



*N.1 pubblicazione tecnico-divulgativa:*

## **ODINet: un innovativo motore di ricerca per i dati pubblicati sul web**

Codice documento: **D4.1.2**

Numero revisione: **3**

Data revisione: **19/03/2015**

*Il Progetto ODINet è realizzato con il determinante contributo della Regione Toscana a valere sul Programma Operativo Regionale cofinanziato dal FESR per l'obiettivo "Competitività regionale e occupazione" anni 2007-2013*



Regione Toscana



REPUBBLICA ITALIANA



Unione Europea  
Fondo Europeo di Sviluppo Regionale

**Le ali alle tue idee**

## Sommario

ODINet - Un innovativo motore di ricerca per i dati pubblicati sul web .....	3
Introduzione .....	3
Le componenti del framework ODINet .....	4
Data Integrator .....	4
Data Services .....	7
Data Viewer .....	7
Data Explorer .....	8
Lo scenario di validazione e gli enti sperimentatori.....	8
Il riuso degli Open Data .....	8
Prospettive future e scenari di utilizzo .....	9
ODINet a... Lisbona, conferenza eKNOW 2015.....	9
Evento finale di presentazione del progetto .....	10
La rete d'impresе "ODINet" .....	10

## ODINet - Un innovativo motore di ricerca per i dati pubblicati sul web

*ODINet è un progetto di ricerca e sviluppo cofinanziato dalla Regione Toscana nell'ambito del programma POR CREO 2007-2013, iniziato a Gennaio 2014 e concluso a Marzo 2015. Pensato e sperimentato nel campo dei Big Data, Open Data, Web Semantico e Linked Data, ODINet ha avuto come obiettivo la definizione di una piattaforma tecnologica per l'accesso e la diffusione online di dati strutturati ed eterogenei, caratterizzato da metodi innovativi per l'individuazione, la ricerca e la rappresentazione di dati sul Web. Il progetto ha previsto la realizzazione prototipale di un innovativo motore di ricerca semantico specializzato su dati relativi al settore sanitario, economico e sociale.*

### Introduzione

Negli ultimi anni, grazie alla diffusione di nuove tecnologie ed all'evoluzione delle relative infrastrutture, stiamo assistendo ad un'esplosione della quantità dei dati disponibili online, il cosiddetto "Big Data Bang". Quest'enorme quantitativo di dati racchiude al suo interno un patrimonio informativo di vasto interesse in numerosi ambiti. Tuttavia, ad una mole di dati in continua crescita, non è corrisposta una altrettanto rapida evoluzione degli strumenti tecnologici a disposizione per l'individuazione e l'analisi di tali dati, così che si è venuto a creare un gap tra l'informazione che abbiamo a disposizione e la capacità di riuscire a sfruttarla pienamente.

Il progetto ODINet si inquadra in questo contesto, al fine di condurre Ricerca e Sviluppo su tecnologie innovative che usate in sinergia consentano di :

- Raccogliere e uniformare dati e open data di interesse individuati presso diverse fonti e in diversi formati.
- Accedere all'informazione tematica, descrittiva e alle relazioni contenute nei dati per caratterizzarli e poterli inserire in una rete ontologica e sociale che favorisca la loro ricerca ed individuazione;
- Esplorare la rete di dati per mezzo di interfacce utente web user-friendly;
- Individuare per mezzo di un motore di ricerca semantico i dataset più rilevanti in base alle esigenze dell'utente.
- Rappresentare in modo uniforme i risultati della ricerca per mezzo di reportistiche tabellari, grafiche e cartografiche;
- Creare servizi web che consentano l'accesso federato a tutti i dati inerenti una determinata tematica, al fine di facilitare la realizzazione di specifiche applicazioni.

I partner di progetto sono la capofila Sistemi Territoriali srl, che si è occupata di tutti gli aspetti tecnologici, Simurg Ricerche, che ha curato la realizzazione dell'ontologia sociale, Rete Sviluppo sc, che ha progettato l'ontologia economica, l'Istituto di Fisiologia Clinica del CNR e CQR srl, che hanno collaborato per la realizzazione dell'ontologia di ambito sanitario. Il progetto si è avvalso inoltre della consulenza della Scuola Superiore Sant'Anna di Pisa.

