



OR3 - Semantic Web in ambiente Enterprise

RI3.1 - STUDIO E ANALISI DI METODI E TECNICHE PER LA RAPPRESENTAZIONE SEMANTICA DELLE INFORMAZIONI

La frammentazione delle informazioni presenti, prodotte, gestite e necessarie ai vari sistemi e processi di un'azienda è una delle problematiche più spinose da gestire nelle fasi di adoption e maintenance di approcci organizzativi e tecnologici di tipo Enterprise 2.0. In questi contesti, infatti, la forte connotazione sociale tipica del modello 2.0, favorisce e incoraggia attori interni ed esterni a produrre informazioni attraverso strumenti di cooperazione e condivisione quali wiki, blog, social network, document management systems, eccetera. Tali strumenti, generalmente, confluiscono in silos informativi indipendenti che, tipicamente, ostacolano l'interoperabilità incrementando i costi aziendali di fruizione del capitale intellettuale derivanti, in particolare, dalla necessità di scoprire, normalizzare, integrare ed aggregare l'informazione rilevante.

In questo documento si analizzano metodologie e tecnologie che, basate su approcci semantici (linguaggi, modelli e meta-modelli ontologici), possono essere utilizzati per la rappresentazione e la gestione della conoscenza disponibile in ambito Enterprise. Tale conoscenza, opportunamente organizzata e strutturata, può essere sfruttata efficacemente per far emergere il valore proprio del capitale intellettuale circolante in un'Organizzazione, per quanto distribuito tra gli svariati silos informativi disponibili e ma troppo spesso non interoperabili su cui è frammentato.

Si vogliono riportare, quindi, i risultati dell'attività di analisi condotta sugli strumenti di rappresentazione dell'informazione allo stato dell'arte, nonché la proposta di metodologie e tecniche semantic-driven ritenute più adeguate alla rappresentazione formale (possibilmente machine-readable) di contenuti informativi prodotti e accessibili ai differenti servizi/applicazioni, al fine di migliorare l'interoperabilità, l'integrazione e la condivisione della conoscenza nell'ambito della Extended Adaptive Enterprise Architecture (EAEA).

Nell'ambito dell'indagine condotta, tra i possibili approcci metodologici, si è fatto particolare riferimento al pattern "Semantic Web Grounding (SWG)", come elemento abilitante la definizione di un layer di descrizione semantica e integrazione delle informazioni e dei contenuti prodotti e gestiti dai diversi servizi/applicazioni che coesistono nell'ecosistema IT aziendale, con particolare riferimento ai sistemi di Enterprise Content Management. Nello specifico, il pattern SWG incoraggia l'uso di schemi ontologici e, conseguentemente, l'adozione di linguaggi standard (de jure o de facto) definiti nell'ambito dei gruppi di lavoro W3C sul Semantic Web. Per questo motivo, quindi, a valle della descrizione del dominio applicativo e della metodologia selezionata e



ragionando per domini applicativi, l'analisi si focalizza principalmente su quegli approcci patrocinati e standardizzati dal W3C stesso, di cui sono alcuni esempi SCOT, SIOC, Dublin Core, FOAF, OWL-S.

Trattati tali tecnologie, si procede all'analisi di metodologie e tecniche di integrazione e correlazione nonché di), orientate a supportare la definizione di uno strato di collective knowledge condivisibile tra gli svariati sistemi aziendali e in grado di supportare attività di knowledge discovery and retrieval. In questa classe rientrano approcci di merging, trasformazione, traduzione e mediazione, nonché approcci di reasoning e formalismi standardizzati per l'integrazione quali SKOS e OWL (con l'ultimo nato OWL2), oltre ai linguaggi di interrogazione della famiglia SPARQL.

Infine, per meglio caratterizzare e sintetizzare il panorama di riferimento connesso allo stack concettuale del Semantic Web, ed in grado di elaborare l'informazione sfruttando la potenza espressiva e l'interpretazione connessa ai modelli matematici di riferimento soggiacenti, si analizzano le tecnologie e i framework più promettenti, tentando un'analisi comparativa tra questi ed ottenendo di conseguenza utili indicazioni per la razionalizzazione del panorama degli approcci ammissibili e, ove possibile praticamente, utilizzabili.

