

***Gli Open Research Data
e la Science & Technology Digital Library***

Seminario

Gli Open Data a scuola e nella ricerca

Liceo classico “Giulio Cesare”
Roma, 17 gennaio 2014

Maurizio Lancia
Ufficio Sistemi Informativi e Documentali, CNR

Progetto Science & Technology Digital Library

I prodotti della ricerca

- *Varietà di prodotti* dell'indagine scientifica → progressiva elaborazione
 - dati della ricerca
 - prodotti intermedi (letteratura grigia)
 - pubblicazioni
- *Differente natura dei dati della ricerca*: grezzi, strutturati, elaborati, ecc.
- *Eterogeneità disciplinare*:
 - varietà di stili cognitivi e comportamenti nella produzione, gestione, condivisione, conservazione e valutazione di dati fra le differenti community
- *Apertura e chiusura dei dati*: distinzione fondamentale, ma non netta
 - Dati sensibili (sicurezza nazionale, salute e dati personali, segreto industriale...)

L'apertura dei dati della ricerca: i vantaggi

- *Fra gli scienziati:*
all'interno delle community e fra le differenti community (interdisciplinarietà, multidisciplinarietà...)
 - *Libero riutilizzo* → maggiore e migliore condivisione, migliore qualità della ricerca, accelerazione del progresso scientifico
- *Nella società*
 - Maggiore trasparenza
 - Partecipazione attiva dei cittadini
 - Innovazione tecnologica e impatto (imprese)

L'Open Science e il nuovo paradigma scientifico: le caratteristiche essenziali

- Progressiva **apertura dell'accesso** (*Open Access*) alla produzione scientifica e ai risultati della ricerca:
 - cooperazione fra gli scienziati & partecipazione attiva dei cittadini
 - trasparenza & disintermediazione
- L'ascesa della **scienza in rete** (*Networked Science*):
 - la dimensione comunitaria e collaborativa e la condivisione
 - il ruolo chiave delle nuove tecnologie relazionali e collaborative e delle piattaforme partecipative → le reti virtuali fra gli scienziati
 - la *citizen science*
- La *Open Notebook Science* e l'in fieri del lavoro scientifico:
 - nascita di un **modello alternativo**, ancora in via di definizione, **di scienza e comunicazione scientifica**

L'apertura dei dati della ricerca: questioni aperte e linee di indirizzo

I differenti attori/community parlano lingue *diverse* con interessi e obiettivi *diversi*

- *Questioni aperte, a molteplici livelli, fortemente interconnesse:*
 - *politiche, sociali ed economiche, culturali, giuridiche, organizzative, tecnico-tecnologiche*
- ***Necessità di una visione d'insieme e di un approccio coerenti*** → **interventi organici multilivello**

A livello internazionale:

- *OECD Principles and Guidelines for Access to Research Data from Public Funding 2007*
<http://www.oecd.org/sti/sci-tech/38500813.pdf>

L'Unione Europea:

- *Raccomandazione della Commissione del 17 luglio 2012 sull'accesso all'informazione scientifica e sulla sua conservazione*
<http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:IT:PDF>
- *Horizon 2020 (2014-2020) - 16 dicembre 2013:*
 - *Pilot on Open Research Data in Horizon 2020* http://europa.eu/rapid/press-release_IP-13-1257_en.htm
 - *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020*
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
 - *Guidelines on Data Management in Horizon 2020*
http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

L'apertura dei dati della ricerca: questioni aperte e linee di indirizzo

L'Italia:

- Position statement sull'accesso aperto ai risultati della ricerca scientifica in Italia - 21 marzo 2013
http://www.cnr.it/sitocnr/Iservizi/Biblioteche/Position_statement_OA_IT.pdf
- Linee programmatiche del MIUR presentate dal Ministro Carrozza alle Commissioni riunite del Senato e della Camera dei Deputati - impegno da parte del Ministero per lo «sviluppo della politica di Open Access dei risultati e dei dati della ricerca, sulla base delle raccomandazione della Commissione Europea...» - 6 giugno 2013
http://hubmiur.pubblica.istruzione.it/alfresco/d/d/workspace/SpacesStore/07d0c056-2a92-4ca9-9187-77d1107b9559/audizione_min_carrozza_060613.pdf
- Legge 7 ottobre 2013, n. 112 (in G.U. n. 236 dell'8 ottobre 2013) - Conversione in legge, con modificazioni, del decreto-legge 8 agosto 2013, n. 91, recante disposizioni urgenti per la tutela, la valorizzazione e il rilancio dei beni e delle attività culturali e del turismo (articolo 4 «Disposizioni urgenti per favorire lo sviluppo delle biblioteche e degli archivi e per la promozione della recitazione e della lettura» - comma 2 «I soggetti pubblici preposti all'erogazione o alla gestione dei finanziamenti della ricerca scientifica adottano, nella loro autonomia, le misure necessarie per la promozione dell'accesso aperto ai risultati della ricerca finanziata per una quota pari o superiore al 50 per cento con fondi pubblici, quando documentati in articoli pubblicati su periodici a carattere scientifico che abbiano almeno due uscite annue....omissis....»
http://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2013-10-10&atto.codiceRedazionale=13A08109&elenco30giorni=false

