

I Università IUAV di Venezia
U Scuola di Dottorato
A Dottorato di ricerca in Nuove Tecnologie & Informazione Territorio e Ambiente
V

**Dottorato in Nuove Tecnologie e Informazione
Territorio e Ambiente**

Vincenzo Giannotti

Resoconto dell'attività del secondo anno di dottorato

**ecologia del web
recommender systems e semantic web**

Dottorando: Vincenzo Giannotti
gennaio 2013

Indice generale

Premessa.....	3
L'argomento della ricerca.....	4
Attività di studio e formazione.....	6
Partecipazione a seminari e conferenze Iuav.....	6
Incontri, conferenze e convegni extra Iuav.....	9
Attività di ricerca.....	10
Il progetto Adriamuse.....	11
Il progetto Recommending to people.....	13
Pubblicazioni relative alle attività del 2012.....	17

Premessa

Nel 2010 un alto funzionario del Ministero delle Comunicazioni finlandese Olli-Pekka Rantala, affermava che “la banda larga dovrebbe essere considerata un bene primario”

Nel 2011 “[...] una ricerca commissionata da Cisco alla sua seconda edizione, il Cisco Connected World Technology Report 2011, ha messo di evidenza proprio la tendenza tra le giovani generazioni di studenti e lavoratori che la Rete rappresenta un bene fondamentale nella loro vita, a cui difficilmente rinuncerebbero, preferendola a beni che fino a qualche anno fa probabilmente non sarebbero stati barattati per una connessione a Internet. Lo studio di Cisco ha coinvolto studenti universitari e giovani lavoratori tra i 20 e i 30 anni [.....] emerge dalla ricerca che il 33% degli intervistati considera Internet una risorsa fondamentale per l'uomo tanto quanto aria, acqua, riparo. Circa metà di loro (49% degli studenti e 47% dei lavoratori) ritiene che la sua importanza si avvicini molto a quella di questi beni essenziali. Complessivamente, quattro persone su cinque ritengono che Internet abbia una importanza vitale e faccia parte dei “mezzi di sostentamento” necessari nel quotidiano.”

Nel 2012, secondo la sentenza di un giudice di Trieste, “quando il provider lascia i clienti senza collegamento, quest'ultimi hanno diritto al risarcimento per danni patrimoniali ed esistenziali causati dallo stress del non avere la connessione.”

In questi giorni (gennaio 2013) un giudice tedesco ha affermato che “Siamo in molti a considerare Internet una parte essenziale della nostra vita, un bene primario di cui - che ci piaccia o meno - non siamo in grado di fare a meno. Lo usiamo per lavoro, per svago, per informazione e per studio e rimanere improvvisamente senza connessione è qualcosa che non si augura nemmeno al proprio peggior nemico.”

Internet è dunque un bene primario o almeno viene considerato tale soprattutto dai più giovani, per i quali è un mezzo insostituibile per socializzare, informarsi e rimanere in contatto col mondo.

C'è da chiedersi se in un contesto come l'attuale, l'espressione "navigare nel web" abbia ancora un senso. Navigare vuol dire muoversi sopra qualcosa e noi oggi il web non lo navighiamo: lo respiriamo. Siamo immersi in una atmosfera

modificata dove, oltre ad azoto ed ossigeno, respiriamo informazioni di ogni tipo: buone e cattive. Come respiriamo aria buona e aria cattiva. Un po' cominciamo ad accorgercene. È evidente che la nostra civiltà della rete comincia ad evolvere e a maturare. Come nelle civiltà giovani, in via di sviluppo, non si va per il sottile e non si bada troppo all'inquinamento se non dopo un po', quando l'euforia dello sviluppo comincia a scemare e si ha il tempo per prendersi una pausa, dare un'occhiata al cielo e fare un respiro profondo, allo stesso modo noi, ora, abbiamo la sensazione che qualcosa non va, vediamo che il cielo non è più così limpido e che l'aria ha uno strano sapore e forse dobbiamo cominciare a ripulirla.

Ecco, questa premessa, forse un po' lunga, mi serve per dire di cosa mi voglio occupare in questa mia ricerca: voglio vedere cosa si sta facendo e cosa si può fare per ripulire un po' il web.

Chiamiamola Ecologia del Web.

Ecologia, però, non dal punto di vista energetico, ma da quello della qualità dell'informazione.

