

PROGETTO DI UNA UNITÀ DI RICERCA - MODELLO B  
Anno 2005 - prot. 2005095190\_002

**PARTE I**

**1.1 Programma di Ricerca afferente a**

*Area Scientifico Disciplinare 09: Ingegneria industriale e dell'informazione 100%*

---

**1.2 Durata del Programma di Ricerca**

*24 Mesi*

---

**1.3 Coordinatore Scientifico del Programma di Ricerca**

**CATARCI**                      **TIZIANA**                      *catarci@dis.uniroma1.it*

*ING-INF/05 - Sistemi di elaborazione delle informazioni*

*Università degli Studi di ROMA "La Sapienza"*

*Facoltà di INGEGNERIA*

*Dipartimento di INFORMATICA E SISTEMISTICA "Antonio Ruberti"*

---

**1.4 Responsabile Scientifico dell'Unità di Ricerca**

**TANCA**                                      **LETIZIA**  
*Professore Ordinario*                      *09/05/1955*                      *TNCLTZ55E49H703J*

*ING-INF/05 - Sistemi di elaborazione delle informazioni*

*Politecnico di MILANO*

*Facoltà di INGEGNERIA dell'INFORMAZIONE*

*Dipartimento di ELETTRONICA E INFORMAZIONE*

*02/23993531*                                      *02/23993411*                                      *tanca@elet.polimi.it*  
*(Prefisso e telefono)*                                      *(Numero fax)*                                      *(Indirizzo posta elettronica)*

---

**1.5 Curriculum scientifico del Responsabile Scientifico dell'Unità di Ricerca**

**Testo italiano**

*Letizia Tanca si è laureata in Matematica all'Università di Napoli, e ha poi lavorato per 4 anni nell'industria delle Telecomunicazioni. Letizia Tanca ha ottenuto il dottorato di ricerca in Matematica Applicata e Informatica nel 1988, ed ha in seguito lavorato presso il Politecnico di Milano come ricercatore e poi come professore associato. Dal 1995 è stata professore di I fascia presso l'Università di Verona, dove è rimasta fino all'ottobre del 1998, svolgendo tra l'altro il ruolo di Presidente del Consiglio di Corso di Laurea in Informatica. Letizia Tanca è attualmente professore ordinario presso il Politecnico di Milano; durante la sua carriera ha insegnato principalmente in corsi di Informatica di base e di Basi di Dati. Letizia Tanca è autore di diverse pubblicazioni internazionali sulle basi di dati e sulla teoria delle basi di dati, e del libro "Logic Programming and Databases", scritto con S. Ceri e G. Gottlob. Ha partecipato a vari progetti nazionali e internazionali.*

I suoi interessi di ricerca riguardano tutta la teoria delle basi di dati, in particolare le basi di dati deduttive, attive e orientate agli oggetti, i linguaggi a grafi per basi di dati, la semantica delle basi di dati evolute, la rappresentazione e l'interrogazione di informazione semistrutturata, in particolare l'informazione disponibile sul World Wide Web. Negli anni più recenti si è occupata di basi di dati dipendenti dal contesto per dispositivi mobili. E' stata revisore di varie riviste internazionali e membro del comitato di programma di varie conferenze nazionali e internazionali.

E' attualmente presidente del Consiglio di Corso di Studi di Ingegneria Informatica del Politecnico di Milano, campus Milano Leonardo.

#### **Testo inglese**

Letizia Tanca studied Mathematics at the University of Naples, where she obtained her Master's degree in Mathematical Logic; she then worked for four years as a software engineer in telecommunication firms.

Letizia Tanca obtained her Ph.D. in Computer Science in 1988. Later, she remained with Politecnico di Milano first as a research associate and then as an associate professor. She joined the University of Verona in 1995 as a full professor, and stayed there as the Dean of the Computer Science course until 1998. Letizia Tanca is currently with Politecnico di Milano again, as a full professor.

During her career she has taught and teaches courses on Databases and the Foundations of Computer Science. She is the author of several papers on databases and database theory, published in international

journals and conferences, and of the book "Logic Programming and Databases",

coauthored with S. Ceri and G. Gottlob. She has taken part to several national and international projects.

Her research interests range over all database theory, especially on deductive, active and object oriented databases, graph-based languages for databases, the semantics of advanced database and information systems,

representation and querying of semistructured information, particularly information on the World Wide Web. More recently, her research has been concentrated on context aware data management for mobile computing.

Letizia Tanca has been a referee of several international journals, and a member of the program committee of a large number of international conferences. She is currently chairperson of the degree and master Courses in Computer Engineering at the Politecnico di Milano, Leonardo campus.

