



Forum PA

Roma, 16-19 maggio 2012

Open Data INPS

**Modelli di sviluppo per l'Open Government
nella PA**

STATO DELL'ARTE

VALUTAZIONI EU
MODELLO INPS
DOMINIO
TIPOLOGIA
PREPARAZIONE
METADATI
LICENZA
SEMANTICA E
INTEROPERABILITA'



Portale INPS OPEN DATA

Una piattaforma per i dati del settore pubblico:

AFFIDABILI

AGGIORNATI

CERTIFICATI

15 marzo 2012

l'**Inps** pubblica sul proprio portale una sezione dedicata agli **Open Data**, in linea con la Direttiva del Parlamento Europeo 2003/98/CE.



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

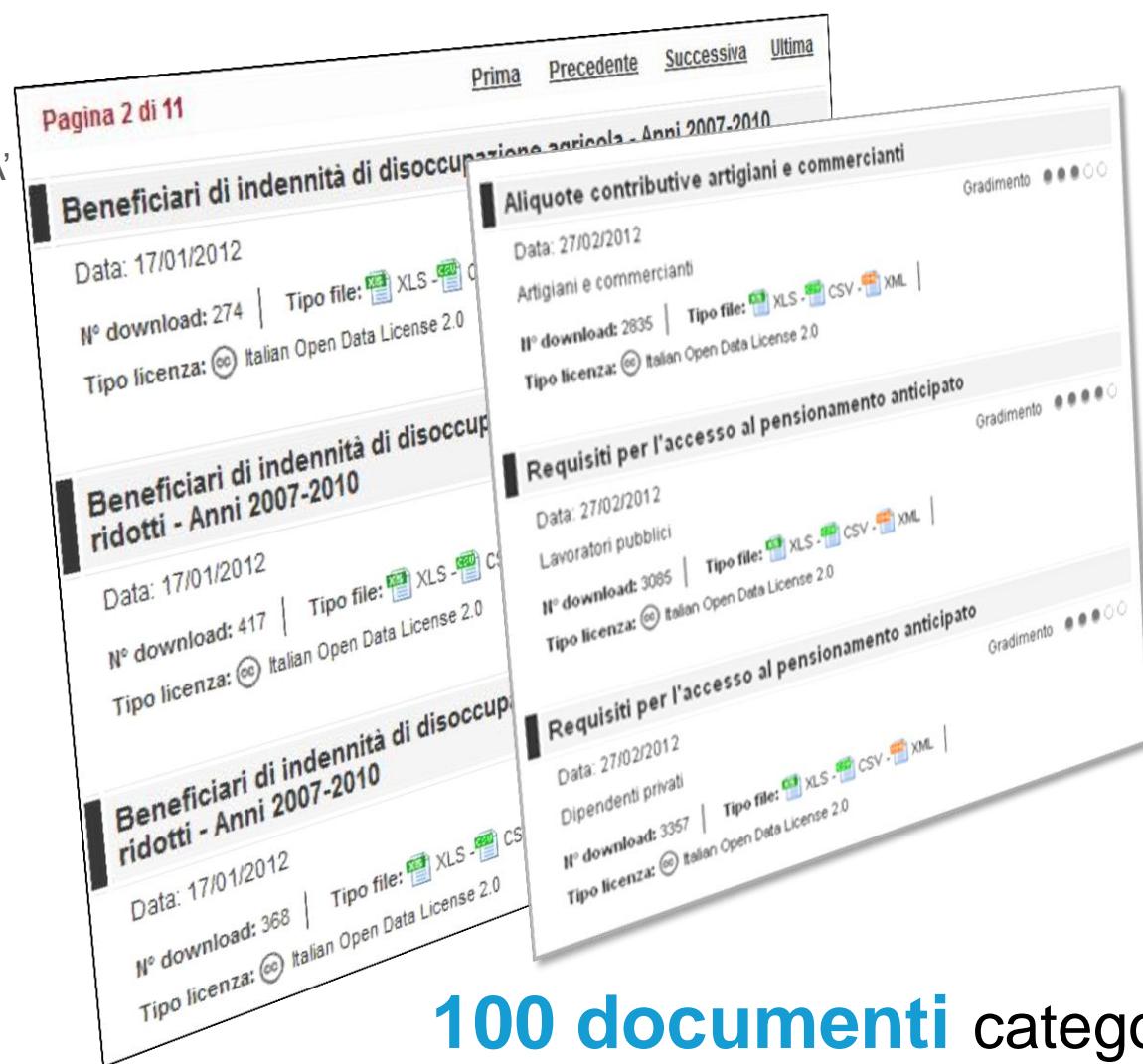
SEMANTICA E

INTEROPERABILITA'



Obiettivo ?

- Favorire il **riuso sicuro di dati e informazioni** utili, verificate, accessibili e consultabili in modo rapido.
- Agevolare la **condivisione di informazioni** tra enti e istituzioni pubbliche nazionali e comunitarie.



100 documenti categorizzati per argomento e fonte.
consultabili dai cittadini e **utilizzati da operatori pubblici e privati** per condurre ricerche e produrre soluzioni informatiche.

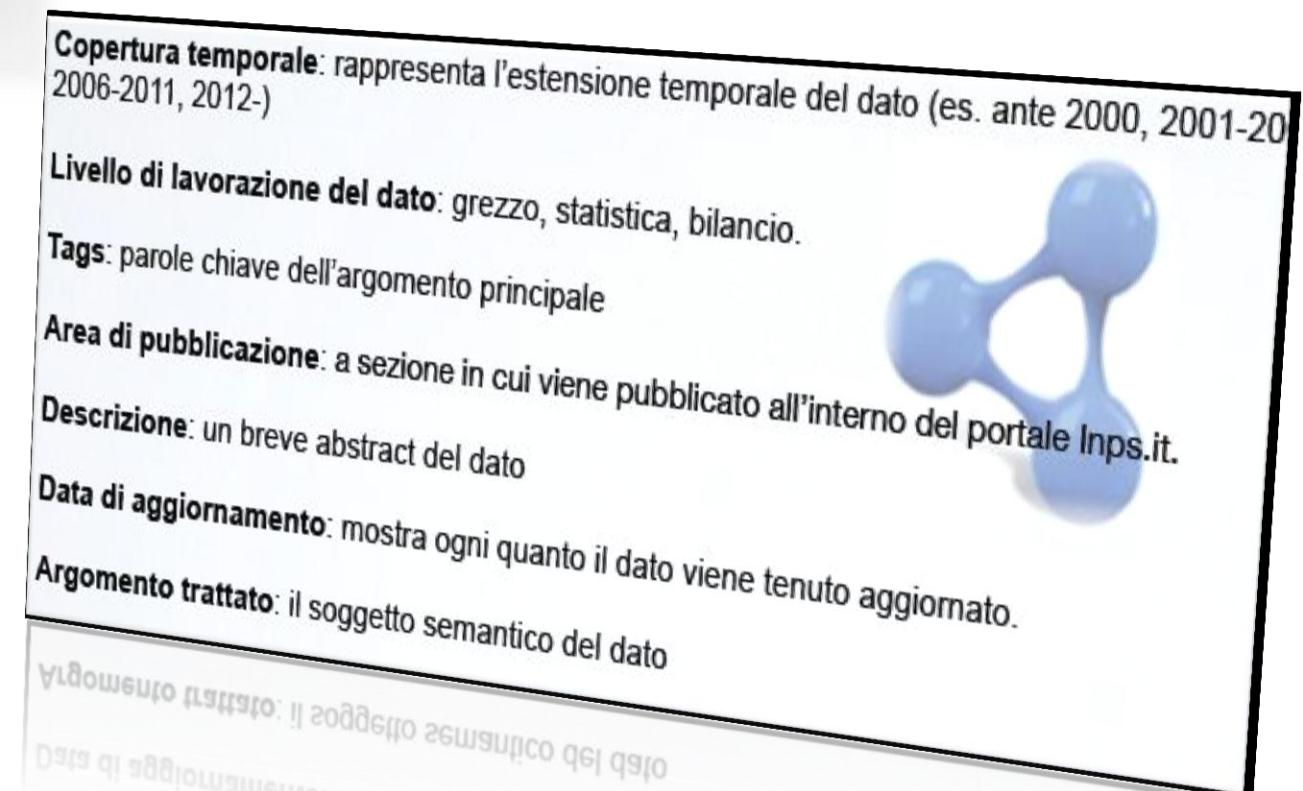


STATO DELL'ARTE

VALUTAZIONI EU
 MODELLO INPS
 DOMINIO
 TIPOLOGIA
 PREPARAZIONE
 METADATI
 LICENZA
 SEMANTICA E
 INTEROPERABILITA'



I **contenuti** variano dalle **aliquote contributive ai requisiti pensionistici**, dai dati relativi alle **prestazioni a sostegno del reddito** agli **osservatori sulle tipologie di lavoratori**, solo per fare alcuni esempi.

