Linked Open Data e Semantic Web: Fondamenti e Linguaggi di Interrogazione Parte Prima

Cristiano Longo longo@dmi.unict.it

Università di Catania, 2014-2015

Argomenti

Questa presentazione tratterà i seguenti argomenti:

- Motivazioni del Web Semantico
- Definizione formale di Ontologie
- Interrogazioni sulle Ontologie
- Vocabolari

World Wide Web Consortium

Le tecnologie del *Web Semantico* rispondono ad un insieme di standard e protocolli promossi e mantenuti dal *World Wide Web Consortium* (in breve *W3C*, vedi http://www.w3.org).

II W3C è un consorzio di standardizzazione per il Web che conta 403 membri tra aziende e organizzazioni governative: CNR, Microsoft Corporation, Apple Inc., Intel Corporation, Facebook, Google Inc., . . .

Altri standard sviluppati in seno al W3C sono: URL, HTTP, XML, HTML, CSS, SOAP, WSDL, Javascript.

Limiti del World Wide Web (1/4)

The Web was designed as an information space, with the goal that it should be useful not only for human-human communication, but also that machines would be able to participate and help. One of the major obstacles to this has been the fact that most information on the Web is designed for human consumption, and even if it was derived from a database with well defined meanings (in at least some terms) for its columns, that the structure of the data is not evident to a robot browsing the web.

Semantic Web Roadmap, Tim Berners-Lee, 1998.

Limiti del World Wide Web (2/4)

Alcuni problemi nell'interpretazione di testi derivano da:

Lingue Differenti e.g. Parigi e Paris possono indicare la stessa città.

Omonimie e.g. esistono svariate città chiamate *Paris* nel mondo (Arkansas, Idaho, Illinois, Kentucky, Maine, Michigan, Missouri, New York, . . .);

Limiti del World Wide Web (3/4)

La situazione si complica in presenza di contenuti multimediali.



Limiti del World Wide Web (4/4)

Come conseguenza, spesso è impossibile eseguire su web ricerce *complesse* ottenendo risultati accurati. Ad esempio, cercando sul web *"Federico II places"* non si ottengono risultati in prima pagina su Federico II, ma solo sull'omonima università:

- Università degli Studi di Napoli "Federico II" OPEN Places
- AOU Policlinico "Federico II" Napoli, Italy Hospital Facebook
- Federico II Ingegneria Via Claudio College and University Facebook
- MARIA CATERINA FONTE www.docenti.unina.it

II Web Semantico (1/2)

[...] the Semantic Web approach instead develops languages for expressing information in a machine processable form.

Semantic Web Roadmap, Tim Berners-Lee, 1998.



Figure: Federico II su dbpedia.org

II Web Semantico (1/2)

I *linguaggi di rappresentazione* usati nel Web semantico hanno una *sintassi rigorosa* e sono dotati di una *semantica formale*.

Questo rende possibile effettuare interrogazioni complesse sui dataset, ottenendo dei risultati precisi anche se a volte parziali:

Q = "Luoghi di nascita di Federico II e dei suoi parenti stretti" .

Linked Open Data Cloud (1/2)

The Semantic Web is a web of data, in some ways like a global database.

Semantic Web Roadmap, Tim Berners-Lee, 1998.

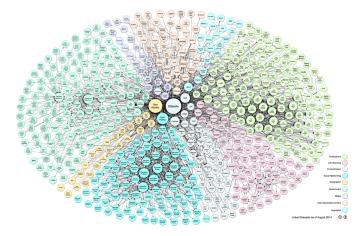


Figure: Linked Open Data Cloud

Linked Open Data Cloud (2/2)

Nel Linked Open Data Cloud sono presenti 365 dataset (fonte http://stats.lod2.eu/).

