D4.1.1 – Linked Open Data – Casi d'uso, PAL e dati

Informazioni deliverable

Titolo del deliverable	D4.1.1 – Linked Open Data – Casi d'uso, PAL e dati
Sigla di classificazione del documento	D4.1.1
Responsabile del Deliverable	*CNR*
Obiettivo Realizzativo	*OR4: Linked Open Data e Community Trust nel Cloud PAL*
Attività relativa	*AR4.1 – Estrazione e arricchimento di dati di un campione di PAL secondo il paradigma dei Linked Open Data*
Autori	*CNR (Aldo Gangemi, Andrea G. Nuzzolese, Silvio Peroni, Valentina Presutti, Daria Spampinato)*
Approvazione	
Autorizzazione	

Abstract

Il presente Rapporto Tecnico descrive i metodi e i passi iniziali per definire i casi d'uso e lo sviluppo di alcuni campioni di schemi e dati, nell'ambito dell'attività di estrazione e arricchimento dei dati di alcune Pubbliche Amministrazioni Locali.

I metodi sono basati sugli standard del W3C, su buone pratiche internazionali, sulle linee-guida emanate dall'Agenzia per l'Italia Digitale e su esperienze vicine già avute dai partecipanti a questa attività, in particolare lo sviluppo di data.cnr.it e del Semantic Scout, nonché dei linked open data dell'Indice delle Pubbliche Amministrazioni.

Due ontologie sono state sviluppate a partire da documenti di specifica sui servizi relativi agli asili nido e all'assistenza domiciliare integrata.

I contenuti del documento sono da intendersi "confidenziali" di proprietà esclusiva delle parti coinvolte nel progetto "PRISMA – PiattafoRme cloud Interoperabili per SMArt-government".

Né il documento né le sue parti possono essere pubblicate, riprodotte, copiate o comunque divulgate senza autorizzazione scritta delle parti coinvolte nel progetto oltre quanto previsto dalla lista di distribuzione

Lista di distribuzione

Stato del deliverable

Ver.	Data	Autore della modifica	Note	Validazione

INDICE DEGLI ARGOMENTI

INDICE DELLE FIGURE **5**
INDICE DELLE TABELLE **6**
ACRONIMI **7**
NOTAZIONE **8**
EXECUTIVE SUMMARY **9**
INTRODUZIONE **10**
1. TECNOLOGIE SEMANTICHE: METODI PER L'ESTRAZIONE, LA PUBBLICAZIONE E L'INTERROGAZIONE DI DATI DELLA PA **12**
1.1 Standard W3C **12**
1.2 Buone pratiche di *ontology design* e pubblicazione/consumo di *linked data* **13**
1.2.1 Linee-guida dell'Agenzia per l'Italia Digitale 13
1.2.2 Gli ontology design pattern e l'eXtreme Design (XD) 13
1.3 Esperienze **18**
1.3.1 I metodi di data.cnr.it e del Semantic Scout 18
2. ANALISI DELLO SCENARIO DI RIFERIMENTO E DEI REQUISITI **20**
3. MATERIALE E RISORSE OTTENUTE E ANALIZZATE FINORA **22**
4. DESCRIZIONE DELLE ONTOLOGIE-CAMPIONE FINORA ESTRATTE **25**
4.1 Ontologia degli asili nido. **25**
4.2 Ontologia estratta da uno schema DB sull'Assistenza Domiciliare Integrata **27**

Indice delle Figure

Figure 1. Due diagrammi UML che mostrano la codifica OWL del design pattern *time interval*. 13

Figure 2. L'idea di progettazione di ontologie basata su *pattern*. 15

Figure 3. (a) Il processo XD;. (b) l'iterazione eseguita da coppie di progettisti.....16

Figura 4*. Lo stack di metodi usati per la realizzazione di* data.cnr.it *e del Semantic Scout*. 19

Figure 5. *Entità principali dell'ontologia Asili Nido. Le frecce spesse indicano relazioni di sottoclasse, quelle fini e nere o blu relazioni possibili fra individui di classi diverse e frecce rosse per relazioni di disgiuntezza. Gli attributi all'interno dei rettangoli sono relazioni possibili con i relativi vincoli di cardinalità. 26*

Figure 6. *Un grafo dei dati relativi a un asilo nido specifico (profilo OWL per UML). Le relazioni fra i nodi rappresentano relazioni fra individui o di appartenenza a una classe. Gli attributi all'interno dei rettangoli sono relazioni fra individui e valori simbolici (ex. nomi, commenti, etc.). 27*

Figure 7. *Entità principali dell'ontologia Assistenza Domiciliare Integrata.*.....28

Indice delle Tabelle

Non è stata trovata alcuna voce dell'indice delle figure.

Acronimi

PRISMA	PiattafoRme cloud Interoperabili per SMArt-government

Notazione

Executive Summary

Il presente Rapporto Tecnico descrive i metodi e i passi iniziali per definire i casi d'uso e lo sviluppo di alcuni campioni di schemi e dati, nell'ambito dell'attività di estrazione e arricchimento dei dati di alcune Pubbliche Amministrazioni Locali.

I metodi sono basati sugli standard del W3C, su buone pratiche internazionali, sulle linee-guida emanate dall'Agenzia per l'Italia Digitale e su esperienze vicine già avute dai partecipanti a questa attività, in particolare lo sviluppo di data.cnr.it e del Semantic Scout, nonché dei linked open data dell'Indice delle Pubbliche Amministrazioni.

Due ontologie sono state sviluppate a partire da documenti di specifica sui servizi relativi agli asili nido e all'assistenza domiciliare integrata.

Introduzione

L'AR 4.1 consiste nell'applicare metodi e componenti software per l'estrazione, l'arricchimento, la pubblicazione e il riuso di *linked open data* per le Pubbliche Amministrazioni Locali (PAL), supportando così anche con tecnologie semantiche l'attività 4.3 per la valutazione esterna della PA mediante la rilevazione della community trust (fiducia della comunità).

Il Codice dell'Amministrazione Digitale recepisce molte delle esperienze internazionali sulla pubblicazione di dati aperti delle pubbliche amministrazioni. La pubblicazione, il riuso e l'interoperabilità di dati della Pubblica Amministrazione in modo semplice e il più possibile aperto è diventata negli ultimi anni una priorità in paesi dove le amministrazioni e la politica sono attente a funzioni come trasparenza, tracciabilità, e accessibilità dei dati.

La Pubblica Amministrazione soggetto sperimentatore del progetto previsto dal paragrafo 4.10 della seconda parte del progetto esecutivo PRISMA è il Comune di Catania.

Il Comune di Catania in quanto Ente Locale capofila della Coalizione Territoriale 14 - PIANO INTEGRATO DI SVILUPPO URBANO (PISU) "Catania – Città Metropolitana", nell'ambito dell'Asse 6 "Sviluppo Urbano Sostenibile" PO FESR 2007-2013, ha definito alcuni progetti esecutivi per la riorganizzazione di alcuni ambiti di sperimentazione: i servizi sociali e socio-sanitari on-line, la gestione del traffico e dei trasporti e il decoro urbano.