L'argomento della ricerca

Il primo anno di attività di dottorato l'ho speso principalmente a definire le condizioni per effettuare una scelta meditata del tema di ricerca sul quale orientare i miei interessi. Non che avessi dei dubbi su cosa mi interessasse veramente fare, ma le numerose attività di ricerca che mi hanno visto coinvolto in questi ultimi anni e il tema specifico della smart city, posto al centro delle attività di studio del dottorato in nuove tecnologie, hanno imposto una riflessione attenta.

Nel rapporto conclusivo del I anno, avevo ipotizzato di potermi occupare di “...tecniche e di tecnologie rivolte a “mettere ordine” nella enorme mole di dati e di informazioni da cui veniamo sommersi oggi. ... un tema che mi ha appassionato in questi ultimi anni e che ho avuto modo di approfondire è quello del Semantic Web.”

Sullo specifico tema dello studio e della implementazione del semantic web ho avuto modo di realizzare delle interessanti esperienze, a partire dal 2008-2009, con il progetto SISC (Sistema Informativo Semantico per la Cultura) realizzato per conto della Regione Veneto e successivamente col progetto Adriamuse

(Cross-Border Cooperation Programme IPA Adriatic 2007-2013) tuttora in corso.

Oltre al tema del semantic web, sempre nel rapporto del I anno, affermavo che “Un altro tema al quale ho cominciato ultimamente ad interessarmi, ... , riguarda lo sviluppo e l'utilizzazione dei Recommender System.” e anche in questo caso ho avuto modo di fare una interessante esperienza.

A questo punto, ricollegandomi a quando citato in premessa, credo di poter affermare con sufficiente certezza che l'argomento della ricerca riguarderà la “ecologia del web”, ovvero lo studio di tecniche e tecnologie in grado di aiutarci ad utilizzare il web in maniera efficiente, evitando di sprecare energia, tempo, risorse. Tutto ciò a partire dai sistemi di raccomandazione, e dall'uso della semantica sul web.

Attività di studio e formazione

Partecipazione a seminari e conferenze luav

Ho partecipato ad una buona parte dei seminari organizzati nell'ambito del dottorato, sia a quelli a mio giudizio più attinenti a quello che ritengo possa essere il mio percorso di ricerca, sia a quelli che ho ritenuto maggiormente interessanti per il mio aggiornamento professionale.

Tra le iniziative a cui ho partecipato, fisicamente o connettendomi via web, cito:

[17/12/2012](#)

Seminario: La piattaforma City Knowledge

ore:11.00 - 13.00 - Fabio Carrera - Worcester Polytechnic Institute (WPI, Massachusetts), - VEGA - edificio Pegaso I piano - sede Unisky

[13/12/2012](#)

Seminario: "L'utilizzo di tecnologia RFID come strumento di indagine delle reti sociali all'interno delle aziende"

ore:10.30 - 13.00 - Mariachiara Barzotto - VEGA - edificio Pegaso I piano - sede Unisky

[15/11/2012](#)

Conferenza: Esry Italia per le Università. Principali prodotti e novità: City Engine

ore 14:30 – 17:00 – Relatore Pietro Coffaro

Presso Iuav – Cotonificio, Venezia.

[12/09/2012](#)

Conferenza: Design e riduzione alla fonte dei rifiuti.

ore:11.00 - 13.00 – Relatore: prof. Medardo Chiapponi

presso Iuav – Ca' Tron, Venezia.

[27/09/2012](#)

Conferenza: Filiere corte, territorio e agroalimentare

ore:11.00 - 13.00 - Relatore: Giorgio Piazza (Presidente di Coldiretti veneto).

Presso l'Auditorium della Provincia di Venezia in via Forte Marghera n.191 - Mestre

[19/07/2012](#)

Conferenza: Software open source e open data per la PA

ore:11.00 - 13.00 - Relatore: Luca Menini (Direttore della Scuola ASA di Arpav). Presso l'Auditorium della Provincia di Venezia in via Forte Marghera n.191 – Mestre

[19/07/2012](#)

Seminario: Geospatial Web - ingegnerizzazione

ore:14.00 - 18.00 - Relatore: Stefano Menegon - (Dottorato di ricerca in Nuove Tecnologie IUAV). Presso l'Auditorium della Provincia di Venezia in via Forte Marghera n.191 – Mestre