Copertura temporale: rappresenta l'estensione temporale del dato (es. ante 2000, 2001-2006-2011, 2012-)

Livello di lavorazione del dato: grezzo, statistica, bilancio.

Tags: parole chiave dell'argomento principale

Area di pubblicazione: a sezione in cui viene pubblicato all'interno del portale Inps.it.

Descrizione: un breve abstract del dato

Data di aggiornamento: mostra ogni quanto il dato viene tenuto aggiornato.

Argomento trattato: il soggetto semantico del dato

A **supporto** dell'iniziativa, una **organizzazione semantica dei dati**, a tutte le risorse sono associati dei **metadati** che descrivono il dato attraverso le sue caratteristiche semantiche, ovvero le informazioni sul contenuto. Tali aspetti consentono di individuare univocamente le risorse.

STATO DELL'ARTE

VALUTAZIONI EU
MODELLO INPS
DOMINIO
TIPOLOGIA
PREPARAZIONE
METADATI
LICENZA
SEMANTICA E
INTEROPERABILITA'



Metadati?

Una marcia (di ricerca) in più..



Le tecniche di **analisi semantica dei contenuti** permettono la creazione delle **facchette per il motore di ricerca.**

*Le facchette rappresentano i diversi aspetti o dimensioni attraverso i quali si può descrivere una stessa risorsa e aiutano gli utenti a **mettere in relazione** i dati tra di loro.*

PER PERIODO

- ▶ 2001-2005 (25)
- ▶ 2006-2011 (102)
- ▶ 2012- (7)

PER FONTE

- ▶ Osservatori statistici (49)
- ▶ Rapporti (52)

PER ARGOMENTO

- ▶ Contributi (7)
- ▶ Lavoro (78)
- ▶ Pensione (23)
- ▶ Attività Ispettive (2)
- ▶ Prestazioni Assistenziali (58)



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'



Marzo 2012

l'**Inps** pubblica sul proprio portale una sezione dedicata agli **Open Data**, in linea con la Direttiva del Parlamento Europeo 2003/98/CE.

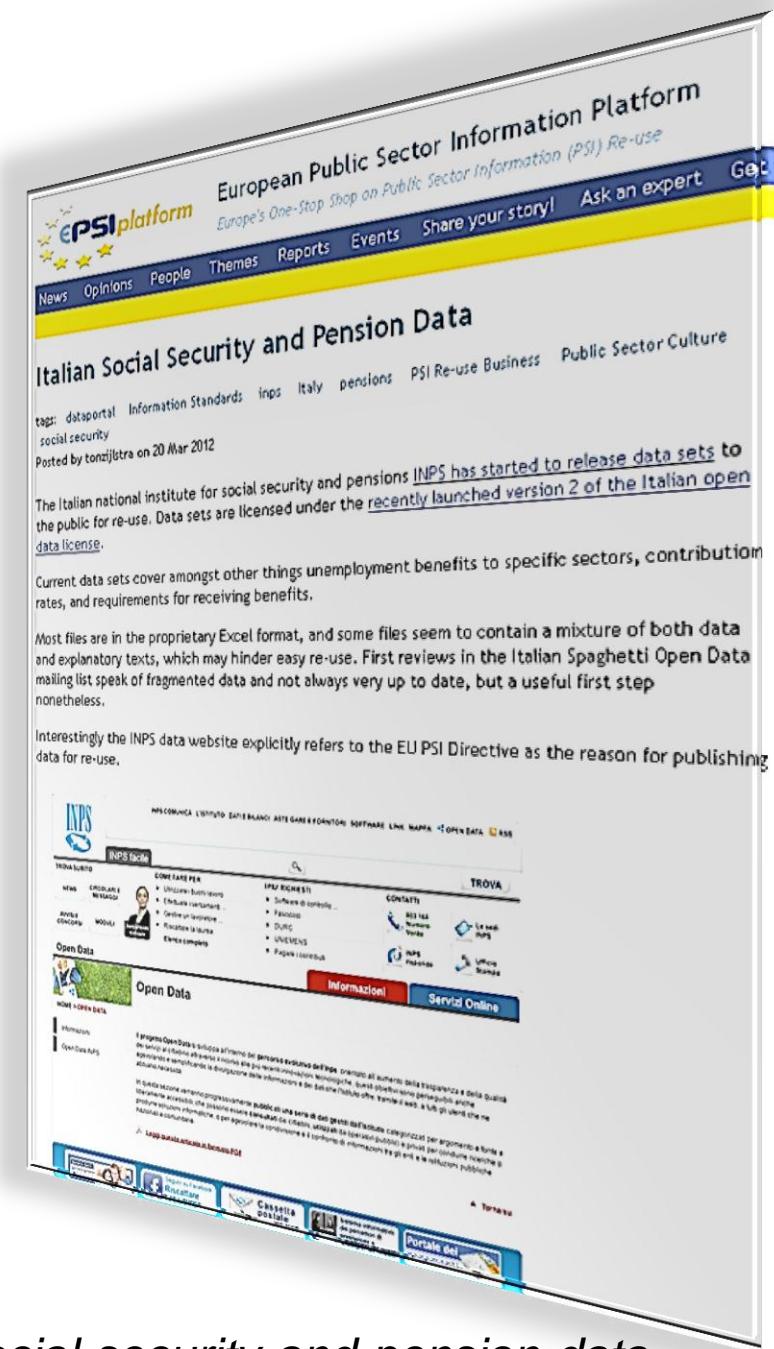
Aprile 2012

quasi **11.000 download** effettuati nel primo mese di pubblicazione.

Importanti consensi internazionali in un articolo dalla **European Public Sector Information Platform** (EPSI), punto di riferimento europeo per gli Open Data per tecnici e amministratori pubblici.



Visibile al link <http://epsiplatform.eu/content/italian-social-security-and-pension-data>





STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

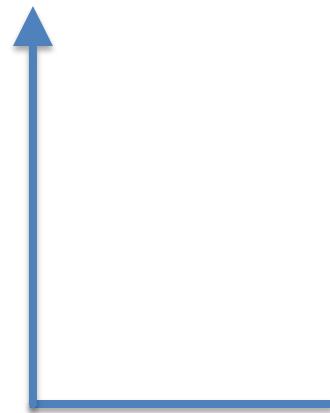
SEMANTICA E

INTEROPERABILITA'

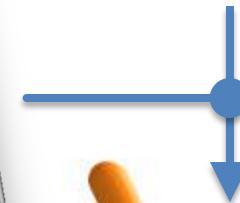
UN'OPPORTUNITÀ DI SVILUPPO DEI SERVIZI AL CITTADINO



Maggiore Sviluppo di servizi di terze parti basati su dati certificati



Maggiore Trasparenza, accesso a dati aggiornati, Interoperabili, riutilizzabili





STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

... l'Europa corre veloce

Dicembre 2011 - Nuova OpenData Strategy

Neelie Kroes, Vice-Presidente Commissione europea per l'Agenda digitale.

Marzo 2012- ePSI platform Conference

Taking re-use to the next level

Aprile 2012 – EU JoinUp ADMS

Semantic Interoperability solution for EU public administrations



STATO DELL'ARTE

VALUTAZIONI EU
MODELLO INPS
DOMINIO
TIPOLOGIA
PREPARAZIONE
METADATI
LICENZA
SEMANTICA E
INTEROPERABILITA'



Aggiornamenti e direttive

- Tutti i dati resi disponibili dal PSI devono poter essere riutilizzabili per uso commerciale e non commerciale.
- Rendere i dati in formato machine readable (RDF, JSON, XML..).
- Gli stati membri devono introdurre un sistema di supervisione e regolamentazione al fine di rispettare le direttive UE sul riuso nei tempi stabiliti.
- Lo scopo delle direttive sarà esteso alle informazioni contenute nelle digital libraries, musei e archivi.

Aggiornamenti ISA - SEMIC

- *Interoperability solution for European public administration* -
- *Semantic interoperability center Europe* -

- La pubblicazione di E-government Metadata.
- L'implementazione dei formati verso i Linked OpenData.

