

DANIELE MARIA BRAGA - Curriculum Vitae et Studiorum

1. INFORMAZIONI PERSONALI

1.1 Dati biografici

<i>Luogo e data di nascita</i>	Albenga (SV) - 7 Ottobre 1975
<i>Cittadinanza</i>	Italiana
<i>Codice Fiscale</i>	BRG DLM 75R07 A145 L
<i>Stato civile</i>	Celibe
<i>Obblighi militari</i>	Congedato
<i>Indirizzo di residenza</i>	Via Valsesia, 44 - 20152 - Milano (MI)
tel.	casa: 02-4565204 ufficio: 02-23993661 cell.: 328-6883118
e-mail:	braga@elet.polimi.it
<i>Titolo di Studio</i>	<i>Dottorato in Ingegneria dell'Informazione</i> , Politecnico di Milano, aprile 2005
<i>Posizione Attuale</i>	<i>Assegnista di Ricerca</i> presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano
<i>Lingue straniere parlate</i>	Inglese fluente

1.2 Percorso formativo

<i>Luglio 1994</i>	Diploma di Maturità Classica con votazione 60/60 presso il Liceo Ginnasio G. Parini di Milano.
<i>Aprile 2001</i>	Laurea in Ingegneria Informatica , presso il Politecnico di Milano, votazione 94/100, discutendo la tesi " <i>Studio e Realizzazione di un Traduttore dal linguaggio XML-GL ai linguaggi Lorel e XQuery</i> ", relatore il prof. Stefano Ceri.
<i>Maggio 2001</i>	Inizia a collaborare alle attività di ricerca del gruppo di Basi di Dati del Politecnico di Milano.
<i>Settembre 2001</i>	Vince un Assegno di ricerca finanziato da Microsoft Research® presso il Politecnico di Milano, denominato " <i>Tecnologie e metodi per applicazioni Web applicati ad un Campus Virtuale</i> ".
<i>Agosto 2002</i>	Partecipa alla "EDBT '02 Summer School: Distributed databases on the Net, Models, Languages, and Infrastructures" svoltasi a Cargese (FR).
<i>Settembre 2004</i>	Partecipa come speaker alla "EDBT '04 Summer School: XML & Databases" svoltasi a Pula (CA).
<i>Aprile 2005</i>	Vince un Assegno di ricerca a carico del Dipartimento di Elettronica e Informazione del Politecnico di Milano, denominato " <i>Architettura dei sistemi di elaborazione: sistemi di collaudo e sistemi di valutazione delle prestazioni</i> ".
<i>Maggio 2005</i>	Consegue il Dottorato di Ricerca in Ingegneria dell'Informazione presso il Politecnico di Milano, discutendo la tesi dal titolo " <i>XQBE: XQuery By Example</i> ", relatore il prof. Stefano Ceri.

2. ATTIVITÀ SCIENTIFICA

2.1 Attività di ricerca

L'attuale ricerca di Daniele Braga si concentra principalmente sull'estensione delle potenzialità dei linguaggi per l'interrogazione, la manipolazione e l'analisi di dati semi-strutturati, e più specificamente dati rappresentati in XML (eXtensible Markup Language). I linguaggi di riferimento sono XPath, XQuery e XSLT (<http://www.w3.org>).

Un'area di ricerca più applicata è la progettazione di piattaforme e ambienti per il content management e la fruizione di contenuti didattici multicanale, con particolare riferimento alla loro archiviazione e indicizzazione e alla specifica delle loro proprietà tramite opportuni metadati.

La ricerca si articola pertanto in diverse aree:

2.1.1 Interfacce visuali (XQBE)

XQuery è un linguaggio testuale pensato da programmatori per programmatori, inadatto agli utenti privi di un background informatico specifico. Per contro, XML è usato nei campi applicativi più disparati, e un numero sempre crescente di utenti ha (e avrà in futuro) necessità di interrogare dati semi-strutturati codificati in XML, tipicamente senza conoscere XQuery.

XQBE è un linguaggio grafico che funge da interfaccia visuale a XQuery, ispirato al paradigma di QBE (un linguaggio visuale per interrogazioni su dati relazionali) e basato su una metafora visuale di interrogazione ereditata da XML-GL, un altro linguaggio di interrogazione visuale per XML definito nel '98 al Politecnico di Milano. I costrutti visualisono stati progettati in modo da corrispondere al meglio con i costrutti di XQuery e coprirne il massimo sottoinsieme, senza però rinunciare alla semplicità e alla intuitività - criteri di progetto fondamentali di XQBE.

