



SMART SERVICES

Citizen-centered Smart Services in Big Data Framework

Responsabili di Ricerca:

- Valeria De Antonellis (DII) - coordinatore
- Maurizio Carpita (DEM), Giorgio Pedrazzi (DIGI), Angela Tincani (DSCS)
- Maria Grazia Speranza (DEM)
- Alfonso Gerevini (DII)

Il contesto: Smart Cities OPEN & BIG DATA



Migliorare e sviluppare nuovi servizi ai cittadini

Obiettivi di Ricerca

OR1. Organizzazione, accessibilità e fruizione di dati e servizi digitali (Prof. Valeria De Antonellis - DII).

OR2. Integrazione di basi dati per la valutazione dei servizi socio-assistenziali e sanitari (Prof. Maurizio Carpita - DEM, in collaborazione con Prof. Angela Tincani - DSCS e Prof. Giorgio Pedrazzi - DIGI).

OR3. Pianificazione e monitoraggio del trasporto passeggeri e uso di nuove fonti di dati (Prof. Maria Grazia Speranza - DEM).

OR4. Estrazione di informazioni e conoscenza medica da referti testuali (Prof. Alfonso Gerevini - DII) Estrazione ed analisi automatica di informazioni diagnostiche da referti testuali per supporto all'attività clinica e gestionale.

Co-Finanziamenti e Collaborazioni

- **Regione Lombardia**, Progetto ODE (Open Data Exploration) – Prof. Devis Bianchini, DII
- **Associazione Bresciana Artrite Reumatoide (ABAR)**
- Centro di Ricerca ICT della **Fondazione Bruno Kessler (FBK)** di Trento

In particolare:

- Ufficio di Statistica del **Comune di Brescia**
- U.O. Radiologia Diagnostica 2 degli **Spedali Civili** – Direttore Prof. Maroldi

OR1. Organizzazione, accessibilità e fruizione di dati e servizi - DII

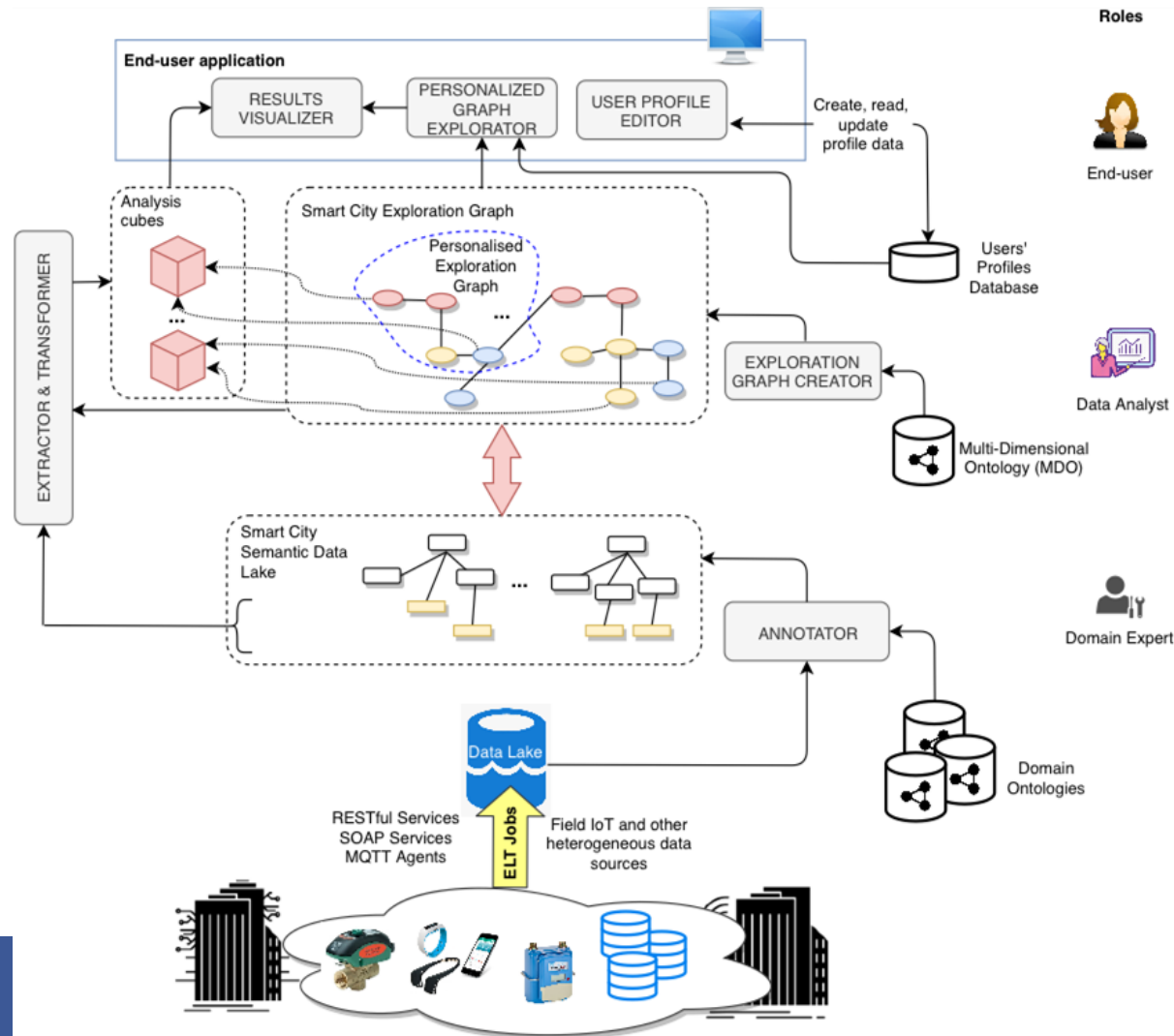
Modelli, metodi e strumenti ICT per:

- **Descrizione dei dati eterogenei orientata all'utente**, facilmente comprensibile.
- **Accesso ed esplorazione guidata dei dati**, per diverse categorie di utenti.
- **E-participation** attraverso il Web e tramite l'utilizzo di moderne tecnologie di rete e dispositivi "smart", per coinvolgere cittadini in iniziative sociali.

Problematiche

- **Gestire dati eterogenei** – I diversi tipi di dati (strutturati, semi-strutturati, non strutturati) devono essere facilmente interpretabili (**modello semantico**).
- **Consentire esplorazione efficace dei dati** – I dati raccolti devono essere opportunamente organizzati per l'esplorazione (**modello multidimensionale dei dati**).
- **Esporre i dati in rete sotto forma di servizi** (**data services**, accessibili attraverso opportune **Web API** - Application Programming Interface).
- **Considerare esigenze di diversi attori** – Cittadini, Fornitori di servizi, Amministrazione Pubblica...

Un approccio basato su Semantica - 3-Livelli



Personalised Views/Exploration graphs are generated for different categories of end-users

Multi-Dimensional Representation in terms of Semantic Models Concepts

Semantic Models are created for heterogeneous data sources within a Data Lake

OR2. Integrazione di basi dati per la valutazione dei servizi socio-assistenziali e sanitari - DEM

Progettazione, sviluppo e applicazione di **metodi e modelli statistici per Intelligent Big Data Analysis e Data Mining** per:

- analizzare big data provenienti da basi dati di vario formato, qualità, provenienza, frequenza e dimensione.
- sintetizzare, rappresentare e comunicare informazioni per **supportare decisioni relative ad organizzazione, accessibilità e fruizione dei servizi.**

Analisi delle vulnerabilità delle famiglie bresciane

- **Integrazione e gestione di basi di dati di natura amministrativa**, per l'analisi delle vulnerabilità delle famiglie bresciane.
- **Costruzione di indicatori di povertà e vulnerabilità** utilizzando dati dal 2008 al 2015 presenti nell'Anagrafe del Comune e nel sistema delle dichiarazioni annuali dei redditi del Ministero delle Finanze.
- **Sviluppo di una prima sperimentazione** utilizzando base dati Istat. Effettuate alcune elaborazioni per **ottenere la distribuzione delle famiglie povere e vulnerabili residenti a Brescia**.

Analisi del grado di affollamento in zone della città

- **Analisi di big data digitali geo-referenziati ad alta frequenza temporale per il monitoraggio dell'utilizzo della città** per valutare il grado di affollamento in diverse zone della città e in diversi momenti (ore del giorno, giorni della settimana, mesi dell'anno). Fornita dal Servizio di Statistica del Comune di Brescia una base dati dell'operatore di telefonia mobile TIM contenente informazioni sulle presenze di telefoni cellulari nella Provincia di Brescia da aprile 2014 ad agosto 2016.
- **Analisi effettuata con metodologie statistiche avanzate per la riduzione della dimensionalità di immagini da mappe geografiche (HOG: Histogram of Oriented Gradients) e il raggruppamento di profili spazio-temporali (FDA: Model-Based Clustering Functional Data Analysis).** Utilizzata per **stimare l'impatto dei grandi eventi culturali sul territorio**, e.g., The Floating Piers di Christo's sul lago d'Iseo a giugno del 2016.

DSCS ed Associazione Bresciana Artrite Reumatoide (ABAR)

- Raccolta tramite questionario di informazioni riguardo all'utilizzo e alla valutazione dei servizi socio-sanitari e assistenziali da parte di malati reumatici cronici residenti nel Comune di Brescia.
- Risposte da 374 malati, di cui circa l'80% donne prevalentemente over 65.
- Emerso un quadro di «moderata» soddisfazione per i servizi socio-sanitari.
- Rilevato **profilo del malato reumatico "fragile" nel Comune di Brescia**: donna over 65, con un basso livello di istruzione e con una lunga durata di malattia.

