

Go Live Infrastruttura Linked Open Data

Manuale utente

Versione: 1.0

Data: 26/07/2018



UNIONE EUROPEA
Fondo Sociale Europeo
Fondo Europeo di Sviluppo Regionale



Agencia per la Coesione Territoriale



GOVERNANCE
E CAPACITÀ
ISTITUZIONALE
2014-2020



Dipartimento
dell'Amministrazione Generale,
del Personale e dei Servizi

INDICE

1	DESCRIZIONE DEL PROGETTO.....	3
1.1	OBIETTIVI DEL DOCUMENTO	3
1.2	ACRONIMI	4
1.3	SPARQL ENDPOINT DI INTERESSE.....	4
1.4	ONTOLOGIE E VOCABOLARI CONTROLLATI DI INTERESSE	4
2	NOIPA LINKED OPEN DATA	5
2.1	RISORSE TIPOLOGICHE: INTERVALLI DI ETÀ	6
2.2	RISORSE TIPOLOGICHE: ENTI.....	6
2.3	RISORSE TIPOLOGICHE: PLACE	7
2.4	RISORSE TIPOLOGICHE: UNITÀ ORGANIZZATIVA	7
2.5	RISORSE TIPOLOGICHE: ENTE SPT	8
2.6	ENTRIES: ENTRY IN PLACE	8
2.7	ENTRIES: ENTRY ANAGRAFICHE	8
2.8	ENTRIES: ENTRY INQUADRAMENTI	9
2.9	ENTRIES: ENTRY RETRIBUZIONI	10
2.10	ENTRIES: ENTRY RIDUZIONE PART-TIME	10
3	PROCESSO DI ARRICCHIMENTO SEMANTICO	11
3.1	ONTOLOGIA PRODOTTA	11
3.2	RIUTILIZZO DELLE RISORSE ESTERNE	12
3.3	CLASSIFICAZIONE DEGLI ENTI.....	13
3.4	MAPPING DELLE CATEGORIE NOIPA-DAF.....	15
4	QUERY SPARQL	17
4.1	LISTA DEI GRAFI	17
4.2	DATASET #1: ENTRYAMMINISTRATI	17
4.3	DATASET #2: ENTRYRESIDENTI.....	17
4.4	DATASET #3: ENTRYINQUADRAMENTI.....	18
4.5	DATASET #4: ENTRYRETRIBUZIONI	18
4.6	DATASET #5: ENTRYRIDUZIONEPARTTIME.....	18
5	APPENDICE – FIGURE.....	19

1 DESCRIZIONE DEL PROGETTO

I motori di ricerca hanno indubbiamente contribuito al miglioramento della fruizione del web da parte degli utenti, fornendo riscontri rapidi ed esaurienti sviluppando un semplice processo di *matching* tra parole chiave indicate. Tuttavia questo metodo implica necessariamente l'intervento umano per interpretare e combinare i risultati proposti. Questi, infatti, non sono facilmente accessibili da altri strumenti ed il significato del contenuto non recepibile alla macchina.

Il *web semantico* si pone nella prospettiva di andare oltre questo tipo di limitazione. L'obiettivo è rappresentare il web utilizzando tecniche intelligenti, affinché anche le macchine possano interpretare e correlare le informazioni: da un web di documenti collegati tra di loro ad uno di dati interconnessi. Si associa quindi ad ogni documento un'etichettatura intelligente (*metadato*), al fine di correlare una semantica ai dati ed integrarla per ricerche estremamente precise.

Vi è infatti l'esigenza di esplicitare i collegamenti tra le informazioni, anche se questi sono di differente formato e provengono da fonti eterogenee. Per realizzare il collegamento tra i dati si ricorre alla definizione di un link diretto che colleghi le informazioni accessibili a tutti (da qui il concetto di *Linked Open Data*) che si riferiscono allo stesso argomento, oppure che risultano simili.

Nell'ambito del portale NoiPA sono stati aperti dati in formato Linked Open Data *5 stelle*.

1.1 OBIETTIVI DEL DOCUMENTO

La classificazione è definita come LOD 5 stelle se i dati aperti:

1. sono disponibili su Web con una licenza aperta;
2. sono in formato strutturato;
3. non sono in un formato proprietario;
4. fanno uso di identificatori univoci;
5. includono link ad altre sorgenti dati.

I dati del portale non sono isolati, ma aperti come aggregati (economici ed anagrafici) e riguardano gli impiegati della Pubblica Amministrazione.

La strategia di apertura dei dati si focalizza su interoperabilità e disponibilità.

Il presente documento intende fornire le istruzioni per la gestione operativa dell'intera infrastruttura Linked Open Data.

1.2 ACRONIMI

Acronimo/Termine	Definizione
LOD	Linked Open Data
RDF	Resource Description Framework
SKOS	Simple Knowledge Organization System
SPARQL	acronimo ricorsivo di SPARQL Protocol and RDF Query Language, si pronuncia /'spa:kl/ come la parola inglese sparkle, "scintillare"
OWL	Web Ontology Language
RDFS	Resource Description Framework Schema

1.3 SPARQL ENDPOINT DI INTERESSE

Organizzazione	Link
ISPRA	http://dati.isprambiente.it/sparql
ISTAT	http://datiopen.istat.it/sparql
SPC-DATA	http://spcdata.digitpa.gov.it:8899/sparql

1.4 ONTOLOGIE E VOCABOLARI CONTROLLATI DI INTERESSE

Nome	Link
ISPRA	http://dati.isprambiente.it/ontologie-e-thesauri/ispra-core-ontology
ISTAT	http://datiopen.istat.it/ontologie.php
SPC-DATA	http://spcdata.digitpa.gov.it/data/ontology.owl
Vocabolari controllati DAF	https://github.com/italia/daf-ontologie-vocabolari-controllati

2 NOIPA LINKED OPEN DATA

Le triple che popolano il database sono raccolte nei seguenti grafi:

1. risorse tipologiche (anagrafiche) (<https://sparql-noipa.mef.gov.it/tipologiche>);
2. dati (entries) (<https://sparql-noipa.mef.gov.it/entries-201806>);
3. dati (entries) (<https://sparql-noipa.mef.gov.it/entries-201807>);
4. ...