Alcuni dataset:

- DBPedia (dbpedia.org) corrispondente a wikipedia.org;
- Linked Movie Database (http://linkedmdb.org/) controparte sul Web Semantico di Internet Movie Database (http://www.imdb.com/);
- Linked GeoData (http://linkedgeodata.org) contiene i dati di OpenStreetMap (http://www.openstreetmap.org/);
- AGROVOC (http://aims.fao.org/agrovoc) è il dataset della FAO (http://fao.org);
- Europeana (http://pro.europeana.eu/linked-open-data) contiene dati su beni culturali e tradizioni Europee.

Linked Data (1/2)

I dataset nel Web Semantico possono essere *collegati* tra loro. Ad esempio, una stessa risorsa può essere descritta sotto diversi aspetti in dataset differenti.

Ad esempio, la città di Catania è presente:

- come pubblica amministrazione nel dataset del sistema pubblico di connettività e cooperazione¹ con la url http://spcdata.digitpa.gov.it/Amministrazione/c_c351;
- come divisione amministrativa nel dataset http://www.geonames.org;
- come area territoriale nel dataset dell'ISTAT http://linkedstat.spaziodati.eu/.



Linked Data (2/2)

È possibile effettuare interrogazioni che coinvolgano diversi dataset (anche eterogenei).

Ad esempio, la seguente query può essere eseguita interrogando un data set contenente dati storici ed uno sulle strutture ricettive:

Q = "Strutture ricettive nei luoghi di nascita di Federico II e dei suoi parenti stretti."

Ontologie

I dataset del Web Semantico vengono spesso definiti ontologie.

Una ontologia è una descrizione parziale del mondo:

- descrive una porzione del mondo, spesso è limitata ad un'unico dominio di conoscenza;
- non si assume che i fatti non esplicitamente presenti nell'ontologia siano falsi (Open World Assumption).

Essa è costituita da un insieme finito di affermazioni. Ad esempio:

- Tutti gli esseri umani sono mortali;
- Socrate è mortale;
- Alice è la madre di Roberto.

Ontologie - Affermazioni

Le affermazioni contenute in una ontologia sono di tre tipi:

Constraints: impongono dei vincoli semantici sul dominio di conoscenza che si va a rappresentare. La notazione richiama quella insiemistica;

 $HumanBeing \sqsubseteq Mortal$

Property Assertions: impongono una relazione tra due elementi del dominio;

Alice motherOf Bob

Class Assertions: indicano l'appartenenza di un elemento ad un insieme.

HumanBeing(Socrate)

Ontologie - Reasoning

Con il termine *reasoning* si intende l'attività di estrazione di conoscenza *implicita* in una ontologia.

$$\left\{ \begin{array}{l} \textit{HumanBeing} \sqsubseteq \textit{Mortal}, \\ \textit{HumanBeing}(\textit{Socrate}) \end{array} \right\} \quad \Longrightarrow \quad \textit{Mortal}(\textit{Socrate})$$

Le attività di reasoning sono rese possibili dalle *semantiche formali* associate ai linguaggi di rappresentazione utilizzati.

Ontologie - Definizione

Siano N_C , N_P , N_I tre insiemi infiniti, numerabili e a due a due disgiunti di nomi di classe, proprietà e individuo, rispettivamente.

Una ontologia è un insieme finito di asserzioni dei seguenti tipi:

	Sintassi	Semantica
Constraints	$C \sqsubseteq D$ $R \sqsubseteq S$ $dom(R) \sqsubseteq C$ $range(R) \sqsubseteq C$	$(\forall x)(x \in C \to x \in D)$ $(\forall x, y)([x, y] \in R \to [x, y] \in S)$ $(\forall x, y)([x, y] \in R \to x \in C)$ $(\forall x, y)([x, y] \in R \to y \in C)$
Class Assertions	C(a)	$a \in C$
Property Assertions	a P b (equivalente $P(a, b)$)	$[a,b] \in P$

dove $C, D \in N_C$, $R, S \in N_P$ e $a, b \in N_I$.