In questo contesto di riferimento, l'azione dell'OR4 si è al momento concentrata su un ambito di sperimentazione: i servizi sociali e socio-sanitari on-line. Al fine del recupero dei dati e dei processi sono state realizzate alcune riunioni conoscitive con la Dirigenza della Direzione Famiglia e Politiche sociali del Comune di Catania.

Uno dei progetti esecutivi del PISU "Catania – Città Metropolitana" è il progetto esecutivo "Città policentrica" mira alla riorganizzazione della Direzione Famiglia e Politiche sociali e degli uffici ad essa afferenti, in modo da dotarsi di un sistema informativo dei servizi sociali. Sono altresì individuati i servizi sociali offerti ai cittadini che verranno prodotti a regime per effetto degli interventi realizzati; tali servizi sono raggruppati in tre aree welfare: Area Famiglia e sostegno delle responsabilità familiari, Area Inclusione Sociale e Area Domiciliarità e residenzialità anziani e disabili.

La fase di riorganizzazione della Direzione del Comune di Catania ha portato a un ritardo nella raccolta delle risorse informative della PAL e dei requisiti necessari all'AR4.1.

Di concerto con il Comune di Catania sono stati comunque individuati due settori per i casi di studio e sulla base di dati raccolti sono state sviluppate ontologie di base per la gestione degli asili nido e per il piano di assistenza individuale – cure domiciliari.

Questo deliverable presenta in modo preliminare I metodi che si utilizzeranno in AR4.1 e alcuni esempi per la costruzione di ontologie, la definizione di requisiti e l'estrazione e la pubblicazione dei dati, lasciandone l'articolazione completa (inclusa la discussione rispetto allo stato dell'arte internazionale e una valutazione) in riferimento ai requisiti completi del Comune di Catania nel deliverable principale D4.1.2.

1. Tecnologie semantiche: metodi per l'estrazione, la pubblicazione e l'interrogazione di dati della PA

I metodi sono basati sugli standard semantici del W3C, su buone pratiche internazionali, sulle linee-guida emanate dall' Agenzia per l'Italia Digitale e su esperienze vicine già avute dai partecipanti a questa attività, in particolare lo sviluppo di data.cnr.it e del Semantic Scout, nonché dei *linked open data* dell'Indice delle Pubbliche Amministrazioni.

1.1 Standard W3C

Il W3C rappresenta il punto di riferimento privilegiato per il Web Semantico in generale e per i Linked Data in particolare. A tal riguardo, sono stati istituiti diversi gruppi di lavoro dal W3C. I risultati dei gruppi sull'RDF, sul Semantic Web Deployment, e sui Government Linked Data sono da considerarsi di riferimento nel contesto della PA ai fini dell'apertura interoperabile dei dati pubblici.

Il W3C e la Comunità Europea hanno di fatto costruito l'infrastruttura e la cultura del Web Semantico fin dal 1999, attraverso svariati gruppi di lavoro, standard e diversi progetti finanziati. Per quanto riguarda il W3C, oltre ai gruppi di lavoro su RDF e sul Semantic Web Deployment, meritano di essere citati i gruppi di lavoro su OWL (Web Ontology Language) e su RIF (Rule Interchange Format), che hanno creato le condizioni per poter fare ragionamento automatico su schemi e dati nel Web; il gruppo su SPARQL che ha permesso di realizzare un linguaggio di interrogazione molto flessibile e ricco per RDF; il gruppo su SKOS (Simple Knowledge Organization System) che ha definito una semplice ontologia per esportare strutture di metadati (classificazioni, tesauri, ecc.) in RDF; il gruppo su RDFa che ha definito formati per "incorporare" informazione semantica direttamente in HTML, così abilitando di fatto annotazioni semantiche del testo interoperabili con RDF. Partendo da tali raccomandazioni, sono state poi intraprese iniziative volte alla definizione di specifiche per implementare una federazione di "repository" di "asset" semantici sul sito Joinup dalla Commissione Europea.

1.2 Buone pratiche di *ontology design* e pubblicazione/consumo di *linked data*

1.2.1 *Linee-guida dell'Agenzia per l'Italia Digitale*

L'Agenzia per l'Italia Digitale ha pubblicato a fine 2012 delle linee-guida definite dapprima da un comitato di esperti e poi sottoposto a consultazione pubblica. Tali lnee-guida contengono soprattutto buone pratiche e costituiscono un ottimo sistema di riferimento per lo sviluppo di *linked open data* nelle PAL, in quanto coprono aspetti legislativi, sociali, tecnici, riassumono esperienze nazionali e internazionali e menzionano strumenti e componenti (soprattutto *open-source*) utlizzabili nel processo di produzione.

Lasciamo al deliverable D4.1.2 il dettaglio di come tali linee-guida verranno recepite nel lavoro concreto con le PAL del Comune di Catania e ci concentriamo invece nella presentazione del metodo usato per la produzione delle due ontologie-campione (sezione 4), metodo sviluppato durante il progetto europeo NeOn e implementato nel componente XDTools del NeOn Toolkit.

1.2.2 *Gli* ontology design pattern *e l'*eXtreme Design *(XD)*

Un *ontology design pattern* (ODP) è una soluzione ad un problema ricorrente di modellazione ontologica (Gangemi & Presutti, Ontology Design Patterns, 2009). Ci sono diversi tipi di ODP, in questo contesto si fa riferimento ai "Content ODP" (CP). Un esempio di CP chiamato "time interval" è raffigurato in Figura 1.

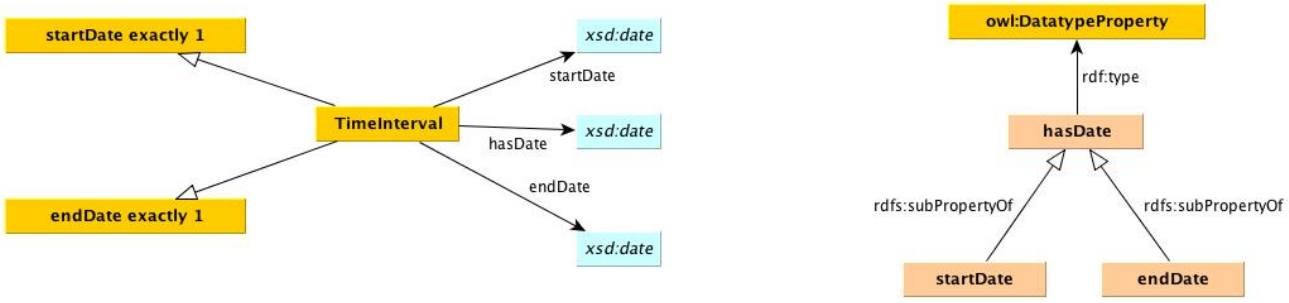

Figure 1. Due diagrammi UML che mostrano la codifica OWL del design pattern *time interval*.

