



## Ciclo di eventi: **“Internet. Le parole chiave”**

L'associazione Nuova Civiltà delle Macchine, con la sponsorizzazione di FORMULA SERVIZI e in collaborazione con il COSMONAUTA, organizza:

**Lunedì 11 maggio ore 20,45**  
presso il Cosmonauta, via Regnoli 41

una serata dedicata all'approfondimento della tematica dei:

### **“BIG DATA”**

Il tema sarà trattato dalla dott.ssa **Sonia Bergamaschi**, docente di “Tecnologia delle basi di dati” c/o Dipartimento di Ingegneria - Università degli studi di Modena e Reggio Emilia.

*La mera creazione di dati (“datafication”) avrebbe scarso valore se non avesse un impatto sulla vita quotidiana. Parlare di Big Data implica non solo l'esistenza di vasti database ma soprattutto la definizione di un nuovo uso che di questi è possibile fare. La disponibilità di dati permette di correlare fenomeni diversi e valutarne l'evoluzione nel breve e nel lungo periodo. Ciò può implicare la rinuncia all'accuratezza dei dettagli e alla minimizzazione degli errori - propria delle tecniche statistiche basate sull'analisi di piccoli campioni – ma permette di fare previsioni che tengano conto di numerose variabili e per ricomporre una visione d'insieme altrimenti non ottenibile.*

La serata è aperta a tutti. E' pensata come un momento divulgativo e prevedendo la possibilità di ospitare una interazione con i presenti.

L'evento organizzato dalla Associazione Nuova Civiltà delle Macchine è attuato con:

- la sponsorizzazione di **FORMULA SERVIZI**
- con la collaborazione del **COSMONAUTA**

-con il sostegno del **Comune di Forlì** e della **Fondazione della Cassa dei Risparmi di Forlì**

[www.nuovaciviltadellemacchine.it](http://www.nuovaciviltadellemacchine.it)

## **BIG DATA: tendenze e sfide.**

BIG DATA è un termine popolare per descrivere la crescita esponenziale, la disponibilità e l'uso, delle informazioni, sia strutturate che non strutturate. Molto è stato scritto sulla grande tendenza della crescita dei dati e sulle potenzialità per l'innovazione e la crescita delle imprese. IDC, una importante società di consulenza specializzata nella tecnologia dell'informazione, consiglia alle organizzazioni e responsabili IT di concentrarsi sul crescente "VOLUME, VARIETA', VELOCITA', VERIDICITA' e COMPLESSITA'" delle informazioni che danno forma ai Big Data.

- **VOLUME.** L'aumento del volume dei dati è dovuto a molti fattori: dati di transazioni ripetute nel tempo e memorizzati nel corso degli anni, dati di testo costantemente in streaming da social media, incremento dei dati raccolti da sensori (internet of things), etc. In passato, un eccessivo volume di dati ha creato un problema di stoccaggio, ma oggi, con i costi unitari della memoria che continuano a calare, altri problemi emergono e, accanto al gestire grandi volumi di dati, diventa rilevante la pertinenza tra i grandi volumi di dati e il come creare valore dai dati che vengono gestiti.
- **VARIETA'.** I dati oggi disponibili provengono da tanti tipi di formati diversi: da database tradizionali, da archivi di dati gerarchici creati dagli utenti finali, dai sistemi OLAP (On-Line Analytical Processing), dai documenti di testo, da e-mail, dalla raccolta dati effettuabile con contatori - strumenti di misura - sensori, da video, da audio, da dati di borsa e da transazioni finanziarie, ecc..
- **VELOCITA'.** Secondo Gartner, la velocità riguarda sia quanto velocemente i dati vengono prodotti sia quanto velocemente i dati devono essere trattati per soddisfare la domanda. Reagire abbastanza rapidamente per affrontare la velocità è una sfida per la maggior parte delle organizzazioni.
- **VERIDICITA' (Esattezza, Accuratezza).** La scarsa qualità dei dati porta ad affrontare i problemi con informazioni non affidabili.
- **COMPLESSITA'.** Quando si tratta di enormi volumi di dati provenienti da fonti multiple è un impegno rilevante il collegare, l'accoppiare, il pulire e trasformare i dati tra sistemi.

Metodologie tradizionali di Business Intelligence sono basate su datawarehouse (magazzini di dati), adottando Data Base Management System "Relazionali", e le tecnologie statistiche applicate ai dati privati.