## Le componenti del framework ODINet

Il framework ODINet è composto da quattro diversi moduli integrati tra loro, ovvero:

- Data Integrator
- Data Services
- Data Viewer
- Data Explorer

### Data Integrator

Il modulo “Data Integrator” ha come principale funzionalità quella di alimentare il catalogo dei dati del sistema per mezzo di tecniche di “data harvesting”, ovvero tramite l’individuazione e l’acquisizione di dati pubblicati sul web in modo automatizzato. E’ stato quindi implementato un sistema modulare di connettori software capaci di acquisire ad intervalli di tempo configurabili metainformazione e dati dai principali portali nazionali ed europei di open data, ovvero:

- Portali sviluppati con la tecnologia CKAN e SOCRATA, tra cui DatiGov.IT, il catalogo nazionale degli open data della pubblica amministrazione.
- Dati esposti secondo l’OpenData Protocol – ODATA ed lo Statistical Data and Metadata Exchange – SDMX , adottato da enti come ISTAT ed Eurostat.

Oltre ai dati la base di conoscenza di ODINet è composta di ontologie tematiche che consentono di classificare e ricercare le informazioni; in informatica un'ontologia è una rappresentazione formale, condivisa ed esplicita della concettualizzazione di un dominio.

Lo studio ontologico del progetto focalizza la propria attenzione sui domini della **sanità**, del **sociale** e dell'**economia** e delle loro connessioni partendo dall’assunto che l’innovazione tecnologica su questi tre ambiti può avere un impatto importante dal punto di vista organizzativo, socio-economico e clinico. La forte interrelazione che esiste fra i tre ambiti in termini di politiche di sviluppo, investimenti e ricadute sociali, incrementa la necessità di creare un patrimonio di conoscenza ‘open’, condiviso e accessibile per soddisfare le esigenze informative delle Pubbliche Amministrazioni e del singolo cittadino, mettendolo a disposizione della comunità scientifica.

La creazione dell’ontologia è avvenuto attraverso uno studio mirato degli strumenti scientifici disponibili, tenendo presente che la possibilità di scegliere ontologie già validate ed utilizzate per altri dataset di domini simili, permette di creare applicazioni aperte ed interoperabili. Le tematiche scelte coprono buona parte delle tipologie di dati disponibili e prodotti dalla PA nel contesto italiano e quindi si prestano per costruire dei casi d’uso concreti e significativi. Inoltre, queste tematiche, essendo strettamente collegate fra di loro, offrono l’opportunità di fare indagini incrociate soprattutto da un punto di vista conoscitivo, tramite l’individuazione di ontologie che descrivono in modo interconnesso concetti che legano le diverse tematiche. Le stesse ontologie, costituiscono poi una base messa a fattor comune di chiunque voglia in futuro usarle per la categorizzazione di banche dati di dominio e creare un linguaggio condiviso che decodifica le entità e le relazioni che stanno alla base della maggior parte delle esigenze informative delle PA e del singolo cittadino.

Le tre ontologie tematiche sono state definite come estensione di Eurovoc, un thesaurus multilingue pluridisciplinare redatto dall’Ufficio delle pubblicazioni dell’Unione europea che comprende la terminologia dei settori d’attività dell’Unione europea. Disponibile nelle 23 lingue ufficiali dell’Unione europea, è utilizzato dal Parlamento Europeo, dall’Ufficio delle pubblicazioni ufficiali dell’Unione europea, dai

parlamenti nazionali e regionali in Europa, da alcune amministrazioni nazionali e dalle organizzazioni europee. Esso costituisce di fatto la base per i nomi di dominio utilizzati nel database terminologico dell'Unione europea "Inter-Active Terminology for Europe".

### ***Ontologia del dominio "Salute"***

L'obiettivo principale dell'ontologia di dominio salute è quello di fornire informazioni sulle malattie croniche con particolare focus sulle malattie cardiovascolari. Partendo dal presupposto che la malattia è un processo in evoluzione, l'ontologia si basa su tre prospettive che convergono e si sovrappongono in modo naturale nel dominio di conoscenza: il rischio, la malattia ed il suo trattamento.

Oltre ad EuroVoc, sono state valutate e riutilizzate altre risorse specializzate disponibili in rete quali l'Unified Medical Language System (UMLS), un meta-tesauro di concetti biomedici, ed il MEDical Subject Headings (MESH), dizionario standardizzato della National Library of Medicine utilizzato per l'indicizzazione degli articoli scientifici pubblicati sulle riviste biomediche.