I grafi <https://sparql-noipa.mef.gov.it/entries-201807> e <https://sparql-noipa.mef.gov.it/entries-201806> raccolgono i dati e metadati prodotti con cadenza mensile (ad esempio 06/2018, 07/2018), mentre il grafo delle risorse anagrafiche viene aggiornato (anch'esso con frequenza mensile)

1) Il grafo delle risorse tipologiche contiene le risorse delle seguenti classi:

- <https://sparql-noipa.mef.gov.it/ontology/AgeRange>
- <https://sparql-noipa.mef.gov.it/ontology/Ente>
- <https://sparql-noipa.mef.gov.it/ontology/EnteSPT>
- <https://sparql-noipa.mef.gov.it/ontology/Place>
 - <https://sparql-noipa.mef.gov.it/ontology/StatoEstero>
 - <https://sparql-noipa.mef.gov.it/ontology/Comune>
- <https://sparql-noipa.mef.gov.it/ontology/UnitaOrganizzativa>

2) Il grafo dei dati (entries) contiene le risorse delle seguenti classi:

- <https://sparql-noipa.mef.gov.it/ontology/EntryInPlace>
- <https://sparql-noipa.mef.gov.it/ontology/EntryAnagrafiche>
 - <https://sparql-noipa.mef.gov.it/ontology/EntryAmministrati>
 - <https://sparql-noipa.mef.gov.it/ontology/EntryResidenti>
- <https://sparql-noipa.mef.gov.it/ontology/EntryInquadramenti>
- <https://sparql-noipa.mef.gov.it/ontology/EntryRetribuzioni>
- <https://sparql-noipa.mef.gov.it/ontology/EntryRiduzionePartTime>

Nel seguito si farà riferimento alle classi, alle risorse e alle property utilizzando i seguenti prefissi:

npont	https://sparql-noipa.mef.gov.it/ontology/
npres	https://sparql-noipa.mef.gov.it/resource/

2.1 RISORSE TIPOLOGICHE: INTERVALLI DI ETÀ

Le risorse AgeRange descrivono gli intervalli di età presi in considerazione.

Nell'esempio seguente, l'intervallo compreso tra 25 e 34 anni è così identificato dalla risorsa:

```
https://sparql-noipa.mef.gov.it/resources/AgeRange/2

<resources/AgeRange/2>
  a                               npont:AgeRange ;
  rdfs:label                       "La seguente risorsa descrive un intervallo di età" ;
  npont:maxAge                     "34"^^xsd:int ;
  npont:minAge                     "25"^^xsd:int .
```

Le property *npont:maxAge* e *npont:minAge* definiscono l'età massima e l'età minima dell'intervallo: in questo caso si rappresenta l'intervallo di età tra 25 e 34 anni.

2.2 RISORSE TIPOLOGICHE: ENTI

L'ente di appartenenza dell'unità organizzativa è una risorsa collegata mediante la property *npont:inEnte*, anche le singole Entry possono far riferimento agli Enti mediante la property *npont:inEnte*.

Per ciascun ente abbiamo una risorsa:

```
https://sparql-noipa.mef.gov.it/resources/Ente/3

<resources/Ente/3> a           npont:Ente ;
  rdfs:comment                 "La seguente risorsa descrive l'ente con codice:3" ;
  rdfs:label                   "MINISTERO DELLA DIFESA" ;
  npont:codiceEnte             "3" ;
  npont:codiceNatura           "C01" ;
  npont:hasLegalStatus         legalstatus:2220 .
  owl:sameAs                 <resources/EnteSPT/14>
```

La classificazione degli enti segue la nomenclatura di NoiPA (*npont:codiceNatura*), ma è racciardata anche al tesoro delle organizzazioni del Data & Analytics framework (*npont:hasLegalStatus*), dove – ad esempio – *legalstatus:2220* (<http://w3id.org/italia/controlled-vocabulary/classifications-for-organizations/legal-status/2220>) è uno *skos:Concept* (per approfondimento vedi sezione 3.2 e seguenti).

2.3 RISORSE TIPOLOGICHE: PLACE

La classe *Place* è costituita da due sottoclassi, che identificano i Comuni e gli Stati esteri.