[12/07/2012](#)

Seminario: Le tecnologie ICT per dispositivi mobili: progettare LBS sul proprio cellulare

ore:14.00 - 18.00 - Relatore: Andrea Prati (Università IUAV di Venezia). Presso l'Auditorium della Provincia di Venezia in via Forte Marghera n.191 – Mestre

[12/07/2012](#)

Conferenza: Servizi localizzati: definizione, disegno, integrazione, interoperabilità

ore:11.00 - 13.00 - Relatore: Robert Laurini (INSA Lyon). Presso l'Auditorium della Provincia di Venezia in via Forte Marghera n.191 - Mestre

[05/07/2012](#)

Seminario: Nuove piattaforme geoweb: il programma geoSDI

ore:14.00 - 18.00 - Relatore: Dimitri Dello Buono (CNR-IMAA Protezione Civile Nazionale). Presso l'Auditorium della Provincia di Venezia in via Forte Marghera n.191 – Mestre

[05/07/2012](#)

Conferenza: Società, Politica e Piano

ore:11.00 - 13.00 - Relatore: Domenico Patassini (Università IUAV di Venezia), Ugo Baldini (CAIRE-Unisky). Presso l'Auditorium della Provincia di Venezia in via Forte Marghera n.191 – Mestre

[13/04/2012](#)

Seminario: Workshop il manifesto per la città intelligente

ore:9.30 - 11.30 - Parco scientifico VEGA

[12/04/2012](#)

Seminario: Geo-FreeDOM: impatti di Free, Digital, Open, Mobile sul business geomatico, tendenze evolutive e prospettive future. Parte 1: il concetto di Free.

ore:14.30 - 16.30 - Paolo Dosso - Tolentini sede Unisky

[05/04/2012](#)

Seminario: Smart cities dissertazione delle tesi di dottorato e lezione magistrale di Luca De Biase

ore:14.00 - 18.00 - Parco scientifico tecnologico VEGA - Marghera (VE)

[26/01/2012](#)

Seminario: Elaborazioni di immagini e video per l'estrazione della conoscenza- Andrea Prati

ore:14.00 - 16.00 - Sede Tolentini -Unisky

[23/01/2012](#)

Seminario: Wikicrazia: le frontiere delle politiche pubbliche collaborative al tempo della crisi

ore:15.00 - 17.00 - Alberto Cottica - Palazzo Badoer - aula Tafuri

Incontri, conferenze e convegni extra Iuav

Ho partecipato anche ad alcune iniziative extra Iuav che ho ritenuto utili sia per il mio aggiornamento professionale sia, in maniera più specifica, per il dottorato. Tra le più significative cito:

- Partecipazione ad un ciclo di seminari sui servizi mobile e wireless e sul marketing digitale, tenuti allo SMAU di Padova nell'aprile 2012.
- Partecipazione al Workshop "EGNOS and Galileo for maritime and marine applications" organizzato da NEREUS (Network of European Regions Using Space Technologies) in collaborazione con la Regione Veneto e il consorzio di ricerca CORILA. Venezia 16-17 April 2012.
- Partecipazione al convegno "Small Business Act: il futuro dell'Europa e dell'Italia attraverso le PMI", nel quale sono stati presentati gli obiettivi specifici della Commissione Europea per lo sviluppo delle piccole e medie imprese. Convegno organizzato da Apindustria Venezia - Venezia 11 Maggio 2012
- Partecipazione ai meeting Adriamuse con presentazione dei risultati della ricerca. Di particolare interesse il meeting tenutosi a Scutari (Albania) a metà ottobre 2012 nel corso del quale è stata presentata una prima versione della piattaforma di comunicazione *Adriamuse.com* e il successivo meeting di Bologna (gennaio 2013) nel quale la piattaforma è stata presentata in versione definitiva.

Attività di ricerca

Nel periodo 2011-2012 ho realizzato alcune attività che ritengo possano essere considerate propedeutiche e in parte ascrivibili al tema di ricerca di cui ho accennato in premessa e nel primo capitolo.

Una prima attività riguarda la partecipazione ad Adriamuse: un progetto transfrontaliero del Cross-Border Cooperation Programme IPA Adriatic 2007-2013, cofinanziato dalla Unione Europea, che ha come obiettivo lo sviluppo del turismo culturale attraverso la promozione di una rete museale europea.