I CP forniscono una soluzione per la modellazione di classi e proprietà che popolano un'ontologia, quindi sono dipendenti dal dominio di conoscenza considerato. Ogni CP è associato ad un insieme di requisiti che rappresenta il problema per cui fornisce una soluzione. Questi requisiti sono

rappresentati in tre forme: (i) domande di competenza (CQ), (ii) asserzioni di contesto e (iii) requisiti di inferenza. Le CQ indicano tipiche query a cui una base di conoscenza deve essere in grado di rispondere se fosse basata su quel CP.

Le asserzioni di contesto indicano le condizioni che valgono su, e tra, i concetti codificati dal CP. Infine, i requisiti di inferenza indicano le conseguenze logiche che devono essere assicurate dal CP.

Ad esempio, il CP raffigurato in Figura 1 è associato alle seguenti CQ:

1. When does a certain time interval start?

2. When does a certain time interval end?

3. What are the points in time that belong to a certain time interval?

Tipicamente un CP e associato anche ad un insieme di query SPARQL che forniscono una formalizzazione delle CQ. Questa informazione è molto utile ad un progettista di ontologie che vuole validare un ontologia che contiene un CP con dei dati di esempio.

Il CP “time interval” è associato alla seguente asserzione di contesto.

“A time interval always has one starting and one ending point and no more than them.”

Questo requisito è soddisfatto dalla restrizione di cardinalità sulla proprietà che identifica l’inizio e la fine dell’intervallo temporale.

Infine, questo CP è associato con i seguenti requisiti di inferenza:

1. The starting and ending date of a time interval belongs to it.

2. Two time intervals with the same start and end dates should be recognized to be the same.

Il primo requisito è soddisfatto attraverso la definizione di due proprietà “startDate” e “endDate” come sotto-proprietà di “hasDate”. Il secondo requisito è soddisfatto con la definizione dell’assioma “hasKey[startDate, endDate]” sulla classe “TimeInterval”.

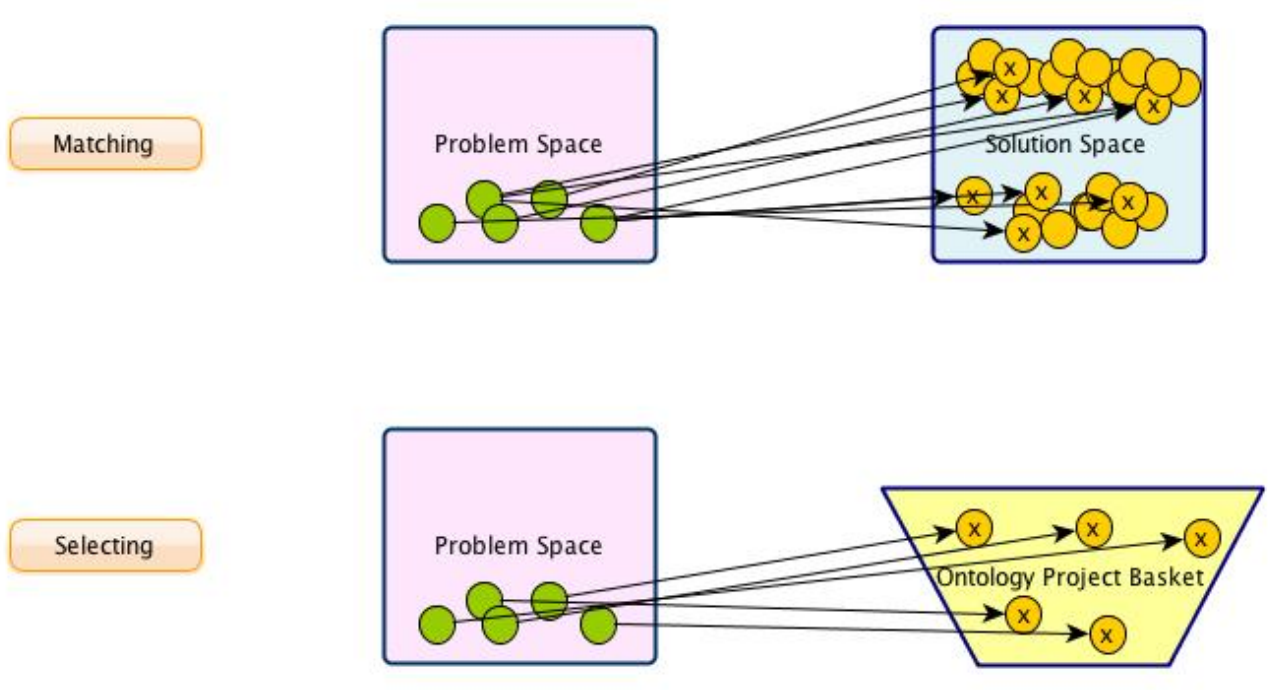

Figure 2. L'idea di progettazione di ontologie basata su *pattern*.

Progettazione di ontologie basata sul riuso dei pattern

Idealmente, un progetto ontologico può essere completamente sviluppato attraverso il riuso di soluzioni esistenti cioè CP, combinandoli in modo appropriato. Per identificare i candidati CP da riusare in un certo progetto ontologico, la descrizione del problema specifico deve essere espressa sotto forma di CQ in modo da essere confrontabile con i CP. La Figura 2 mostra l'idea di progettazione di ontologie basata sul riuso di pattern. Il progetto di sviluppo di un'ontologia è diviso in due parti: (i) il "Problem space", che contiene un insieme di requisiti che descrivono esplicitamente i problemi di modellazione da risolvere e (ii) il "Solution space" che contiene tutti i CP disponibili.

Prima i requisiti sono confrontati con i CP e un insieme di CP candidati sono identificati, poi i CP più appropriati sono selezionati per il riuso.