Il riconoscimento avviene univocamente dal codice catastale, un esempio di URI è riportato di seguito:

```

Comune di Roma: https://sparql-noipa.mef.gov.it/resources/Place/H501

<resources/Place/H501>
  a                               npont:Comune ;
  rdfs:label                       "ROMA" ;
  geonames:locatedIn              <http://spcdata.digitpa.gov.it/Provincia/RM> ;
  owl:sameAs
    <http://spcdata.digitpa.gov.it/Comune/H501>,
    <http://datiopen.istat.it/odi/risorsa/territorio/comuni/Roma_58091>,
    <http://dati.isprambiente.it/id/place/58091>;
  npont:hasCodiceCatastale        "H501" ;
  npont:hasCodiceIstat            "058091" .

Guyana e Guyana francese: https://sparql-noipa.mef.gov.it/resources/Place/Z607

<resources/Place/Z607>
  a                               npont:StatoEstero
  rdfs:label                       "Guyana"^^<http://www.w3.org/2001/XMLSchema#string>
  npont:hasCodiceCatastale        "Z607"^^<http://www.w3.org/2001/XMLSchema#string>

```

I Comuni sono accompagnati da una descrizione, dalla provincia di appartenenza e dal codice catastale (*rdfs:label*, *geonames:locatedIn* e *npont:hasCodiceCatastale*). Per tutti i Comuni ancora in attività (secondo il registro Istat) sono disponibili altri metadati come il codice Istat, il link alle risorse Istat, Spc-data, ed Ispra. I link agli SPARQL Endpoint di ISPRA, ISTAT e SPC-DATA sono disponibili nella sezione 1.3.

La presenza di questi collegamenti permette di arricchire i dataset con risorse recuperabili da diversi endpoint (per dettagli far riferimento a “Query federate e costruito service” <https://www.w3.org/TR/sparql11-federated-query>).

2.4 RISORSE TIPOLOGICHE: UNITÀ ORGANIZZATIVA

Le Unità Organizzative sono identificate con il codice Ente di appartenenza ed il codice dell’Unità Organizzativa, pertanto una risorsa avrà la seguente URI:

```

https://sparql-noipa.mef.gov.it/resources/UnitaOrganizzativa/3_556

<resources/UnitaOrganizzativa/3_556>
  a                               npont:UnitaOrganizzativa
  rdfs:label                       "A.I.D. FIRENZE"^^xsd:string
  npont:inEnte                     <resources/Ente/3>
  npont:codiceUnitaOrganizzativa  "556"^^xsd:string

```

Il numero “3” identifica il codice dell’ente, mentre “556” è il codice dell’unità organizzativa.

2.5 RISORSE TIPOLOGICHE: ENTE SPT

Concettualmente questi Enti sono gli stessi descritti dalle risorse "Ente", infatti è stato realizzato un collegamento di tipo *owl:sameAs* per collegare le risorse che descrivono lo stesso Ente (risorse di tipo *Ente* ed *EnteSPT*). La motivazione per cui esiste un'altra risorsa Ente è dovuta al fatto che i dati provengono da database che utilizzano codici diversi per descrivere lo stesso Ente.

2.6 ENTRIES: ENTRY IN PLACE

Tutte le risorse Entry hanno delle property finalizzate a descrivere la data di produzione della risorsa: guarda *npont:temporalID* e *npont:hasDate*. Le *EntryInPlace* sono risorse generiche che potrebbero far riferimento al luogo di appartenenza di qualche Unità Organizzativa o al luogo di residenza di un raggruppamento di amministrati. Questa classe viene estesa e specializzata da tutte le altre classi di tipo Entry.

2.7 ENTRIES: ENTRY ANAGRAFICHE

Le EntryAnagrafiche sono EntryInPlace che hanno il vincolo di riportare informazioni anagrafiche riguardo raggruppamenti di amministrati, come l'intervallo di età ed il genere sessuale.

Sottoclassi di EntryAnagrafiche sono EntryAmministrati ed EntryResidenti. Di seguito vengono riportate le descrizioni di ciascuna sottoclasse mediante esempi di risorse nella sintassi Turtle:

Turtle ("Terse RDF Triple Language") è un formato ideato per esprimere dati di tipo RDF con una sintassi adatta a SPARQL. Secondo le convenzioni RDF, le informazioni sono rappresentate per mezzo di "triple", ciascuna delle quali consiste di un soggetto, un predicato e un oggetto. Ognuno di questi elementi è espresso come URI. Nella sintassi turtle si compongono le triple utilizzando le URI di ciascuna risorsa in modo molto intuitivo, ad esempio:

```
@prefix npres: <https://noipa.mef.gov.it/resources/>
@prefix npont: <https://noipa.mef.gov.it/ontology/>
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
npres:Risorsa1
  rdfs:label "descrizione human-readable" ;
  npont:Property npres:Risorsa2 .
```

L'uri del soggetto è *npres:Risorsa1*, la property (o predicato) *rdfs:label* ha come oggetto una stringa, mentre la property *npont:Property* ha come oggetto la risorsa *npres:Risorsa2*. I prefissi *rdfs*, *npont*, e *npres* definiti all'inizio con il costrutto *prefix* ci permettono di non ripetere la radice dell'URI ogni volta nella definizione della tripla.

I prefissi utilizzati nei seguenti esempi sono:

```
@prefix npont: <https://noipa.mef.gov.it/ontology/>
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
@prefix geonames: <http://www.geonames.org/ontology#>
@prefix owl: <http://www.w3.org/2002/07/owl#>
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>
@prefix npres: <https://noipa.mef.gov.it/resources/>
```

Le risorse EntryAmministrati consentono di osservare l'andamento della distribuzione degli amministrati per Comune dell'Unità Organizzativa in cui prestano servizio, per fascia di età e per sesso. Più precisamente, le risorse contengono i dati relativi a: gli amministrati delle Amministrazioni in convenzione (esclusa Sanità) per cui sono state elaborate; emissioni nel mese di riferimento; aggregati su base territoriale, per Comune di residenza di ciascuna partita.

```
<https://sparql-noipa.mef.gov.it/resources/EntryAmministrati/EntryAmministrati_9572_2018-05>
a npon:EntryAmministrati ;
npon:hasGender <http://w3id.org/italia/controlled-vocabulary/classifications-for-people/gender/M> ;
npon:hasDate "2018-05-03"^^xsd:date ;
npon:hasAgeRange <resources/AgeRange/2> ;
npon:hasPlace <resources/Place/A052> ;
npon:howMany "3"^^xsd:int ;
npon:temporalID "201807"^^xsd:int ;
npon:inUO <resources/UO/4> .
```