DIGI

- Analisi di obblighi e adempimenti per i titolari del **trattamento di dati personali** per la fornitura di servizi alla persona e le categorie di dati personali.
- Studio dei diversi elementi alla base del modello di **cittadinanza digitale** secondo cui il cittadino può fruire dei servizi della pubblica amministrazione:
 - lo SPID (Sistema Pubblico di Identità Digitale), il domicilio digitale, le firme elettroniche, la disciplina dei pagamenti digitali, l'attivazione della Carta d'identità elettronica, l'istituzione del difensore civico digitale unico e il raccordo dell'Anagrafe della popolazione residente.

OR3. Pianificazione e monitoraggio del trasporto passeggeri e uso di nuove fonti di dati - DEM

- Sistema per la **gestione integrata dell'informazione relativa alla mobilità pubblica** dei cittadini (orari dei sistemi di trasporto pubblico, numero di biciclette disponibili in ogni postazione bike sharing, numero di posti disponibili nei parcheggi, ...).
- Definizione di possibili indicatori e sviluppo di cruscotti per:
 - rilevazione della performance della **sostenibilità** del sistema per la mobilità dei passeggeri
 - erogazione di un **servizio di trasporto passeggeri personalizzato** in grado di sfruttare al meglio l'integrazione dei diversi sistemi di trasporto
 - **valutazione economica e ambientale e impatto sociale** di diversi gradi di intervento

Modelli per la distribuzione del traffico

- **Diversi modelli di ottimizzazione** per la distribuzione del traffico su una rete stradale. Dimostrato sperimentalmente che distribuendo le auto con la stessa origine e la stessa destinazione su diversi percorsi di poco più lunghi di quello più corto (o più veloce) è possibile in molti casi **eliminare gli ingorghi e sempre ridurli notevolmente**.
- **Modello a variabili miste intere**, con generazione dei percorsi che rispettano il limite fissato a priori sulla qualità dei percorsi accettabili per l'utente. Sperimentazioni su dati simulati e reali hanno dimostrato la capacità di **trovare distribuzioni del traffico che eliminano la congestione e sono vicine all'ottimo di sistema** con limitato disagio per gli utenti.

Servizio di trasporto passeggeri personalizzato

- Sistema a chiamata con mezzi di capacità limitata, da 4 a 8 posti, (**shared taxi**)
- **Simulazione di una rete cittadina**, con matrice origine-destinazione, in cui ogni richiesta di trasporto può essere soddisfatta da uno dei modi previsti, bus o metro, mezzo proprio, shared taxi.
- Risultati computazionali mostrano la **superiorità del sistema a chiamata che potrebbe attrarre la grande maggioranza delle richieste di trasporto, perché più rapido e meno costoso**. Le ricadute comprendono la **riduzione del numero dei mezzi propri su strada**, in particolare in città, con benefici economici e sociali.

OR4. Estrazione di conoscenza e informazioni mediche da referti di esami diagnostici - DII

- Tecniche innovative di elaborazione del linguaggio naturale, **data/text mining e machine learning** per la **classificazione automatica di referti di esami radiologici** attualmente disponibili in forma testuale.
- Strumenti innovativi per la **creazione di database di informazioni diagnostiche strutturate**, che il medico potrà consultare molto più facilmente degli attuali database con referti in formato testuale (ad esempio per valutare l'utilità di certi tipologie di esami, in generale, o per determinate fasce della popolazione).

Annotazione di referti ed estrazione di conoscenza

- **Strumento informatico** che consente al medico di fare annotazioni al testo attraverso un'interfaccia utente appositamente progettata.
- **Tecniche di NLP e machine learning** che, a partire da un data set di testi annotati, consentono di identificare automaticamente le frasi del testo che sono rilevanti per le informazioni che si intendono estrarre.

Con U.O. Radiologia Diagnostica 2 degli Spedali Civili

- Schema gerarchico di **categorie di informazioni inerenti la TAC al torace**, che costituisce la base con cui classificare i referti.
- Attraverso lo strumento di annotazione manuale dei referti è stato creato un primo **data base di testi annotati**, da usarsi come training set per la generazione di modelli di classificazione automatica dei referti.
- Creati alcuni primi **modelli di classificazione** di cui si è avviata una sperimentazione basata sui referti esaminati.

Principali Applicazioni ed Evoluzione dei risultati conseguiti

- **Modellazione ed Esplorazione Big Data**
 - MIUR: Progetto Cluster Fabbrica Intelligente - Progetto [Smart Manufacturing 2020](#);
 - Regione Lombardia: Progetto [Smart4CPPS](#) (Smart Factory), Progetto [FISVAL](#) (Digital Twin), Progetto [Mo.So.Re](#) (Mobilità Sostenibile e Resiliente)
- **Utilizzo Big Data Digitali Geo-referenziati**
 - Regione Lombardia: Progetto [Mo.So.Re](#)
- **Mobilità Sostenibile**
 - Regione Lombardia: Progetto [Mo.So.Re](#)
- **Annotazione e Classificazione Automatica Testi**
 - DII- DSMC-Spedali Civili: Progetto [AI4COVID19](#)

