Lal Bhasin 2006
58. TENE-BIG-DATA-FOR-ALL-PRIVACY AND USER CONTROL IN THE AGE OF ANALYTICS, Omer Tene and Jules Polonetsky
59. The limits of privacy in automated profiling and data mining, Bart W. Schermer
60. Vedder, A. (1999), KDD: The challenge to individualism. In: Ethics and Information Technology, Volume 1, Number 4/December, 1999.
61. Schermer, B.W. (2007), Software Agents, Surveillance, and the Right to Privacy: a Legislative Framework for Agent-enabled Surveillance, PhD. Thesis, Leiden University, p. 42
62. The Promise and Peril of Big Data, David Bollier 201
63. Big Data: The End of Privacy or a New Beginning?, Ira S. Rubinstein, October 2012 New York University School of Law
64. EDPS Opinion on coherent enforcement of fundamental rights in the age of big data, European Data Protection Supervisor (EDPS) Opinion 8/2016
65. Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection, Alessandro Mantelero Nanjing University of Information Science & Technology (NUIST), School of Public Administration, Jiangsu, China 2016
66. C-12/03 Commission v Tetra Laval BV [2005] ECR I-987.
67. [http://ec.europa.eu/competition/publications/cpn/2005\\_1\\_19.pdf](http://ec.europa.eu/competition/publications/cpn/2005_1_19.pdf)[http://ec.europa.eu/information\\_society/media\\_taskforce/doc/pluralism/media\\_pluralism\\_swp\\_en.pdf](http://ec.europa.eu/information_society/media_taskforce/doc/pluralism/media_pluralism_swp_en.pdf) [accessed 17.09.2016].
68. COM (2016)288, Communication from the Commission, 'Online Platforms and the Digital Single Market, Opportunities and Challenges for Europe'.
69. Resolution of The Spring Conference of data protection authorities, May 2016, available at: <http://www.naih.hu/budapest-springconf/files/Resolution—new-frameworks.pdf> [accessed 17.09.2016].
70. Hijmans H., The European Union as a constitutional guardian of Internet privacy and data protection, pp.63 - 65, <http://hdl.handle.net/1245/1.511969> [accessed 17.09.2016]. (A modified version of this thesis was to be published in Summer 2016 by Springer International Publishing *The European Union as Guardian of Internet Privacy*).
71. JRC Technical Reports, Institute For Prospective Technological Studies Digital Economy Working Paper 2016/05, An Economic Policy Perspective On Online Platform, pp.42-43; [https://ec.europa.eu/jrc/sites/jrcsh/files/JRC\\_101501.pdf](https://ec.europa.eu/jrc/sites/jrcsh/files/JRC_101501.pdf) [accessed 17.09.2016].
72. Consumer Justice Enforcement Forum (CoJEF) Enforcement of Consumer rights: strategies and recommendations, May 2016.
73. The full report of the study is available at: <http://fbrno.climg.no/wp-content/uploads/2014/02/2014-05-14-Unfair-cloud-storage-terms-report.pdf> [accessed 17.09.2016].
74. C-191/15, Verein fur Konsumenteninformation v Amazon EU Sarl, *judgment of 28 July 2016*.
75. President Juncker's Political Guidelines, 15.7.2014, [https://ec.europa.eu/priorities/publications/president-junckers-political-guidelines\\_en](https://ec.europa.eu/priorities/publications/president-junckers-political-guidelines_en) [accessed 16.09.2016]; BEUC response to consultation 'Empowering the National Competition Authorities to be more effective enforcers', 2016
76. Kamiran, F., Calders, T., Pecehnikziy, M. (2009), Building Classifiers with Independency Constraints. In: IEEE ICDM Workshop on Domain Driven Data Mining. IEEE press 1
77. The Effect of FinTech on the Traditional Norwegian Banking Sector , NORWEGIAN SCHOOL OF ECONOMICS 2017
78. The EU General Data Protection Regulation (GDPR) in the banking industry, An impact analysis on banks and wealth managers with the focus on Switzerland (PwC).
79. Daniel J Solove, The Digital Person: Technology and Privacy in the Information Age, New York University Press, 2006, p.60
80. Case C-131/12 Google Spain SL and Google Inc. v Agencia Espanola de Proteccion de Datos (AEPD) and Mario Costeja Gonzalez [2014] EU:C:2014:317.

81. *Mario Viola de Azevedo Cunha, Market Integration Through Data Protection - An Analysis of the Insurance and Financial Industries in the EU, Springer, 2013, p. 50*
82. *The General Data Protection Regulation: A New Opportunity and Challenge for the Banking Sector, Andra Giurgiu and Thierry Lallemand, August 2017 University of Luxembourg*
83. For an interpretation of the concept of "establishment" by the CJEU in its recent jurisprudence see *Triggering Jurisdiction for DPAs and Limitations of Their Sanctioning Powers (Case C-230/14, Weltimmo)*", *European Data Protection Law Review* no. 1/2015, pp. 310-315.  
 also Mark D. Cole and Andra Giurgiu, *"The 'Minimal' Approach: the CJEU on the Concept of 'Establishment' in*
84. *Krugman, Paul & Obstfeld, Maurice, International Economics. Theory and Policy, 6th ed., Pearson, Boston, 2003, p. 142.*
85. *Kotler, Philip & Armstrong, Gary, Principles of marketing, Pearson, Boston, 2nd ed., 2016, p. 325.*
86. *Shapiro, Carl & Varian, Hal, Information rules: a strategic guide to the network economy, Harvard Business School Press, Boston, 1999, pp. 39-40 [cit. Shapiro & Varian].*
87. *Ibid.*, p. 44.
88. *Shapiro & Varian*, pp. 44-45.
89. *Shapiro & Varian*, pp. 53-54.
90. *Ibid.*
91. *Shapiro & Varian*, pp. 40-44.
92. *Maggiolino*, p. 7.
93. *Personalized Pricing through Profiling, Wictor Bjorklund FACULTY OF LAW, Stockholm University 2017*