---

### **1.6 Pubblicazioni scientifiche più significative del Responsabile Scientifico dell'Unità di Ricerca**

1. BOLCHINI C., SCHREIBER F. A., TANCA L. (2004). A context-aware methodology for very small data base design. SIGMOD RECORD. vol. 33 pp. 71-76 ISSN: 0163-5808
2. BOLCHINI C., SALICE F., SCHREIBER F. A., TANCA L. (2003). LOGICAL AND PHYSICAL DESIGN ISSUES FOR SMART CARD DATABASES. ACM TRANSACTIONS ON INFORMATION SYSTEMS. vol. 21 pp. 254-285 ISSN: 1046-8188
3. CORTESI A., DOVIER A., QUINTARELLI E., TANCA L. (2002). Operational and abstract semantics of the query language G-Log. THEORETICAL COMPUTER SCIENCE. vol. 275(1-2) pp. 521-560 ISSN: 0304-3975
4. COMAI S., TANCA L. (2001). Termination and Confluence by Rule Prioritization. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING. ISSN: 1041-4347 (accettato per la pubblicazione).
5. OLIBONI B., TANCA L. (2001). A Visual Language should be easy to use: A step forward for XML-GL. INFORMATION SYSTEMS. ISSN: 0306-4379 accettato per la pubblicazione.
6. CERI S., COMAI S., DAMIANI E., FRATERNALI P., PARABOSCHI S., TANCA L., PARABOSCHI S. (1999). XML-GL: a graphical language for querying and restructuring XML documents. COMPUTER NETWORKS. vol. 31 pp. 1171-1187 ISSN: 1389-1286
7. BARALIS E., GARZA P., QUINTARELLI E., TANCA L. (2004). Summarizing XML Data by Means of Association Rules. EDBT Workshops. (pp. 260-269).
8. COMAI S., MARRARA S., TANCA L. (2004). XML Document Summarization: Using XQuery for Synopsis Creation. DEXA Workshops. (pp. 928-932).
9. COMAI S., MARRARA S., TANCA L. (2003). Representing and Querying Summarized XML Data. DEXA. (pp. 171-181).
10. DAMIANI E., OLIBONI B., QUINTARELLI E., TANCA L. (2003). Modeling Semistructured Data by Using Graph-Based Constraints. OTM Workshops. (pp. 20-21).
11. DAMIANI E., TANCA L., ARCELLI FONTANA F. (2000). Fuzzy XML Queries via Context Based Choice of Aggregations. KYBERNETIKA. vol. 36:6 pp. 635-655 ISSN: 0023-5954
12. DAMIANI E., TANCA L. (2000). Blind Queries to XML Data. DEXA. (pp. 345-356).
13. PAREDAENS J., PEELMAN, P., TANCA L. (1998). Merging graph-based and Rule-based Computation: the language g-log. DATA & KNOWLEDGE ENGINEERING. vol. 25(3) pp. 267-300 ISSN: 0169-023X
14. FRATERNALI, P., TANCA L. (1995). A structured Approach for the definition of the semantics of Active databases. ACM TRANSACTIONS ON DATABASE SYSTEMS. vol. 20(4) pp. 414-471 ISSN: 0362-5915
15. PAREDAENS, J., PEELMAN, P., TANCA L. (1995). G-log: a graph based query language. IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING. vol. 7(3) pp. 436-453 ISSN: 1041-4347
16. CERI, S., FRATERNALI, P., PARABOSCHI S., TANCA L. (1994). Automatic generation of Production rules for integrity maintenance. ACM TRANSACTIONS ON DATABASE SYSTEMS. vol. 19(3) pp. 367-422 ISSN: 0362-5915
17. BOLCHINI C., CURINO C., GIORGETTA M., GIUSTI A., MIELE A., SCHREIBER F.A., TANCA L. (2004). PoLiDBMS:

*Design and Prototype Implementation of a DBMS for Portable Devices. SEBD. (pp. 166-177).*

18. COMAI S., TANCA L. (2002). *Editorial message: special track on web and E-business applications. SAC. (pp. 1086-1087).*
19. DAMIANI E., LAVARINI N., MARRARA S., OLIBONI B., PASINI D., TANCA L., VIVIANI G. (2002). *The APPROXML Tool Demonstration. EDBT. (pp. 753-755).*
20. OLIBONI B., QUINTARELLI E., TANCA L. (2001). *Temporal aspects of semistructured data. TIME. (pp. 119-127).*
21. CERI S., COMAI S., DAMIANI E., FRATERNALI P., TANCA L. (2000). *Complex Queries in XML-GL. SAC. (pp. 888-893).*
22. PAREDAENS J., PEELMAN P., TANCA L. (1998). *MERGING GRAPH-BASED AND RULE-BASED COMPUTATION: THE LANGUAGE G-LOG. DATA & KNOWLEDGE ENGINEERING. vol. 25 pp. 267-300 ISSN: 0169-023X*
23. COMAI S., TANCA L. (1997). *Using the Properties of Datalog to Prove Termination and Confluence in Active Databases. Rules in Database Systems. (pp. 100-117).*
24. DI NITTO E., TANCA L. (1996). *DEALING WITH DEVIATIONS IN DBMSS: AN APPROACH TO REVISE CONSISTENCY CONSTRAINTS. FMLDO 1996. (pp. 11-24).*

## **1.7 Risorse umane impegnabili nel Programma dell'Unità di Ricerca**

### **1.7.1 Personale universitario dell'Università sede dell'Unità di Ricerca**

#### **Personale docente**

n°	Cognome	Nome	Dipartimento	Qualifica	Settore Disc.	Mesi Uomo	
						1° anno	2° anno
1.	TANCA	Letizia	Dip. ELETTRONICA E INFORMAZIONE	Prof. Ordinario	ING-INF/05	5	5
2.	QUINTARELLI	Elisa	Dip. ELETTRONICA E INFORMAZIONE	Ricercatore Universitario	ING-INF/05	6	6
3.	SCHREIBER	Fabio Alberto	Dip. ELETTRONICA E INFORMAZIONE	Prof. Ordinario	ING-INF/05	4	4
4.	SBATTELLA	Licia	Dip. ELETTRONICA E INFORMAZIONE	Prof. Associato	ING-INF/05	4	4
<b>TOTALE</b>						<b>19</b>	<b>19</b>

#### **Altro personale**

*Nessuno*

### **1.7.2 Personale universitario di altre Università**

#### **Personale docente**

*Nessuno*

#### **Altro personale**

*Nessuno*

### **1.7.3 Titolari di assegni di ricerca**

*Nessuno*

---

### **1.7.4 Titolari di borse**

*Nessuno*

---

### **1.7.5 Personale a contratto da destinare a questo specifico programma**

n° Qualifica	Costo previsto	Mesi Uomo		Note
		1° anno	2° anno	
1. <i>Borsista</i>	40.000	11	11	
2. <i>Altre tipologie</i>	9.000		5 progettista software	
<b>TOTALE</b>	<b>49.000</b>	<b>11</b>	<b>16</b>	

---

### **1.7.6 Personale extrauniversitario indipendente o dipendente da altri Enti**

*Nessuno*

## **PARTE II**

### **2.1 Titolo specifico del programma svolto dall'Unità di Ricerca**

#### **Testo italiano**

*Accesso e gestione dell'informazione su sistemi mobili per la collaborazione semantica in comunità virtuali*