Le biblioteche digitali nel mondo della ricerca

- Le **biblioteche digitali** sono dei sistemi informativi complessi e devono gestire:
 - servizi informativi e documentali ad alto valore aggiunto
 - risorse digitali scientifiche e culturali provenienti da sorgenti informative eterogenee (service provider, Repository open access, Repository di dati della ricerca, Banche dati citazionali, Opac e MetaOpac, Current Research information system, sistemi di open data, ecc...)
- Sono contraddistinte da sistemi politico-organizzativi complessi e da strutture interne di gestione molto diversificate, che si riflettono nella struttura e nel tipo di risorse prodotte e conservate.
- Le biblioteche digitali devono quindi garantire l'accesso, la condivisione, la conservazione e il reperimento delle risorse prodotte da queste organizzazioni, gestendone l'eterogeneità.
- Il grado di complessità e la ricchezza dei patrimoni informativi e culturali richiede azioni improntate a logiche di 'sistema' e a forte cooperazione.

Il Progetto 'Science & Technology Digital Library'

- **Il quadro istituzionale:**

- è un'iniziativa dell'*Agenda digitale italiana* per lo sfruttamento delle ICT allo scopo di favorire crescita, innovazione e competitività → Agenda digitale europea (*Strategia Eu 2020*)
- è oggetto di un'apposita *Convenzione siglata il 17-07-2012* tra il **Consiglio Nazionale delle Ricerche** e il Dipartimento per la digitalizzazione della pubblica amministrazione e l'innovazione tecnologica della **Presidenza del Consiglio dei Ministri**, nell'ambito di un protocollo tra il MIUR e il CNR

- **L'obiettivo principale:**

sviluppare un *sistema integrato per l'accesso all'informazione sulla R&S* a beneficio di diverse comunità di utenti

- **I destinatari:**

- in primo luogo la **comunità scientifica** nelle sue **molteplici componenti**...
- successivamente una più **vasta platea di utenza** che costituisce il tessuto sociale e produttivo del Paese (istituzioni, imprese, società civile)

Gli obiettivi del sistema

La Digital Library sarà **integrata** e **interoperabile** con i maggiori sistemi nazionali ed internazionali della R&S e garantirà:

- *l'accesso permanente, certificato ed efficace* a risorse informative bibliografiche e documentarie nonché a dati scientifici, tecnici, statistici, attività/programmi di ricerca, expertise, ecc.
- *l'integrazione* di sistemi di gestione e di erogazione di servizi bibliografici, bibliotecari e biblioteconomici
- *l'integrazione* della documentazione, dei prodotti e dei dati tecnico-scientifici resi disponibili da istituzioni che operano nel settore della R&S nella logica dell'*Open Government* e dell'*Open Access*
- *lo sviluppo di servizi avanzati* profilati sulle esigenze informative dell'utenza
- *l'implementazione* di attività e servizi finalizzati alla *valorizzazione* e alla *conservazione della produzione scientifica nazionale* (deposito legale)