In questo progetto Iuav svolge un ruolo di partner tecnologico e si occupa, in particolare, della realizzazione del work package 5 “enlarging the network: IT tools for promoting museum and events”.

Gli obiettivi principali di questo WP sono due:

- facilitare l'accesso e la consultazione dell'informazione relativa agli eventi culturali organizzati dai musei dell'area adriatica;
- disseminare questa informazione nel web, utilizzando tecnologie allo stato dell'arte, incluse quelle di tipo semantico e collaborativo.

Questo progetto mi ha consentito di realizzare alcuni interessanti risultati che si stanno concretizzando sia in prodotti sia in pubblicazioni e che ritengo siano perfettamente attinenti al tema di ricerca.

Una seconda attività che ho realizzato ha riguardato invece il mondo dei Recommender Systems.

I Recommender Systems sono strumenti in grado di “raccolgere e selezionare” l'informazione del web, con lo scopo di presentarla ad un utilizzatore o a una categoria di utilizzatori in una forma personalizzata in base a alle loro personali aspettative e preferenze.

È accaduto dunque che nella seconda metà del 2011 sono venuto a conoscenza, tramite il prof. Luigi Di Prinzio, del fatto che il MIUR aveva riproposto un sostanzioso finanziamento (in base al DM593 art.11) teso a favorire progetti di start-up e di spin-off universitari su attività di ricerca industriale e di sviluppo precompetitivo. All'inizio del 2012 ho avuto modo di valutare questa opportunità e di scoprire che il finanziamento poteva essere concesso anche a progetti presentati da dottorandi di ricerca.

Mettendo quindi a fattore comune tutte queste cose ho pensato di proporre, assieme ad alcuni amici e colleghi, tra cui il Prof. Andrea Prati, un progetto di

spin-off proprio sul tema dei sistemi di raccomandazione per il turismo culturale.

La compilazione del progetto ha richiesto un notevole impegno e circa tre mesi di tempo ed è stato presentato al MIUR lo scorso giugno col titolo “Recommending to people”.

Purtroppo il progetto non ha ottenuto il finanziamento e tuttora siamo in attesa del verbale con le motivazioni della mancata accettazione.

In teoria, alla luce delle motivazioni, il progetto potrebbe anche essere corretto e ripresentato ma pare che questa linea di finanziamento sia stata soppressa per cui, eventualmente, si dovranno attendere chiarimenti in merito.

Resta una esperienza di grande valore (almeno per me) e un buon lavoro che forse, anche in altri contesti, può essere recuperato.

Nel due capitoli seguenti ho voluto fare degli approfondimenti sui due progetti che ho citato. In particolare del progetto “Recommending to people” ho allegato un estratto delle prime pagine, sia per documentare, seppure parzialmente, l'attività che ho svolto, sia per rendere evidente il fatto che, a meno che non si vogliano considerare i sistemi di raccomandazione meramente da un punto di vista ingegneristico e tecnico, l'affinità e la complementarità tra questi e i sistemi semantici è notevole, soprattutto per quanto riguarda l'orientamento generale che, dal mio punto di vista, è proprio quello della ecologia del web.

Il progetto Adriamuse

AdriaMuse è un progetto transfrontaliero del Cross-Border Cooperation Programme IPA Adriatic 2007-2013, cofinanziato dalla Unione Europea. Il progetto comprende 11 partner di 5 paesi su entrambi i lati del Mare Adriatico: Provincia di Rimini, capofila, Istituto per i Beni Artistici, Culturali e Naturali (IBACN) della Regione Emilia-Romagna, Regione Veneto, Università IUAV di Venezia, Provincia di Pesaro e Urbino, Provincia di Campobasso, Skupa s.r.l., Museo Nazionale di Centro Montenegro (Montenegro), Comune di Scutari (Albania), Servizio di Governo del Cantone Zenica-Doboj (Bosnia-Erzegovina) e Regione Istriana (Croazia).