#### **Testo inglese**

*TIMELY AND UBIQUITOUS INFORMATION ACCESS AND MANAGEMENT IN COLLABORATIVE MULTI-KNOWLEDGE ENVIRONMENTS*

---

### **2.2 Settori scientifico-disciplinari interessati dal Programma di Ricerca**

*ING-INF/05 - Sistemi di elaborazione delle informazioni*

---

### **2.3 Parole chiave**

#### **Testo italiano**

*GESTIONE DEI DATI IN SISTEMI MOBILI ; SISTEMI SENSIBILI AL CONTESTO ; SISTEMI PER LAVORO COOPERATIVO ; INTEGRAZIONE DEI DATI GUIDATA DA ONTOLOGIE ; BASI DI DATI PER PICCOLISSIMI SISTEMI ; SINCRONIZZAZIONE BASATA SULLA SEMANTICA ; DISSEMINAZIONE MULTICANALE DELL'INFORMAZIONE*

#### **Testo inglese**

*UBIQUITOUS INFORMATION MANAGEMENT ; CONTEXT-AWARE SYSTEMS ; COOPERATIVE INFORMATION SYSTEMS ; ONTOLOGY DRIVEN DATA INTEGRATION ; VERY SMALL DATABASES ; SEMANTIC BASED SYNCHRONISATION ; MULTICHANNEL DELIVERY*

---

### **2.4 Base di partenza scientifica nazionale o internazionale**

#### **Testo italiano**

*L'ambiente comunicativo che ci circonda quotidianamente*

*è sempre più caratterizzato da dispositivi mobili, che permettono lo scambio di informazioni provenienti da media differenti e l'accesso a servizi multimediali complessi.*

*L'accesso ad informazioni complesse è attualmente affidato a sistemi basati sia su sorgenti informative eterogenee e fisse che su dispositivi mobili; di conseguenza i concetti di cooperazione, dipendenza da contesto, adattività, aleatorietà, distribuzione multi-canale e mobilità sono fondamentali. Questi aspetti influiscono sia sull'interazione uomo-macchina che sulla gestione delle informazioni, cioè sulla specifica di informazioni, sulla loro presentazione e sullo scambio.*

*In particolare, agli aspetti di mobilità e di distribuzione multi-canale è stata prestata particolare attenzione anche grazie agli aumenti di prestazioni di dispositivi come telefoni cellulari, palmari e computer portatili.*

*Negli ultimi anni la ricerca si è concentrata notevolmente su aspetti riguardanti l'adattività e la personalizzazione dell'interazione dell'utente [18, 1].*

*I risultati di questo lavoro approfondito permettono ai progettisti e agli utenti di estendere l'accesso alle informazioni e di migliorare l'efficienza dello stesso, ma richiedono una sostanziale modifica dei paradigmi di gestione e utilizzo delle informazioni, coinvolgendo concetti quali il contesto, il canale di comunicazione coinvolto nel dialogo, e il ruolo svolto dall'ambiente.*

*Il termine contesto è stato studiato approfonditamente nell'area dell'interazione uomo-macchina; è associato principalmente al concetto di posizione, ma è molto più complesso di questo; alcuni lavori hanno evidenziato la presenza di diverse categorie di contesto, quali,*

*ad esempio, il contesto computazionale, quello legato all'utenze, il contesto fisico e temporale [7, 8, 12, 14, 16, 17]. Nell'area delle basi di dati e dei sistemi informativi il ruolo del contesto è stato considerato approfonditamente solo recentemente [3, 6].*

*Il problema dell'integrazione di molteplici sorgenti eterogenee è stato affrontato e pubblicato in letteratura già dalla fine degli anni '70.*

*L'integrazione è tradizionalmente concepita come il problema di combinare dati memorizzati in sistemi informativi diversi in modo da fornire all'utente una vista unificata e consistente degli stessi [15].*

*La presenza di molte sorgenti informative (in termini sia di fonti di dati che di*

servizi) che possono anche sovrapporsi parzialmente, nelle quali lo schema e/o l'interfaccia di accesso possono essere sconosciuti, e la presenza di fornitori di informazioni diversi con vocabolari e criteri di classificazione distinti (anche all'interno dello stesso dominio applicativo) portano alla creazione di un sistema informativo le cui potenzialità non sono conosciute apriori dall'utente.

Questo concetto di contesto richiede tecniche di integrazione più sofisticate (come quelle proposte in [13]), coinvolgendo l'utilizzo di ontologie di dominio [8] e di servizi informativi generali [2].

In questo scenario le caratteristiche dei dispositivi in termini di risorse (spesso limitate se si considerano la capacità di memoria, la durata della batteria, la potenza di calcolo, ...) hanno un impatto significativo sia su come selezionare la porzione di dati di cui disporre localmente, sia sul modo in cui tali dati sono gestiti e processati. Per far fronte a questi aspetti l'unità di ricerca ha in passato sviluppato il prototipo di un dbms relazionale "leggero", che sfrutta tecniche innovative di memorizzazione ed accesso ai dati per ottimizzare le prestazioni e ridurre il consumo di potenza, orientato ai dispositivi mobili [4]. In futuro si estenderà tale approccio al caso di dati semistrutturati basati su XML. D'altra parte, anche gli aspetti relativi alla scelta dinamica della porzione di dati da sostituire con i dati nuovi dovrà essere basata su lavori relativi a strategie di caching basate su criteri semantici [9].

Questo ambito evolve in modo naturale verso le recenti tematiche di ricerca inerenti la condivisione di risorse e la

condivisione coordinata di risorse e rilevazione di informazioni in comunità dinamiche, virtuali multi-istituzioni, e le proposte di semantic grid [11]. Le applicazioni tipiche che trovano un beneficio immediato da tali ricerche sono la sanità [19], il turismo elettronico [6], "e-science" [10].