Il microtesauro salute di Eurovoc ed una porzione delle malattie cardiovascolari estratta dal MESH, sono stati dunque la base per la costruzione della ontologia successivamente ampliata per inserire una sezione dedicata alla classificazione dei farmaci e delle malattie.

Per quel che riguarda i farmaci si è scelto di riutilizzare lo standard internazionale Anatomical Therapeutic Chemical Classification system (ATC), mentre per quanto concerne le malattie si è adottato il dizionario International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM).

Poiché i dati di tipo clinico, normalmente storicizzati in cartella clinica/fascicolo sanitario elettronico ed i dati di tipo amministrativo disponibili attraverso flussi informativi ministeriali, presentano requisiti di privacy tali da non essere fruibili in modalità open, è stata effettuata una ulteriore estensione della ontologia sulla base dei dati open disponibili in forma aggregata messi a disposizione principalmente dalla Agenzia Regionale di Sanità (ARS Toscana) e da altre fonti istituzionali quali l'ISTAT. Sono stati mappati all'interno della ontologia una serie di indicatori di supporto alla predisposizione dei piani integrati di salute (Indicatori PIS) ed i concetti presenti nella banca dati delle malattie croniche degli assistibili Toscani (Macro) entrambi guidati dalla risorsa web mARSupio ([www.ars.toscana.it/marsupio/](http://www.ars.toscana.it/marsupio/)).

Il processo di definizione delle relazioni tra i concetti è stato guidato da esperti di dominio utilizzando il formalismo SKOS. Allo stato attuale l'ontologia di dominio salute si compone di 1381 concetti e relative relazioni e 1378 label di tipo *seeAlso* per la definizione di sinonimi dei concetti.

### ***Ontologia del dominio "Economia"***

L'ontologia di dominio economico ha inteso fornire una rappresentazione formale della conoscenza di tale ambito, con particolare riferimento all'utilizzo di quei concetti strettamente correlati ai domini sociale e sanitario, coerentemente con lo scenario di validazione del progetto ODINet.

La costruzione dell'ontologia è partita dal thesaurus EUROVOC, con l'utilizzo di 43 microthesauri afferenti 6 diversi settori EUROVOC:

- Settore 16 – Vita economica
- Settore 20 – Scambi economici e commerciali
- Settore 24 – Finanze
- Settore 40 – Impresa e concorrenza
- Settore 44 – Occupazione e lavoro

- Settore 68 - Industria

A partire da tale base, l'ontologia di dominio è stata successivamente arricchita con il Glossario economico statistico multilingua di Istat, una risorsa fondamentale poiché ha consentito di raccogliere i metadati relativi all'intero repertorio delle statistiche ufficiali prodotte dall'Istat e da altri enti del SISTAN. Il Glossario economico ha rappresentato un valore aggiunto nella sistematizzazione della conoscenza tematica poiché essa è una risorsa strettamente connessa alla principale fonte di dati open in Italia, l'Istat. Il Glossario – nelle sue 726 voci – fornisce inoltre i termini corrispondenti nelle principali lingue dell'Unione Europea, ovvero l'inglese e il francese.

Attraverso il coinvolgimento degli esperti di dominio, in aggiunta alle risorse citate sono stati aggiunti 42 nuovi concetti correlati al patrimonio di open data attualmente disponibile in Italia per l'ambito economico.

Il lavoro di 'affinamento' è avvenuto attraverso il matching con le altre ontologie, individuando anzitutto i punti di contatto tra ambito economico e 1) ambito sanitario (costi della salute, spese farmaceutiche, spese ospedaliere, politica sanitaria, spesa pubblica sanitaria, relazione tra patologie mediche e struttura economica dei territori) e 2) ambito sociale (welfare, costi della protezione sociale, occupazione e lavoro).

Per quanto concerne le 'dimensioni' dell'ontologia di dominio economico, essa si compone di 1.271 concetti, arricchiti – attraverso l'utilizzo del formalismo SKOS – da proprietà e relazioni quali broader, narrower, isRelated, hiddenLabel, seeAlso. Attraverso le annotazioni 'isDefinedBy', inoltre, sono state associate 572 descrizioni più dettagliate a concetti SKOS esistenti.