Obiettivo strategico del progetto AdriaMuse è quello di rafforzare le relazioni tra le organizzazioni partner per favorire uno sviluppo sostenibile dell'area adriatica. Da un punto di vista generale esso viene perseguito attraverso l'armonizzazione delle attività dei partner nel settore del turismo culturale,

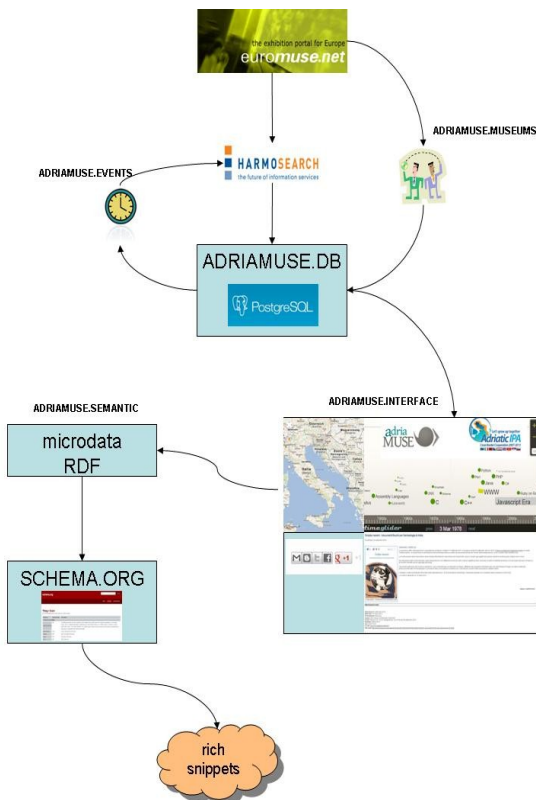
cercando di sviluppare l'accessibilità ai musei che si trovano nella regione adriatica. Più nel dettaglio:

- stimolando i musei dall'area adriatica ad entrare a far parte della rete Euromuse.net, vetrina internazionale che raccoglie le informazioni sui musei e sugli eventi più importanti organizzati a livello europeo e che consente di raggiungere una vasta gamma di potenziali visitatori;
- organizzando delle attività museali “beyond their walls”, con lo scopo di proporre l'offerta museale anche a quei potenziali visitatori che non sono frequentatori abituali dei musei.

Entrambi questi obiettivi di progetto prevedono che la rete sia rafforzata e integrata attraverso lo sviluppo di un insieme di strumenti informativi e di comunicazione e di servizi innovativi, realizzando una sorta di tavolo comune per la condivisione delle conoscenze, delle esperienze e delle idee tra le regioni partner.

E' proprio in questo specifico ambito che va ad inserirsi la parte più significativa delle attività che Iuav svolge nell'ambito del progetto.

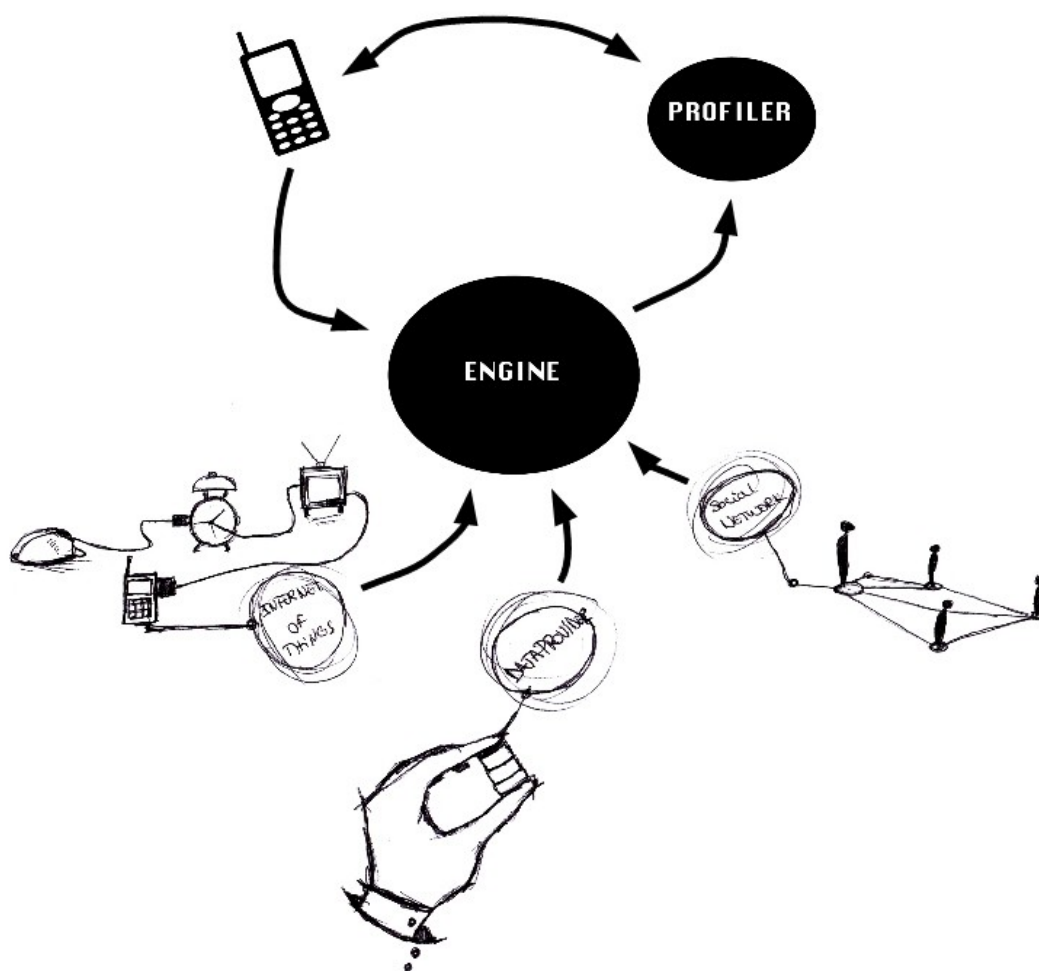
In particolare, è compito di Iuav facilitare l'accesso e la consultazione dei dati relativi agli eventi promossi dai Musei dell'area adriatica che entrano a far parte della rete Euromuse.net. Inoltre deve consentire la diffusione di tali dati nel web, utilizzando strumenti e tecnologie allo stato dell'arte, anche di tipo collaborativo.