STATO DELL'ARTE

VALUTAZIONI EU
MODELLO INPS
DOMINIO
TIPOLOGIA
PREPARAZIONE
METADATI
LICENZA
SEMANTICA E
INTEROPERABILITA'



Obiettivo

Valutare la qualità dei dataset

Open (Government Data)

Tutte le tipologie di dati di interesse pubblico

Valutare la qualità dei processi di Governance

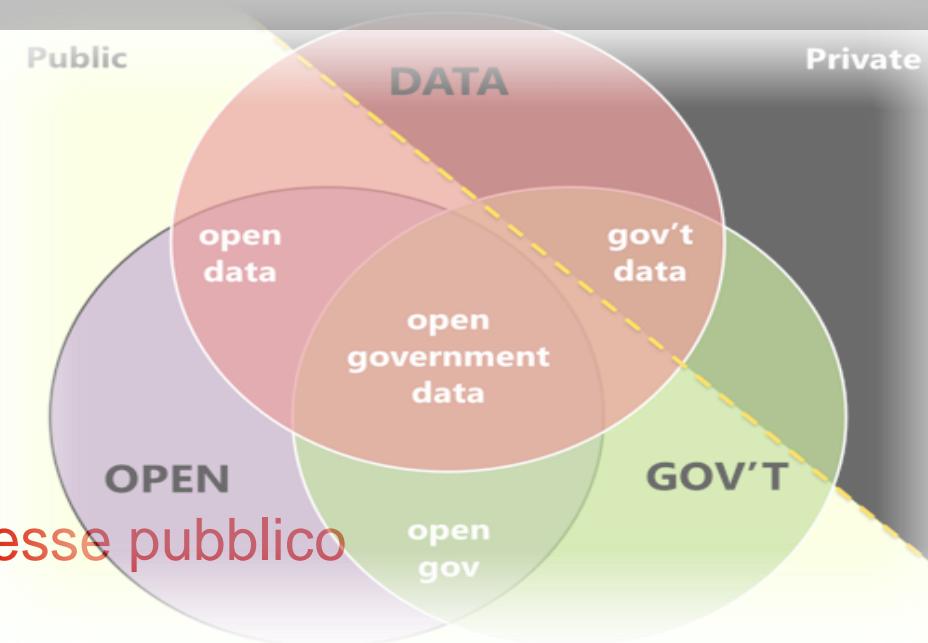
(Open Government) Data

Tutte le tipologie di dati che contribuiscono alla trasparenza delle amministrazioni.

Favorire l'interoperabilità dei dati PSI

(ADMS)

Asset description metadata schema: vocabolario per l'interoperabilità semantica.





STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

Valutazioni EPSI platform

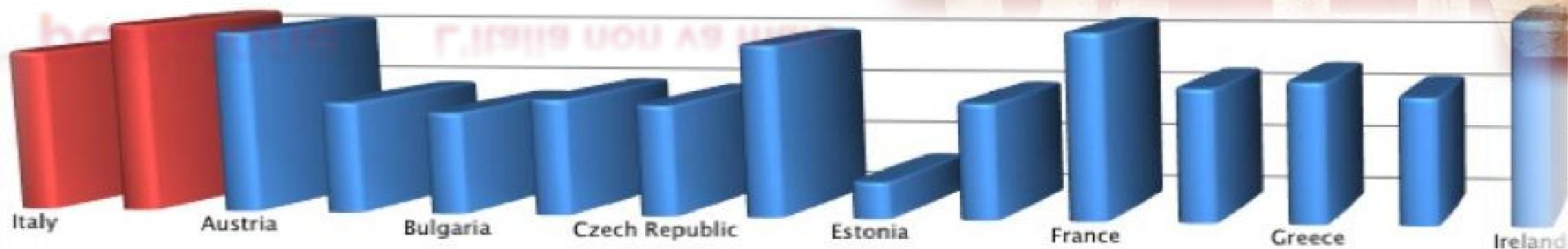
European Public Sector information

Criteria

- ⦿ Implementazione delle direttive EU
- ⦿ Sviluppo dei formati
- ⦿ Eventi ed attività correlati
- ⦿ Dati delle amministrazioni locali
- ⦿ Pratiche di riuso

8° posizione

L'Italia non va male





STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

Punti di forza!

L'analisi degli indicatori mostra che non siamo lontani da paesi già avviati in materia di Open Government (UK)

Countries	Overall Score	Implementation of the PSI Directive	The practice of re-use
Italy	345	100	40
United Kingdom	405	100	60

Formats	Local PSI	Events and activities
50	1	75
30	40	75

- ✓ Implementazione delle direttive EU
- ✓ Sviluppo dei formati
- ✓ Eventi ed attività correlate



Cosa migliorare?

- ✗ Pratiche di riuso
- ✗ Rilascio dei dati dalle amministrazioni locali



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

La correlazione tra **pratiche di riuso** e rilascio dei **dataset degli enti locali** suggerisce di implementare un piano di rilascio OpenData a partire dal modello degli **Enti pubblici** più importanti.



OBIETTIVO?

Costruire un framework di principi di sviluppo condivisi, facilitando la comprensione e la gestione di grandi e complessi volumi di dati provenienti da fonti pubbliche differenti.

- Definire vincoli di qualità**
- Standardizzare i processi**
- Condividere le risorse**

STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'



L'INPS e l'Europa

Un canale aperto

*“Interestingly the **INPS** data website explicitly refers to the EU PSI Directive as the reason for publishing data for re-use.”*

*“..continuous dialogue between data holders and re-users is needed, as we are learning how to do this by doing. The **INPS** data release is a good step forward, and creates appetite for more.”*

The screenshot shows the INPS Open Data website interface. At the top, there is a navigation bar with 'Informazioni' and 'Servizi Online'. The main header features the INPS logo and the text 'Open Data INPS'. Below this, a breadcrumb trail reads 'Istituto Nazionale > OPEN DATA > OPEN DATA INPS Previdenza Sociale'. A sidebar on the left contains a search bar and a list of categories, including 'Open Data INPS'. The main content area displays two search results. The first result is 'Aliquote contributive artigiani e commercianti', dated 27/02/2012, with 946 downloads, an XLS file type, and an Italian Open Data License 2.0. The second result is 'Requisiti per l'accesso al pensionamento anticipato', also dated 27/02/2012, with 1030 downloads, an XLS file type, and an Italian Open Data License 2.0. On the right side, there is a 'SELEZIONE CATEGORIE OPEN DATA' section with a list of categories: Contributi (1), Lavoro (78), Pensione (27), Prestazioni sociali (58), Reddito e tasse (2), and Trasparenza e valutazione (2). Below this is a 'PER FONTE' section with 'Allegati, circolari e messaggi (7)'. At the bottom left, there is a logo for 'EPSI platform' with a starburst graphic.



- STATO DELL'ARTE
- VALUTAZIONI EU
- MODELLO INPS**
- DOMINIO
- TIPOLOGIA
- PREPARAZIONE
- METADATI
- LICENZA
- SEMANTICA E
- INTEROPERABILITA'

LIBERARE I DATI



Da dove partire?

Quali dati vanno aperti?

E i metadati?



Come mettere insieme i pezzi?

Come produrre un dataset di qualità?



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

Metodologia INPS

OPEN

DATA

- *Analisi di dominio e individuazione dei dati*
- *Analisi tipologica dei dati*
- *Preparazione dei dati e definizione dei formati*
- *Definizione dei metadati*
- *Attribuzione della licenza*





STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

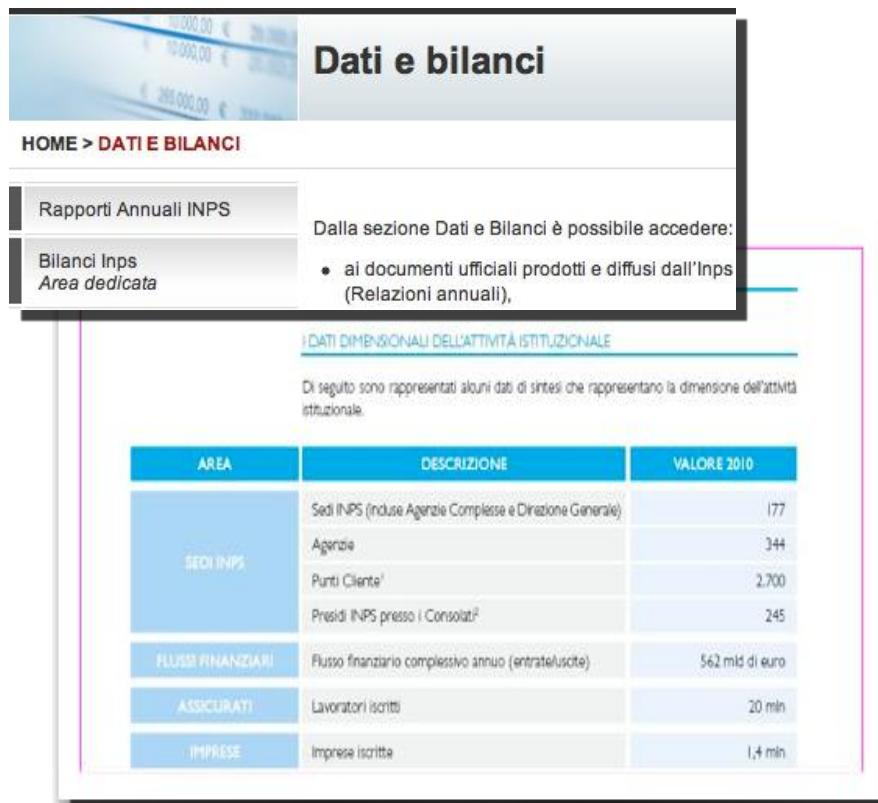
Analisi di dominio e individuazione dei dati

Obiettivo

Identificare i contenuti già esposti all'interno delle relative aree di pubblicazione.