Attraverso l'attività condotta si ritiene di essere pervenuti all'identificazione dell'insieme di modelli, metodologie e tecnologie di rappresentazione e gestione semantica della conoscenza allo stato dell'arte tecnico-scientifico che trovano adeguata applicazione al dominio d'indagine del Progetto MODERN, in particolare, a supporto della definizione di un modello architetturale che favorisca la cooperazione tra applicazioni eterogenee attraverso schemi semantici condivisi, oltre che di un modello basato su strumenti semantici in grado di supportare un'efficace rappresentazione e organizzazione della conoscenza.

RI3.2 - DEFINIZIONE DI UN MODELLO SEMANTICO PER LA COOPERAZIONE APPLICATIVA

Il presente RI ha avuto come obiettivo la descrizione dei modelli semantici e dei pattern che rendono possibile la cooperazione applicativa tra i sistemi legacy già presenti nelle diverse organizzazioni.

La cooperazione applicativa tra i sistemi ereditati costituisce un punto focale dell'architettura MODERN, in quanto consente l'utilizzo combinato di servizi applicativi legacy in un modello unificato, preservando la naturale indipendenza dei sistemi legacy e consentendo l'erogazione di nuovi servizi e funzionalità.

La natura semantica che costituisce il pilastro su cui regge l'architettura MODERN rende naturale una estensione semantica del modello per la cooperazione applicativa, contribuendo quindi alla definizione del Semantic Middleware layer dell'architettura EAEA e mantenendo i principi cardine di pariteticità ed indipendenza dei sistemi legacy.



Di per sé la cooperazione applicativa deve necessariamente affrontare i problemi di non cooperazione: l'eterogeneità dei sistemi applicativi, la necessità di un formato di interscambio dei dati e la necessità di modificare gli applicativi dovuta alla mutabilità del significato dei dati sono tutti esempi di problematiche concrete che è necessario risolvere per poter ottenere un modello unico di cooperazione.

Durante il corso dell'attività sono stati discussi i sistemi che allo stato dell'arte affrontano i problemi relativi alla cooperazione tra applicazioni, mostrando pregi e limiti di ogni approccio descritto e descrivendo successivamente una visione dello stato dell'arte per ciò che riguarda l'approccio generico ai processi di cooperazione applicativa. All'analisi di tali sistemi allo stato dell'arte è seguita una descrizione dei principali pattern per la cooperazione applicativa. In particolare sono stati discussi i pattern:

- Message, per lo scambio di messaggi;
- Message Endpoint, per lo scambio di messaggi attraverso un canale di messaging;
- Splitter, per la suddivisione del contenuto di un messaggio nelle diverse parti di cui è composto;
- Aggregator, per l'aggregazione di dati provenienti da diverse elaborazioni in un unico messaggio;
- Envelope Wrapper, per adattare i dati applicativi all'interno di un pacchetto in modo che sia compatibile con l'infrastruttura di messaggistica che lo deve trasportare;
- Content Enricher, per l'utilizzo delle informazioni del messaggio in input al fine di recuperare informazioni aggiuntive da una sorgente dati esterna;
- Pattern Content Filter, per il filtraggio dei dati irrilevanti;
- Mashup, per l'aggregazione e trasformazione di dati estratti da diverse sorgenti al fine di utilizzarli in contesti diversi;
- Content-Based Router, per gestire correttamente l'instradamento basandosi sulle informazioni contenute all'interno del messaggio;
- Message Translator, per la traduzione di un messaggio da un formato a un altro;
- Mediator, per la coordinazione tra diverse componenti e la negoziazione del formato dei messaggi scambiati;
- Channel Adapter, per consentire alle applicazioni di riuscire a pubblicare e ricevere messaggi senza bisogno di intervenire sulla logica applicativa;
- Publish-Subscribe channel, per la pubblicazione di messaggi e la loro divulgazione agli utenti che hanno effettuato la sottoscrizione al canale;
- Pattern Event-Driven Consumer, per il processing dei messaggi non appena questi sono disponibili;
- Remote Procedure Invocation, per l'invocazione di funzioni all'interno di applicazioni differenti, passando i dati da condividere e invocando la funzione indicata sfruttando il paradigma di incapsulamento dei dati;
- Shared Data, per l'accesso condiviso ai dati da parte di diverse applicazioni.