Riportiamo un esempio di ontologia. Siano HumanBeing, $Mortal \in N_C$, $teacherOf \in N_P$, Socrate, $Platone \in N_I$.

```
 \mathcal{O} = \begin{cases} \textit{HumanBeing} \sqsubseteq \textit{Mortal}, \\ \textit{range}(\textit{teacherOf}) \sqsubseteq \textit{HumanBeing}, \\ \textit{HumanBeing}(\textit{Socrate}), \\ \textit{Socrate teacherOf Platone} \end{cases}
```

Riportiamo un esempio di ontologia. Siano HumanBeing, $Mortal \in N_C$, $teacherOf \in N_P$, Socrate, $Platone \in N_I$.

Mediante reasoning è possibile esplicitare ulteriori affermazioni.

```
 \mathcal{O} = \{ \begin{aligned} & \mathsf{HumanBeing} \sqsubseteq \mathsf{Mortal}, & \mathcal{O}' = & \{ \mathsf{Mortal}(\mathsf{Socrate}), \\ & \mathsf{range}(teacherOf) \sqsubseteq \mathsf{HumanBeing}, & \Longrightarrow \\ & \mathsf{HumanBeing}(\mathsf{Socrate}), \\ & Socrate\ teacherOf\ Platone \} \end{aligned}
```

Riportiamo un esempio di ontologia. Siano HumanBeing, $Mortal \in N_C$, $teacherOf \in N_P$, Socrate, $Platone \in N_I$.

Mediante reasoning è possibile esplicitare ulteriori affermazioni.

```
 \begin{array}{lll} \mathcal{O} = & \{ \textit{HumanBeing} \sqsubseteq \textit{Mortal}, & \mathcal{O}' = & \{ \textit{Mortal}(\textit{Socrate}), \\ & \textit{range}(\textit{teacherOf}) \sqsubseteq \textit{HumanBeing}, & \Longrightarrow & \textit{HumanBeing}(\textit{Platone}), \\ & \textit{HumanBeing}(\textit{Socrate}), & & \\ & \textit{Socrate teacherOf Platone} \} \\ \end{array}
```

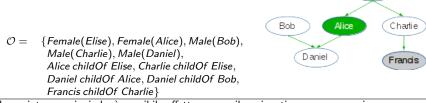
Riportiamo un esempio di ontologia. Siano HumanBeing, $Mortal \in N_C$, $teacherOf \in N_P$, Socrate, $Platone \in N_I$.

Mediante reasoning è possibile esplicitare ulteriori affermazioni.

```
 \mathcal{O} = \begin{cases} \mathsf{HumanBeing} \sqsubseteq \mathsf{Mortal} & \mathcal{O}' = \\ \mathsf{range}(\mathsf{teacherOf}) \sqsubseteq \mathsf{HumanBeing}, & \Longrightarrow & \mathsf{HumanBeing}(\mathsf{Platone}) \\ \mathsf{HumanBeing}(\mathsf{Socrate}), & \mathsf{Mortal}(\mathsf{Platone}) \end{cases}
```

Interrogazioni

Il metodo più immediato per ottenere informazioni da una ontologia è il *Conjunctive Query Answering*. Consideriamo ad esempio la seguente ontologia:



Alcune interrogazioni che è possibile effettuare con il conjunctive query answering sono:

- "Trova tutti gli individui maschi."
- "Chi sono gli individui con almeno un figlio maschio?"
- "Chi sono i figli di Alice?"
- "Chi sono gli individui con almeno un figlio maschio ed una femmina?"
- "Chi sono gli individui maschi con almeno un figlio maschio?"

Elise

Formule Atomiche

Per definire in maniera rigorosa le query congiuntive è necessario definire preliminarmente l'insieme delle *formule atomiche*.

Sia $V = \{x, y, z, ...\}$ l'insieme infinito, numerabile e disgiunto da N_C , N_P e N_I delle variabili. Le formule atomiche sono espressioni dei due seguenti tipi:

$$C(x)$$
, $P(x,y)$

con $x, y \in N_I \cup V$, $C \in N_C$ e $P \in N_P$.