Conoscere il dominio permette di:

- **Selezionare i dati rilevanti alla pubblicazione.**
- **Individuare i contenuti in relazione alle aree in cui sono esposti**
Mantenendo l'autenticità della fonte
- **Individuare le macroaree semantiche**
Utili alla definizione dei metadati
- **Individuare subito i dati che possono essere aperti.**
Rendere parallele le filiere di lavoro sui dati, riadattando le fasi complessive del processo di liberazione alla specifiche esigenze dell'ente produttore



Dati e bilanci

HOME > DATI E BILANCI

Rapporti Annuali INPS

Bilanci Inps
Area dedicata

Dalla sezione Dati e Bilanci è possibile accedere:

- ai documenti ufficiali prodotti e diffusi dall'Inps (Relazioni annuali).

I DATI DIMENSIONALI DELL'ATTIVITÀ ISTITUZIONALE

Di seguito sono rappresentati alcuni dati di sintesi che rappresentano la dimensione dell'attività istituzionale.

AREA	DESCRIZIONE	VALORE 2010
SEDI INPS	Sedi INPS (Incluse Agenzie Complesse e Direzione Generale)	177
	Agenzie	344
	Punti Clienti ¹	2.700
	Presidi INPS presso i Consorzi ²	245
FLUSSO FINANZIARI	Flusso finanziario complessivo annuo (entrate/uscite)	562 mld di euro
ASSICURATI	Lavoratori iscritti	20 mln
IMPRESE	Imprese iscritte	1,4 mln



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

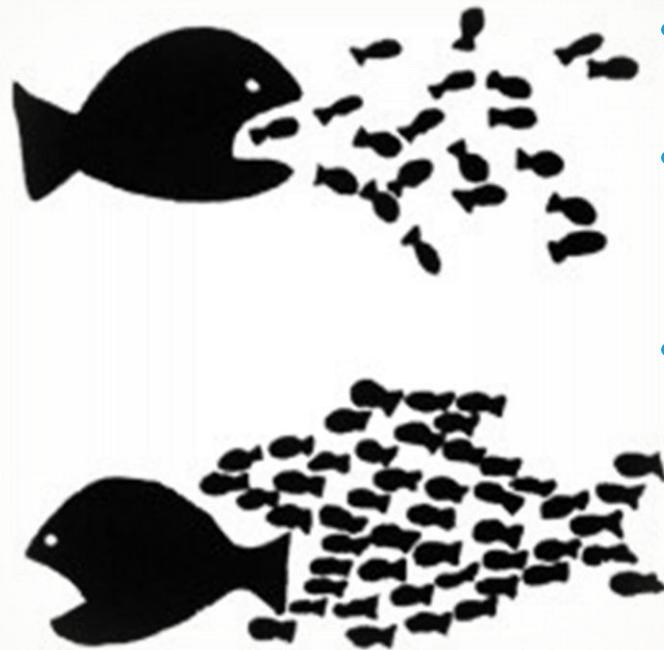
INTEROPERABILITA'

Analisi di dominio e individuazione dei dati

Il vantaggio di partire da un'analisi accurata del dominio?

Permette di affrontare fin dall'inizio tutti i punti critici dell'intero processo fin dalla sua prima fase.

In questo modo è possibile già da subito:



- ***valutare costi e tempi dell'operazione complessiva.***
- ***Individuare eventuali punti critici del processo e pianificare immediatamente la loro risoluzione.***
- ***Individuare le esigenze specifiche dell'ente produttore e modificare di conseguenza la realizzazione del processo di apertura.***



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

Analisi tipologica dei dati

1° Livello di analisi

individua il tipo in base al modo in cui è stato prodotto:

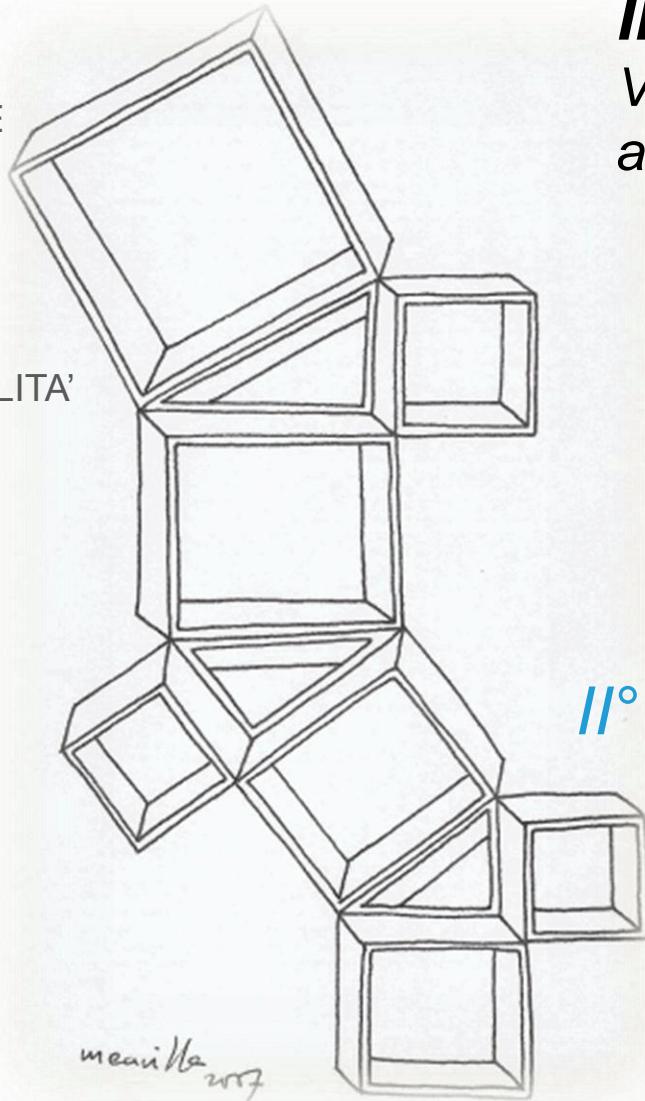
Dati grezzi: (da non confondere con la definizione "Raw" trad. Grezzo relativa al formato) raccolte di dati semplici, nei quali di solito è presente una sola chiave con un valore associato.

Ad esempio, la tabella "Nuclei familiari con entrambi i genitori e almeno un figlio minore" individua l'importo dell'ANF in relazione al reddito familiare.

NUCLEI FAMILIARI CON ENTRAMBI I GENITORI E ALMENO UN FIGLIO MINORE IN CUI NON SIANO PRESENTI COMPONENTI INABILI													
Importo complessivo mensile dell'assegno per livello di reddito e numero componenti il nucleo													
<i>Reddito familiare annuo di riferimento valido dal 1° luglio 2010</i>													
Reddito familiare annuo (euro)		Importo dell'assegno per numero dei componenti il nucleo familiare											
		1	2	3	4	5	6	7	8	9	10	11	12
fino a	13.211,14			137,50	258,33	375,00	500,00	625,00	773,75	922,50	1.071,25	1.220,00	1.368,75
13.211,15 -	13.316,82			136,73	257,25	374,04	499,58	624,38	773,04	921,69	1.070,35	1.219,01	1.367,67
13.316,83 -	13.422,51			135,95	256,17	373,08	499,17	623,75	772,31	920,88	1.069,44	1.218,00	1.366,56
13.422,52 -	13.528,20			135,18	255,08	372,13	498,75	623,13	771,60	920,07	1.068,54	1.217,01	1.365,48
13.528,21 -	13.633,89			134,40	254,00	371,17	498,33	622,50	770,88	919,25	1.067,63	1.216,00	1.364,38
13.633,90 -	13.739,58			133,63	252,92	370,21	497,92	621,88	770,16	918,44	1.066,73	1.215,01	1.363,29
13.739,59 -	13.845,27			132,85	251,83	369,25	497,50	621,25	769,44	917,63	1.065,81	1.214,00	1.362,19
13.845,28 -	13.950,96			132,08	250,75	368,29	497,08	620,63	768,72	916,82	1.064,91	1.213,01	1.361,10
13.950,97 -	14.056,65			131,30	249,67	367,33	496,67	620,00	768,00	916,00	1.064,00	1.212,00	1.360,00



- STATO DELL'ARTE
- VALUTAZIONI EU
- MODELLO INPS
- DOMINIO
- TIPOLOGIA**
- PREPARAZIONE
- METADATI
- LICENZA
- SEMANTICA E
- INTEROPERABILITA'



Analisi tipologica dei dati

II° Livello di analisi

Verificare il grado di complessità strutturale dei dataset e di annidamento dei dati

I° livello: semplice

	Maschi	Femmine
	Numero	Numero
2005	71.207	43.344
2006	59.640	40.884
2007	54.268	40.338
2008	54.417	40.263
2009	75.955	40.904