### **Testo inglese**

*The communication environment surrounding our daily experience is more and more characterized by mobile devices, which can exchange multimedia information and provide access to multimedia services of complex nature. Access to complex information is today entrusted to systems based both on heterogeneous fixed sources of information and mobile devices, where cooperation, context-dependence, adaptability, transiency, multichannel delivery and ubiquity are key concepts.*

*This affects both human-computer interaction and information management, i.e., information specification, presentation, exchange and delivery.*

*In particular, mobility and multichannel delivery have received great attention also due to the increasing performance of devices such as cellular phones, PDAs and notebooks.*

*In recent years a consistent research effort has been devoted to the issues of adaptiveness and personalization in user interaction [18, 1]. The results of this extensive work allow designers and users to extend information access and efficiency, but require a deep revision of the paradigms of information management and use, in terms of issues such as the context, the sensorial channels involved in dialogue and the role of the environment.*

*The term context has been extensively studied in the area of human-computer interaction; it is mainly associated to the concept of location, but is far richer than that; some works have underlined different categories of context, such as computational, user, physical and temporal context [7, 8, 12, 14, 16, 17]. In the area of databases and information systems, however, the role of the context has been largely unexplored until recent times [3, 6].*

*Integration of multiple heterogeneous sources has been studied and published in the literature since the end of the 70's. Integration is traditionally conceived as the problem of combining data stored in different systems and providing the user with a unified and consistent view [15].*

*The presence of many information sources (in terms of both data repositories and services) which may partially overlap, where schema and/or interface may be unknown, and the presence of many information providers with different vocabularies and classification criteria (even in a same application domain) leads to an information system whose information potentialities are apriori unknown to the user.*

*Such a context calls for more sophisticated integration techniques (as the ones presented in [13]), involving the use of domain ontologies [8] and generalized information services [2].*

*In this scenario the technological characteristics of the devices in terms of resources (often limited when considering data storage, battery life, computational power, ...) significantly affect both how the data to be stored locally is selected and how it can be managed and processed. To cope with these critical aspects, a prototype "light" relational dbms has already been*

*developed by the research unit exploiting innovative data storage and access policies aimed at optimizing performance and minimizing power consumption, targeting mobile devices [4]. Future work will*

*extend such an approach to semistructured data, in particular based on XML. On the other hand, research on aspects related to the dynamic choice of the data portion to be replaced with the new data found will be based on previous work on semantic-based caching strategies [9].*

*This framework naturally evolves towards recent researches envisaging coordinated resource sharing and information discovery in dynamic,*

multi-institutional virtual organizations, such as the semantic grid proposal [11]. Typical applications which take immediate benefit from such researches are e-health [19], e-tourism [6], e-science [10].

## 2.4.a Riferimenti bibliografici

- [1] T. Barbieri, A. Bianchi, L. Sbatella, *Multimodal Communication for Vision and Hearing Impairments. Conference and Workshop on Assistive Technologies for Vision and Hearing Impairment (CVHI 2004), Granada (Spain) 2004.*
- [2] D. Bianchini, V. De Antonellis, B. Pernici and P. Plebani, *Ontology based Methodology for e-Service discovery, Journal of Information Systems, Special Issue, 2004.*
- [3] C. Bolchini, F. A. Schreiber, and L. Tanca. *A context-aware methodology for very small data base design. SIGMOD Rec., 33(1):71-76, 2004.*
- [4] C. Bolchini, C. Curino, M. Giorgetta, A. Giusti, A. Miele, F. A. Schreiber, L. Tanca. *PolLiDBMS: Design and Prototype Implementation of a DBMS for Portable Devices. Proc. Dodicesimo Convegno Nazionale sui Sistemi Evoluti per Basi di Dati S. Margherita di Pula, I, pp. 166-177, 2004*
- [5] S. Castano, A. Ferrara, S. Montanelli. *Methods and Techniques for Ontology-based Semantic Interoperability in Networked Enterprise Contexts. CAiSE Workshops (3): 261-264, 2004*
- [6] A. Celentano, F. A. Schreiber, and L. Tanca. *Requirements for context-dependent mobile access to information services. Proceedings Tenth International Workshop on Multimedia Information Systems, College Park (MA) 2004, pp.60-65.*
- [7] G. Chen and D. Kotz. *A survey of context-aware mobile computing. Technical Report TR2000-381, Dartmouth College, Department of Computer Science, 2000.*
- [8] H. Chen, T. Finin, and A. Joshi. *An ontology for contextaware pervasive computing environments. Special Issue on Ontologies for Distributed Systems, Knowledge Engineering Review, 18(3):197-207, 2004.*
- [9] S. Dar, M.J. Franklin, B. Jonsson, D. Srivastava, M. Tan. *Semantic data Caching and Replacement, Proc. 22nd VLDB Conference. Bombay, India, 1996.*
- [10] D.C. De Roure, Y. Gil, J.A. Hendler (eds). *IEEE Intelligent Systems, Special Issue on E-Science, Volume 19, Issue 1 (January/February), 2004. ISSN: 1094-7167.*
- [11] D. De Roure, N.R. Jennings, N.R. Shadbolt. *The Semantic Grid: Past, Present and Future. IEEE Proceedings 93:3, March 2005.*
- [12] A. K. Dey and G. D. Abowd. *Towards a better understanding of context and context-awareness. In Proc. CHI 2000 Workshop on the What, Who, Where, When, and How of Context-Awareness, The Hague, The Netherlands, 2000.*
- [13] A. Doan, J. Madhavan, R. Dhamankar, P. Domingos A. Halevy. *Learning to match ontologies on the Semantic Web. The VLDB Journal. 12(4): 303-319, 2003.*
- [14] A. Harter, A. Hopper, P. Stegges, A. Ward, and P. Webster. *The anatomy of a context-aware application. Wireless Networks, 8(2/3):187-197, 2002.*
- [15] M. Lenzerini. *Data integration: a theoretical perspective. In Proceedings of the twenty-first ACM SIGMODSIGACT-SIGART symposium on Principles of database systems, pages 233-246. ACM Press, 2002.*
- [16] W. N. Schilit, N. Adams, and R. Want. *Context-aware computing applications. In IEEE Workshop on Mobile Computing Systems and Applications, pages 85-90, Santa Cruz, CA, US, 1994. IEEE Computer Society Press.*
- [17] A. Schmidt, M. Beigl, and H.-W. Gellersen. *There is more to context than location. Computers & Graphics, 23(6):893-901, 1999.*
- [18] C. Stephanidis. *Adaptive techniques for universal access. User Modeling and User Adapted Interaction, 11(1-2):159-179, 2001.*
- [19] *E-health european conference. [http://europa.eu.int/information\\_society/europe/ehealth/conference/2003/index\\_en.htm](http://europa.eu.int/information_society/europe/ehealth/conference/2003/index_en.htm)*

## 2.5 Descrizione del programma e dei compiti dell'Unità di Ricerca

### Testo italiano

Lo scopo della ricerca svolta da PoliMI è lo sviluppo di metodologie avanzate, metodi e tecniche per i) gestione dell'informazione con caratteristiche di mobilità, per supportare utenti in movimento che necessitano di un accesso coordinato a servizi e risorse dati condivise; ii) progettazione di basi di dati dipendenti dal contesto, semistrutturate e di piccole dimensioni da memorizzare su supporti ridotti, che vanno dai computer palmari, ai cellulari, alle smartcard o targhette RFID; iii) sincronizzazione basata su aspetti semantici tra diverse sorgenti di informazione e supporti di dimensioni ridotte.

