

Η μελλοντική διαχείριση της Γνώσης

Βασίλειος Νικολόπουλος

Διπλ. Ηλεκτρολόγος Μηχ. & Μηχ. Η/Υ Hons.Eng MSc DIC Maj.Ing(X99)

Υπ.Δρ. ΕΜΠ

Σύμβουλος Μηχανικός, European Dynamics SA

Email: vnikolop@iee.org

<http://www.medialab.ntua.gr/vnikolop>

Η έννοια της Πληροφορίας και η Γνώση

Η έννοια της πληροφορίας έχει αποκτήσει μεγάλη σημασία στην σημερινή ψηφιακή εποχή. Η καλή πληροφορία από μόνη της είναι μία κενή έννοια, η οποία αν μείνει ανεκμετάλλευτη δεν εξυπηρετεί κανέναν και τίποτα. Η πληροφορία γίνεται όμως εξαιρετικά ενδιαφέρουσα, εάν μέσω μιας κατάλληλης επεξεργασίας εξάγουμε από αυτή πολυδιάστατη και συνδυαστική γνώση, η οποία σε συνδυασμό με κάποιο ή κάποια μαθηματικά ή εννοιολογικά μοντέλα και κανόνες, δημιουργεί το φαινόμενο του συμπερασμού και της λογικής λήψης απόφασης (decision reasoning). Η λήψη απόφασης βασισμένη σε κάποια πληροφορία, είναι κάτι αρκετά πολύπλοκο, αν πρόκειται μάλιστα για λήψη βέλτιστης απόφασης. Και αυτό γιατί οι περισσότερες τεχνικές που ακολουθούνται σήμερα προσπαθούν να προσεγγίσουν την ανθρώπινη λογική ή κάποιο μέρος της (πχ. ανθρώπινος εγκέφαλος και νευρώνες) και να εξομοιώσουν την διαδικασία του ανθρώπινου συμπερασμού (πχ. νευρωνικά δίκτυα, ασαφής λογική κοκ). Ένας σημαντικός παράγοντας της ανθρώπινης λογικής και εξαγωγής συμπεράσματος είναι η μνήμη και οι ατομικές εμπειρίες. Ο ανθρώπινος εγκέφαλος και η διαδικασία λήψης απόφασης στηρίζεται πολύ πάνω στην προηγούμενη εμπειρία. Αυτό προσπαθούν να κάνουν τα Τεχνητά Νευρωνικά δίκτυα με κάποιες μεθόδους εκπαίδευσης (training sets), που προσπαθούν να δημιουργήσουν μία τεχνητή τράπεζα γνώσης, πάνω στην οποία να στηριχτούν οι διάφοροι κανόνες συμπερασμού (neural weights, threshold functions). Επίσης, η ανθρώπινη λογική δεν είναι διακριτή (crisp logic) αλλά ασαφής. Η ανθρώπινη γλώσσα, που εκφράζει τον ανθρώπινο συμπερασμό, περιέχει λέξεις όπως πολύ, λίγο, αρκετά και έτσι η θεωρία των ασαφών συνόλων (fuzzy groups) πλησίασε ακόμα πιο κοντά τις διάφορες τεχνικές επεξεργασίας και λήψης απόφασης (fuzzy logic). Τέλος, άλλες τεχνικές εξομοίωσης λήψης αποφάσεων, ειδικά σε επιχειρηματικά, κοινωνικά ή οικονομικά προβλήματα (πχ. θεωρία παιγνίων και Nash equilibrium) αποτελεί ένα τρόπο εξομοίωσης των εκάστοτε εφαρμοσμένων προβλημάτων και στρατηγικών (trade-production, prisoner's dilemma etc) και όχι τόσο γενικευμένη μέθοδο ανάλυσης, αναπαράστασης και εξαγωγής γνώσης και απόφασης.