D'altra parte, big data sociali non strutturati sono emersi dal web, e hanno spinto i grandi giocatori web (Google, Facebook, Twitter, Amazon, ecc.) ad inventare nuove tecnologie di gestione dei dati basate sul parallelismo massiccio e sul paradigma "map-reduce" e nuove tecniche di previsione/analisi: il riferimento è alla Web Intelligence. Web Intelligence denota l'insieme di tecniche di Intelligenza Artificiale a scala Web per rendere operativa una intelligenza predittiva. Ad esempio, prevedere l'intenzione e l'interesse per la pubblicità online, la misurazione della fiducia dei consumatori e prevedere il comportamento, migliorare l'interazione uomo-macchina come in "Watson" (progetto IBM), categorizzando e riconoscendo i luoghi, i volti, le persone, la medicina genomica personalizzata del futuro, la creazione di servizi pubblici più intelligenti, etc.

L'impatto dei dati resi disponibili sul web e delle tecnologie ha spinto verso un più potente scenario di "Big Data Analitici". L'idea è quella di fondere tecniche di "social intelligence" e di "business intelligence" in un mix di dati privati e dati del web e sfruttare le combinazioni di una tecnologia di gestione dati sviluppata da società di web per le loro attività di web-intelligence. Inoltre, stanno emergendo altre tecnologie di gestione dei dati: DBMS (Data Base Management System) in memoria centrale, DBMS verticali, ecc..

## Sonia Bergamaschi. Sintetica biografia

Sonia Bergamaschi è nata a Modena (Italia) e ha ricevuto la sua laurea in Matematica presso l'Università di Modena nel 1977. Attualmente è professore ordinario di Ingegneria Informatica presso il Dipartimento di Ingegneria "Enzo Ferrari" - Università di Modena e Reggio Emilia e conduce il "DBGGroup", il gruppo di ricerca sui database ([www.dbgroup.unimo.it](http://www.dbgroup.unimo.it)).

La sua attività di ricerca è stata dedicata principalmente alla rappresentazione e gestione della conoscenza nel contesto di grandi database affrontando aspetti sia teorici che di implementazione. Dal 1985 è stata molto attiva nel settore delle tecniche di Intelligenza Artificiale (Description Logics) e nelle tecniche di database per sviluppare sistemi di database intelligenti. Su questo argomento sono stati ottenuti risultati teorici importanti ed è stato sviluppato un sistema "ODB-Tools" che effettua un controllo di consistenza e l'ottimizzazione semantica delle interrogazioni Object Oriented Database. Dal 1999, i suoi sforzi di ricerca sono stati dedicati al settore della ricerca Intelligent Information Integration. È stato sviluppato un sistema di integrazione dei dati, chiamato MOMIS, che fornisce un accesso integrato a dati strutturati e semistrutturati e che permette di presentare una singola query e ricevere una singola risposta unificata. Nel 2009 ha partecipato alla fondazione della start-up di origine universitaria chiamata "DataRiver", il cui obiettivo è stato lo sviluppo di una versione open source del sistema MOMIS (prima release su aprile 2010 ([www.datariver.it](http://www.datariver.it))). Dal 2010 (e fino ad ora) la sua attività di ricerca è stata estesa a: Parole chiave di ricerca su database, Semantic Web e annotazione automatica delle fonti di dati. Recentemente, i suoi sforzi di ricerca sono stati dedicati ai Big Data e Big Data Analytics.

Sonia Bergamaschi è stata coordinatrice e partecipe di molti progetti europei in ambito ICT: SEWASIE (2002-2005), WINK (2002-2003), STASI (2006-2009), FACIT-SME (2010-2012), Keystone (2013-2017). E' stata coordinatrice del progetto MURST FIRB "NeP4B (Networked. Peers for Business)" (2006-2009).

I suoi lavori sono stati pubblicati in oltre un centinaio di occasioni: in riviste internazionali e negli atti di conferenze. Le sue ricerche sono state finanziate dal MURST italiana, CNR, ASI e progetti comunitari. Ha partecipato a comitati internazionali e conferenze, internazionali e nazionali, inerenti le problematiche dei Data Base e dell'AI. E' membro della IEEE Computer Society e di ACM. Per una descrizione dettagliata dell'attività di ricerca e dei sistemi sviluppati vedere: [www.dbgroup.unimo.it](http://www.dbgroup.unimo.it).