Lo studio metodologico produrrà inoltre un applicativo di supporto, mentre i metodi e le tecniche per la gestione di basi di dati di piccole dimensioni verranno applicati in un prototipo dimostrativo, coordinato con l'architettura di riferimento P2P middleware definita nel progetto.

Come scenario applicativo considereremo personale medico sparso per il mondo che necessita di condividere informazioni

mediche (ad esempio, siti di ospedali, basi di dati cliniche, rapporti di ricerca, come anche dispositivi mobili del personale medico) al fine di migliorare il processo decisionale e di ridurre il tasso di errori medici. Un dottore, potenzialmente equipaggiato di un dispositivo portatile come un PDA o un telefono cellulare, può essere interessato a inserire i dati relativi allo stato del paziente e ottenere tutte le informazioni in proposito, come i sintomi della malattia e i possibili trattamenti. Questo richiede l'integrazione di sorgenti disparate, spesso dal web, che non sono note a priori.

Lo scenario prevede pertanto sistemi aperti, collegati in rete, e peer-based, secondo il paradigma del semantic grid, dove non si ipotizza una conoscenza a priori è alcuna relazione tra le parti, le quali possono essere costituite da dispositivi mobili o fissi. L'interazione tra tali dispositivi è transiente, poiché è condizionata dalla disponibilità di reti e supporti: infatti, la natura dei dispositivi considerati e la diversità dell'ambiente influiscono pesantemente sulla loro interazione e di conseguenza sui metodi e le tecniche di cooperazione.

Verranno approfondite tecniche avanzate di caching per affrontare i problemi di disponibilità di rete sopra citati, permettendo ai singoli partecipanti di mantenere memorizzata la porzione di dati necessaria, mentre gli altri frammenti di informazione, memorizzati in nodi diversi, verranno interrogati solo in presenza di una connessione disponibile.

Il contesto mobile e dinamico all'interno del quale cooperano i dispositivi determina la porzione di dati locale, che, a causa della quantità limitata di memoria disponibile, deve essere continuamente aggiornata, in base a criteri semantici e dipendenti dal contesto. Inoltre, devono essere considerate tecniche di accesso ai dati attente al consumo di potenza per gestire il problema del tempo di vita limitato delle batterie dei dispositivi.

Pertanto nell'applicazione precedentemente descritta il dispositivo del medico deve ospitare differenti porzioni di conoscenza, che dipendono dal tempo e dal contesto specifico, ed è scopo del sistema selezionare queste porzioni dinamicamente.

Lo scenario risultante pone le basi per lo studio di una metodologia per lo sviluppo di basi di dati di dimensioni molto piccole, dipendenti dal contesto e sensibili al consumo di potenza, considerando anche algoritmi e tecniche appropriate relative all'architettura di riferimento P2P middleware definita nel progetto.

#### *Fase 1*

La fase 1 è composta da cinque Attività principali, descritte in dettaglio nel seguito; ognuna di esse sarà preceduta da uno studio approfondito della letteratura esistente.

##### *Attività 1.1*

In questa Attività PoliMI svilupperà la metodologia che guida la selezione della porzione di dati che deve essere mantenuta nel dispositivo mobile; diverse dimensioni di analisi [3] per determinare il contesto forniscono le diverse prospettive rispetto alle quali il dispositivo mobile può essere visto e che sono usate per stabilire il suo ambiente. Le dimensioni identificate e i relativi valori determinano la scelta delle informazioni da memorizzare sul micro-dispositivo, che verranno in realtà selezionate a tempo di esecuzione. Allo scopo di formalizzare e successivamente ottenere in modo automatico questa vista (l'ambiente del dispositivo), le dimensioni verranno modellate in una struttura gerarchica, basata su un grafo diretto aciclico, che permette di considerare una specifica ontologica dei concetti considerati e di modellare inoltre vincoli semantici tra le diverse dimensioni.

##### *Risultati dell'Attività 1.1*

*D4.1: Rapporto tecnico che descrive uno studio comparativo e la prima versione della metodologia sopra citata.*

##### *Attività 1.2*

Questa Attività verrà svolta da PoliMI congiuntamente alle altre unità operative. L'attività consisterà nel raccogliere i requisiti per lo scenario applicativo di ESTEEM. La raccolta dei requisiti sarà svolta attraverso interviste e questionari. Il risultato sarà quindi utilizzato per verificare sulla carta i prodotti delle altre attività.

##### *Risultati dell'Attività 1.2*

*DALL.1: (sviluppato in coordinamento con tutte le unità operative): Rapporto tecnico che definisce i requisiti dello scenario applicativo.*

##### *Attività 1.3*

Verranno sviluppati i metodi e le tecniche per la selezione automatica citati nell'Attività 1.1, inseriti in un ambiente per lo più ignoto; il grafo orientato aciclico delle dimensioni contribuirà alla selezione automatica e all'interpretazione delle risorse condivise da importare sul dispositivo. Si consideri l'esempio già illustrato di comunità semantica costituita allo scopo di permettere collaborazioni scientifiche nel contesto medico; in questo caso i nodi mettono a disposizione diverse strutture e strumenti relativi alle risorse disponibili. Oltre a selezionare ed integrare in modo semanticamente coerente le risorse più appropriate fornite dai diversi nodi, lo scopo specifico delle tecniche considerate sarà quello di soddisfare i vincoli imposti dal contesto del dispositivo. Ad esempio, durante una visita a domicilio, un medico che necessita di informazioni riguardanti i sintomi di una malattia rara può ricercare, utilizzando un palmare o un telefono cellulare, basi di dati mediche e rapporti di ricerca via web per ottenere un supporto alla diagnosi. Ovviamente non dovrà essere disorientato dai diversi formati dei documenti ottenuti, dalla discrepanza lessicale dei relativi contenuti e dalle informazioni inutili per l'ambiente nel quale sta operando (sintomi da malattie tropicali in un paziente eschimese), invece, l'informazione più significativa dovrà essere presentata nella forma più adeguata per il contesto e il dispositivo in uso. L'obiettivo dei metodi e delle tecnologie sviluppati in questa Attività è di conseguenza la ricerca e la possibile

memorizzazione delle informazioni necessarie da tenere sul dispositivo portatile, adeguando le stesse al contesto locale.