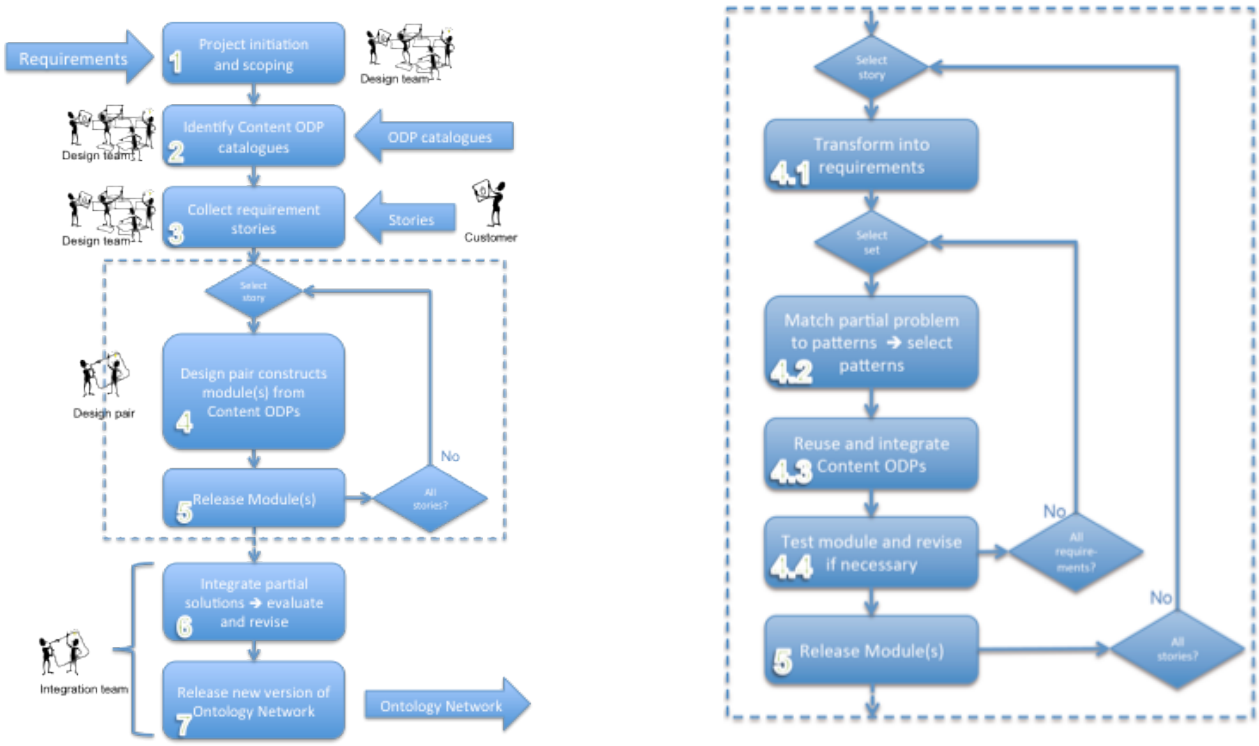

Figure 3. (a) Il processo XD;. (b) l'iterazione eseguita da coppie di progettisti.

eXtreme Design (XP)

XP è un metodo "agile", ispirato da eXtreme Programming (XP) (Shore & Warden, 2007), e da "experience factory" (Basili, Caldiera, & Rombach, 1994), per lo sviluppo di ontologie attraverso il riuso di ontology design pattern. XD è un metodo "test-driven" cioè il testing è una parte centrale dello sviluppo, e promuove il "pair-design" analogamente al "pair-programming".

La Figura 3 mostra l'intero processo XD (a) e i dettagli di una singola iterazione eseguita da una coppia di progettisti (b). Il processo inizia con il team, che include gli esperti di dominio (clienti), che attraverso delle interviste acquisisce familiarità con il dominio allo scopo di identificare la portata del progetto (in termini di copertura del dominio), dipendentemente dall'obiettivo applicativo (Step 1.). In seguito, il team XD identifica le sorgenti di CP da utilizzare durante il progetto (Step 2). I rappresentanti del cliente sono quindi invitati a scrivere storie, preferibilmente ispirate a situazioni reali, che servono come esempi di tipici fatti che devono essere rappresentati dalla ontologia e di fornire degli esempi su come questi fatti sono utilizzati e collegati tra loro (Step 3). Tutte le storie

sono organizzate in ordine di priorità e possibili inter-dipendenze sono identificate e messe in evidenza. Ogni storia è descritta con una “small card” che include il titolo della storia, una lista di altre storie da cui essa dipende, una descrizione in linguaggio naturale e un valore di priorità. Il cliente teoricamente può aggiungere storie durante l’intero ciclo di vita del progetto.

Quando un numero sufficiente di storie per iniziare lo sviluppo sono state raccolte, ogni coppia di progettisti seleziona una storia che sarà il focus del loro lavoro per l’iterazione successiva (Step 4). La selezione è basata sull’esperienza e competenza della coppia e sulla priorità della storia. Eseguendo questo attività una coppia entra un’iterazione di sviluppo (Descritta in dettaglio in Figura 3 (b)). Quando una storia è stata completamente modellata è rilasciata internamente per l’integrazione (Step 5.), questa attività costituisce la fine di una iterazione per una coppia.

Quando un nuovo modulo è rilasciato, esso deve essere integrato con il resto dei moduli, che insieme costituiscono la versione corrente dell’ontologia (Step 6). Di solito una coppia di progettisti è dedicata all’integrazione.

Quando la fase di integrazione di un nuovo modulo è completa, una nuova versione (incrementale) dell’ontologia viene rilasciata (Step 7).

Lo Step 4 del processo XD identifica l’iterazione eseguita da una coppia di progettisti.

La Figura 3 (b) mostra i passi principali di una singola iterazione:

- Task 4.1: La coppia di progettisti analizza la storia selezionata e deriva un insieme di CQ, asserzioni di contesto e requisiti di inferenza;

- Task 4.2: L’iterazione continua con la selezione di una delle CQ, o di un piccolo insieme di CQ che costituisca un problema di modellazione coerente, e il suo confronto con le CQ associate con i CP disponibili, allo scopo di identificare candidati per il riuso;

- Task 4.3: I CP selezionati vengono importati nel modulo corrente, raffinati ed estesi fino a completare la modellazione della storia;

- Task 4.4: Il modulo viene validato sulla base dei requisiti che deve soddisfare. Le CQ vengono formalizzate sotto forma di query SPARQ, le asserzioni di contesto e i requisiti di inferenza vengono validati attraverso test di stress e di inferenza. Il modulo è revisionato fino a quando tutti i test sono passati.

1.3 Esperienze

La spinta principale alla pubblicazione di *linked open data* per le PAL è venuta da grandi iniziative negli Stati Uniti (data.gov) e Regno Unito (data.gov.uk), tuttavia esistono alcune esperienze/iniziative degne di nota che sono in linea con lo stato dell'arte internazionale. Il gruppo di lavoro è unanime nel ritenere che prendendo esempio anche da queste iniziative sia possibile incentivare, nel medio termine, politiche e strategie di apertura in formato Linked dei dati della PA italiana.

Una di queste è "data.cnr.it" il progetto di dati aperti del Consiglio Nazionale delle Ricerche, progettato e mantenuto dal Laboratorio di Tecnologie Semantiche dell'ISTC-CNR, condiviso con l'unità Sistemi Informativi del CNR. Inoltre, altri Linked Open Data sono stati pubblicati dal comune di Firenze, da DigitPA, dalla regione Piemonte, dalla Camera dei Deputati.