Ένας πολύ αφαιρετικός ορισμός που ταιριάζει στον ψηφιακό κόσμο της Πληροφορικής, μπορεί να είναι ότι πληροφορία ονομάζεται οτιδήποτε είναι επεξεργάσιμο από μία μηχανή και μπορεί να συμβάλει στην ανάπτυξη και διαμόρφωση γνώσης. Τί όμως είναι γνώση και πως αυτή περιγράφεται? Και πώς μπορούμε να επεξεργαστούμε κάτι που λέγεται πληροφορία, ώστε να διαμορφώσουμε γνώση? Όπως λοιπόν και η ενέργεια στην φυσική, η πληροφορία βρίσκεται σε διαφορες μορφές στην φύση και περιγράφεται με διάφορους τρόπους, όπως ακριβώς και οι διάφορες ετερογενής ενεργειακές καταστάσεις (θερμότητα, κίνηση, δυναμικό) που περιγράφονται με Η/Μ κύματα, φωτόνια, μποζόνια,

ισοδυναμία De Broglie κοκ. Διάφορες μορφές πληροφορίας μπορεί να είναι το Internet, αρχεία word, η μνήμη ενός υπολογιστή, οι εμπειρίες ενός ανθρώπου, τα χαρακτηριστικά του και η συμπεριφορά του, η εμφανής καταπόνηση κάποιων υλικών, οι ψυχολογικές καταστάσεις των ανθρώπων κα. Για παράδειγμα, η φορτισμένη ψυχολογική κατάσταση ενός ανθρώπου, οδηγεί στην γνώση (συμπέρασμα) ότι θα ήταν επικίνδυνο να οδηγήσει .. Η θεωρία του λογικού συμπερασμού (reasoning) είναι μία πολύ σημαντική αλγοριθμική διαδικασία, η οποία αποτελεί το συνδετικό κρίκο μεταξύ της πληροφορίας και της γνώσης, όπως σε ένα φυσικό σύστημα χρειαζόμαστε μία μετατροπή ενέργειας από μία μορφή σε μία άλλη, για να την εξηγήσουμε ή να την χρησιμοποιήσουμε (μαθηματικός δυϊσμός). Όπως αναφέραμε, πολλές τεχνικές συμπερασμού και γενικώς Τεχνητής Νοημοσύνης έχουν δημιουργηθεί, για να μπορούμε να επεξεργαζόμαστε πληροφορία και αναλόγως να βγάζουμε συμπεράσματα (Νευρωνικά Δίκτυα, Ασαφής λογική, Γενετικοί αλγόριθμοι κοκ) και κάθε τεχνική χρησιμοποιείται ανάλογα με την εφαρμογή και το πρόβλημα. Στην ουσία όμως, έχουμε πάλι πολλές μορφές πληροφορίας και πολλές ετερογενείς τεχνικές επεξεργασίας της, όπως και με τα διαφορετικά πεδία στην φύση. Για να εξηγήσουμε την μεταφορά H/M ενέργειας σε μακροσκοπικό επίπεδο χρησιμοποιούμε τις εξισώσεις Maxwell, σε μικροσκοπικό (πχ. κυματοσυνάρτηση μέσα στο πυρίτιο) την εξίσωση Schrodinger ενώ για την βαρυτική μεταφορά ενέργειας την Νευτώνιο μηχανική (μέχρι την θεμελίωση των βαρυτικών κυμάτων). Η πληροφορία που εμπεριέχεται σε αυτές τις 3 μορφές μπορεί να είναι και η ίδια (πχ. ταχύτητα ενός σωματιδίου). Χρειαζόμαστε λοιπόν μία ενοποιημένη τεχνική για να περιγράψουμε ετερογενείς πληροφορία και να την επεξεργαστούμε. Κάτι τέτοιο όμως φαίνεται να συμβαίνει στην πληροφορική, όπου αναπτύχθηκαν κάποιες πλατφόρμες για την ομοιόμορφη αναπαράσταση της πληροφορίας, η οποία μπορεί να είναι ετερογενής.

XML και η ενοποίηση της Επιχειρησιακής Πληροφορίας

Αυτή λοιπόν η πολυδιάστατη προσέγγιση της πληροφορίας μας κάνει αμέσως να αναρωτηθούμε πώς μπορούμε να επεξεργαστούμε αυτήν την πληροφορία όταν βρίσκεται σε διαφορετικές μορφές και μπορεί να περιγράφει κάτι πολύ απλό. Για παράδειγμα, μία ακριβή τιμή μιας σύνδεσης ADSL αποτελεί πληροφορία η οποία βρίσκεται στο εταιρικό site του πάροχου και είναι προσπελάσιμη από έναν explorer ανταλλάσσοντας πακέτα HTTP, υπάρχει στην μνήμη ενός υπαλλήλου της εταιρίας που είναι ο πάροχος και είναι προσπελάσιμη απλά είτε ρωτώντας τον υπάλληλο είτε προσπελώνοντας την μνήμη του (!) ή υπάρχει μέσα σε ένα αρχείο word το οποίο είναι ο τιμοκατάλογος της εταιρίας και είναι προσπελάσιμη αν ανοίξουμε το αρχείο. Άλλο είδος πληροφορίας μπορεί να ενσωματώνεται και στην ψυχολογική η οπτική κατάσταση του ατόμου (πχ. κούραση, άγχος) η οποία είναι χρήσιμη διότι σε συνδιασμό με άλλη πληροφορία να μπορέσει να δημιουργήσει γνωστικό συμπέρασμα. Για παράδειγμα η ανίχνευση άγχους σε ένα άτομο σε συνδιασμό με την γνώση του ιατρικού του ιστορικού, το οποία μπορεί να είναι προσβάσιμο μέσω Internet και web services, θα μπορεί να συμβάλλει σε ένα online e-health σύστημα αρκετά αποδοτικό.