Le risorse EntryResidenti consentono di osservare l'andamento della distribuzione degli amministrati per Comune di residenza, oltre alla loro variazione per fascia di età e per sesso. Anche in questo caso, le risorse contengono i dati relativi a: gli amministrati delle Amministrazioni in convenzione (esclusa Sanità) per cui sono state elaborate; emissioni nel mese di riferimento; aggregati su base territoriale, per Comune di residenza di ciascuna partita.

```
<https://sparql-noipa.mef.gov.it/resources/EntryResidenti/3_2018-06>
a npon:EntryResidenti ;
npon:hasGender <http://w3id.org/italia/controlled-vocabulary/classifications-for-people/gender/M> ;
npon:hasDate "2018-05-03"^^xsd:date ;
npon:hasAgeRange <resources/AgeRange/2> ;
npon:hasPlace <resources/Place/A052> ;
npon:howMany "3"^^xsd:int ;
npon:temporalID "201807"^^xsd:int ;
npon:inEnte <resources/Ente/4> .
```

2.8 ENTRIES: ENTRY INQUADRAMENTI

Le risorse di questo tipo consentono di osservare l'andamento della distribuzione degli amministrati per qualifica, distinti per Comune dell'Unità Organizzativa. Più precisamente, le risorse contengono i dati relativi a: gli amministrati delle Amministrazioni in convenzione (esclusa Sanità) per cui sono state elaborate; emissioni nel mese di riferimento; aggregati su base territoriale, per Comune di residenza di ciascuna partita.

```
<https://sparql-noipa.mef.gov.it/resources/EntryInquadramenti/3_2018-06>
a npon:EntryInquadramenti ;
npon:temporalID "201807"^^xsd:int ;
npon:hasDate "2018-05-03"^^xsd:date ;
npon:hasPlace <resources/Place/H501> ;
npon:howMany "1"^^xsd:int ;
npon:inUO <resources/UnitaOrganizzativa/2020_701054> ;
npon:inEnte <resources/Ente/2020> ;
npon:compartoIGOP "COMPARTO FORZE DI POLIZIA DELLO STATO"^^xsd:string ;
npon:qualificaContrattuale "COLLABORATORE"^^xsd:string .
```

2.9 ENTRIES: ENTRY RETRIBUZIONI

Le risorse di questo tipo consentono di osservare l'andamento della spesa stipendiale delle Amministrazioni di NoiPA, distinta nelle diverse componenti e per Comune dell'Unità Organizzativa. Nel dettaglio, le risorse contengono i dati relativi ai seguenti costi: emissioni effettuate nel mese di riferimento; Amministrazioni in bilancio e fuori bilancio per cui sono effettuati i pagamenti telematici; restanti Amministrazioni ed Enti per cui sono effettuate le elaborazioni CBI; spesa aggregata per Unità Organizzativa su base territoriale, a partire dal Comune dell'Unità Organizzativa di ciascuna partita.

```
<https://sparql-noipa.mef.gov.it/resources/EntryAmministrati/EntryRetribuzioni_7245_2018-07>
a
  npont:EntryRetribuzioni ;
  rdfs:label "EntryRetribuzioni 7245_2018-07"^^xsd:string ;
  npont:temporalID "201807"^^xsd:int ;
  npont:hasPlace <resources/Place/F158> ;
  npont:howMany "155"^^xsd:int ;
  npont:hasDate 2018-07-01 ;
  npont:addizionaleComune 3486.05 ;
  npont:contributiDipendente 40455.8 ;
  npont:tipologiaEnte "BILANCIO"^^xsd:string ;
  npont:inEnte <resources/EnteSPT/26> ;
  npont:tipologiaPagamento "CBI NO"^^xsd:string ;
  npont:altreRitenute 11136.5 ;
  npont:irpef 60555.2 ;
  npont:lordoComplessivo 379904.0 ;
  npont:cessioneQuinto 0.0 ;
  npont:irap 32262.5 ;
  npont:netto 256716.0 ;
  npont:addizionaleRegione 7554.0 ;
  npont:contributiDatore 114469.0 .
```

2.10 ENTRIES: ENTRY RIDUZIONE PART-TIME

Le risorse di questo tipo consentono di osservare l'andamento dei risparmi per riduzione part-time delle Amministrazione di NoiPA, distinta per percentuali di part-time e per Comune dell'Unità Organizzativa. Specificatamente, le risorse contengono i dati relativi ai risparmi di gestione: risultanti dalle elaborazioni dalle emissioni effettuate nel mese di riferimento; Amministrazioni in bilancio e fuori bilancio per cui sono effettuati i pagamenti telematici; restanti Amministrazioni ed Enti per cui sono effettuate le elaborazioni CBI; aggregati per Unità Organizzativa su base territoriale, a partire dal Comune dell'Unità Organizzativa di ciascuna partita.

```
<https://sparql-noipa.mef.gov.it/resources/EntryAmministrati/EntryRiduzionePartTime_6520_2018-07>
a
  npont:EntryRiduzionePartTime ;
  npont:percentualePartTime 50.0 ;
  npont:temporalID "201807"^^xsd:int ;
  npont:hasPlace <resources/Place/H509> ;
  npont:inEnte <resources/EnteSPT/26> ;
  npont:hasDate 2018-07-01 ;
  npont:tipologiaEnte "BILANCIO"^^xsd:string ;
  npont:tipologiaPagamento "CBI NO"^^xsd:string ;
  npont:codiceSpesa "7984"^^xsd:string ;
  npont:ufficio "ROMANO DI LOMBARDIA E.FERMI"^^xsd:string ;
  npont:risparmioPartTime 2206.86 .
```