II° livello: medio

Anno	ARTIGIANI-COMMERCianti			
	Titolari di qualunque età e coadiuvanti /coadiutori di età superiore ai 21 anni		Coadiuvanti /coadiutori di età inferiore ai 21	
	I fascia di reddito	II fascia di reddito	I fascia di reddito	II fascia di reddito
2011	20,00%	21,00%	17,00%	18,00%
2012	21,30%	22,30%	18,30%	19,30%
2013	21,75%	22,75%	18,75%	19,75%
2014	22,20%	23,20%	19,20%	20,20%
2015	22,65%	23,65%	19,65%	20,65%
2016	23,10%	24,10%	20,10%	21,10%

Tavola III.3.1.2 - Beneficiari di disoccupazione non agricola requisiti ordinari e disoccupazione speciale mensile 2007-2009; 1° semestre 2010 (valori assoluti)

MESI	Nord			Centro		
	Numero Beneficiari	Entrate	Uscite	Numero Beneficiari	Entrate	Uscite
	2007					
gennaio	86.272	-	15.785	37.866	-	6.128
febbraio	80.645	10.158	15.533	36.532	4.794	5.817
marzo	75.625	10.513	20.788	35.355	4.640	8.172
aprile	68.898	14.061	15.803	31.553	4.370	7.841
maggio	62.917	9.822	15.554	27.481	3.769	6.852
giugno	58.650	11.287	13.167	25.353	4.724	4.880
luglio	74.009	28.526	10.113	33.159	12.686	4.150
agosto	76.158	12.262	19.247	34.484	5.475	9.076
settembre	71.388	14.477	23.965	30.584	5.176	9.487
ottobre	73.481	26.058	11.186	31.049	9.952	5.132
novembre	87.194	24.899	13.056	36.117	10.200	5.094
dicembre	94.310	20.172	16.503	38.213	7.190	4.543
Media annua beneficiari	75.796	-	-	33.146	-	-

III° livello: difficile

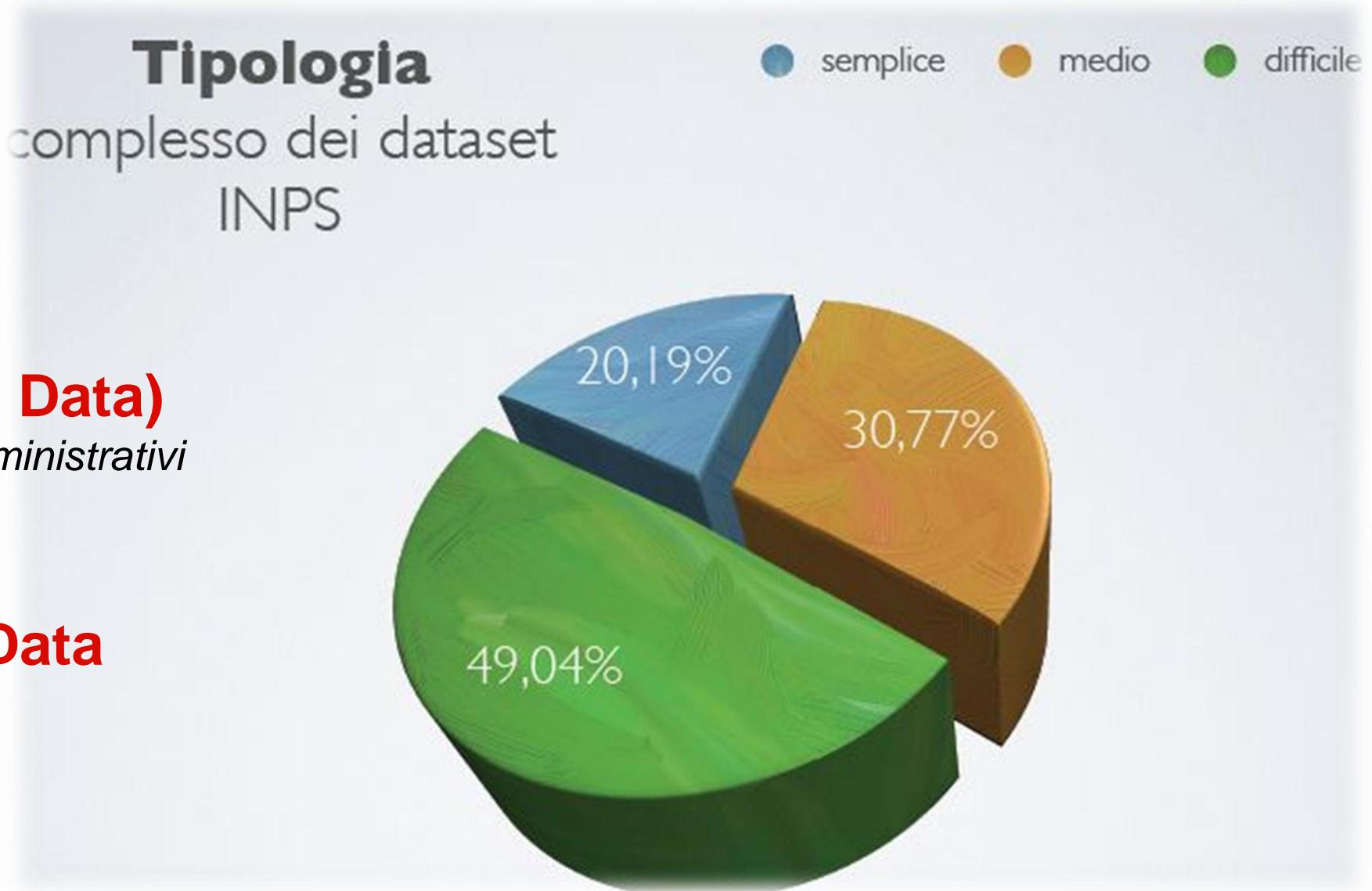


- STATO DELL'ARTE
- VALUTAZIONI EU
- MODELLO INPS
- DOMINIO
- TIPOLOGIA**
- PREPARAZIONE
- METADATI
- LICENZA
- SEMANTICA E
- INTEROPERABILITA'

Analisi tipologica dei dati

III° Livello di analisi

E' ora possibile distinguere ulteriormente i dati: quelli che contribuiscono alla trasparenza da quelli di più generale interesse pubblico.



Open (Government Data)

Es. posizione degli uffici amministrativi

(Open Government) Data

Es. dati sugli appalti pubblici

STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'



Preparazione dei dati e definizione dei formati

Poiché molte delle fonti presentano le informazioni in maniera non strutturata **è opportuno effettuare un'operazione di pulizia sulla struttura e sul contenuto delle tabelle.**

Creare i primi formati (.xls per lo più) in modo preciso e puntuale consente di:

- **Implementare in modo più veloce i formati successivi** (XML, RDF fino ai Linked Open Data).
- **Offrire dati più facilmente leggibili sia dagli utenti che dalle macchine.** In questo modo si favorisce sia la comunicazione e la trasparenza con i cittadini, sia un riuso basato su servizi e applicazioni ad integrazione di dataset (mashup, openservice).



Ad esempio, una buona formattazione .XLS consente una più facile conversione XML tramite XSD (XLSschema) e CSV



STATO DELL'ARTE
 VALUTAZIONI EU
 MODELLO INPS
 DOMINIO
 TIPOLOGIA
PREPARAZIONE
 METADATI
 LICENZA
 SEMANTICA E
 INTEROPERABILITA'



Preparazione dei dati e definizione dei formati

CURE FORME VIE RESPIRATORIE						
ELENCO ALBERGHI CONVENZIONATI PER LA STAGIONE TERMALE 2011 (V. N.B. IN FONDO)						
LOCALITA' e Provincia	NOME ALBERGO	n° stelle	Cat. tariff. Terme	Indirizzo albergo n° telef. e n° fax	Distanza in metri dalle terme e mezzo di collegamento (solo alberghi non dotati di terme interne)	Trattamenti particolari (non termali) praticati gratuitamente
Abruzzo						
CARAMANICO TERME (PE)	MAIELLA E DELLE TERME	3	1S	Via Roma, 29 - 65023 Caramanico Terme (PE) tel. 085/92301 fax 085/9230710 e-mail: info@albergomaiella.it	INTERNE	Albergo annesso allo Stabilimento termale; parcheggio riservato non custodito; camere con servizi privati; tv color sat e telefono diretto; piccola colazione a buffet; frutta a buffet; menù a scelta; dessert (due volte alla settimana); animazione; pianobar
CARAMANICO TERME (PE)	VINCENZELLA	3	1S	Via Roma, 21 - 65023 Caramanico Terme (PE) tel. 085/922104	30 metri	Dolce a pranzo la domenica e il giovedì, escursioni naturalistiche
CARAMANICO TERME (PE)	VIOLA	3	1S	Via della libertà, 9 - 65023 Caramanico Terme (PE) Tel. 085/922292 - fax 085/9231818	50 metri	Una serata danzante settimanale
CARAMANICO TERME (PE)	IACOBUCCI	3	1S	Via della Libertà 7 - 65023 Caramanico Terme (PE) Tel. 085-922103	15 metri	
Basilicata						
LATRONICO (POTENZA)	TERME LUCANE S.r.l. - Hotel Monte Alpi	3	a	Contrada Calda 85043 - Latronico (PZ) telef. 0973/859238 fax 0973/859856	250 metri - bus navetta	Sconto del 30% sulle tariffe in vigore
Calabria						
ACQUAPPESA-GUARDIA PIEMONTESE (CS)	GRAND HOTEL	4	1a SUPER	LOCALITA' TERME LUIGIANE.GUARDIA PIEMONTESE N.TELEFONO 0982-94052, N.FAX 0982.94478	TERME INTERNE	SERVIZIO DI ANIMAZIONE-SERVIZIO NAVETTA DA E PER LA SPIAGGIA-INGRESSO PARCO TERMALE IN OCCASIONE DELLE MANIFESTAZIONI DELL'ESTATE CULTURALE-SCONTO DEL 50% SUL PREZZO DI INGRESSO AL PARCO TERMALE.
ACQUAPPESA-GUARDIA PIEMONTESE (CS)	ALBERGO TERMALE MODERNO	3	1a SUPER	LOCALITA' TERME LUIGIANE.GUARDIA PIEMONTESE N.TELEFONO 0982-94048, N.FAX 0982.94478	250 metri di strada carrabile e pedonale,collegamento con servizio di navetta dedicato.	SERVIZIO DI ANIMAZIONE-SERVIZIO NAVETTA DA E PER LA SPIAGGIA-INGRESSO PARCO TERMALE IN OCCASIONE DELLE MANIFESTAZIONI DELL'ESTATE CULTURALE-SCONTO DEL 50% SUL PREZZO DI INGRESSO AL PARCO TERMALE.