#### *Risultati dell'Attività 1.3*

*D4.2: Rapporto tecnico che comprende: i) uno studio comparato per; ii) una specifica per; iii) i metodi e le tecniche proposti per l'individuazione automatica e l'estrazione delle risorse informative da importare nei micro-dispositivi. Particolare attenzione verrà posta nel mantenere la coerenza con il modello semantico e con l'architettura proposti nel progetto.*

#### *Attività 1.4*

*Lo scopo di questa Attività è lo sviluppo di un protocollo di sincronizzazione. Quando i concetti di interesse vengono scoperti e selezionati dalle sorgenti di dati, può risultare utile raccogliere sul dispositivo molti altri concetti che sono in relazione con questi. Tuttavia, poiché i dispositivi portatili hanno memoria limitata, strategie di risoluzione dinamica dei conflitti devono essere concepiti, basandosi su varie nozioni di vicinanza, eventualmente semantica, al fine di decidere quali dati debbano essere tenuti e quali scartati.*

#### *Risultati dell'Attività 1.4*

*D4.3: Rapporto tecnico contenente: uno studio comparativo e una descrizione formale del protocollo di sincronizzazione.*

#### *Attività 1.5*

*Questa attività si occuperà di definire la specifica di alto livello dell'architettura ESTEEM, in collaborazione con le altre unità operative. L'architettura offrirà un serie di servizi per un accesso integrato e "fidato" ai dati distribuiti tra i membri di una comunità semantica, e per la formazione e gestione di comunità semantiche basate sul consenso. Tutti questi servizi risulteranno accessibili attraverso un strato di software multicanale e adattativo, abilitando la localizzazione e la selezione di risorse informative di interesse che dovranno essere accessibili dai micro-dispositivi.*

#### *Risultati dell'Attività 1.5*

*DALL.2 (sviluppato in coordinamento con tutte le unità operative): Rapporto tecnico per la definizione della architettura ESTEEM ad alto livello, e per la specifica WSDL di tutti i servizi offerti. UNIROMA sarà il coordinatore dei prodotti e sarà responsabile della specifica del livello di cooperazione P2P e dei servizi di integrazione dei dati.*

### *Fase 2*

*Lo scopo principale delle attività della seconda fase sarà lo sviluppo dei prototipi e la loro integrazione nell'architettura di riferimento.*

#### *Attività 2.1*

*In questa attività la metodologia per la selezione della porzione dei dati da far risiedere sul dispositivo verrà completata e formalizzata e lo strumento di supporto alla metodologia sarà sviluppato in accordo con i risultati dell'Attività 1.1.*

#### *Risultati dell'Attività 2.1*

*D4.4: Rapporto definitivo sulla metodologia.  
D4.5: Prototipo dello strumento di supporto alla metodologia*

#### *Attività 2.2*

*Questa attività produrrà il modulo di supporto della sincronizzazione dell'architettura del progetto, che in accordo con i risultati dell'Attività 1.3, integrerà la selezione automatica delle risorse sulla base del contesto all'interno del livello di cooperazione P2P, mediante l'implementazione del protocollo di sincronizzazione presentato nell'Attività 1.4. Dal lato dispositivo, il modulo di sincronizzazione sarà integrato nel prototipo del DBMS per dispositivi mobili presentato in [4].*

#### *Risultati dell'Attività 2.2*

*D4.6: Implementazione prototipale del modulo di sincronizzazione.*

#### *Attività 2.3*

*Il raggiungimento degli scopi del progetto richiede la definizione di una architettura omnicomprensiva per ESTEEM, che prenda in considerazione tutti gli aspetti precedentemente evidenziati. PoliMI sarà partecipe nella progettazione dei servizi di localizzazione e selezione dei dati, del modulo di sincronizzazione semantica, e della loro integrazione nell'architettura globale.*

*Al fine di supportare la progettazione dell'architettura ESTEEM e far emergere tutti i problemi relativi all'interazione dei servizi sviluppati dalle varie unità operative, l'attività di progettazione includerà lo sviluppo di un prototipo mock-up per i servizi sviluppati*

dal PoliMI, come descritto nell'Attività 2.2.

### Risultati dell'Attività 2.3

DALL.3: (sviluppato in coordinamento con tutte le unità operative) Rapporto tecnico sull'architettura.

DALL.4: (sviluppato in coordinamento con tutte le unità operative) Prototipo mock-up e dati collezionati durante la fase di testing.

### Attività 2.4

PoliMI svolgerà questa Attività congiuntamente alle altre unità operative. L'Attività consisterà nell'implementazione e integrazione del dimostratore, per illustrare l'uso e la fattibilità della piattaforma ESTEEM su un caso reale. L'implementazione del dimostratore di ESTEEM sarà basata sui servizi dell'architettura ESTEEM. PoliMI sarà responsabile dello sviluppo dei moduli del dimostratore relativi ai prototipi menzionati precedentemente.

### Risultati dell'Attività 2.4

DALL.5: (sviluppato in coordinamento con tutte le unità operative) Rapporto tecnico che illustra il dimostratore integrato, così come i risultati della sperimentazione nel contesto dello scenario applicativo.

### Testo inglese

Goal of the research activity of PoliMI unit is the development of advanced methodologies, methods and techniques for: i) ubiquitous information management, for supporting mobile users requiring coordinated access to shared service and data resources; ii) context-aware design of very small, semistructured databases stored on small devices, ranging from palm computers to cell phones, smartcards or RFID tags; iii) semantic based synchronisation between the different information sources and the small devices.

The methodological study will also produce a design support tool, while the methods and techniques for very small database management will be included in a prototype demonstrator, coordinated with the reference P2P middleware architecture defined in the project.