1.3.1 *I metodi di data.cnr.it e del Semantic Scout*

Per quanto riguarda data.cnr.it, realizzato nel 2010, si tratta di un progetto motivato non solo dalla volontà di pubblicare dati aperti della ricerca, ma anche dalla necessità di rispondere a requisiti di trasferimento tecnologico richiesti dal CNR. Il requisito principale era quindi quello di far incontrare l'offerta di risultati della ricerca con l'eventuale domanda. Questo requisito è analogo a quelli che sussistono in molte altre amministrazioni, sia nei rapporti con il cittadino, sia con le aziende o con altre amministrazioni. Sul piano tecnico, il progetto è stato sviluppato dedicando particolare attenzione alla completezza e all'accuratezza degli schemi (ontologie) e all'arricchimento dei dati con l'integrazione di tecnologie d'inferenza e di analisi del linguaggio naturale. Analogamente, particolare attenzione è stata posta all'implementazione di interfacce di navigazione ed esplorazione dei dati innovative, accessibili dal portale Semantic Scout. I metodi riguardano trasformazione, analisi, bonifica, *linking* dei dati ecc., sintetizzati nella Figura 4 di questo documento.

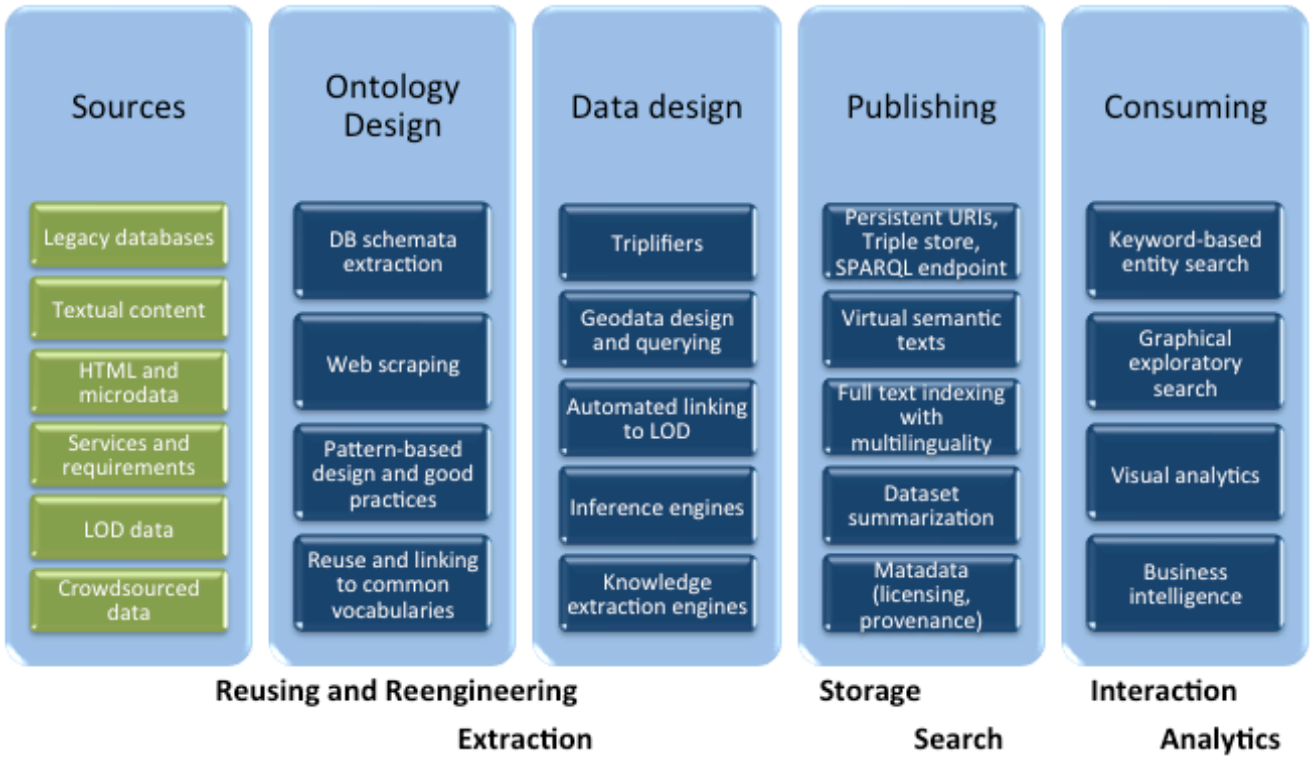

Figura 4. ***Lo stack di metodi usati per la realizzazione di*** **data.cnr.it** ***e del Semantic Scout.***

2. Analisi dello scenario di riferimento e dei requisiti

Per l'analisi dei requisiti è stato definito un insieme di domande da utilizzare per le interviste con le persone chiave dei settori per la gestione degli asili nido e per il piano di assistenza individuale – cure domiciliari del Comune di Catania. Lo scopo delle interviste è stato quello di far emergere cosa il Comune di Catania offrisse in termini di dati e servizi implementati, sotto quali licenze essi fossero rilasciati e come intervenire sulla reingegnerizzazione questi in ottica generazione Linked Data. Inoltre, lo scopo è stato quello di individuare le persone di riferimento (*champion*) all'interno del Comune di Catania con le quali interagire.

Di seguito sono riportate le domande delle interviste organizzate in una sorta di albero decisionale

- Quali sono le tipologie di dati disponibili?
  Per ogni tipologia:
  - Chi sono i *champion* (persone di riferimento per la loro attività *di fatto*) per questa tipologia di dati?
  - Quale è interessato a collaborare direttamente con noi?
  - In quali formati è disponibile questa tipologia di dati?
    Per ogni formato:
    - Chi sono i *champion* per questo formato?
    - Quale è interessato a collaborare direttamente con noi?
    - Che query vengono effettuate sui dati in questo formato?
  - Quali sono le modalità di accesso alla tipologia di dati?
    Per ogni modalità di accesso:
    - Chi sono i *champion* che gestiscono le modalità accesso ai dati?
    - Quale è interessato a collaborare direttamente con noi?
  - È possibile avere accesso diretto (online) o a dump (offline) a questa tipologia di dati?
    - Chi sono i *champion* che possono abilitare l'accesso ai dati?
    - Quale è interessato a collaborare direttamente con noi?
  - Quali servizi sono implementati attualmente su questa tipologia di dati?
    Per ogni servizio:
    - Chi sono i *champion* per questo servizio?
    - Quale è interessato a collaborare direttamente con noi?
    - Come sono usate le interfacce utente di questo servizio?
    - Qual è la tipologia di utenti che usano il servizio?
    - Quali sono le difficoltà che sono state riscontrate per la realizzazione del servizio? In particolare, in relazione a:
      - aggregazione dei dati
      - conversione dei dati
      - duplicazione dei dati
  - Quali servizi si vogliono realizzare su questa tipologia di dati?
    Per ogni servizio:

- Chi sono i *champion* per questo servizio?
- Quale è interessato a collaborare direttamente con noi?
- Qual è la tipologia di utenti che dovrebbero usare il servizio?
- Quali sono i partner del progetto interessati a questo servizio?
  Per ogni partner:
  - Qual è il ruolo del partner relativamente a questo servizio?
  - Chi è il referente del partner?