Per tutti questi pattern è stata compilata una matrice che indica quali pattern coprono ognuno dei requisiti della cooperazione applicativa.

Dall'analisi dello stato dell'arte prende spunto il modello di cooperazione applicativa secondo MODERN, che fonde i principi della cooperazione con i concetti relativi alla semanticizzazione delle informazioni, definendo il metodo di integrazione dei sistemi eterogenei per l'erogazione di nuovi servizi, descrivendo i concetti di cooperazione a livello funzionale e a livello di dati e mostrando i principali pattern che possono essere combinati per ottenere in modo pragmatico la cooperazione applicativa.

In ultima istanza sono mostrati alcuni esempi per quanto riguarda i possibili scenari di cooperazione, con la descrizione dei relativi pattern di cooperazione utilizzati. In questa fase è stata effettuata una descrizione degli scenari di cooperazione (e.g. "Semantic Mashup", "Business Process e Social Network") che è possibile costruire sulla base del modello di cooperazione applicativa illustrato nella fase precedente. Per ciascun esempio di scenario sono stati descritti i benefici apportati dalle tecnologie semantiche in termini di integrazione, riuso, interoperabilità e serendipity.

Il ruolo assunto dalle applicazioni durante gli scambi cooperativi può appartenere alle seguenti tipologie:

- partecipazione attiva – l'applicazione partecipa al processo di cooperazione traendo beneficio dagli scambi cooperativi, ovvero fruisce dei servizi erogati da altre applicazioni oppure recupera nuovi dati e informazioni;
- partecipazione passiva – l'applicazione partecipa al processo di cooperazione esponendo i suoi servizi e i suoi dati ad altre applicazioni.

RI3.3 - DEFINIZIONE DI UN MODELLO PER LA RAPPRESENTAZIONE DELLA CONOSCENZA

Il presente RI si pone l'obiettivo di presentare il modello in grado di gestire la conoscenza collettiva e condivisa all'interno di un'organizzazione, definendo il dominio applicativo, il modello di rappresentazione della conoscenza ed il deployment.

Il modello di rappresentazione della conoscenza fornisce gli strumenti per modellare la conoscenza presente nell'azienda. L'approccio metodologico utilizzato prende spunto ed estende la metodologia NeOn, alla cui base troviamo l'approccio *divide et impera* che considera come main problem la costruzione del modello ontologico. Secondo NeOn le ontologie sono collegate le une alle altre, costruendo così una rete utile a NeOn stesso per la realizzazione dell'intero processo legato allo sviluppo e all'evoluzione di un'ontologia.

Per scegliere la metodologia e l'ontologia sono state fatte diverse considerazioni. Per quanto riguarda la prima infatti, deve essere una soluzione generica e non specifica, oltre ad essere indipendente dalla tecnologia e consentire un apprendimento veloce. Per quanto concerne l'ontologia, invece, la scelta è stata effettuata prendendo in considerazione diverse caratteristiche ovvero: consistenza, efficienza, ambiente, finitezza, flessibilità, prospettiva, trasparenza ed usabilità.



La gestione della conoscenza è fondamentale in quanto permette di incrementare l'efficienza dei processi delle organizzazioni dando modo al capitale intellettuale dell'azienda di crescere.

Sono diversi gli studi e i modelli definiti per la suddivisione della conoscenza; la suddivisione qui è quella fornita dal modello di Nonaka che prevede la divisione in conoscenza "tacita e implicita", risultato dell'apprendimento personale, e conoscenza "esplicita", resa disponibile dall'organizzazione.