REGIONE	NOME ALBERGO	n° stelle	Cat. tariff. Terme	VIA	CAP	LOCALITA'	Provincia
Abruzzo	MAIELLA E DELLE TERME	3	1S	Via Roma, 29	65023	Caramanico terme	PE
Abruzzo	VINCENZELLA	3	1S	Via Roma, 21	65023	Caramanico Terme	PE
Abruzzo	VIOLA	3	1S	Via della libertà, 9	65023	Caramanico Terme	PE
Basilicata	TERME LUCANE S.r.l. - Hotel Monte Alpi	3	1S	Contrada Calda	85043	Latronico	PZ



STATO DELL'ARTE
VALUTAZIONI EU
MODELLO INPS
DOMINIO
TIPOLOGIA
PREPARAZIONE
METADATI
LICENZA
SEMANTICA E
INTEROPERABILITA'

Preparazione dei dati e definizione dei formati

*In questo modo è possibile produrre XML e CSV di qualità, comprensibili da uomini e macchine **senza dover ripetere ad ogni conversione l'operazione di pulizia e controllo**. Un esempio di XML INPS:*

```
<data>
  <datum Anno="2007" Regione_Area_Geografica="Piemonte">
    <Maschi>2595</Maschi>
    <Femmine>1674</Femmine>
    <Maschi_e_Femmine>4269</Maschi_e_Femmine>
  </datum>
  <datum Anno="2007" Regione_Area_Geografica="Valle d'Aosta / Vallée d'Aoste">
    <Maschi>826</Maschi>
    <Femmine>161</Femmine>
    <Maschi_e_Femmine>987</Maschi_e_Femmine>
  </datum>
  <datum Anno="2007" Regione_Area_Geografica="Lombardia">
    <Maschi>3898</Maschi>
    <Femmine>1447</Femmine>
    <Maschi_e_Femmine>5345</Maschi_e_Femmine>
  </datum>
</data>
```

STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'



Definizione dei metadati

*Tali informazioni rappresentano delle proprietà più significative fissate e consentono la ricerca e il recupero dei dati stessi con lo scopo di **migliorarne la visibilità e facilitarne l'accesso**.*

Tuttavia i vocabolari dei metadati sono molto ampi e richiedono una selezione di quelli che effettivamente possono arricchire di informazioni i dati in questione. Anche qui è quindi necessario adattare le specificità di ciascun ente agli standard trovati.

La creazione e l'implementazione dei metadati richiede attenzione alle risorse informative che tali metadati devono riflettere.

Molta attenzione deve essere data alla qualità, perché un pessimo metadato è peggio dell'assenza di metadati.

Per l'INPS è stata realizzato un set che include gli standard Dublin Core ma li arricchisce con alcuni tipi di metadati proprio del dominio INPS



- STATO DELL'ARTE
- VALUTAZIONI EU
- MODELLO INPS
- DOMINIO
- TIPOLOGIA
- PREPARAZIONE
- METADATI**
- LICENZA
- SEMANTICA E
- INTEROPERABILITA'

Definizione dei metadati

Metadati sintattici: *descrivono il dato attraverso le sue caratteristiche strutturali, in cui sono comprese le informazioni sul processo di creazione, organizzazione e pubblicazione. Tali metadati non aggiungono informazioni sul contenuto del dataset.*

Titolo: titolo dei dati

Data di pubblicazione: Questa informazione riguarda la data di pubblicazione del dato all'interno del portale/minisito OpenData Inps.

Formato: l'estensione del file (.pdf, .xls, .rdf, .doc)

Licenza : il tipo di licenza a cui è legato il dato

Dimensione in byte: la dimensione fisica del dato

Identificativo : numero progressivo univoco del dato

Link: (URL-permanent link-, link annidati, short URL)

Fonte: opzionale, chi ha prodotto il dato (Direzione, altro soggetto)



- STATO DELL'ARTE
- VALUTAZIONI EU
- MODELLO INPS
- DOMINIO
- TIPOLOGIA
- PREPARAZIONE
- METADATI**
- LICENZA
- SEMANTICA E
- INTEROPERABILITA'

Definizione dei metadati

Metadati semantici: *descrivono il dato attraverso le sue caratteristiche semantiche, ovvero le informazioni sul contenuto. Tali aspetti consentono di individuare univocamente le risorse.*

Copertura temporale: rappresenta l'estensione temporale del dato (es. ante 2000, 2001-2006-2011, 2012-)

Livello di lavorazione del dato: grezzo, statistica, bilancio.

Tags: parole chiave dell'argomento principale

Area di pubblicazione: a sezione in cui viene pubblicato all'interno del portale Inps.it.

Descrizione: un breve abstract del dato

Data di aggiornamento: mostra ogni quanto il dato viene tenuto aggiornato.

Argomento trattato: il soggetto semantico del dato





STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'



In questo modo ritorniamo una lista metadati:

- **coerente con quelle internazionali.**
- **Adeguate agli standard di interoperabilità europea.**
- **specifiche per l'ente nazionale.**

Definizione dei metadati

Una volta descritta la risorsa attraverso opportuni metadati è necessario rappresentare il livello di descrizione anche nei successivi formati di pubblicazione.

Dal formato XML in poi è utile **implementare l'interoperabilità dei dati**. E' importante, quindi, fare nuovamente un match con gli standard metadata Dublin Core.

Metadati INPS	Adeguamento Dublin Core
Titolo	dc: title
Data di Pubblicazione	dcTerms: issued
Licenza	dcTerms: License
Link	dc: Identifier
Fonte	dc: Publisher
Copertura Temporale	dcTerms: Temporal
Area	dc: Source
Descrizione	dc: Abstract
Argomento trattato	dc: Subject



STATO DELL'ARTE
 VALUTAZIONI EU
 MODELLO INPS
 DOMINIO
 TIPOLOGIA
 PREPARAZIONE

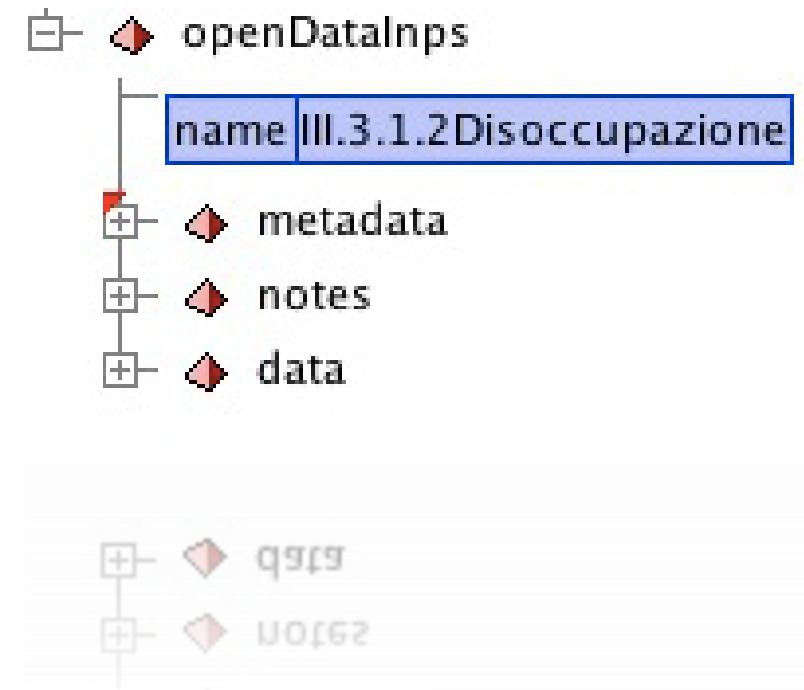
METADATI

LICENZA
 SEMANTICA E
 INTEROPERABILITA'