- Esistono delle licenze associate a questa tipologia di dati?
  - Chi sono i *champion* che gestiscono le licenze dei dati?
  - Quale è interessato a collaborare direttamente con noi?
- Che tipo di licenze bisogna prevedere per eventuali nuovi dati per questa tipologia?
  - Chi sono i *champion* per la gestione delle licenze dei dati?
  - Quale è interessato a collaborare direttamente con noi?

Alcuni risultati finora raggiunti:

La persona di riferimento per acquisire dati sui servizi sociali gestiti dal comune è stata individuata nella persona dell'ing. Persico, in qualità di funzionario dirigente della direzione servizi sociali. In particolare è stata appurata l'esistenza di due database di circa 1000 record con dati sui minori ex legge 285. Inoltre esisterebbe un applicazione in versione alpha per la gestione dei medesimi dati che permetterebbe l'estrazione di query. Tutti i dati citati sono contenuti in DB proprietari.

Un'altra persona chiave è stata individuata nella persona dell'assessore prof. Pennisi che è stato il direttore del Laposs che si è occupato di problematiche di interoperabilità dei dati e risulterebbe particolarmente interessato agli Open Data.

Un problema importante emerso riguardi ai dati disponibili riguarda l'eterogeneità dei formati, in particolare dei dati legati alle statistiche dell'abbandono scolastico. Il Laposs sembra essere il possessore di questi dati. Inoltre attualmente molti dei dati e delle procedure gestite dai servizi sociali del comune sono disponibili solo in cartaceo o in formati non strutturati o senza linee guida particolari.

3. Materiale e risorse ottenute e analizzate finora

Le riunioni conoscitive svolte con la Dirigenza della Direzione Famiglia e Politiche sociali del Comune di Catania sono servite per avere un'idea dello stato in cui versa la Direzione, nonché dei servizi da essa prestati e per individuare i dati e le risorse della PAL al fine di attivare il processo di apertura e linking dei dati.

Un quadro dello stato della Direzione e dei servizi si ha analizzando la relazione dell'Assessore alla Famiglia e Politiche sociali sullo stato della riorganizzazione dei servizi e delle politiche sociali del 12 gennaio 2013 e i documenti del Piano Integrato di Sviluppo Urbano (PISU) “Catania – Città Metropolitana” (i progetti esecutivi Polo Librino, Servizi On-Line, Assegni maternità e nucleo familiare e soprattutto il progetto esecutivo “Città policentrica”).

Infatti dall'analisi dei documenti si evince che la Direzione deve riorganizzare sia la struttura che i servizi che essa presta, in considerazione delle esigenze normative, dei mutamenti intervenuti nel welfare locale e per ottimizzare le risorse economiche. Il progetto PISU mette a bando la ri-funzionalizzazione e riorganizzazione dei servizi della Direzione prevedendo di dotarsi di un sistema informativo dei servizi sociali. Vengono individuati tre aree welfare: Area Famiglia e sostegno delle responsabilità familiari, Area Inclusione Sociale e Domiciliarità e residenzialità anziani e disabili in cui raggruppare tutti i servizi che a regime verranno prodotti per effetto degli interventi realizzati. Tale riorganizzazione costituirà un radicale cambiamento rispetto a pratiche di lavoro consolidate negli anni e una completa informatizzazione dei servizi.

La legge quadro per la realizzazione del sistema integrato di interventi e servizi sociali (n. 328/2000) ha ridefinito il profilo delle politiche sociali apportando una serie di elementi di novità. Ha individuato il fruitore dei servizi sociali come una persona nella sua totalità e quindi è necessario che gli Enti preposti istituiscano “un sistema informativo dei servizi sociali per assicurare una compiuta conoscenza dei bisogni sociali, del sistema integrato degli interventi e dei servizi sociali e poter disporre tempestivamente di dati ed informazioni necessari alla programmazione, alla gestione e alla valutazione delle politiche sociali“

I dati individuati in questa prima fase rispondono alle priorità espresse dalla PAL, da un lato a esigenze di rispetto di vincoli normativi e dall'altro da riorganizzazione del servizio per ottimizzazione

di risorse. In particolare sono stati analizzati i casi di studio riguardanti la gestione degli asili nido e il piano di assistenza individuale – cure domiciliari.

La normativa di riferimento per gli asili nido è data dalla Legge regionale n. 214/79 e dal Decreto presidenziale 29 giugno 1988 “Standards strutturali ed organizzativi dei servizi e degli interventi socio assistenziali previsti dalla legge regionale 9 Maggio n. 22”, a cui si aggiunge il regolamento di gestione degli asili nido adottato dal Comune di Catania.

I dati riguardanti gli asili nido del Comune di Catania sono stati estrapolati dalla pagina degli Asili nido del sito web del Comune di Catania (http://www.comune.catania.it/il_comune/organizzazione/uffici_comunali/direzioni/istruzione/asili-nido/).

Per operare l'apertura dei dati in RDF secondo schemi Linked Data, questi devono subire un processo di bonifica e poi un processo di reingegnerizzazione e refactoring della base informativa. Infatti i dati provenienti dalle PAL sono spesso “sporchi” e presentano problemi, quali l'incompletezza, i formati diversi, i significati ambigui, i datatype inconsistenti, la mancanza di corrispondenza tra nomi usati e i dati effettivamente contenuti.

A partire dalle informazioni contenute nei documenti: descrizione_asilo_nido_20081023.pdf e circoscrizioni_asili_nido.pdf è stato sviluppato un modello concettuale degli asili nido del Comune di Catania.

La legge 328/2000 prevede, inoltre, che sia redatto un Piano di Zona tra i soggetti territoriali operanti nell'ambito della solidarietà sociale dei Comuni dell'area metropolitana e l'Azienda Sanitaria Provinciale per individuare tra gli altri: gli “obiettivi strategici e le priorità di intervento nonché gli strumenti e i mezzi per la relativa realizzazione; le modalità organizzative dei servizi, le risorse finanziarie, strutturali e professionali, i requisiti di qualità; le forme di concertazione con l'azienda unita' sanitaria locale. Il Piano di Zona 2010/2012 del Distretto Socio Sanitario di riferimento e il Protocollo Operativo per l’attuazione del servizio di assistenza domiciliare integrata hanno previsto l’“Assistenza Domiciliare Integrata”da gestire in accordo con la ASP. Tale ADI mira a favorire l’autonomia e la permanenza a domicilio della persona non autosufficiente rafforzando l'integrazione tra i servizi sanitari e socio-assistenziali.

Per quanto concerne i dati del settore socio-sanitario abbiamo, quindi, operato la bonifica e l'analisi e modellazione di uno schema di database per la scheda dell'Assistenza Domiciliare Integrata.