3 PROCESSO DI ARRICCHIMENTO SEMANTICO

Parlando di open-data, spesso si pone l'accento sulla disponibilità ed accessibilità dei dati. Oltre a queste caratteristiche fondamentali è opportuno sottolineare anche che i dati devono essere forniti in un formato tale da garantirne un più facile riutilizzo e redistribuzione. Riepilogando:

- **Disponibilità e accesso:** i dati devono essere disponibili nel loro complesso, per un prezzo non superiore ad un ragionevole costo di riproduzione, preferibilmente mediante scaricamento da Internet. I dati devono essere disponibili in un formato utile e modificabile.
- **Riutilizzo e redistribuzione:** i dati devono essere forniti a condizioni tali da permetterne il riutilizzo e la redistribuzione. Ciò comprende la possibilità di combinarli con altre basi di dati.

Al fine di garantire un alto livello di interoperabilità, è stato progettato uno *Schema OWL* con alcune accortezze, quali l'assegnazione di metadati XSD Schema alle costanti, il riutilizzo di tesauri e standard di classificazione e normalizzazione implementati in SKOS, includendo infine alcuni collegamenti a risorse esterne come nel caso dei Comuni. Il vantaggio di utilizzare SKOS è dato dal fatto che tutti i dataset che utilizzano questi tesauri saranno interoperabili. Pertanto, sarà possibile operare interrogazioni e regole di inferenza a cavallo di più dataset, aumentando considerevolmente il potenziale delle analisi.

3.1 ONTOLOGIA PRODOTTA

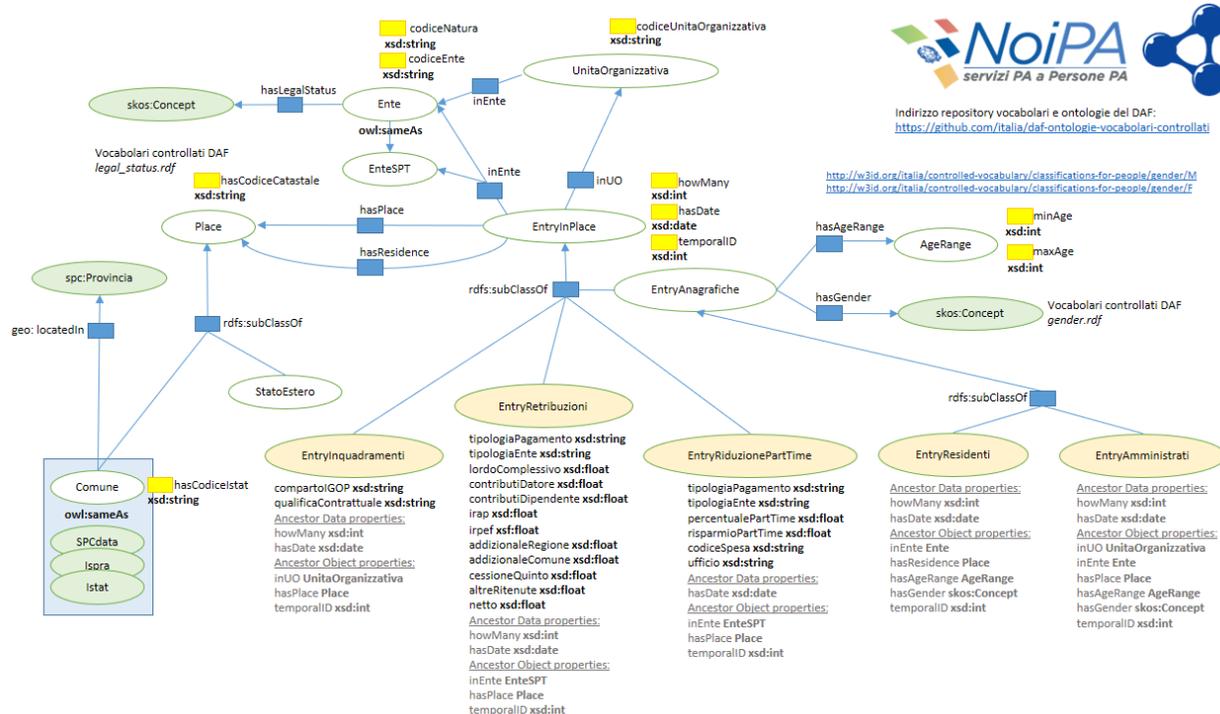


Figura 1 – Dettaglio dello Schema OWL

Le classi Entry sono quelle che raccolgono i dati già presenti nei dataset open-data del portale NoiPA, le altre classi (tipologiche) sono anch'esse provenienti dalle basi di dati di NoiPA, ma arricchite e collegate con delle risorse esterne. Nella fattispecie, le risorse riutilizzate sono tesauri, vocabolari controllati e, per quanto riguarda i dati territoriali, è stato realizzato il collegamento alle risorse gestite da SPC-data, Istat ed Ispra.