Εδώ και αρκετό καιρό λοιπόν πολλές επιχειρήσεις αναγνωρίζοντας την σημασία της γνώσης, αναπτύξανε συστήματα επεξεργασίας πληροφοριών, για δική τους χρήση. Τα συστήματα αυτά, γνωστά και ως OLAP (On-line analytical processing) ή OLTP (On-line transaction processing) όταν πρόκειται

για συστήματα συνδιαλλαγών (πχ. τραπεζικά), αποτελούν τον πυρήνα μιας νέας φιλοσοφίας και τεχνικής που ονομάζεται Επιχειρηματική Ευφυΐα (Business Intelligence). Η βάση του πληροφοριακού συστήματος που εφαρμόζει τις παραπάνω τεχνικές δεν είναι άλλη από μια Βάση δεδομένων που σε συνδυασμό με την markup γλώσσα XML, ενοποιεί και επεξεργάζεται ετερογενή πληροφορία. (πχ. CRM σύστημα -> Knowledge Management -> Business Intelligence -> profit -> best practices). Εδώ λοιπόν άρχισε η χρησιμοποίηση της XML και των XML-based γλωσσών και τεχνικών (XPath κα), ως μίας αποτελεσματικής markup γλώσσας για την περιγραφή ετερογενούς ημιδομημένης πληροφορίας. Οι τεχνικές ανάλυσης της πληροφορίας είναι πολλές (MOLAP, ROLAP, HOLAP Cubes etc) και τα αποτελέσματα πολύ ικανοποιητικά, αλλά ακόμα η πλήρης και αποτελεσματική ενοποίηση κάθε εξωγενούς πληροφορίας και η δυναμική της επεξεργασία για εξαγωγή απόφασης είναι μακριά. Ο λόγος ήταν ότι και πάλι με τα OLAP & BI συστήματα και την επιχειρηματική ευφυΐα, η βάση γνώσης που αναλύεται (CRM, ERP, MRP κα) είναι σε πλήρη δομημένη μορφή ή ημιδομημένη μορφή (XML, PMML κα) ακολουθώντας τα κλασικά πρότυπα των σχεσιακών βάσεων δεδομένων (RDBMS), πράγμα που καθιστά το σύστημα αρκετά κλειστό, μη ελαστικό και απόλυτο σε κάποιες τεχνολογίες (low interoperability). Για αυτό το λόγο, η σημερινή έρευνα έχει στραφεί προς 3 συνδυαστικές κατευθύνσεις που θα στιγματίσουν την νέα εποχή στην αναπαράσταση και διαχείριση της επιχειρησιακής γνώσης : το αμφίδρομο διαδίκτυο με τα web services και τις νέες τεχνικές Information Retrieval, τις ασύρματες επικοινωνίες και την περιβάλλουσα ευφυΐα (wireless-ambient intelligence) και την σημασιολογία (semantic networks).