In questo contesto le direzioni di ricerca specifica sono innumerevoli:

- la formulazione di interrogazioni agevolata dalla disponibilità di una specifica (anche parziale) dello schema dei documenti oggetto dell'interrogazione;
- l'approfondimento e la formalizzazione rigorosa delle tecniche di traduzione di XQBE verso vari linguaggi destinazione (tra cui XQuery e XSLT, ma non solo);
- l'estensione del potere espressivo, con il supporto per viste e funzioni;
- la generazione automatica della versione XQBE di una generiche query testuali;
- l'uso della stessa interfaccia per la generazione di comandi di update;
- la rappresentazione visuale di interrogazioni approssimate;
- il supporto ad altri servizi sui dati, dai comportamenti reattivi alla specifica intensionale di regole di riscrittura e di vincoli per il mantenimento dell'integrità dei dati stessi.

I risultati di tale ricerca sono stati presentati in diverse conferenze [5, 7, 8, 9]. Inoltre a [8] è seguito l'invito a pubblicare una versione estesa su rivista, mentre è stata pubblicato un articolo per le transaction di riferimento del settore [1] che definisce formalmente la sintassi e la semantica consolidate di XQBE.

Un prototipo dell'interfaccia visuale è scaricabile liberamente dal sito che descrive i vari progetti, (<http://dbgroup.elet.polimi.it/xquery/XQBE.html>), accessibile anche dalla pagina ufficiale del W3C che riporta le iniziative collegate a XQuery (<http://www.w3.org/XML/Query#products>).

2.1.2 Comportamenti attivi e reattivi in repository XML

Alla specifica di XQuery come linguaggio di interrogazione è stata affiancata da subito una proposta di linguaggio di update. Traendo ispirazione al noto paradigma delle regole attive per

SQL è stata proposta [3] una sintassi per la specifica dichiarativa dei comportamenti reattivi, derivata dalla combinazione della sintassi di XQuery con quella dei trigger in SQL-3. La principale difficoltà affrontata è la definizione della semantica della loro esecuzione, dovuta al fatto che il modello dei dati di XML è gerarchico: un evento che occorre ad un elemento complesso si ripercuote su tutti i suoi sotto-elementi, e bisogna gestire l'interazione concorrente di tutte le regole attivate da tale evento su tutti i sotto-elementi. Si stanno studiando le potenzialità di questo paradigma per l'imposizione automatica di vincoli di integrità dei dati e per la personalizzazione e configurazione dinamica dei Web Services in base a un profilo degli utenti.

2.1.3 Data mining

L'estensione di XQuery alle tecniche di data mining si propone di ridefinire nel contesto gerarchico dei dati semi-strutturati il concetto di *regola di associazione*, tradizionalmente sviluppato nell'ambito delle basi di dati relazionali, e di adattare opportunamente la sintassi di un noto operatore relazionale (MINE RULE) alla specifica dichiarativa delle proprietà delle regole di associazione tra frammenti di documenti XML. La difficoltà fondamentale è rappresentata dall'assenza di una base univoca per il conteggio degli "elementi" (item) rilevanti per l'analisi statistica, quale era il computo delle tuple nel modello tradizionale. La definizione del contesto di ricerca dei frammenti rilevanti per un particolare *mining task* diviene pertanto una parte determinante della specifica dichiarativa del task stesso. I risultati sono stati pubblicati e presentati nel settembre e novembre 2002 [12, 4] e nel marzo 2003 [6].

2.1.4 Interrogazioni approssimate e ranking dei risultati (FuzzyXPath)

FuzzyXPath è una versione estesa di XPath (a sua volta un sottoinsieme di XQuery), arricchito di costrutti capaci di esprimere interrogazioni approssimate, mediante una formulazione dichiarativa e intuitiva che utilizzi (ma nel contempo mascheri il più possibile) i principi della logica fuzzy. L'obiettivo è la definizione di un insieme minimo di operatori di base che forniscano la possibilità di risolvere e computare l'ambiguità derivante dall'approssimazione sia sui valori (prossimità nei domini) che sulla struttura (metriche di prossimità tra schemi), e di realizzare di conseguenza delle librerie di funzioni XQuery che implementino i predicati fuzzy di base e la loro composizione in operatori di livello via via più elevato. Basando l'estensione su funzioni scritte direttamente in XQuery si ottiene il vantaggio di poter effettuare la valutazione in modo "trasparente" su qualsiasi XQuery engine esistente. I risultati preliminari di questa ricerca sono stati pubblicati e presentati nel settembre 2002 [12].