SKOS è una famiglia di linguaggi formali creata per rappresentare glossari, classificazioni, tassonomie e qualsiasi tipo di vocabolario strutturato. SKOS si basa su RDF e RDFS e il suo obiettivo principale è di consentire una facile pubblicazione di vocabolari strutturati per il Semantic Web. SKOS è progettato come una famiglia di linguaggi modulari ed estensibili. Il suo uso e la sua implementazione dovrebbe essere il più semplice possibile.

3.2 RIUTILIZZO DELLE RISORSE ESTERNE

Le risorse esterne riutilizzate sono finalizzate all'arricchimento e *metadattazione* delle risorse che descrivono informazioni tipologiche quali Comuni, Enti, Unità organizzative ed informazioni anagrafiche degli amministratori:

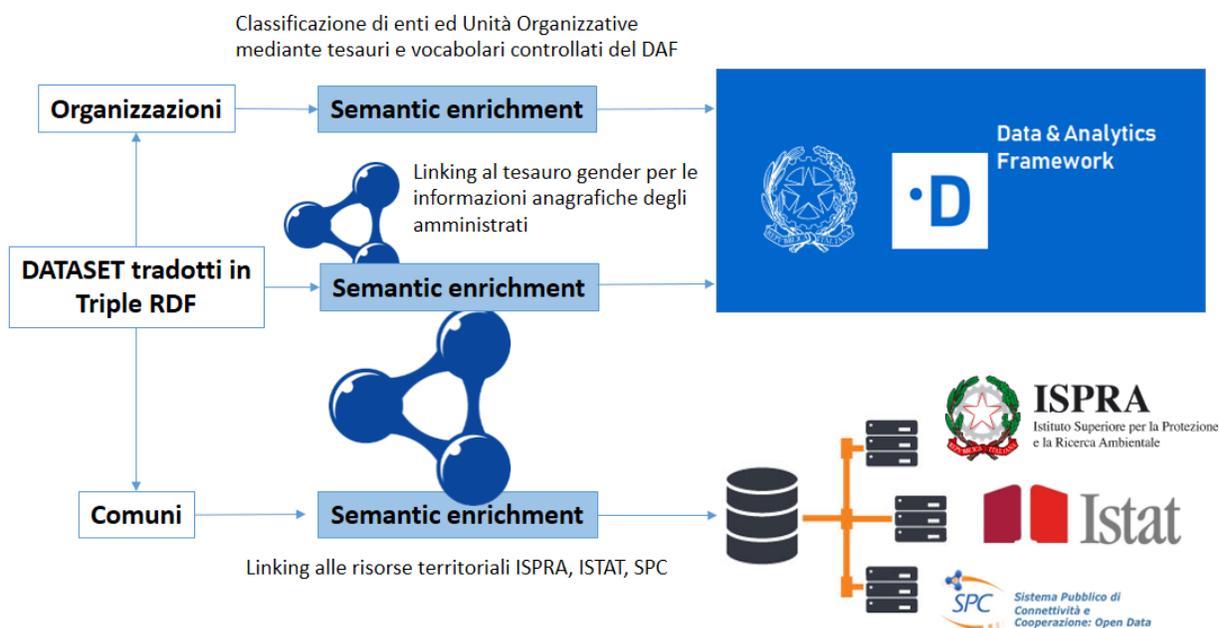


Figura 2 – Semantic Enrichment

Il Data & Analytics Framework (DAF) ha l'obiettivo di migliorare e semplificare l'interoperabilità e lo scambio dei dati pubblici tra PA, standardizzare e promuovere la diffusione degli Open Data, ottimizzare i processi di analisi dati e generazione di sapere. L'idea è quella di aprire il mondo della Pubblica Amministrazione ai benefici offerti dalle moderne piattaforme per la gestione e l'analisi dei Big Data. Grazie alle tecnologie Big Data sarà possibile operare lungo tre direttrici principali:

1. *Amplificare sensibilmente il valore del patrimonio informativo della PA mediante la predisposizione e l'uso di strumenti di analisi finalizzati a sintetizzare conoscenza per i decision makers, e diffusione di informazioni verso cittadini e imprese.*
2. *Ottimizzare lo scambio dei dati tra le PA e la diffusione degli Open Data, minimizzando i costi transattivi per l'accesso e l'utilizzo dei dati.*
3. *Favorire l'analisi e la gestione dei dati da parte di team di data scientist in seno alle PA, al fine di migliorare la conoscenza dei fenomeni descritti dai dati e sviluppare applicazioni "intelligenti", nonché condurre iniziative utili a promuovere attività di ricerca scientifica su tematiche applicative di interesse per la PA.*

<https://teamdigitale.governo.it/it/projects/daf.htm>

3.3 CLASSIFICAZIONE DEGLI ENTI

Di seguito i concetti contenuti nel tesoro SKOS *legal-status* del DAF:

ID LEGAL-STATUS	DESCRIZIONE
legal-status/2240	Archivio notarile
legal-status/2100	Organo costituzionale o a rilevanza costituzionale
legal-status/2220	Ministero
legal-status/2230	Agenzia dello Stato
legal-status/2300	Autorità indipendenti
legal-status/2410	Regione
legal-status/2420	Provincia
legal-status/2430	Comune
legal-status/2450	Unione di Comuni
legal-status/25	Azienda o ente del servizio sanitario nazionale
legal-status/2620	Università pubblica
legal-status/2711	Istituto o ente pubblico di ricerca
legal-status/2751	Ente parco
legal-status/2753	Ente di sviluppo agricolo regionale o di altro ente locale
legal-status/2755	Ente ambientale regionale
legal-status/2756	Ente per la ricerca e per l'aggiornamento educativo
legal-status/2790	Altro ente pubblico non economico nazionale
legal-status/1	Forme disciplinate dal diritto privato
legal-status/11	Imprenditore individuale, libero professionista e lavoratore autonomo
legal-status/1110	Imprenditore individuale agricolo
legal-status/1120	Imprenditore individuale non agricolo
legal-status/1130	Libero professionista
legal-status/1140	Lavoratore autonomo
legal-status/12	Società di persone
legal-status/1210	Società semplice