Ovviamente la gestione della conoscenza esiste perché permette di raggiungere diversi obiettivi all'interno dell'organizzazione, come consentire agli utenti di trovare la giusta informazione così da raggiungere il proprio obiettivo, condividere la conoscenza attraverso l'intera organizzazione, integrare le informazioni tra loro correlate ma che sono memorizzate in punti differenti.

Con l'evoluzione del web si è passati da una prima definizione del World Wide Web la cui idea era di fornire un modo per consultare documenti statici, con il Web 1.0, per passare al Web 1.5 grazie all'introduzione del concetto di database, fino ad arrivare al Web 2.0 in cui l'utente può interagire in maniera diretta arricchendo il contenuto presente nel web stesso.

Espandendo il concetto di web con il termine semantico, otteniamo la definizione del WWW come un ambiente dove i documenti sono associati a dati e informazioni che ne specificano il contesto semantico adatto all'elaborazione automatica. Il termine enterprise si riferisce alle piattaforme che hanno l'obiettivo di creare una collaborazione ed un dialogo tra i dipendenti e i clienti, pertanto dall'unione di questi tre termini è stato coniato il concetto di Enterprise Semantic Web 2.0.

Affinché questo risulti possibile è necessario considerare i sei concetti descritti dall'acronimo SLATES e successivamente estesi con l'aggiunta di altri nell'acronimo FLATNESSES. Nel dettaglio abbiamo quindi Freeform, Links, Authorship, Tagging, Network-Oriented, Extension, Search, Social, Emergence e Signal.

Il modello di rappresentazione della conoscenza deve favorire la creazione e la gestione della conoscenza collettiva e condivisa; ciò consentirà di orientare Modern verso il paradigma dell'Enterprise 2.0 attraverso la visione FLATNESSES.

Il modello è stato progettato per rispondere alle esigenze reali che possono emergere in un contesto aziendale dinamico e fortemente evolutivo. In tal senso, è strutturato in due macro blocchi:

schemi ontologici formali - per modellare i key concept che caratterizzano il patrimonio conoscitivo dell'organizzazione (conoscenza che cambia raramente);

lightweight ontology - per modellare le istanze dei concetti chiave, ad esempio i topic relativi a un progetto (conoscenza che cambia frequentemente).



Seguendo un approccio basato sugli scenari, viene mostrato il corpo del modello e considerando uno degli scenari viene presentata un'istanziatura del modello.

Infine, sono presentate le soluzioni proposte a problematiche quali deployment, gestione e manutenzione del modello.

RI3.4 - REALIZZAZIONE DI UNO O PIÙ DIMOSTRATORI

Le attività condotte hanno portato alla realizzazione di un unico dimostratore in grado di validare i modelli individuati in RI3.2 e RI3.3. Lo sviluppo applicativo si avvale della metafora applicativa del Personal Environment introdotta in RI2.4 usufruendo, in questo modo, anche, dell'infrastruttura tecnologica predisposta per tale attività.

Il dimostratore è stato progettato ed implementato sulla base dello scenario applicativo individuato per la verifica, con il duplice obiettivo di dimostrare:

il valore aggiunto fornito dalle tecnologie semantiche per la rappresentazione e ed accesso alla conoscenza applicate in un ambiente Enterprise,

i vantaggi della cooperazione applicativa tra sistemi enterprise sia essi legacy che esterni all'azienda, attraverso strumenti che supportino l'interoperabilità semantica.

Il contesto individuato rientra nell'ambito del marketing pubblicitario, ed in particolare si focalizza sul processo relativo alla configurazione di una campagna di marketing associata al lancio di un nuovo prodotto commerciale, concentrandosi sulla fase di identificazione della tipologia degli utenti potenzialmente interessati al prodotto e verso i quali indirizzare l'intera campagna di commercializzazione.

Sfruttando le potenzialità modellate nel Personal Environment realizzato in OR2 ed i meccanismi di cooperazione applicativa definiti in OR3, lo scenario implementa un flusso operativo che, all'atto di creazione del nuovo pacchetto e/o offerta commerciale (App CRM), individua, in modo automatico, gli utenti potenzialmente interessati (interazione con lo strato di conoscenza semantica) per procedere a contattarli automaticamente attraverso canali classici del mondo enterprise (App e-mail) oppure via strumenti più orientati al web 2.0 (Twitter).