2.2 Partecipazione a progetti di ricerca

IST cInQ (consortium on discovering knowledge with Inductive QQueries - **Progetto Europeo**) - L'obiettivo del consorzio è studiare il processo di 'scoperta della conoscenza' nelle basi di dati a partire da dati atomici. La tecnologia più recente per ricavare informazioni dallo strato primitivo dei dati è il *data mining*, che rappresenta la punta estrema dell'attuale tendenza a collezionare e ripulire i dati di transazioni, allo scopo di facilitare le decisioni. I risultati attesi del progetto cInQ sono quelle di facilitare il percorso che porta dai dati alla conoscenza e quindi alle decisioni. Il consorzio coinvolge Institut National des Sciences Appliquées (Lyon), Università di Torino, Politecnico di Milano, Albert-Ludwigs-Universität (Freiburg), Nokia Research Center (Helsinki), Jozef Stefan Institute (Ljubljana, Slovenia). (<http://www.cinq-project.org>)

CNR ECD (Technologies and Services for Enhanced Content Delivery - **Progetto Nazionale**) - Il progetto si concentra sullo sviluppo di tecnologie e strumenti per offrire contenuti arricchiti (enhanced contents delivery) a una vasta gamma di utenti finali, cioè identificare materiale digitale presente su fonti diverse, trasformarlo, organizzarlo, aggiungervi metadati e informazioni utili a qualificarlo, così da offrire agli utenti ciò che più è rilevante per i loro interessi. In

particolare il progetto si propone di ripensare le tecnologie adottate dalle biblioteche digitali e dai sistemi di ricerca e di accesso al Web. I problemi di ricerca connessi sono molteplici, tra cui lo sviluppo di algoritmi per indicizzare e ricercare documenti in formato compresso, le tecniche di High Performance Computing per fronteggiare le moli dei dati e il numero di utenti, le tecniche di Web Mining per determinare il rank o l'autorevolezza delle fonti e classificare i documenti, l'utilizzo di XML per strutturare documenti ed esprimere metadati, la formulazione e la valutazione di interrogazioni su dati e schemi XML. (<http://www-ecd.cnuce.cnr.it>)

Virtual Campus (Progetto del Politecnico di Milano finanziato da **Microsoft Research**[®]) - Il progetto riguarda lo sviluppo di un sistema di **e-learning** aperto per la progettazione, la realizzazione, la fruizione e la validazione di contenuti didattici caratterizzati da una proprietà fondamentale: la possibilità di essere *riusati*. Gli obiettivi fondamentali sono il supporto al progetto, alla composizione e al riuso di Learning Object, la realizzazione di un ambiente per la loro condivisione e fruizione, il loro uso cooperativo, l'analisi a posteriori dei dati di utilizzo per estrarre dati di sintesi sul comportamento degli utenti e personalizzare in modo accurato i servizi offerti a ognuno. (<http://www.elet.polimi.it/res/vcampus>)

2.3 Partecipazione a Comitati di programma di conferenze internazionali

- SAC-ICAR 2005, Santa Fé, New Mexico, USA

2.4 Partecipazione a Conferenze Internazionali

- VLDB 2001, Rome, Italy
- DaWak 2002, Aix-En-Provence, FR, presentando l'articolo [11] (in 4.2)
- Eurofuse 2002, Varenna, Italy
- ACM-SAC 2003, Melbourne, FL, USA, presentando gli articoli [5, 6] (in 4.2)
- ACM-WWW 2003, Budapest, Hungary, presentando l'articolo [7] (in 4.2)
- IEEE-WISE 2003, Rome, Italy, per la demo software legata alla presentazione di [8] (in 4.2)
- EDBT 2004, Heraklion, Greece, presentando l'articolo [13] (in 4.2)
- ACM-WWW 2005, Tokyo, Japan, per la presentazione di [9] (in 4.2)
- ACM-SIGMOD/PODS 2005, Baltimore, MD, USA, per la presentazione di [10] (in 4.2)

2.5 Riconoscimenti Conseguiti

Vincitore del **Best student poster award** alla conferenza internazionale **ACM-WWW 2003** per la seguente presentazione: *D. BRAGA, A. Campi, S. Ceri, E. Augurusa "XQuery By Example"* (<http://www.www2003.org/news.htm>)

2.6 Attività di revisione per conferenze

È stato revisore per le seguenti conferenze internazionali:

- EDBT 2002 - (Extending Data Base Technology)
- ACM-SAC 2002 - (Symposium on Applied Computing)
- IEEE-ICECCS 2002 - (Int'l Conf. on Engineering of Complex Computer Systems)
- ACM-SAC 2003 - (Symposium on Applied Computing)
- EDBT 2004 - (Extending Data Base Technology)
- ICWE 2004 - (Int'l Conf. On Web Engineering)
- VLDB 2004 - (Int'l Conf. on. Very Large Data Bases)
- SEBD 2005 - (Sistemi Evoluti per Basi di Dati)

3. ATTIVITÀ DIDATTICA

Daniele Braga è attualmente titolare di due corsi universitari, svolge e ha svolto esercitazioni e seminari didattici presso il Politecnico di Milano [PoliMi] e presso l'Università degli Studi di Bergamo [UniBg], e in quanto collaboratore del gruppo di Basi di Dati del Politecnico di Milano coordina e supervisiona le attività di tesi svolte dagli studenti all'interno del gruppo.

3.1 Titorarietà di corsi universitari curricolari

- a.a. 05/06: [PoliMi] Corso di **Informatica I**, per il corso di laurea in Ing. dell'Automazione;
[PoliMi] Corso di **Progetto di Basi di Dati** per il corso di laurea in Ing. Informatica;
- a.a. 04/05: [PoliMi] Corso di **Progetto di Basi di Dati** per il corso di laurea in Ing. Informatica;

3.2 Seminari didattici nell'ambito di corsi universitari curricolari

- a.a. 05/06: [PoliMi] Esercitazioni: **Basi di Dati**, prof. Stefano Ceri, sede di **Milano** (INF);
[PoliMi] Esercitazioni: **Basi di Dati 2**, prof. Stefano Paraboschi, sede di **Milano** (INF);
[PoliMi] LaureaOnLine: **Informatica I**, del prof. P. Della Vigna, guida di **1 classe**;
- a.a. 04/05: [PoliMi] Esercitazioni: **Basi di Dati**, prof. Stefano Ceri, sede di **Milano** (INF);
[PoliMi] Esercitazioni: **Basi di Dati**, prof. Alessandro Campi, sede di **Milano** (TLC);
[PoliMi] Esercitazioni: **Basi di Dati 2**, prof. Stefano Paraboschi, sede di **Milano** (INF);
[PoliMi] LaureaOnLine: **Informatica I**, del prof. P. Della Vigna, guida di **2 classi**;
[PoliMi] Esercitazioni: **Informatica II**, prof. Mario Verdicchio, sede di **Milano** (TLC);
[PoliMi] Esercitazioni: **Ingegneria del Software**, prof. L. Baresi, sede di **Milano** (INF);
[PoliMi] Esercitazioni: **Informatica II**, prof. Luciano Baresi, sede di **Cremona** (INF);
- a.a. 03/04: [PoliMi] Esercitazioni: **Basi di Dati**, prof. Stefano Ceri, sede di **Milano** (INF);
[PoliMi] LaureaOnLine: **Informatica I**, del prof. P. Della Vigna, guida di **2 classi**;
[PoliMi] Esercitazioni: **Informatica A** della Prof. Maristella Matera, sede di **Milano**;
[PoliMi] Esercitazioni: **Informatica II**, prof. Luciano Baresi, sede di **Milano** (TLC);
[PoliMi] Esercitazioni: **Informatica II**, prof. Luciano Baresi, sede di **Cremona** (INF);
- a.a. 02/03: [PoliMi] Esercitazioni: **Basi di Dati** (N.O.), prof. Stefano Ceri, sede di **Milano** (INF);
[PoliMi] Esercitazioni: **Basi di Dati** (V.O.), prof. Stefano Paraboschi, sede di **Milano**;
[PoliMi] Esercitazioni: **Informatica II**, prof. Luciano Baresi, sede di **Milano** (TLC);
- a.a. 01/02: [PoliMi] Esercitazioni: **Informatica A+B**, prof. S. Paraboschi, sede di **Cremona** (GES);
[PoliMi] Esercitazioni: **Informatica C**, prof. Giuseppe Psaila, sede di **Milano** (INF);
[PoliMi] Esercitazioni: **Informatica II**, prof. Luciano Baresi, sede di **Milano**;
[UniBg] Esercitazioni: **Sistemi Operativi e Reti di Calcolatori**, prof. S. Paraboschi, sede di **Dalmine**;

3.3 Coordinamento e supervisione di tesi di laurea

È stato relatore o correlatore o è attualmente supervisore di circa 15 tesi e 50 tesine per gli esami di laurea presso il Politecnico di Milano (per la maggior parte relative ai progetti di ricerca a cui collabora)