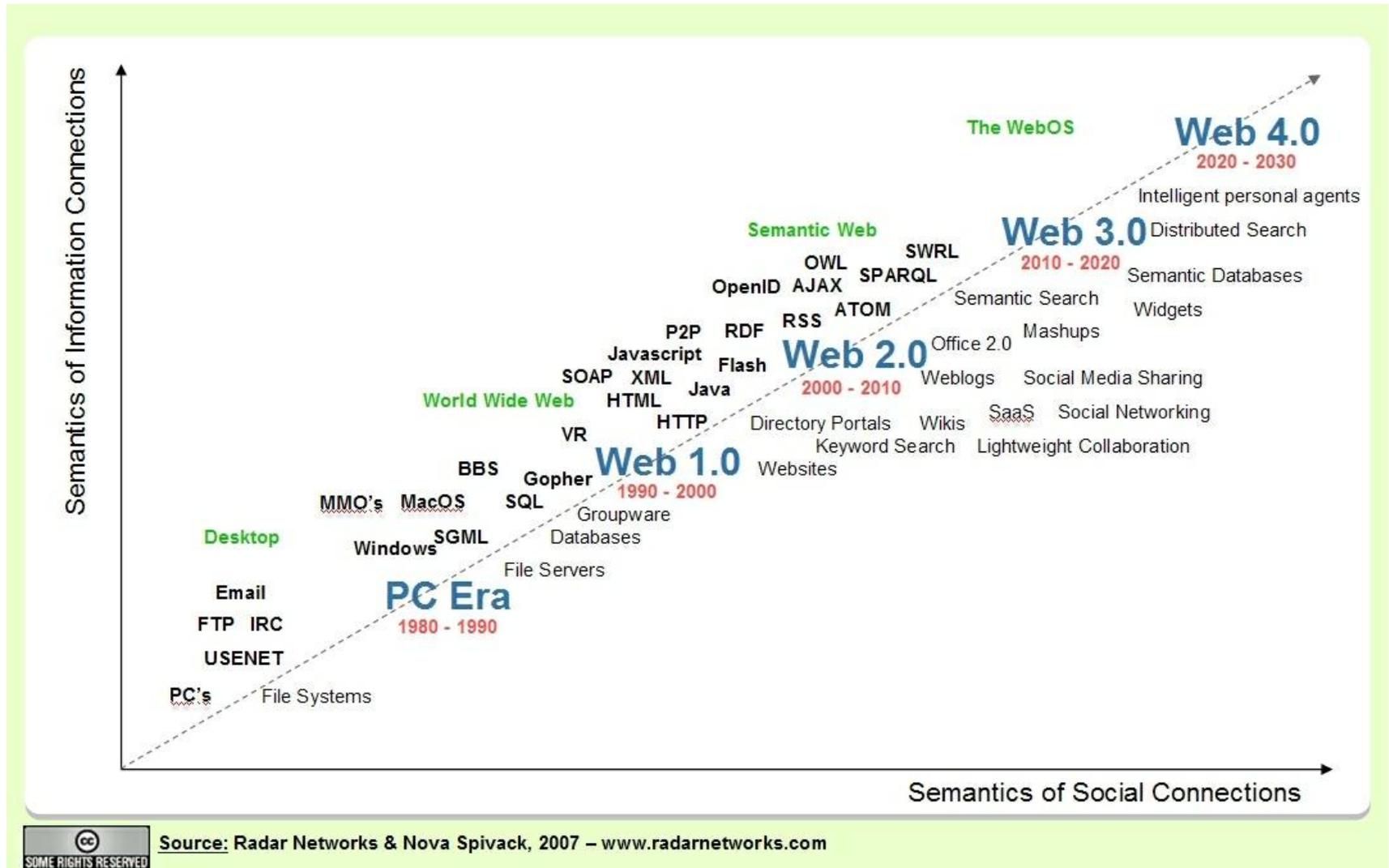
I principi guida

- *Integrazione & apertura*
- *Cooperazione*, condivisione, partecipazione
- *Interoperabilità* multilivello: politico-organizzativa, semantica, tecnica e tecnologica
- *Architettura distribuita e federata*
- *Valorizzazione della partnership:*
 - iniziative di collaborazione interistituzionale fra partner altamente rappresentativi del sistema nazionale della Scienza e della Tecnologia
- *Valorizzazione delle community:*
 - **dal *Semantic Web* al *Social Semantic Web***
- Sviluppo di strumenti di *diffusione e valorizzazione della cultura della R&S* → **mission del Consiglio Nazionale delle Ricerche**

Il ruolo del web semantico: verso il social semantic web

- L'idea di una social semantic digital library della scienza e della tecnologia nasce dall'esigenza di sviluppare ed estendere i sistemi di gestione delle biblioteche digitali (DLMS) integrando tecnologie proprie del web semantico e del social networking (2.0).
- Il S2W enfatizza l'importanza della semantica creata dall'uomo come mezzo per realizzare il web semantico. Invece di basarsi esclusivamente su semantiche automatizzate mediante l'elaborazione di ontologie formali e di inferenza e l'utilizzo del modello RDF, utilizza strumenti e sistemi di creazione e gestione delle meta-informazioni di tipo collaborativo, costruendo sistemi informativi socio-semantici.
- Il web socio-semantico sviluppa un'interfaccia nei processi di creazione, gestione e condivisione della conoscenza, consentendo un'interoperabilità partecipativa tra i sistemi informativi e i loro utenti, mediante l'utilizzo di tecniche, tecnologie e standard aperti, consolidati e largamente diffusi.