Esempi di formule atomiche sono:

- HumanBeing(x),
- x childOf Alice,
- Bob childOf x,
- x childOf y,
- Mortal(Socrate),
- Alice childOf Elise

con HumanBeing, Mortal $\in N_C$, childOf $\in N_P$, Alice, Bob, Elise $\in N_I$ e $x, y \in V$.

Formule Atomiche Chiuse

Una formula atomica nella quale non compaiano variabili si dice chiusa.

Negli esempi che seguono sono evidenziate le formule atomiche chiuse:

- HumanBeing(x),
- x childOf Alice,
- Bob childOf x,
- x childOf y,
- Mortal(Socrate),
- Alice childOf Elise

 $con \ \textit{HumanBeing}, \textit{Mortal} \in \textit{N}_{\textit{C}}, \ \textit{childOf} \in \textit{N}_{\textit{P}}, \ \textit{Alice}, \textit{Bob}, \textit{Elise} \in \textit{N}_{\textit{I}} \ \textit{e} \ \textit{x}, \textit{y} \in \textit{V}.$

Formule Atomiche Chiuse

Una formula atomica nella quale non compaiano variabili si dice chiusa.

Negli esempi che seguono sono evidenziate le formule atomiche chiuse:

- HumanBeing(x),
- x childOf Alice,
- Bob childOf x,
- x childOf y,
- Mortal(Socrate),
- Alice childOf Elise

 $con\ \textit{HumanBeing}, \textit{Mortal} \in \textit{N}_{\textit{C}}, \ \textit{childOf} \in \textit{N}_{\textit{P}}, \ \textit{Alice}, \textit{Bob}, \textit{Elise} \in \textit{N}_{\textit{I}} \ e \ x, y \in \textit{V}.$

Le asserzioni presenti nelle ontologie sono formule atomiche chiuse.

Query Congiuntive

Una query congiuntiva è una congiunzione finita di formule atomiche $T_1 \wedge \ldots \wedge T_n$.

Alcuni esempi di query congiuntive:

"Trova tutti gli individui maschi."

"Chi sono gli individui con almeno un figlio maschio?"

$$y \ childOf \ x \ \land \ Male(y)$$

• "Chi sono i figli di Alice?"

con $x, y \in V$, $Male, Female \in N_C$, $childOf \in N_P$ e $Alice \in N_I$.

Per definire le soluzioni (risposte) delle query congiuntive introduciamo la nozione di sostituzione.

Una sostituzione $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ $(x_1, \dots, x_n \in V, a_1, \dots, a_n \in N_I)$ è una mappa finita che associa nomi di individui a variabili.

Sia T una formula atomica e $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ una sostituzione. L'applicazione $T\sigma$ di σ a T è la formula atomica che si ottiene sostituendo in T ad ogni occorrenza della variabile x_i il corrispondente nome di individuo a_i , per ogni $1 \le i \le n$.

Alcuni esempi:

$$\begin{array}{lll} \mathit{Male}(x)[x \to \mathit{Bob}] & = & \mathit{Male}(\mathit{Bob}) \\ \mathit{Male}(x)[y \to \mathit{Bob}] & = & \\ (x \, \mathit{childOf} \, y)[x \to \mathit{Alice}] & = & \\ (x \, \mathit{childOf} \, y)[x \to \mathit{Alice}, y \to \mathit{Elise}] & = & \\ \end{array}$$

Per definire le soluzioni (risposte) delle query congiuntive introduciamo la nozione di sostituzione.

Una sostituzione $\sigma = [x_1 \to a_1, \dots, x_n \to a_n] \ (x_1, \dots, x_n \in V, \ a_1, \dots, a_n \in N_I)$ è una mappa finita che associa nomi di individui a variabili.

Sia T una formula atomica e $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ una sostituzione. L'applicazione $T\sigma$ di σ a T è la formula atomica che si ottiene sostituendo in T ad ogni occorrenza della variabile x_i il corrispondente nome di individuo a_i , per ogni $1 \le i \le n$.