4. Descrizione delle ontologie-campione finora estratte

4.1 Ontologia degli asili nido.

URI (indirizzo dell'ontologia):

http://www.ontologydesignpatterns.org/ont/prisma/asili.owl

Namespace (indirizzo di default delle entità definite nell'ontologia):

http://www.ontologydesignpatterns.org/ont/prisma/asili.owll#

Live OWL Documentation Environment (documentazione dell'ontologia):

http://www.essepuntato.it/lode/http://www.ontologydesignpatterns.org/ont/prisma/asili.owl

L'ontologia fornisce classi e proprietà per rappresentare le Unità Funzionali di Asilo Nido, i servizi da esse offerti e la loro collocazione geografica. Un'Unità Funzionale generale (UF) di Asilo Nido è costituita da almeno due unità funzionali dedicate – una relativa ai lattanti e una relativa ai divezzi – da servizi specifici e da servizi generali. Per tale motivo sono state introdotte tre classi nell'ontologia: una per rappresentare le UF di Asili Nido (NurserySchoolFunctionalUnit), una per rappresentare le UF per gli Infanti (InfantFunctionalUnit) ed una per le UF per Divezzi (WeanedFunctionalUnit). Tutte e tre le classi che rappresentano le UF sono sottoclassi di FuncionalUnit. Alla classe NurserySchoolFunctionalUnit sono state aggiunte delle restrizioni che permettono di rappresentare un'UF Asilo Nido come composta da almeno due unità funzionali, una relativa ai lattanti e una relativa ai divezzi, da servizi specifici e da servizi generali. Segue uno snippet di esempio OWL che definisce la classe NurserySchoolFunctionalUnit:

```
asili:NurserySchoolFunctionalUnit
      a       owl:Class ;
      rdfs:subClassOf asili:FunctionalUnit ;
      asili:providesNurserySchoolService some asili:NurserySchoolService ;
      asili:hasSubFunctionalUnit some asili:WeanedFunctionalUnit ;
      asili:hasSubFunctionalUnit min 2 asili:FunctionalUnit ;
     asili:hasSubFunctionalUnit some asili:InfantFunctionalUnit
```

Inoltre le UF sono collocate in determinati edifici a cui è associato un indirizzo specifico. Gli edifici sono rappresentati dalla classe Building di cui NurserySchoolBuilding è la sottoclasse specifica per rappresentare gli edifici dedicati agli asili nido. Le UF Asili Nido sono collegate alle strutture che ospitano fisicamente gli asili attraverso la proprietà hasLocation che prevede come dominio della

relazione la classe NurserySchooFunctionalUnit e come range la classe NurserySchoolBuilding. A ogni edificio è associato un indirizzo che è rappresentato come individuo della classe vcadr:Address. Tale associazione è realizzata tramite la proprietà vcard:hasAddress che prevede come range vcadr:Address. Inoltre, all'indirizzo è associata una municipalità, rappresentata come individuo della classe Municipality, attraverso la proprietà hasMunicipality che ha come dominio vcard:Address e come range la classe Municipality. Infine alle UF di Asilo Nido sono associati i servizi di Asilo Nido erogati. Questi ultimi sono rappresentati come individui della classe NurserySchoolService. L'associazione tra UF di Asilo Nido e servizi è realizzata attraverso la materializzazione della proprietà providesNurserySchoolService.

In Fig. 5 è riportato un diagramma UML (con profilo per OWL) delle classi e delle relazioni costituenti l'ontologia.

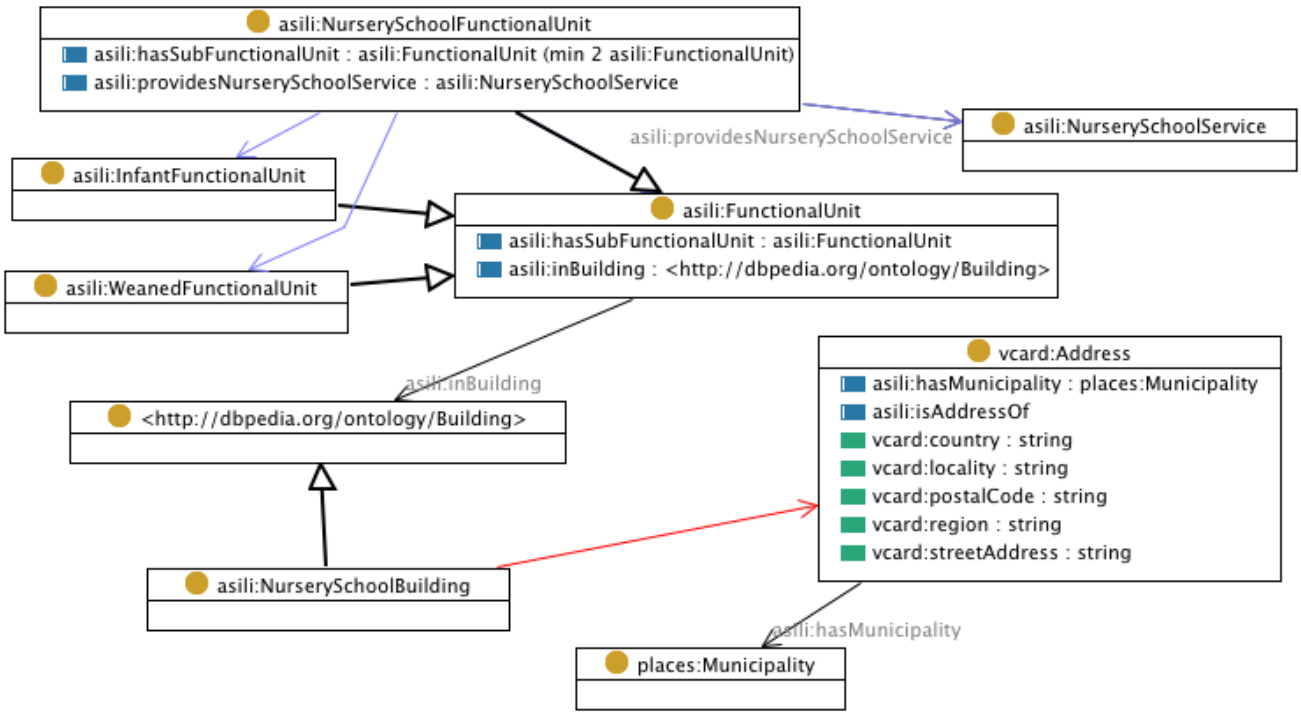

Figure 5. ***Entità principali dell'ontologia Asili Nido. Le frecce spesse indicano relazioni di sottoclasse, quelle fini e nere o blu relazioni possibili fra individui di classi diverse e frecce rosse per relazioni di disgiuntezza. Gli attributi all'interno dei rettangoli sono relazioni possibili con i relativi vincoli di cardinalità.***

In Fig. 6 è riportato un diagramma UML delle relazioni per un individuo (asilo nido) specifico.