### *Ontologia Dominio Sociale*

La costruzione dell'ontologia per l'ambito sociale è partita dalla necessità preliminare di definire il termine "sociale". A questo proposito, sono stati valutati diversi approcci conoscitivi, partendo da un'impostazione di tipo epistemologico che ha condotto ad identificare l'ambito sociale come quello delle "scienze sociali", ovvero tutte le scienze che studiano il comportamento dell'uomo come individuo e nell'ambito delle forme organizzate che storicamente si è dato. Il campo delle scienze sociali, anche limitando il campo di interesse ad alcune scienze sociali (sociologia, demografia), è però molto vasto e si è posta quindi la necessità di delimitare in modo più specifico e pragmatico il dominio conoscitivo oggetto del lavoro. Dal momento che l'obiettivo pratico di ODINET è costruire un'applicazione (motore di ricerca) che consenta di trovare (principalmente) dati statistici pertinenti alle richieste dell'utente, si è quindi scelto di fondare la costruzione dell'ontologia "sociale" sui concetti e le classificazioni già definite dal Sistema della statistica ufficiale internazionale e europea ed italiana in particolare.

In questo ambito è in corso da almeno 20 anni un lavoro di standardizzazione a livello internazionale che ha portato alla definizione di glossari, classificazioni, metodologie e concettualizzazioni condivise dai principali sistemi di produzione statistica internazionali. I sistemi presi come riferimento sono il lavoro di UNECE (United Nations Economic Commission for Europe), che ha elaborato il DISA (Database of International Statistical Activities in the UNECE Region) e un thesaurus delle attività di produzione statistica (Classification of statistical activities - CSA 2009) che ha portato all'identificazione di 5 domini principali:

- Domain 1: Demographic and social statistics
- Domain 2: Economic statistics
- Domain 3: Environment and multi-domain statistics
- Domain 4: Methodology of data collection, processing, dissemination and analysis

- Domain 5: Strategic and managerial issues of official statistics

Questa classificazione è stata confrontata con le classificazioni utilizzate presso Eurostat e Istat, individuando le differenze. Si tratta di classificazioni e concetti per lo più omogenei, con piccole differenze terminologiche che sono state recuperate in sede di costruzione dell'ontologia attraverso le alternate labels. L'ambito di lavoro è stato delimitato all'interno del Dominio 1 e del Dominio 3, individuando i temi sui quali costruire l'ontologia sociale.

### Data Services

Il modulo Data Service è il cuore del motore di ricerca e ha in primo luogo l'obiettivo di derivare collegamenti logici tra i dati acquisiti dal modulo "Data Integrator" e le ontologie tematiche. Ispirandosi ed estendendo i principi dei Linked Open Data, i dati sono infatti connessi tra loro e con le ontologie in base a tematiche comuni, costituendo di fatto una "rete ontologica" che connette dati e concetti di interesse dei vari domini. Per aumentare i collegamenti della base di conoscenza e arricchire l'informazione semantica del grafo si è fatto ricorso al patrimonio informativo dei Big Data di Wikipedia, analizzando le descrizioni e legami tra le varie pagine descrittive dei concetti delle ontologie.

Per individuare i dataset più rilevanti rispetto ad una determinata tematica è quindi analizzato il grafo utilizzando algoritmi di Social Network Analysis che consentono di determinare le caratteristiche delle diverse entità della rete per mezzo di indicatori quali indici di centralità, di flusso, coefficienti di clustering, etc. Sulla base di tali indici è stato adattato l'algoritmo PageRank, utilizzato da tutti i principali motori di ricerca tra cui Google, per individuare i dataset più rilevanti per una ricerca effettuata da un utente. Proprio l'utilizzo degli algoritmi di Social Network Analysis, standard de-facto nel campo delle reti sociali ma mai applicate in un contesto come quello della ricerca semantica, costituisce uno degli aspetti più innovativi del progetto di ricerca.

Tutte queste funzionalità sono rese disponibili come web services, grazie ai quali applicazioni esterne possono condurre ricerche nel grafo ed essere notificati di nuovi dati o di aggiornamenti di dati già presenti nel sistema.

### Data Viewer

La fase di rappresentazione dei dati è una delle più critiche e decisive: l'informazione per diventare conoscenza deve infatti poter essere fruita nel modo più comprensibile possibile e gli utenti devono essere messi nelle condizioni di poter ricavare dal dato le informazioni di cui necessitano in modo intuitivo, veloce ed efficace.