Definizione dei metadati

Ecco quindi l'esempio di una struttura XML per un dataset INPS, completa di metadati, note ai dati e dati. Qui la struttura complessiva dell'XML:

E' necessario separare e distinguere i metadati, i dati ed eventuali note integrative al dato all'interno della struttura.



```
<metadata>
<notes>
  <note>(*) Soggetti che hanno presentato la domanda entro il mese di marzo dell'anno
  <note>Fonte: INPS - Coordinamento Generale Statistico Attuariale</note>
</notes>
<data>
<data>
</data>
<note>Fonte: INPS - Coordinamento Generale Statistico Attuariale</note>
```



STATO DELL'ARTE
 VALUTAZIONI EU
 MODELLO INPS
 DOMINIO
 TIPOLOGIA
 PREPARAZIONE

METADATI

LICENZA
 SEMANTICA E
 INTEROPERABILITA'

Definizione dei metadati

I metadati nell'XML

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<openDataInps xmlns="http://www.inps.it/OpenData/1.0" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/1.1/">
  <metadata>
    <dc:publisher>INPS</dc:publisher>
    <dc:title>III.3.1.6 - Beneficiari di indennità di disoccupazione non agricola (*) con requisiti ridotti</dc:title>
    <dcterms:abstract>Rapporto sulla coesione sociale - Dati sui beneficiari di indennità di disoccupazione non agricola</dcterms:abstract>
    <dcterms:issued>17/01/2012</dcterms:issued>
    <dcterms:license>IODL 2.0</dcterms:license>
    <dcterms:LicenseDocument>http://www.dati.gov.it/iodl/2.0/</dcterms:LicenseDocument>
    <dc:identifier>http://www.inps.it/docallegati//Mig/OpenData/III.3.1.6_Disoccupazione.xls</dc:identifier>
    <facets>
      <argomento>
        <dc:subject>Lavoro.LavoroDipendente</dc:subject>
        <dc:subject>Lavoro.AltriTipiDiLavoro</dc:subject>
        <dc:subject>PrestazioniAssistenziali.Disoccupazione</dc:subject>
      </argomento>
      <fonte>
        <dcterms:source>Rapporti</dcterms:source>
      </fonte>
      <periodo>
        <dcterms:temporal>2006-2011</dcterms:temporal>
      </periodo>
    </facets>
  </metadata>
</openDataInps>
```



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

Attribuzione della licenza

Obiettivo?

Scegliere una licenza che consenta all'utente un riutilizzo più facile e libero possibile

La **IODL 2.0** prevede che l'utente possa liberamente:

1-Consultare, estrarre, copiare e pubblicare i dati

2-Creare un lavoro derivato integrando diversi dataset.

..rispetto alla IODL 1.0?

Non prevede l'obbligo dell'utente di pubblicare o condividere i lavori con la stessa licenza. E' chiesto solo di indicare la fonte!

Italian Open Data License v2.0

Preambolo

La "Italian Open Data License" (IODL) è un contratto di licenza che ha lo scopo di consentire agli utenti di condividere, modificare, usare e riutilizzare liberamente la banca di dati, i dati e le informazioni con essa rilasciati, garantendo al contempo la stessa libertà per altri. La presente licenza mira a facilitare il riutilizzo delle informazioni pubbliche nel contesto dello sviluppo della società dell'informazione.

Le banche di dati, i dati e le informazioni sono protetti dalle leggi applicabili in materia di diritto d'autore (incluso il diritto sui generis del costituente di banche di dati) e/o dalle altre leggi applicabili. I diritti concessi dalla presente licenza non implicano alcun trasferimento di diritto di titolarità sulle banche di dati, sui dati e sulle informazioni pubbliche.

Gli utenti possono utilizzare e riutilizzare la banca di dati ed i dati in essa contenuti in base ai termini della presente licenza; ogni utilizzazione che non sia espressamente autorizzata ai sensi della presente licenza o delle leggi vigenti è proibita.

Con l'esercizio di uno qualunque dei diritti qui previsti, accetti e ti obblighi a rispettare integralmente i termini della presente licenza.

1. Definizioni

Ai fini e per gli effetti della presente licenza, si intende per:

- **"Informazioni"**, la banca di dati ed i dati (testi, data base e data set, immagini, video, ecc.) offerti per l'uso secondo i termini di questa licenza;
"Lavoro derivato", ogni traduzione, adattamento, trasformazione, diversa disposizione ed ogni altra modifica delle Informazioni o di una parte sostanziale della banca di dati;
- **"Licenziante"**, il soggetto (Ente, persona fisica o persona giuridica) che fornisce le Informazioni secondo i termini e le condizioni della presente licenza;
- **"Tu/Te"**, il soggetto (Ente, persona fisica o persona giuridica) che usa le Informazioni ed esercita i diritti derivanti dalla presente licenza, impegnandosi a rispettarne i termini.



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

E' davvero necessaria tutta questa attenzione alla qualità dei dati?

*Arricchire semanticamente i dati significa stabilire uno standard di **qualità** a livello di **pubblicazione**, di **utenza** e di **interoperabilità** nella Pubblica Amministrazione.*

*In più, lo sviluppo dei LOD consente di **connettere e condividere i dataset** attraverso standard già disponibili e implementati*

Molti portali internazionali forniscono risorse mal rappresentate frammentarie e poco collegate

Technical requirements of a good dataset

Computer Science Dept. – Dresden University
 In collaborazione con
Open Knowledge Foundation
 (2012)

Dataset		
Discoverability	Format	
<ul style="list-style-type: none"> - existence of - descriptions - tags 	<ul style="list-style-type: none"> - machine readable - non-proprietary - separation of metadata and data 	
Validity	Quality	Granularity
<ul style="list-style-type: none"> - area (e.g. state, county) - time of validity 	<ul style="list-style-type: none"> - clear attribute semantics - no missing values 	<ul style="list-style-type: none"> - raw data instead of highly aggregated data

Figure 1: Possible Features of an Open Data Platforms and Datasets



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

Un dataset di qualità è la base dell'interoperabilità nella PA

I **metadati** per i dataset **INPS** sono stati creati seguendo tali punti:

- *Analisi degli standard metadata più importanti (DublinCore, ADMS –asset metadata description)*
- *Costruzione di un framework di metadati coerente con il dominio INPS.*
- *Adattamento dei metadati INPS così ottenuti agli standard di interoperabilità europei.*

In questo modo i metadati risultanti sono specifici per l'ente, ma già pronti per essere non solo **comprensibili agli utenti** ma anche **machine-readable**.