An application scenario is considered where medical personnel around the world needs to share medical information (e.g. hospital sites, clinical databases, research reports, as well as other scientists' mobile devices), in order to enhance decision-making processes and decrease medical errors. A doctor, possibly equipped with a portable device such as a PDA or a cell telephone, would like to input a patient's status and obtain all relevant information about it, e.g. disease symptoms and possible treatments. This requires the integration of disparate sources, often from the web, which are not known in advance.

Thus we consider open, networked, peer-based systems, according to paradigms like the semantic grid, where there is no previous knowledge and relationship among the parties, that may be mobile as well as fixed devices. The interaction among such devices is transient, since it is subject to network and device availability: indeed the nature of these devices and the variety of ambient strongly affect the cooperation methods and techniques.

Advanced caching techniques are exploited to cope with the above mentioned network availability problems, always allowing the single party to retain the appropriate portion of needed data, while other fragments, stored at different peers, can be queried only when a connection is available.

The mobile and dynamic context where the devices cooperate determines the fraction of data located on board, which, due to the limited amount of memory available, must be continuously refreshed according to semantic context-based criteria. On the other hand, power aware data access techniques must be employed to manage the problem of limited battery life.

Thus, in the above described application, the doctor's device must host different chunks of knowledge, depending on the specific time and context, and it is the system's goal to dynamically select it.

The resulting scenario quests for a methodology for context- and power- aware design of very small databases, together with a number of appropriate algorithms and techniques embedded in the reference P2P middleware architecture defined in the project.

### Phase 1

Phase 1 envisages five main activities, detailed in the sequel; each of them will be preceded by a thorough review of the existing literature.

#### Activity 1.1

In this activity, PoliMI will develop the methodology driving the selection of the device-residing portion of data; analysis dimensions [3] for the detection of the context provide the different perspectives the mobile device is viewed from, and are used to set out its ambient. The identified dimensions and their current values drive the choice of the information to be kept on the microdevice, to be actually selected at run time. In order to formalize and then automatically obtain this view (the device ambient), we model the dimensions as a hierarchical, DAG-shaped, structure which allows us to consider an ontological specification of each considered concept, and to model semantic constraints between different dimensions as well.

#### Results of Activity 1.1

D4.1: Technical report illustrating a comparative study and the first version of the above mentioned methodology.

#### Activity 1.2

PoliMI will perform this activity jointly with other partners. The activity will consist in the collection of detailed requirements for the

ESTEEM application scenario. The collection of requirements will be done by interviews and questionnaires and will be used for testing the outputs of the other activities on paper.

#### Results of Activity 1.2

DALL.1: (jointly developed by all partners): Technical Report illustrating all the requirements collected for the application scenario.

#### Activity 1.3

The methods and techniques for the dynamic selection mentioned in Activity 1.1 in a largely unknown environment will be developed, where the dimension DAG contributes to the automatic selection and interpretation of the shared resources to be imported to the device. Consider the cited example of a semantic community formed to enable scientific collaboration in the medicine context; here, different resource structures and meanings are provided by the community peers. Besides selecting and semantically integrating the most appropriate resources provided by the various peers, it will be the special goal of such techniques to obey the constraints imposed by the device context.

For instance, during a home visit, a doctor in need of information about the symptoms of a rare disease can search, through her PDA or cell phone, clinical databases and research reports on the web looking for assistance in her diagnosis. Obviously she must not be disoriented by the different format of the retrieved documents, by the possible lexical discrepancies in their contents and by information which would be useless in the environment she operates (tropical diseases symptoms in an Eskimo patient), but the most valuable information must be presented in the most suitable form for the operational context and the available device. Thus it is the task of the methods and techniques developed in this activity to search for, to adapt to the local context and semantics, and possibly to store on the portable device the required information.

#### Results of Activity 1.3

D4.2: Technical report containing: i) a comparative study of; ii) a formal framework for; iii) the proposed methods and techniques for the automatic location and extraction of the information resources to be imported to the microdevices. Particular attention will be posed in keeping the defined frameworks and algorithms coherent with the project semantic model and architecture.

#### Activity 1.4

The development of a synchronization protocol is the aim of this activity. When interesting concepts are found and selected from the available data sources, it is also useful to collect on the device many other concepts which are possibly related to them. However, since small devices have a limited memory, dynamic conflict resolution strategies must be devised, possibly based on various notions of semantic nearness, to decide which data must be retained and which can be discarded.

#### Results of Activity 1.4

D4.3: Technical report containing a comparative study and a formal description of the synchronization protocol.

#### Activity 1.5

This activity will concern the high-level specification of the ESTEEM architecture, to be defined in cooperation with the other partners of the project. The architecture will provide a set of services for integrated and trust-aware access to data distributed over peers of a semantic community, a set of services for the discovery and matching of web services within the community, and a set of services for consensus-driven formation and management of semantic communities. Such services will be accessible through a multi-channel and adaptive software layer, and allow the location and selection of interesting information resources to be made available on the microdevices.

#### Results of Activity 1.5

DALL.2 (jointly developed by all partners): Technical report defining the high-level architecture of ESTEEM, and the WSDL specification of all the provided services.

UNIROMA coordinates the deliverable and is responsible for the specification of the P2P cooperation layer and of the data integration services.

#### Phase 2

The main goal of the second phase activities will be the development of the prototypes and their integration in the reference architecture.

#### Activity 2.1

This activity will complete and formalize the methodology for the selection of the device-residing portion of data, and produce the methodology support tool, developed also according to the results of Activity 1.1

#### Results of Activity 2.1

D4.4 Final report on the methodology.

D4.5: Prototype of the methodology support tool.