Il modulo di progetto "Data Viewer" si è quindi focalizzato in questo ambito, al fine di garantire una visualizzazione flessibile, potente, semplice e uniforme delle informazioni individuate dal motore di ricerca del Data Service. Sono stati quindi progettati ed implementati moduli software specializzati per garantire ove possibile una rappresentazione tabellare, grafica e cartografica su web (HTML 5) ed una interfaccia utente che permetta di interagire con le reportistiche nel modo più semplice possibile. In fase di progettazione dell'interfaccia utente si è inoltre tenuto in considerazione il fatto che allo stato attuale della tecnologia l'attività di consultazione di media è svolta non solo su dispositivi classici come pc dotati di tastiera e mouse, ma anche da strumenti come smartphone, tablet e, nel prossimo futuro, TV. Per questo motivo abbiamo progettato interfacce di tipo responsivo, ovvero che permettono una visualizzazione uniforme e fruibile indipendentemente dalle dimensioni dello schermo del dispositivo.

## Data Explorer

Al fine di rendere comprensibile agli utenti come sono stati individuati i dati restituiti è stato realizzato il modulo Data Explorer, il quale offre la possibilità di esplorare dinamicamente con un'apposita interfaccia web il grafo di ricerca. L'utente potrà quindi navigare dinamicamente tale grafo, comprendere quali sono state le associazioni semantiche compiute dal motore e, in questo modo, ottenere nuovi spunti e chiavi di lettura per interpretare il fenomeno di studio. La realizzazione di questa interfaccia web, che costitutiva di per sé uno degli obiettivi di ricerca del progetto, mette nelle mani dell'utente uno strumento estremamente flessibile e intuitivo per comprendere le piene potenzialità del motore di ricerca semantica.

## Lo scenario di validazione e gli enti sperimentatori

Per la validazione del progetto sono stati individuati tre istituti che rivestono un ruolo centrale nel panorama toscano negli ambiti di sviluppo del progetto, i quali hanno accolto con entusiasmo la sfida lanciata dal progetto ODINet e hanno contribuito alla sua validazione fornendo dati, sperimentando il prototipo e fornendo le loro impressioni relativamente ad i risultati di progetto.

Gli enti coinvolti sono stati:

- *Dominio Sanitario:* Agenzia Regionale di Sanità della Toscana (**ARS**)
- *Dominio Sociale:* Osservatorio Sociale Regionale (**OSR**)
- *Dominio Economico:* l'Istituto Regionale di Programmazione Economica della Toscana (**IRPET**).

Alla fine i dataset indicizzati nel motore di ricerca sono provenienti da:

- I.Stat
- SoldiPubblici.it
- Open Toscana
- ARS - banca dati mARSupio
- IRPET
- OSR - Osservatorio Sociale Regionale
- Portale Open Data della Provincia di Lucca

## Il riuso degli Open Data

I dati aperti, comunemente chiamati **open data**, sono dati accessibili a tutti, tipicamente via web, senza restrizioni di copyright, brevetti o altre forme di controllo che ne limitino la riproduzione. Ma perché un ente pubblico dovrebbe rendere pubblici i propri dati? E che impatto ciò può avere sulla proprietà dei dati stessi? Che vantaggi possono esserci per i cittadini? E, ancora, cosa stanno facendo le istituzioni per promuovere l'adozione di una politica dei dati che sia veramente «open»?

L'open data si richiama alla più ampia disciplina dell'open government, una dottrina che sostiene che la pubblica amministrazione dovrebbe essere aperta ai cittadini, tanto in termini di trasparenza quanto di partecipazione diretta al processo decisionale: molti enti pubblici, proprio per questi motivi, rilasciano open data seguendo la filosofia che i dati da loro prodotti appartengano di fatto alla comunità e che quindi questi debbano essere distribuiti liberamente alla comunità stessa. In Italia il 18 Dicembre 2012 il Decreto Legge 18.10.2012, n. 179 è stato convertito in legge (17.12.2012 n. 221) e pubblicato nella Gazzetta Ufficiale: l'art. 9 prevede che i dati e le informazioni forniti dalla pubblica amministrazione dovranno essere obbligatoriamente pubblicati in formato aperto (cd. open data).