Il progetto Recommending to people

Obiettivo generale del progetto è quello di realizzare un ecosistema digitale nel quale possano interagire diverse componenti (uomini, macchine, sensori, servizi...) con lo scopo di contestualizzare il turista culturale (ma più in generale qualunque turista) rispetto all'ambiente e allo scenario culturale in cui si trova e fornirgli suggerimenti e informazioni anche in base alle sue normali aspettative.

Il processo che sta alla base del servizio proposto, può essere così schematizzato:



Nel modello proposto i dati di contesto e le informazioni provenienti dalla collective intelligence¹ e dal web sono raccolti, omogeneizzati ed organizzati

¹ Definita come l'intelligenza che emerge dalla collaborazione, dalla competizione e dallo scambio di informazioni di individui all'interno di un gruppo, per esempio all'interno di un social network

utilizzando un *Engine* attraverso il quale essi vengono caratterizzati anche dal punto di vista semantico. L'output del Engine alimenta un modulo *Profiler* che ne caratterizza i contenuti in funzione del profilo e dei comportamenti dell'utilizzatore che ha richiesto il servizio. I suggerimenti e le raccomandazioni sono quindi inviate in maniera asincrona all'utilizzatore stesso che ne può usufruire attraverso una interfaccia multimodale ed eventualmente attraverso diversi dispositivi. Una caratteristica importante del servizio proposto è che le raccomandazioni saranno inviate nella maniera meno invasiva possibile, cercando di non modificare le abitudini e gli usi tipici del dispositivo da parte dell'utente.

Il dominio di applicazione del modello appena descritto rispetto alla ricerca proposta, riguarda il turismo culturale, limitatamente al contesto della città storica di Venezia.

In tale contesto i numeri dei potenziali utilizzatori del servizio sono, in prima battuta, quelli legati ai flussi turistici medi (si veda il rapporto “Turismo sostenibile a Venezia”, redatto nel 2011 dall'Ufficio di Piano del Magistrato alle Acque, sui flussi turistici a Venezia: nel 2008 12 ML di escursionisti e 7 ML di pernottamenti nella Venezia insulare) e più in particolare a coloro che si appoggiano ai servizi recentemente attivati tramite la piattaforma di connettività e servizi wireless “Venice Connected” per la prenotazione e la vendita di servizi pubblici della città (oltre 190.000 ticket commercializzati nel 2011 a turisti provenienti da 188 paesi). Il sistema dovrà essere quindi opportunamente dimensionato per fare fronte al potenziale traffico generato dai numeri appena citati. In ogni caso esso sarà studiato affinché presenti caratteristiche di scalabilità tali da non comprometterne la funzionalità nel caso il numero degli utenti dovesse aumentare in maniera sostanziale rispetto a quanto inizialmente ipotizzato.