legal-status/1220	Società in nome collettivo
legal-status/1230	Società in accomandita semplice
legal-status/1240	Studio associato e Società di professionisti
legal-status/1250	Società di fatto o irregolare, comunione ereditaria
legal-status/13	Società di capitali
legal-status/1310	Società per azioni
legal-status/1320	Società a responsabilità limitata
legal-status/1330	Società a responsabilità limitata con un unico socio
legal-status/1340	Società in accomandita per azioni
legal-status/14	Società cooperativa
legal-status/1410	Società cooperativa a mutualità prevalente
legal-status/1420	Società cooperativa diversa
legal-status/1430	Società cooperativa sociale
legal-status/1440	Società di mutua assicurazione
legal-status/15	Consorzio di diritto privato ed altre forme di cooperazione fra imprese
legal-status/1510	Consorzio di diritto privato
legal-status/1520	Società consortile
legal-status/1530	Associazione o raggruppamento temporaneo di imprese
legal-status/1540	Gruppo europeo di interesse economico
legal-status/16	Ente pubblico economico, azienda speciale e azienda pubblica di servizi
legal-status/1610	Ente pubblico economico
legal-status/1620	Azienda speciale ai sensi del t.u. 267/2000
legal-status/1630	Azienda pubblica di servizi alle persone ai sensi del d.lgs n. 207/2001
legal-status/17	Ente privato con personalità giuridica
legal-status/1710	Associazione riconosciuta
legal-status/1720	Fondazione (esclusa fondazione bancaria)
legal-status/1730	Fondazione bancaria
legal-status/1740	Ente ecclesiastico
legal-status/1750	Società di mutuo soccorso
legal-status/1790	Altra forma di ente privato con personalità giuridica
legal-status/18	Ente privato senza personalità giuridica
legal-status/1810	Associazione non riconosciuta
legal-status/1820	Comitato
legal-status/1830	Condominio
legal-status/1890	Altra forma di ente privato senza personalità giuridica
legal-status/19	Impresa o ente privato costituito all'estero non altrimenti classificabile che svolge una attività economica in Italia
legal-status/1900	Impresa o ente privato costituito all'estero non altrimenti classificabile che svolge una attività economica in Italia
legal-status/2	Forme disciplinate dal Diritto pubblico
legal-status/21	Organo costituzionale o a rilevanza costituzionale

legal-status/22	Amministrazione dello Stato
legal-status/2210	Presidenza del Consiglio
legal-status/23	Autorità indipendente
legal-status/24	Regione e autonomia locale
legal-status/2440	Comunità montana o isolana
legal-status/2460	Città metropolitana
legal-status/2500	Azienda o ente del servizio sanitario nazionale
legal-status/26	Istituto, scuola e Università pubblica
legal-status/2610	Istituto e scuola pubblica di ogni ordine e grado
legal-status/27	Ente pubblico non economico
legal-status/2712	Istituto pubblico di assistenza e beneficenza
legal-status/2720	Camera di Commercio
legal-status/2730	Ordine e Collegio professionale
legal-status/2740	Consorzio di Diritto pubblico
legal-status/2752	Ente o Autorità portuale
legal-status/2754	Ente per il turismo

3.4 MAPPING DELLE CATEGORIE NoiPA–DAF

I codici “natura Ente” sono stati raccordati al tesoro SKOS legal-status, così facendo la classificazione degli Enti adottata da NoiPA sarà collegata a quella del DAF (<https://github.com/italia/daf-ontologie-vocabolari-controllati>).

Di seguito viene riportata una tabella riepilogativa:

CODICE NATURA	ID SKOS CONCEPT	DESCRIZIONE SKOS CONCEPT
X00	2790	Altro ente pubblico non economico nazionale
C01	2220	Ministero
C02	2100	Altro ente pubblico non economico nazionale
C03	2430	Comune
C04	2230	Agenzia dello Stato
C05	2240	Archivio notarile
C06	2300	Autorità indipendenti
C08	2756	Ente per la ricerca e per l’aggiornamento educativo
C09	2711	Istituto o ente pubblico di ricerca
C11	2220	Ministero

C12	2230	Agenzia dello Stato
C13	2230	Agenzia dello Stato
L01	2410	Regione
L02	2420	Regione
L03	2430	Comune
L05	2450	Unione di Comuni
L10	2753	Ente di sviluppo agricolo regionale o di altro ente locale
L11	2753	Ente di sviluppo agricolo regionale o di altro ente locale
L13	2755	Ente ambientale regionale
L15	2711	Istituto o ente pubblico di ricerca
L16	25	Azienda o ente del servizio sanitario regionale
L21	2755	Ente ambientale regionale
L27	2751	Ente parco
L29	2620	Università pubblica
EP	2230	Agenzia dello Stato

4 QUERY SPARQL

Le triple RDF sono salvate in OpenLink Virtuoso, un Triplestore (Database per grafi RDF), il quale espone lo Sparql Endpoint, un'interfaccia di interrogazione basata sul protocollo SPARQL.

SPARQL è elevato a linguaggio di interrogazione per dataset RDF e documentato dal W3C: <https://www.w3.org/TR/sparql11-overview/>

Vengono fornite di seguito alcune query SPARQL.