Questo permette lo sviluppo di **migliori soluzioni di front-end dedicati e maggiore integrazione nei processi di back-end**. L'obiettivo è quindi quello di aiutare gli utenti e mettere in relazione i dati tra di loro attraverso le informazioni che li descrivono.





STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'



Come implementare l'interoperabilità semantica Idee e tecnologie

L'**EIF** (*European interoperability framework*) ha definito l'interoperabilità semantica nel PSI come l'abilità delle organizzazioni di **processare informazioni da una fonte esterna** assicurando **la comprensione e la preservazione del loro significato nello scambio tra le diverse parti.**

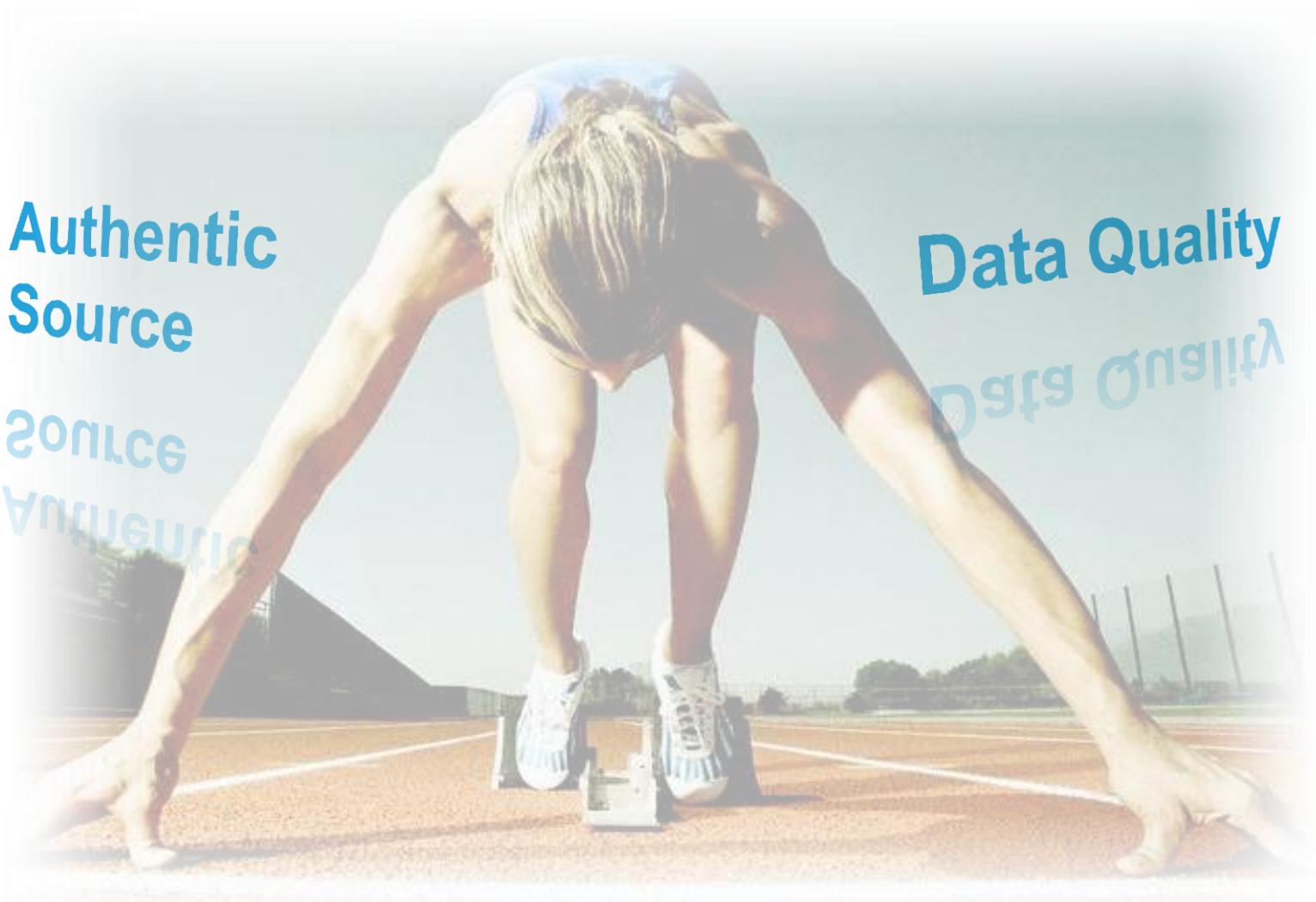
Authentic Source

Data Quality

source

Data Quality

Authentic





STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'

Come implementare l'interoperabilità semantica

Idee e tecnologie

*Una volta definiti i metadati semantici, questi sono utilizzati per la creazione delle **facette per il motore di ricerca**. Le facette rappresentano i diversi aspetti o dimensioni attraverso i quali si può descrivere una stessa risorsa.*

Qui riportate le facets utilizzate nella navigazione dei dati INPS:

PER ARGOMENTO

- ▶ Contributi (7)
- ▶ Lavoro (78)
- ▶ Pensione (23)
- ▶ Attività Ispettive (2)
- ▶ Prestazioni Assistenziali (58)

PER PERIODO

- ▶ 2001-2005 (25)
- ▶ 2006-2011 (102)
- ▶ 2012- (7)
- ▶ Ante 2000 (1)

PER FONTE

- ▶ Osservatori statistici (49)
- ▶ Rapporti (52)
- ▶ Informazioni (7)

Selezionando una **facetta**, si riducono i dataset trovati; selezionandone più di una si opererà una intersezione tra i dati, riducendo ancora di più fino a trovare esattamente quanto cercato.



STATO DELL'ARTE
 VALUTAZIONI EU
 MODELLO INPS
 DOMINIO
 TIPOLOGIA
 PREPARAZIONE
 METADATI
 LICENZA
**SEMANTICA E
 INTEROPERABILITA'**



*Ad esempio, supponiamo di voler trovare **dati statistici sulle prestazioni di disoccupazione rilasciate dall'INPS nell'anno 2008.***

I filtri dinamici consentono agilmente di effettuare un ricerca simile:

Idee e tecnologie Motore a faccette

1

PERCORSO CREATO

- ▶ Prestazioni Assistenziali ✕

SELEZIONE CATEGORIE OPEN DATA

PER ARGOMENTO

- ▶ Prestazioni Assistenziali (58)
 - ▶ ANF (6)
 - ▶ CIG (22)
 - ▶ Disoccupazione (20)
 - ▶ Integrazione sociale (2)
 - ▶ Mobilità (2)
 - ▶ Occupazione (2)

PER FONTE

- ▶ Osservatori statistici (20)
- ▶ Rapporti (38)

PER PERIODO

- ▶ 2001-2005 (16)
- ▶ 2006-2011 (58)
- ▶ Ante 2000 (1)

2

PERCORSO CREATO

- ▶ Prestazioni Assistenziali ✕
- ▶ Prestazioni assistenziali.Disoccupazione

SELEZIONE CATEGORIE OPEN DATA

PER ARGOMENTO

- ▶ Prestazioni Assistenziali (20)
 - ▶ Disoccupazione (20)

PER FONTE

- ▶ Osservatori statistici (6)
- ▶ Rapporti (14)

PER PERIODO

- ▶ 2001-2005 (6)
- ▶ 2006-2011 (20)

3

PERCORSO CREATO

- ▶ Prestazioni Assistenziali ✕
- ▶ Prestazioni assistenziali.Disoccupazione
- ▶ Osservatori statistici ✕
- ▶ 2006-2011 ✕

SELEZIONE CATEGORIE OPEN DATA

PER ARGOMENTO

- ▶ Prestazioni Assistenziali (6)
 - ▶ Disoccupazione (6)

PER FONTE

- ▶ Osservatori statistici (6)

PER PERIODO

- ▶ 2006-2011 (6)



STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E**INTEROPERABILITA'**

La modalità di ricerca **Query expansion** consente di interrogare il sistema utilizzando il linguaggio naturale, piuttosto che inserire parole chiave che l'utente potrebbe non conoscere.

In questo modo il raggiungimento delle risorse sarà notevolmente facilitato per tutte le categorie di utenti (**dai più esperti ai meno esperti**). E' infatti importante adeguare il linguaggio tecnico con cui molto spesso le risorse sono identificate ad un linguaggio semplice, più vicino alle **espressioni comuni degli utenti in ambito di ricerca**.

Idee e tecnologie

Query Expansion

Ad esempio: supponiamo che l'utente digiti la query "**centri per le cure termali**". Utilizzando la query expansion, l'utente sarà indirizzato al dato "**Cure forme vie respiratorie - elenco alberghi convenzionati per lastagione termale**". Oppure, più in generale è possibile eseguire query di ricerca su argomenti generali ed ottenere un risultato coerente con la richiesta.

Dati sulle assunzioni e sul lavoro

Osservatorio sulle aziende e i lavoratori agricoli autonomi - Aziende e lavoratori agricoli autonomi

Data: 17/01/2012 Rating ●●●●○

Aziende e lavoratori agricoli autonomi - Coltivatori diretti per stato del titolare

N° download: 5 | Tipo file: xls | Licenza: Italian Open Data License 2.0

Osservatorio sulle aziende e i lavoratori agricoli autonomi - Aziende e lavoratori agricoli autonomi

Data: 17/01/2012 Rating ●●●●○

Aziende e lavoratori agricoli autonomi - Numero aziende agricole

N° download: 4 | Tipo file: xls | Licenza: Italian Open Data License 2.0

Osservatorio sulle imprese - Imprese e occupati dipendenti del settore artigianato - Anno 2007

Data: 17/01/2012 Rating ●●●●○

N° download: 1 | Tipo file: xls | Licenza: Italian Open Data License 2.0

STATO DELL'ARTE

VALUTAZIONI EU

MODELLO INPS

DOMINIO

TIPOLOGIA

PREPARAZIONE

METADATI

LICENZA

SEMANTICA E

INTEROPERABILITA'



Semantica al servizio della PA



*La sinergia tra E-government INPS metadata e implementazione dei formati verso i LOD consente di **passare dal dato all'informazione** e apre le porte allo sviluppo di quello strato di servizi per il cittadino e per le amministrazioni che, allo stato dell'arte, sta vedendo la luce attraverso applicazioni di mashup tra diverse fonti amministrative.*

Più dati sono arricchiti, esposti e navigabili in modo adeguato, maggiore sarà il riuso che se ne potrà fare, non solo a vantaggio dei cittadini:

- **Riuso nella PA** – permette un **maggior controllo sulle attività e sui processi interni**, come sulle attività di vigilanza esterne su enti, organizzazione e territorio. Consente un **risparmio di risorse in termini di tempo e di spesa** nell'individuazione di fenomeni sociali ed economici di interesse rappresentati nei dati stessi, ma che molto spesso sono in forma di informazione tacita.
- **Riuso commerciale** – consente lo sviluppo di applicazioni in modo più veloce e integrato, favorendo **l'interesse del mondo aziendale verso i modelli economici di sviluppo OpenData** e allargando il bacino aziendale anche al mondo (sempre più in crescita) delle Startup.