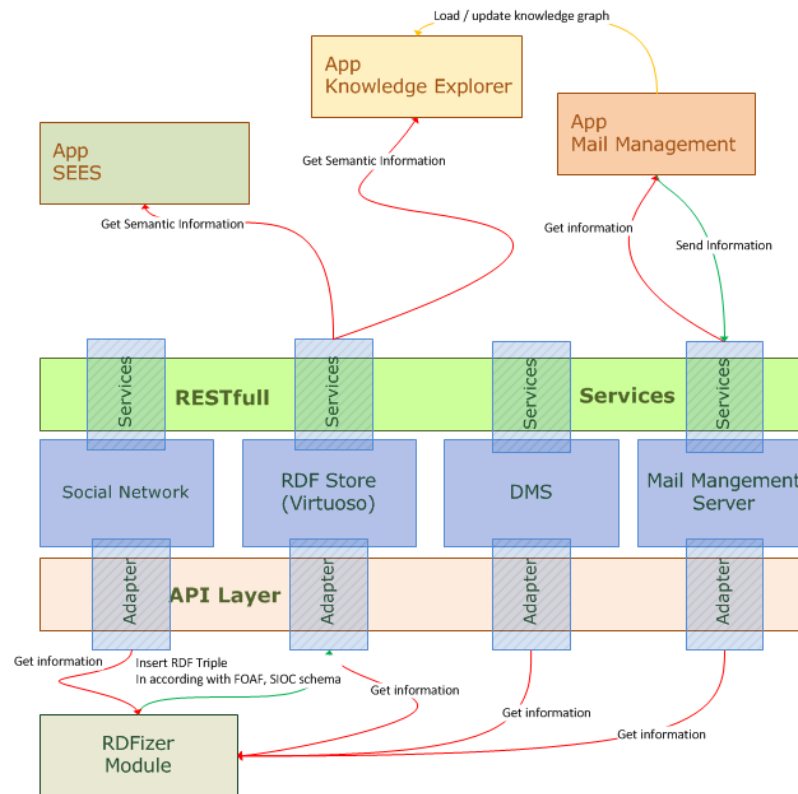


Figura 1 vista architetturale.

In particolare i moduli applicativi realizzati sono di seguito descritti:

- **App Knowledge Explorer (KE):** app che consente di visualizzare e navigare le istanze ontologiche individuate in OR3 (SIOC, FOAF, SCOT, Vocab) a partire da una determinata key (es.: tutti i documenti riconducibili/associabili ad una persona, tutti gli amici di una persona, quali persone hanno partecipato alla stesura di quel documento, etc.). Utilizzando la metafora di navigazione di un grafo attraverso nodi ed archi, l'utente ha la possibilità di partire da un oggetto di conoscenza, come una mail o un documento, e, percorrendo le relazioni definite dal modello di rappresentazione della conoscenza, può risalire a tutti gli altri concetti ed informazioni ad esso collegati (persone, processi, ecc). Ognuno di questi elementi, a sua volta, diventa il punto di partenza di ulteriori connessioni che l'utente può percorrere, risultando così agevolato nel processo di accesso alle informazioni.
- **App Semantic Enriched Enterprise Search (SEES):** app che supporta l'utente nella scrittura di query in linguaggio naturale che si mappano sulle relative query SPARQL applicate al modello semantico.
- **App Blog:** app che consente all'utente di creare un contenuto digitale (post di un Blog) e di arricchirlo semanticamente con tag recuperati direttamente dal modello semantico sottostante. Il contenuto così creato contribuisce ad alimentare la collective knowledge del sistema andando ad integrare il grafo di conoscenza del



modello e diventando di fatto ricercabile e navigabile attraverso le app SEES e Knowledge Explorer.

- **Layer di cooperazione tra le app del sistema:** livello architeturale che consente il discovering e la comunicazione a run time tra le app e che implementa i meccanismi di cooperazione applicativa.
- **Layer semantico** e app di supporto alla realizzazione degli scenari di dimostrazione.