Il Workshop conclusivo del progetto, tenutosi al CNR di Pisa il 24 Marzo 2015, ha ospitato un'ampia sessione sul tema degli open data, in cui rappresentanti degli enti pubblici regionali e nazionali e della pubblica amministrazione si sono confrontati su questo tema di massima attualità, che costituisce uno dei principali temi esplorati dal progetto ODINet.

## Prospettive future e scenari di utilizzo

Calato in un mondo che dal punto di vista tecnologico è in costante evoluzione e che tende ad una apertura dell'informazione sempre maggiore, questo progetto promette rilevanti ricadute nella tecnologia di diffusione dei dati e soprattutto nel panorama italiano dell'open data e della PA, aggiungendo un importante tassello alle diverse iniziative che in questo ambito si muovono da diverse direzioni. Inoltre, con lo sviluppo del Web 3.0 e del Semantic Web i contenuti online tenderanno ad essere sempre più strutturati, permettendo un allargamento delle potenziali fonti di dati ed ulteriori sviluppi delle idee e tecnologie proposte. Parafrasando Tim Berners Lee (co-inventore del World Wide Web e promotore dell'Open Data), i dati diventano aperti quando smettono di essere "scuri, inscatolati e noiosi", ovvero quando smettono di essere inutilizzabili, e questo è quello che si propone il progetto ODINet: innovare per aiutare a rendere utilizzabili i dati pubblicati nel web e per ridurre le attuali asimmetrie informative tra detentore del dato (es. PA) e fruitore (es. cittadino).

ODINet è un progetto di ricerca: adesso occorre dargli gambe per renderlo un vero e proprio "prodotto" disponibile sul mercato. Da più parti – enti pubblici e amministrazioni – già si è mostrato interesse verso questa piattaforma. La sua applicabilità diventa tanto più necessaria quanto più nel prossimo futuro, seguendo le direttive europee, i dati pubblicati on line cresceranno in misura esponenziale. Giunti alla conclusione del progetto occorre inoltre definire una linea guida per gli eventuali sviluppi futuri della piattaforma: ODINet infatti si rivolge attualmente ad un ampissimo spettro di possibili utilizzatori, che vanno dagli enti pubblici ai singoli cittadini. Tuttavia, anche in base alle esigenze di mercato, potrebbe essere individuato in futuro un target di fruitori più ristretto e, in quest'ottica, potrebbero essere arricchite e consolidate le funzionalità della piattaforma che meglio rispondono alle specifiche esigenze di tali utilizzatori. Infine il progetto ODINet presenta ancora un elevato numero di spunti di ricerca esplorati solo in parte, che vanno dal suo utilizzo in ottica smart-cities, al potenziamento del motore di ricerca semantico, a un'internazionalizzazione del servizio prevedendo una versione multilingua della piattaforma. In particolare, dato che i diversi moduli tecnologici da cui ODINet è composto sono completamente disaccoppiati, il motore semantico potrebbe essere utilizzato in contesti completamente diversi, con la sola preconditione di avere a disposizione un'ontologia che modelli la conoscenza del dominio di riferimento. Inoltre una versione multilingua sarebbe realizzabile con la semplice traduzione delle terminologie specifiche introdotte nelle diverse ontologie di dominio, dato che i termini dell'ontologia Eurovoc sono già provvisti di traduzione in 23 lingue. Il Workshop finale ha ospitato diversi interventi sui possibili scenari di utilizzo di ODINet: in particolare i connettori di ODINet hanno da subito suscitato l'interesse di alcuni enti, quale ad esempio Eupòlis Lombardia, che ha adottato tale tecnologia per l'acquisizione automatizzata di tutti i dati statistici di I.Stat ed Eurostat nel suo sistema StatLomb.

## ODINet a... Lisbona, conferenza eKNOW 2015

Un paper relativo all'innovativo motore di ricerca dati semantico alla base del progetto ODINet è stato presentato il 26 Febbraio scorso alla conferenza internazionale eKNOW 2015, che ha avuto luogo a Lisbona dal 22 al 27 Febbraio 2015. Il paper è stato presentato nella sessione *eKNOW 7: Knowledge semantics processing and ontology* ed è risultato tra quelli premiati, tanto che gli autori sono stati invitati a sottometterne una versione estesa da pubblicare su uno degli IARIA Journals.