Le principali problematiche di Ricerca e Sviluppo che dovranno essere affrontate, possono essere elencate in maniera sommaria come segue:

problematiche di natura tecnico-scientifica

- precisione della localizzazione all'interno di una struttura urbana complessa, che risulta particolarmente sfidante e scientificamente rilevante nel contesto urbano di Venezia, con strette strade (calli) ed alti

edifici, dove i segnali (sia satellitari che wireless) faticano a propagarsi correttamente

- acquisizione ed organizzazione di dati disomogenei strutturati/non strutturati provenienti da fonti diverse (*data fusion*)
- armonizzazione di dati strutturati (p.e. dati provenienti da sistemi diversi (trip advisor & google place)) ma in grado di fornire informazioni sovrapponibili (per esempio informazioni sugli hotel)
- *harvesting*² e analisi di dati non strutturati, raccolti dalla collective intelligence (in particolare commenti e opinioni dai Social Network)
- studio di uno strumento per la “cattura” dei comportamenti e delle azioni dell'utente, su base geografica e temporale
- caratterizzazione semantica³ dei dati di contesto, studio di una ontologia⁴ di dominio e impiego di thesauri e vocabolari

problematiche di natura tecnologica

- impiego di architetture di tipo NoSQL per la memorizzazione di grandi moli di dati secondo modelli di scalabilità orizzontale
- impiego di tecniche statistiche di *data summarization*⁵ e database temporali probabilistici per gestire la enorme quantità di dati ridondanti generata
- studio e simulazione delle performance del sistema, con lo scopo di garantire l'erogazione di un servizio in grado di soddisfare i requisiti richiesti
- applicabilità di interfacce multimodali⁶ per una migliore e più completa interazione tra uomo e macchina; tra le interfacce di sicuro interesse si studieranno anche metodi per non-vedenti che possano indirizzarli verso luoghi di interesse mediante comunicazioni sonore/audio

Il progetto, pur presentando caratteristiche generalmente applicabili, sarà realizzato in prima battuta nella città di Venezia, con lo scopo di verificare i

2 raccolta di dati utilizzando procedure software

3 processo di attribuzione di significato ad una risorsa, tramite l'associazione della stessa ad un'ontologia

4 in ambito informatico una ontologia definisce, all'interno di uno specifico tema di interesse (dominio), i concetti utilizzabili e le regole che legano tali concetti. Lo scopo è quello di istruire un computer riguardo al contesto in cui i dati e i concetti vengono utilizzati.

5 il processo di data summarization riguarda l'elaborazione di dati primitivi o derivati con lo scopo di ottenere ulteriori dati derivati

6 sono definite multimodali le interfacce che coinvolgono più di un input di comunicazione, per esempio tastiera e voce

risultati della ricerca e di attivare il servizio a partire dalla scala locale. Questa scelta è motivata dal fatto che Venezia può rappresentare un banco di prova eccellente per l'impianto di un sistema come quello proposto sia per la sua spiccata vocazione turistica, sia per la disponibilità di un patrimonio culturale vastissimo e unico al mondo, sia per la presenza di una infrastruttura tecnologica di prim'ordine (rete in fibra, wi-fi pubblico, servizio Venice Connected) che consente di ipotizzare una effettiva buona riuscita del progetto e la fornitura di un servizio ad elevato valore aggiunto per il turista. Questa tesi è avvalorata dal fatto che questa iniziativa viene sostenuta da Venis spa, società per l'informatica e le telecomunicazioni del Comune di Venezia, che gestisce l'infrastruttura tecnologica della città.

Pubblicazioni relative alle attività del 2012

Attualmente sono in fase di revisione della versione inglese di un *paper* sulla WP5 di Adriamuse dal titolo “Adriamuse.com, a new communication platform” in cui descrivo con discreto dettaglio le attività che ho svolto in questa fase del progetto, assieme al mio collega Ivano Boscolo.

Sto preparando in questi giorni un nuovo articolo sul web semantico come strumento strategico per la cultura, che presenterò in occasione di un convegno previsto a Pesaro per la fine del mese e organizzato dalla provincia di Pesaro-Urbino.