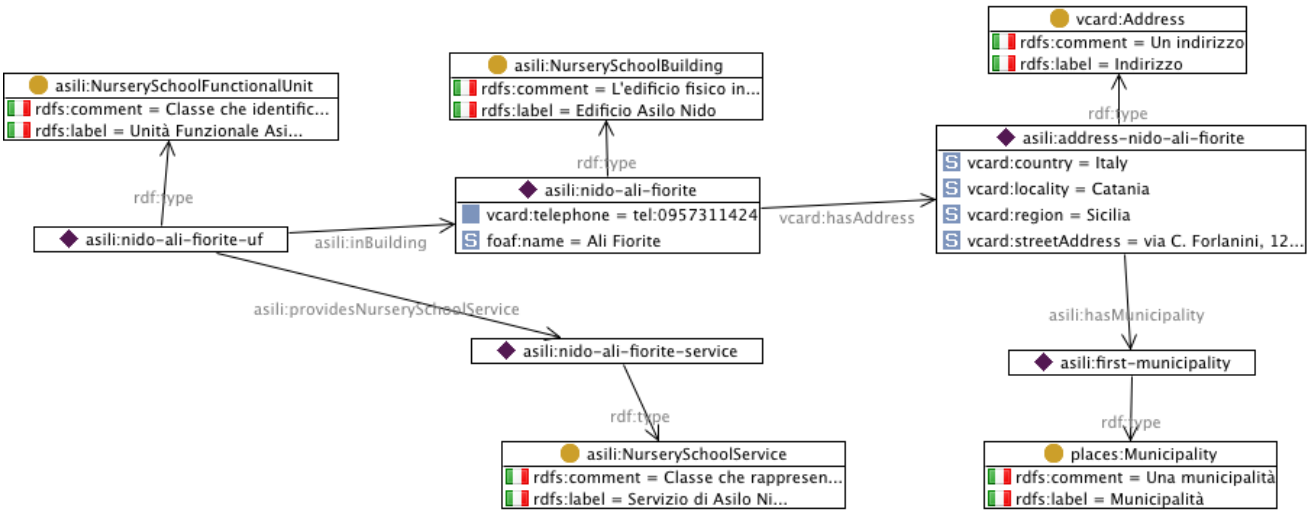

Figure 6. ***Un grafo dei dati relativi a un asilo nido specifico (profilo OWL per UML). Le relazioni fra i nodi rappresentano relazioni fra individui o di appartenenza a una classe. Gli attributi all'interno dei rettangoli sono relazioni fra individui e valori simbolici (ex. nomi, commenti, etc.).***

4.2 Ontologia estratta da uno schema DB sull'Assistenza Domiciliare Integrata

URI:

http://www.ontologydesignpatterns.org/ont/prisma/asp_ct.owl

Namespace:

http://www.ontologydesignpatterns.org/ont/prisma/asp_ct.owl#

Live OWL Documentation Environment:

http://www.essepuntato.it/lode/http://www.ontologydesignpatterns.org/ont/prisma/asp_ct.owl

L'ontologia è stata estratta automaticamente dallo schema relazionale del database sull'Assistenza Domiciliare Integrata e poi arricchita con allineamenti a ontologie in uso su Linked Open Data o a

design pattern generali per la modellazione per esempio di agenti o situazioni, e infine raffinata con assiomi ulteriori, per esempio di disgiuntezza tra classi.

Il cuore dell'ontologia (cf. Fig. 7) è costituito da situazioni rilevanti per l'Assistenza Domiciliare: condizioni cliniche, accettazione e dimissione di pazienti, che includono "agenti" come medici, pazienti, reparti ospedalieri e presidi territoriali. Riportiamo di seguito un diagramma riassuntivo delle classi e delle loro relazioni applicabili.

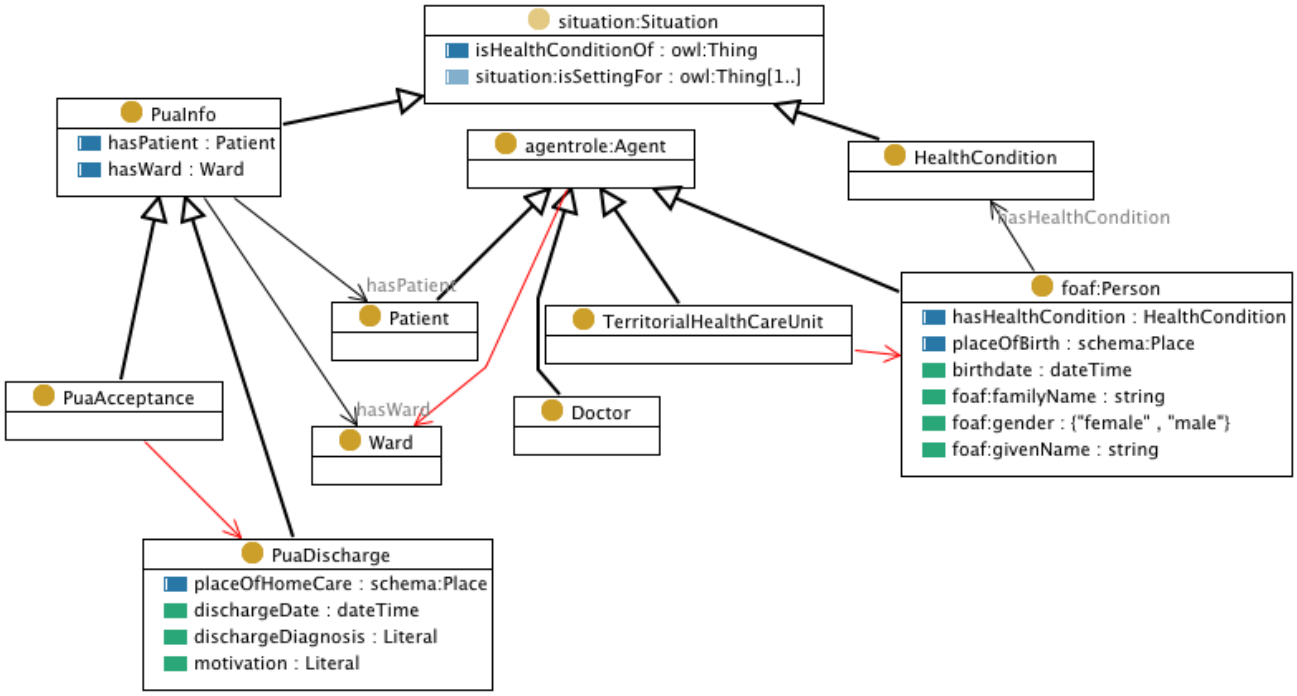

Figure 7. ***Entità principali dell'ontologia Assistenza Domiciliare Integrata.***