Η μελλοντική διαχείριση γνώσης και το Business Process Modeling

Η νέα εποχή της γνώσης θέλει μία ενοποιημένη μαθηματική βάση για να εξηγήσει τα πάντα, μία ενοποιημένη μαθηματική συνάρτηση που να περιγράφει ικανοποιητικά κάποια πληροφορία που αλιεύεται από οποιαδήποτε ετερογενή πηγή, είτε με ενσύρματο (internet, web services) είτε με ασύρματο τρόπο (wireless intelligent LAN). Το πρώτο βήμα, αλλά όχι ακόμα ικανοποιητικό, είναι ότι διεξάγεται αυτή τη στιγμή σημαντική έρευνα πάνω σε οντολογίες, σημασιολογία (semantics), σημασιολογικά δίκτυα, RDF και OWL γλώσσες για τον ιστό, όπου πια η πληροφορία περιγράφεται εννοιολογικά για να είναι πολύ ευκολότερη η ανάκτησή της, η χρησιμοποιησή της και η σύγκρισή της. Το ανθρώπινο μυαλό στην παραγωγή γνώσης και απόφασης δουλεύει και συγκριτικά και για αυτό το λόγο η εξόρυξη και η συσταδοποίηση (clustering) έχει κερδίσει πολύ τα τελευταία χρόνια. Ήδη, τεχνικές βασισμένες στις οντολογίες και τα σημασιολογικά δίκτυα έχουν δημιουργηθεί, για την πλήρη περιγραφή και μοντελοποίηση μιας εταιρίας (Business Process Modeling), όπου η βάση πια είναι τα σημασιολογικά στοιχεία που συνθέτουν μια εταιρία (υπάλληλοι, εμπόρευμα, κέρδος) και οι σημασιολογικές σχέσεις μεταξύ τους (business models ontology). Η εξαγωγή απόφασης με βάση ένα business σημασιολογικό δίκτυο θα είναι η νέα τεχνική που θα χρησιμοποιηθεί για την δημιουργία web-based Decision Support systems, τα οποία θα συσσωρεύουν την εισερχόμενη γνώση από διάφορες εξωγενής πηγές με ενσύρματο ή ασύρματο τρόπο (CRM, ERP βάση, πωλητές, υπάλληλοι, surveys, τηλέφωνα, helpdesk, web crawling, Internet, κα), θα την ενοποιούν και με βάση συγκεκριμένους σημασιολογικούς κανόνες που θα καθορίζονται από την εκάστοτε επιχείρηση, θα εξάγεται

συμπέρασμα. Το Internet θα έχει έναν διπλό ρόλο σε αυτή τη νέα προσέγγιση προσθέτοντας ένα επιπλέον στρώμα (tier) καταλήγοντας σε μία νέα αρχιτεκτονική 4-tier για τα web-based DSS : 1) είναι πια το μέσο (User interface) για την πρόσβαση σε αυτά τα συστήματα δίνοντας μια τρομερή ευκολία στον οποιονδήποτε από οποιονδήποτε μέρος του κόσμου να έχει πρόσβαση όταν και όποτε θέλει σε on-line προχωρημένες υπηρεσίες (e-logistics, Energy Information Systems, Environmental networks, Point of Sales modeling etc) και 2) λειτουργεί σαν ιδεατό τεράστιο γνωσιακό δίκτυο αναζήτησης και μεταφοράς αναγκαίας δομημένης πληροφορίας από οποιονδήποτε μέρος του κόσμου χωρίς μισθωμένες γραμμές, χωρίς MAN και WAN δίκτυα και πολλές φορές χωρίς καν την χρησιμοποίηση στο front-end βάσεων δεδομένων, απλά μέσω των νέων Web services (WSDL, SOAP messages etc). Το πολυδιάστατο σχήμα της σημασιολογίας και η νέα βάση της μαθηματικής τοπολογίας θα αποτελέσει μελλοντικά ένα ενοποιημένο γνωσιακό πεδίο, όπως αυτό που έψαχνε κάποτε ο μέγας Einstein, στην προσπάθειά του να ενοποιήσει τις 4 μεγάλες δυνάμεις στην φύση, κάτω από την θεωρία του ενοποιημένου πεδίου (Unified field theory), κάτι το οποίο δεν κατάφερε. ..