4.1 LISTA DEI GRAFI

```
SELECT DISTINCT ?g
WHERE {
  GRAPH ?g {?s ?p ?o}
}
ORDER BY ?g
```

4.2 DATASET #1: ENTRYAMMINISTRATI

```
SELECT ?date ?ti ?locazione ?unita_organizzativa ?etaMin ?etaMax ?genere ?numero
WHERE{
  ?entry npont:temporalID ?ti . FILTER( ?ti=201807 )
  ?entry a npont:EntryAmministrati .
  ?entry npont:hasDate ?date .
  ?entry npont:howMany ?numero .
  ?entry npont:hasPlace ?p .
  ?p rdfs:label ?locazione .
  ?entry npont:hasGender ?gender .
  ?gender skos:notation ?genere .
  ?entry npont:inUO ?uo .
  ?uo rdfs:label ?unita_organizzativa .
  ?entry npont:hasAgeRange ?ar .
  OPTIONAL{ ?ar npont:minAge ?etaMin . }
  OPTIONAL{ ?ar npont:maxAge ?etaMax . }
}
```

4.3 DATASET #2: ENTRYRESIDENTI

```
SELECT ?date ?ti ?locazione ?etaMin ?etaMax ?genere ?ente_ ?numero
WHERE{
  ?entry npont:temporalID ?ti . FILTER( ?ti=201807 )
  ?entry a npont:EntryResidenti .
  ?entry npont:hasDate ?date .
  ?entry npont:howMany ?numero .
  ?entry npont:hasResidence ?p .
  ?p rdfs:label ?locazione .
  ?entry npont:hasGender ?gender .
  ?gender skos:notation ?genere .
  ?entry npont:inEnte ?ente .
  ?ente rdfs:label ?ente_ .
  ?entry npont:hasAgeRange ?ar .
  OPTIONAL{ ?ar npont:minAge ?etaMin . }
  OPTIONAL{ ?ar npont:maxAge ?etaMax . }
}
```

4.4 DATASET #3: ENTRYINQUADRAMENTI

```
SELECT ?date ?ti ?locazione ?unita_organizzativa ?comparto_igop ?qualifica_contrattuale ?numero
WHERE{
  ?entry npont:temporalID ?ti . FILTER( ?ti=201807 )
  ?entry a npont:EntryInquadramenti .
  ?entry npont:hasDate ?date .
  ?entry npont:howMany ?numero .
  ?entry npont:hasPlace ?p .
  ?p rdfs:label ?locazione .
  ?entry npont:inUO ?uo .
  ?uo rdfs:label ?unita_organizzativa .
  ?entry npont:compartoIGOP ?comparto_igop .
  ?entry npont:qualificaContrattuale ?qualifica_contrattuale .
}
```

4.5 DATASET #4: ENTRYRETRIBUZIONI

```
SELECT ?date ?ti ?addizionale_comune_ ?addizionale_regione_ ?altre_ritenute_ ?cessione_quinto_
?contributi_datore_ ?contributi_dipendente_ ?irap_ ?lordo_complessivo_ ?numero ?ente_ ?locazione
?tipologia_pagamento ?tipologia_ente ?irpef_ ?netto_
WHERE{
  ?entry npont:temporalID ?ti . FILTER( ?ti=201807 )
  ?entry a npont:EntryRetribuzioni .
  ?entry npont:tipologiaPagamento ?tipologia_pagamento .
  ?entry npont:tipologiaEnte ?tipologia_ente .
  ?entry npont:lordoComplessivo ?lordo_complessivo_ .
  ?entry npont:contributiDatore ?contributi_datore_ .
  ?entry npont:contributiDipendente ?contributi_dipendente_ .
  ?entry npont:irap ?irap_ .
  ?entry npont:irpef ?irpef_ .
  ?entry npont:addizionaleComune ?addizionale_comune_ .
  ?entry npont:addizionaleRegione ?addizionale_regione_ .
  ?entry npont:altreRitenute ?altre_ritenute_ .
  ?entry npont:cessioneQuinto ?cessione_quinto_ .
  ?entry npont:hasDate ?date .
  ?entry npont:howMany ?numero .
  ?entry npont:inEnte ?ente .
  ?ente rdfs:label ?ente_ .
  ?entry npont:hasPlace ?place .
  ?place rdfs:label ?locazione .
  ?entry npont:netto ?netto_ .
}
```

4.6 DATASET #5: ENTRYRIDUZIONEPARTTIME

```
SELECT ?date ?ti ?locazione ?ente_ ?codice_spesa_ ?tipologia_ente_ ?tipologia_pagamento_
?risparmio_ ?percentuale_ ?ufficio_
WHERE{
  ?entry npont:temporalID ?ti . FILTER( ?ti=201807 )
  ?entry a npont:EntryRiduzionePartTime .
  ?entry npont:hasDate ?date .
  ?entry npont:hasPlace ?place .
  ?place rdfs:label ?locazione .
  ?entry npont:inEnte ?ente .
  ?ente rdfs:label ?ente_ .
  ?entry npont:codiceSpesa ?codice_spesa_ .
  ?entry npont:percentualePartTime ?percentuale_ .
  ?entry npont:risparmioPartTime ?risparmio_ .
  ?entry npont:tipologiaEnte ?tipologia_ente_ .
  ?entry npont:tipologiaPagamento ?tipologia_pagamento_ .
  ?entry npont:ufficio ?ufficio_ .
}
```

5 APPENDICE – FIGURE

Figura 1 – Dettaglio dello Schema OWL	11
Figura 2 – Semantic Enrichment	12