La conferenza, promossa da IARIA, l'International Academy, Research, and Industry Association, si colloca all'interno del Digital World 2015, un evento che riunisce numerosi convegni internazionali di diverso ambito su tecnologia e mondo digitale. eKNOW 2015 si è concentrata sui temi dei processi di modellazione e gestione della conoscenza: uno dei suoi topic principali è stato quello dell'acquisizione della conoscenza attraverso processi di analisi semantica e l'utilizzo di ontologie.

In questo contesto, il paper di ODINet si è proposto di analizzare i processi di formalizzazione delle conoscenze di dominio attraverso l'uso di ontologie tematiche specifiche e di spiegare in dettaglio il procedimento di analisi semantica alla base del motore di ricerca di dati.

## **Evento finale di presentazione del progetto**

Il progetto si è concluso con un Workshop dal titolo "*Web Semantico, Open Data, Big Data e Social Network Analysis applicate ai settori Sociale, Economico e Sanitario nel contesto Toscano*" tenutosi il 24 Marzo 2015, presso l'Aula Magna del CNR di Pisa. Il Workshop, organizzato dai partner di progetto, ha riunito professionisti, ricercatori, studiosi e decision-maker per una giornata di studio e confronto sui principali ambiti di riferimento del progetto stesso, quali web semantico, open data, big data, Social Network Analysis e loro applicazioni nei settori sociale, economico e sanitario nel contesto della Regione Toscana. Nel corso del Workshop sono stati esposti i principali risultati ottenuti dal progetto di ricerca: la giornata si è articolata con vari interventi sia di carattere scientifico che tematico, per concludersi con una tavola rotonda in cui si sono confrontati esperti e decision-maker. Hanno partecipato il presidente del W3C Italia Oreste Signore, il responsabile per l'open data ISTAT Vincenzo Patrino, il direttore dell'Azienda Regionale Toscana Francesco Cipriani e altre figure di primo piano nel panorama toscano e nazionale negli ambiti di progetto.

Il Workshop ha avuto un'ampia risonanza anche a livello mediatico, tanto che, oltre ad avere grande spazio sui giornali locali e su siti web di settore, è stato menzionato anche su media a diffusione nazionale come l'ANSA e il Corriere della Sera.

## **La rete d'impresе "ODINet"**

Viste le prospettive che si stanno aprendo sui mercati nazionali a seguito del progetto i partner Sistemi Territoriali srl, Rete Sviluppo s.c., Istituto di Fisiologia Clinica del CNR di Pisa, CQR srl e SIMURG snc stanno costituendo la rete di impresa "ODINet". La rete ha come obiettivo quello di offrire strumenti e servizi ad enti pubblici e privati in merito alle attività di analisi e pubblicazione sul web di dati.

La forte complementarità delle competenze degli enti costituenti la rete caratterizza ODINet come un soggetto unico nel panorama Italiano e comunitario, in grado di affiancare con un servizio completo e modulare gli enti in attività inerenti le tematiche degli open data e dello sfruttamento delle potenzialità del web semantico.

I principali servizi che offrirà la rete sono i seguenti:

- Strumenti e tecnologie per la realizzazione e/o evoluzione di portali di dati aperti.
- Consulenza tecnica, metodologica, tematica e giuridica per tutte le attività correlate alla pubblicazione di dati sul web.
- Supporto per la comunicazione delle attività di liberazione dei dati verso i cittadini e lo sviluppo della comunità Open Data .
- Corsi di formazione su Open Data, Linked Data e Web Semantico.

- Strumenti e metodologie per la condivisione ed il riuso di dati tra vari soggetti dello stesso ente.
- Strumenti innovativi basati sul semantic web per la ricerca e la fruizione di dati pubblici e non.

I principali clienti a cui si rivolge la rete sono tutti gli enti pubblici ma anche soggetti privati che volessero intraprendere attività di liberazione di dati in modo efficace, innovativo e conformemente ai massimi livelli di qualità previsti dagli standard. Gli strumenti di ricerca semantica proposti dalla rete hanno come potenziali utenti finali, oltre ai cittadini, anche decisori pubblici e privati ed uffici stampa, in quanto riuscire ad ottenere, in completa autonomia e con tempi molto rapidi, dati certificati su una tematica ed un territorio è per questi soggetti un'esigenza forte ancora non soddisfatta.