Il ruolo del web semantico: verso il social semantic web



Il web semantico: un esempio CNR

Esperienze trasferibili PA



#Buoni esempi #pratiche #casi #riuso #best practice

Home Notizie Documenti Norme Esperienze

Il portale dei dati aperti del Consiglio Nazionale delle Ricerche

Esperienza Martedì, 10 Aprile 2012 - 12:43pm

Il CNR ha intrapreso un'iniziativa per aprire i propri dati al libero accesso.

Il portale dei dati aperti del Consiglio Nazionale delle Ricerche ha l'obiettivo di offrire accesso pubblico alle informazioni del CNR. Il progetto è stato realizzato dal **Laboratorio di Tecnologie Semantiche (STLab)** del CNR, che comprende ricercatori e tecnologi dell'**Ufficio Sistemi Informativi (SI)** e dell'**Istituto di Scienze e Tecnologie della Cognizione (ISTC)**.

I dati pubblicati da data.cnr.it, che è anche il nome del dataset principale, si basano su **RDF** (Resource Description Framework) e **OWL** (Ontology Web Language) e seguono il paradigma del **Linked Data**. I dati disponibili sono un'estrazione in formato RDF di parte della banca dati del CNR, arricchita mediante tecnologie di ragionamento automatico e classificazione automatica. I dati sono concessi sotto la licenza "Attribuzione - Non commerciale - Non opere derivate 3.0 Unported" (CC BY-NC-ND 3.0).

Duplica la via di consultazione:

- seguendo Sfoglia, un [link](#) che consente di ottenere i dati strutturati in RDF o visibili in una pagina web in base alla negoziazione HTTP del tipo di contenuto;
- tramite SPARQL endpoint, un web service per interrogare il grafo usando il linguaggio SPARQL (Simple Protocol and RDF Query Language).

Inoltre è possibile utilizzare il [Semantic Scout](#), un esempio di applicazione di "expert finding" che sfrutta i dati del portale. Infine, i dati sono organizzati mediante un'[ontologia del CNR](#), la cui documentazione in formato OWL-Doc si trova a questo [indirizzo](#).

Social Network: [Twitter](#).

Politica di contesto e inquadramento legale: **Direttiva 2003/98/CE** del Parlamento europeo e del Consiglio, del 17 novembre 2003, relativa al riutilizzo dell'informazione del settore pubblico recepita dal **Decreto Legislativo 24 gennaio 2006, n. 36**; **Decreto Legislativo 7 marzo 2005, n. 82**, Codice dell'amministrazione digitale.

Destinatari: amministrazioni pubbliche, imprenditori, cittadini, enti di ricerca.

Vantaggi: le interfacce sono state realizzate per una fruizione facile sia da parte di una macchina o direttamente dell'utente con delle modalità grafiche con una ricerca semantica.

Il web semantico: un esempio CNR

- **data.cnr.it** ha l'obiettivo di fornire accesso pubblico all'informazione del CNR.
- Parole chiave sono: *trasparenza, interoperabilità, dati aperti*.
- Le tecnologie di data.cnr.it si basano su RDF e seguono il paradigma dei Linked Data: <http://webtemp.src.cnr.it/semanticscouting/semanticscouting.php>

L'esperienza tecnologica maturata e la specializzazione dei servizi sviluppati permettono di mettere in campo soluzioni d'avanguardia da rendere disponibili ad una *platea di utenti estremamente variegata*, dagli specialisti dei domini scientifici e tecnici ai cittadini.

integrazione

apertura

interoperabilità



cooperazione

community-centred approach

condivisione



Portale d'accesso

approcci
innovativi

persistent
identifier

digital right
management

social
semantic web

long-term
digital
preservation



piattaforma repository



**reference
virtuale**



metaOpac



**formazione
a distanza**

Dalla S&TDL all'Open Science