#### Activity 2.2

*This activity will produce the synchronization support module for the project architecture, which, according to the results of Activity 1.3, integrates the context driven automatic selection of resources within the P2P Cooperative Layer, implementing the synchronization protocol of Activity 1.4. On the device side, the synchronization module will be integrated into the prototype [4] of portable light DBMS.*

#### *Results of Activity 2.2*

*D4.6: Prototype implementation of the synchronization support module.*

#### Activity 2.3

*The accomplishment of the project goals requires designing a comprehensive architecture for ESTEEM, which will take into account all the aspects of the problem previously highlighted. PoliMI will be involved in the design and integration of the services for data location and selection, as well as the semantic synchronization modules within the complete architecture.*

*In order to support the design of the ESTEEM architecture, and to bring to the surface all possible problems arising from the interaction between the services to be developed by the various partners, the design activity will involve the development of a mock-up for the services to be developed by PoliMI, as described in Activity 2.2.*

#### *Results of Activity 2.3*

*DALL.3: (jointly developed by all partners) Technical report defining the detailed architecture of ESTEEM.*

*DALL.4: (jointly developed by all partners) Mock-up prototypes and data collected during the testing phase.*

#### Activity 2.4

*PoliMI will perform this activity jointly with other partners. The activity will consist in the implementation of an integrated demonstrator, showing the usage and feasibility of the ESTEEM platform in a real case.*

*The implementation of the ESTEEM demonstrator will be based on the services of the ESTEEM architecture. PoliMI will be in charge of developing the modules of the demonstrator involving the above prototypes.*

#### *Results of Activity 2.4*

*DALL.5: (jointly developed by all partners): Technical report illustrating the integrated demonstrator, as well as the results of the experimentation in the context of the application scenario.*

---

## **2.6 Descrizione delle attrezzature già disponibili ed utilizzabili per la ricerca proposta con valore patrimoniale superiore a 25.000 Euro**

### **Testo italiano**

*Nessuna*

### **Testo inglese**

*Nessuna*

## **2.7 Descrizione delle Grandi attrezzature da acquisire (GA)**

### **Testo italiano**

*Nessuna*

### **Testo inglese**

*Nessuna*

## 2.8 Mesi uomo complessivi dedicati al programma

### Testo italiano

		Numero	Mesi uomo 1° anno	Mesi uomo 2° anno	Totale mesi uomo
<i>Personale universitario dell'Università sede dell'Unità di Ricerca</i>		4	19	19	38
<i>Personale universitario di altre Università</i>		0	0	0	0
<i>Titolari di assegni di ricerca</i>		0			
<i>Titolari di borse</i>	<i>Dottorato</i>	0			
	<i>Post-dottorato</i>	0			
	<i>Scuola di Specializzazione</i>	0			
<i>Personale a contratto</i>	<i>Assegnisti</i>	0			
	<i>Borsisti</i>	1	11	11	22
	<i>Dottorandi</i>	0			
	<i>Altre tipologie</i>	1	0	5	5
<i>Personale extrauniversitario</i>		0			
<b>TOTALE</b>		<b>6</b>	<b>30</b>	<b>35</b>	<b>65</b>

### Testo inglese

		Numero	Mesi uomo 1° anno	Mesi uomo 2° anno	Totale mesi uomo
<i>University Personnel</i>		4	19	19	38
<i>Other University Personnel</i>		0	0	0	0
<i>Work contract (research grants, free lance contracts)</i>		0			
<i>PHD Fellows &amp; PHD Students</i>	<i>PHD Students</i>	0			
	<i>Post-Doctoral Fellows</i>	0			
	<i>Specialization School</i>	0			
<i>Personnel to be hired</i>	<i>Work contract (research grants, free lance contracts)</i>	0			
	<i>PHD Fellows &amp; PHD Students</i>	1	11	11	22
	<i>PHD Students</i>	0			
	<i>Other tipologies</i>	1	0	5	5
<i>No cost Non University Personnel</i>		0			
<b>TOTALE</b>		<b>6</b>	<b>30</b>	<b>35</b>	<b>65</b>

### PARTE III

#### 3.1 Costo complessivo del Programma dell'Unità di Ricerca

##### Testo italiano

Voce di spesa	Spesa in Euro	Descrizione
Materiale inventariabile	5.000	Dispositivi mobili (portatili, palmari, telefoni cellulari, etc)
Grandi Attrezzature		
Materiale di consumo e funzionamento	6.500	cancelleria, accessori per computer non inventariabili, gestione amministrativa del fondo
Spese per calcolo ed elaborazione dati		
Personale a contratto	49.000	un giovane ricercatore (borsa biennale), un progettista software (secondo anno)
Servizi esterni		
Missioni	11.000	cooperazione con altre unità di ricerca, viaggi per partecipazioni a convegni, discussioni con altri ricercatori a livello internazionale
Pubblicazioni		
Partecipazione / Organizzazione convegni	4.500	spese di partecipazione a conferenze
Altro		
<b>TOTALE</b>	<b>76.000</b>	

##### Testo inglese

Voce di spesa	Spesa in Euro	Descrizione
Materiale inventariabile	5.000	mobile devices (portable computers, palm computers, cell phones, etc.)
Grandi Attrezzature		
Materiale di consumo e funzionamento	6.500	stationery, software, computer support consumable equipment, administration costs
Spese per calcolo ed elaborazione dati		
Personale a contratto	49.000	one young researcher (two years' grant), one software designer for the second year
Servizi esterni		
Missioni	11.000	cooperation with other research units, travels for conference attendances, discussions with other international researchers
Pubblicazioni		
Partecipazione / Organizzazione convegni	4.500	conference fee expenses
Altro		
<b>TOTALE</b>	<b>76.000</b>	

#### 3.2 Costo complessivo del Programma di Ricerca

		Descrizione
Costo complessivo del Programma dell'Unità di Ricerca	76.000	
Fondi disponibili (RD + RA) comprensivi dell'8% max per spese di gestione	22.800	Fondo ricerca di Ateneo Anni 2002,2003,2004 (Prof.ssa Tanca - RD), contratto di ricerca ENEL (prof. Ceri - RD), fondi messi a disposizione dall'Ateneo per l'anno

		2005 in caso di accettazione della proposta di progetto (RA)
<b>Cofinanziamento di altre amministrazioni</b>		
<b>Cofinanziamento richiesto al MIUR</b>	53.200	

**3.3.1 Certifico la dichiarata disponibilità e l'utilizzabilità dei fondi di Ateneo (RD e RA)**

SI

*(per la copia da depositare presso l'Ateneo e per l'assenso alla diffusione via Internet delle informazioni riguardanti i programmi finanziati e la loro elaborazione necessaria alle valutazioni; D. Lgs, 196 del 30.6.2003 sulla "Tutela dei dati personali")*

Firma \_\_\_\_\_

Data 05/04/2005 ore 20:31