Υπάρχουν πολλά ακόμα να γίνουν όμως, για να παραχθεί ένα πολύ ικανοποιητικό μαθηματικό-ενοσιολογικό μοντέλο περιγραφής ενοποιημένης πληροφορίας. Η XML, η RDF και η OWL αποτελούν αυτή την στιγμή σημείο αναφοράς για ετερογενή πληροφορία, αλλά η τοπολογία τους (περισσότερο ημιδομημένα ακυκλικά δέντρα) δεν αρκούν για να περιγράψουν την σφαιρικότητα και την τοπολογική πολυπλοκότητα του ανθρώπινου μυαλού και των πολύ σημαντικών **ανθρώπινων συνειρμών**. Τα κλειστά πολυδιάστατα σημασιολογικά δίκτυα μπορούν να λύσουν κάποια από τα προβλήματα αλλά ένας ανθρώπινος στοχαστικός συνειρμός, που αποτελεί σημαντική συνειστώσα της ανθρώπινης ευφυίας και του συμπερασμού είναι κάτι αρκετά πολύπλοκο και χρειάζεται προχωρημένη διαφορική τοπολογία για να εξηγηθεί. Το ερώτημα λοιπόν είναι η εύρεση μίας ικανής μαθηματικής τοπολογίας, η οποία να περιγράψει όσο το δυνατόν καλύτερα μία ετερογενή δομή δεδομένων, η οποία αποτελεί πληροφορία. Και όχι μόνο για την ενοποιημένη περιγραφή της, αλλά και για την επεξεργασία της με σκοπό την εξαγωγή συμπερασματικής γνώσης. Άρα υπάρχουν 2 συνιστώσες στο πρόβλημα: 1) Ικανό μαθηματικό μοντέλο περιγραφής ετερογενούς πληροφορίας και 2) σύνθετοι αλγόριθμοι επεξεργασίας της ετερογενούς πληροφορίας για γνωστικά συμπεράσματα και τον σχηματισμό γνώσης. Αν βρεθεί βέβαια αυτό (που θα βρεθεί και μάλιστα σε μορφή ολοκληρωμένου-σιλικόνης, γνωστό και ως knowledge integrated chip), τότε δε θα χρειαζόμαστε πια ούτε δομημένες βάσεις δεδομένων, ούτε ολόκληρα πληροφοριακά συστήματα με βάσεις, ούτε Συστήματα Λήψης Αποφάσεων, ούτε SQL για να παίρνουμε απαντήσεις από προκαθορισμένη - δομημένη πληροφορία που συχνά δεν βοηθάει πολύ, ούτε OLAP συστήματα για πολυδιάστατη ανάλυση δεδομένων από 3στατους κύβους και Data Warehouses, ούτε XML parsers. Αυτό που θα έχουμε ανάγκη είναι το ενοποιημένο μοντέλο περιγραφής, μία ολοκληρωμένη τοπολογία, γλώσσα και δομή που θα περιγράφει την πληροφορία και τους αλγόριθμους εκείνους που θα εξάγουν απόφαση, μία εξίσωση Maxwell δηλαδή.... Και αν αυτό μάλιστα ακολουθήσει την σύγχρονη τεχνολογία της νανομηχανικής, των οργανικών υπολογιστών και της ασύρματης επικοινωνίας, τότε θα μιλάμε για ένα πραγματικό, αμφίδρομο και κατανεμημένο μοντέλο αναπαράστασης γνώσης, που θα μπορεί μέσω wearable computing και wireless intelligence να γίνει αναπόσπαστο μέρος της ζωής μας, είτε σε μορφή organic wireless grid, είτε σε μορφή DNA knowledge microchips Αξίζει να περιμένουμε..!



* **Short CV** : Ο Βασίλης Νικολόπουλος γεννήθηκε το 1975 στην Αθήνα και αποφοίτησε από την Γαλλική Λεόντειο Σχολή. Είναι διπλ. Ηλεκτρολόγος Μηχ. & Μηχ. Η/Υ αποκτώντας 1st class 4ετής Σκοτσέζικο Honours BEng, Ηλεκτρολόγου Μηχανικού του Πανεπιστημίου Dundee (2000), ερχόμενος 1^{ος} σε 3 τμήματα της Πολυτεχνικής σχολής, MSc και Diploma σε Control systems από το Imperial College του Λονδίνου, Certificates σε Marketing & Management από το LSE και μετά την γαλλική Classe Préparatoire εισήχθη στην Ecole Polytechnique στο Παρίσι από όπου απέκτησε τις ειδικεύσεις Μηχανικού (Majeures Ingenieur) σε Εφαρμοσμένα Μαθηματικά και Πληροφορική (2002). Είναι υποψήφιος Δρ. στο εργαστήριο Τεχνολογίας Πολυμέσων, της σχολής ΗΜΜΗΥ του ΕΜΠ. Η διατριβή του εντάσσεται στο γνωστικό αντικείμενο των Συστημάτων Αποφάσεων και των Ενεργειακών Πληροφοριακών Συστημάτων (Decision Sciences). Από το 2005 είναι Σύμβουλος Μηχανικός στο E-business Division της European Dynamics SA. Είναι μέλος του TEE, IEE, IEEE, BCS, InstMC, IFAC, SEE, Γαλλική Μαθηματική Ένωση, του ΣΕΠΕ και του Ελληνο-Γαλλικού Οικονομικού και Βιομηχανικού Επιμελητηρίου.