4. ELENCO DELLE PUBBLICAZIONI

4.1 Riviste internazionali

- [1] D. BRAGA, A. Campi, S. Ceri
"XQBE (XQuery By Example): a visual interface to the standard XML query language"
ACM-TODS (Transactions on Database Systems), Vol. 30 (2), June 2005

- [2] D. BRAGA, A. Campi
"XQBE: A Graphical Environment to Query XML Data"
SPRINGER - World Wide Web: Internet and Web Information Systems, Vol. 8 (3), Sept. '05
Invited. The issue contains extended versions of the best IEEE-WISE '03 papers (cfr. [8])

4.2 Atti di Conferenze

Conferenze ACM e IEEE

- [3] A. Bonifati, D. BRAGA, A. Campi, S. Ceri
"Active XQuery"
IEEE-ICDE 2002, San José, California - USA, February 2002

- [4] D. BRAGA, A. Campi, S. Ceri, M. Klemettinen, PL. Lanzi
"A Tool for Extracting XML Association Rules from XML Documents"
IEEE-ICTAI 2002, Washington DC, USA, November 2002

- [5] E. Augurusa, D. BRAGA, A. Campi, S. Ceri
"Design and Implementation of a Graphical Interface to XQuery"
ACM-SAC 2003, Melbourne, Florida - USA, March 2003

- [6] D. BRAGA, A. Campi, S. Ceri, M. Klemettinen, PL. Lanzi
"Discovering Interesting Information in XML Data with Association Rules"
ACM-SAC 2003, Melbourne, Florida - USA, March 2003

- [7] D. BRAGA, A. Campi, S. Ceri, E. Augurusa
"XQuery By Example" (poster presentation)
ACM-WWW 2003 (winner of the Best student poster award), Budapest, Hungary, May 2003

- [8] D. BRAGA, A. Campi
"A Graphical Environment to Query XML Data with XQuery"
IEEE-WISE 2003 (Int'l Conf. on Web Inf. Systems Eng.), Rome, Italy, December 2003

- [9] D. BRAGA, A. Campi, R. Cappa, D. Salvi
"Generating XSL Transformations with XQBE"
ACM-WWW 2005 (Poster Track), Chiba, Japan, May 2005

- [10] D. BRAGA, A. Campi, S. Ceri, A. Raffio
"XQBE: a Visual Environment for Learning XML Query Languages"
ACM-SIGMOD 2005 (Demo Session), Baltimore, Maryland (USA), June 2005

Altre conferenze internazionali

- [11] D. BRAGA, A. Campi, S. Ceri, M. Klemettinen, PL. Lanzi
"Mining Association Rules from XML Data"
DEXA 2002 (DaWaK), LNCS 2454, Aix-en-Provence, France, September 2002

- [12] D. BRAGA, A. Campi, E. Damiani, PL. Lanzi, G. Pasi
"XPath: Flexible Querying of XML Documents"
EuroFuse 2002 (Int'l Workshop on Information Systems), Varenna, Italy, September 2002
- [13] D. BRAGA, A. Campi, S. Ceri
"XQBE: A Graphical Interface for XQuery Engines"
EDBT'04 (Software Demo Paper), **LNCS 2992**, Heraklion, Greece, March 2004
- [14] D. BRAGA, A. Campi, S. Ceri
"XML Challenges for the Database Community: Past, Present, and Future"
Invited, **DataX'04** (Inte'l Workshop on Database Technologies for Handling XML information on the Web), Heraklion, Greece, March 2004

Conferenze nazionali

- [15] D. BRAGA, A. Campi, D. Martinenghi, A. Raffio, D. Salvi
"XQBE: the Swiss Army Knife for Semi-structured Data"
SEBD 2005, Bressanone-Brixen, Italy, June 2005

4.3 Riviste Italiane

- [16] D. BRAGA, A. Campi, S. Ceri
"XML: rappresentare e interrogare dati semi-strutturati"
Mondo Digitale (AICA) – Giugno 2005

4.4 Tesi di dottorato e di laurea

D. BRAGA
"XQBE: XQuery By Example"
relatore il prof. Stefano Ceri, Politecnico di Milano – Aprile 2005

D. BRAGA
"Studio e Realizzazione di un Traduttore dal linguaggio XML-GL ai linguaggi Lorel e XQuery"
relatore il prof. Stefano Ceri, Politecnico di Milano – Aprile 2001

Milano, 2 novembre 2005

Daniele M. Braga

Autorizzo il trattamento di questi dati ai sensi della normativa vigente (675/96 e succ. modificazioni e integrazioni)