Alcuni esempi:

$$\begin{array}{lll} \mathit{Male}(x)[x \to \mathit{Bob}] & = & \mathit{Male}(\mathit{Bob}) \\ \mathit{Male}(x)[y \to \mathit{Bob}] & = & \mathit{Male}(x) \\ (x \mathit{childOf} \ y)[x \to \mathit{Alice}] & = & \\ (x \mathit{childOf} \ y)[x \to \mathit{Alice}, y \to \mathit{Elise}] & = & \end{array}$$

Per definire le soluzioni (risposte) delle query congiuntive introduciamo la nozione di sostituzione.

Una sostituzione $\sigma = [x_1 \to a_1, \dots, x_n \to a_n] \ (x_1, \dots, x_n \in V, \ a_1, \dots, a_n \in N_I)$ è una mappa finita che associa nomi di individui a variabili.

Sia T una formula atomica e $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ una sostituzione. L'applicazione $T\sigma$ di σ a T è la formula atomica che si ottiene sostituendo in T ad ogni occorrenza della variabile x_i il corrispondente nome di individuo a_i , per ogni $1 \le i \le n$.

Alcuni esempi:

$$\begin{array}{lll} \mathit{Male}(x)[x \to Bob] & = & \mathit{Male}(Bob) \\ \mathit{Male}(x)[y \to Bob] & = & \mathit{Male}(x) \\ (x \mathit{childOf} \ y)[x \to \mathit{Alice}] & = & \mathit{Alice} \ \mathit{childOf} \ y \\ (x \mathit{childOf} \ y)[x \to \mathit{Alice}, y \to \mathit{Elise}] & = & \end{array}$$

Per definire le soluzioni (risposte) delle query congiuntive introduciamo la nozione di sostituzione.

Una sostituzione $\sigma = [x_1 \to a_1, \dots, x_n \to a_n] \ (x_1, \dots, x_n \in V, \ a_1, \dots, a_n \in N_I)$ è una mappa finita che associa nomi di individui a variabili.

Sia T una formula atomica e $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ una sostituzione. L'applicazione $T\sigma$ di σ a T è la formula atomica che si ottiene sostituendo in T ad ogni occorrenza della variabile x_i il corrispondente nome di individuo a_i , per ogni $1 \le i \le n$.

Alcuni esempi:

$$\begin{array}{lll} \mathit{Male}(x)[x \to Bob] & = & \mathit{Male}(Bob) \\ \mathit{Male}(x)[y \to Bob] & = & \mathit{Male}(x) \\ (x \mathit{childOf} \ y)[x \to \mathit{Alice}] & = & \mathit{Alice} \ \mathit{childOf} \ y) \\ (x \mathit{childOf} \ y)[x \to \mathit{Alice}, y \to \mathit{Elise}] & = & \mathit{Alice} \ \mathit{childOf} \ \mathit{Elise} \end{array}$$

L'applicazione di sostituzioni a query congiuntive si definisce come segue.

Sia
$$\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$$
 una sostituzione e siano T_1, \dots, T_m formule atomiche. Allora
$$(T_1 \wedge \dots \wedge T_m)\sigma =_{\mathsf{Def}} T_1 \sigma \wedge \dots \wedge T_m \sigma.$$

$$\begin{array}{ll} (y \ childOf \ x \ \land \ Male(y))[x \to Alice] &=_{\mathsf{Def}} \\ (y \ childOf \ x \ \land \ Male(y))[x \to Alice, z \to Bob] &=_{\mathsf{Def}} \\ (y \ childOf \ x \ \land \ Male(y))[x \to Alice, y \to Bob] &=_{\mathsf{Def}} \end{array}$$

L'applicazione di sostituzioni a query congiuntive si definisce come segue.

Sia
$$\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$$
 una sostituzione e siano T_1, \dots, T_m formule atomiche. Allora
$$(T_1 \wedge \dots \wedge T_m)\sigma =_{\mathsf{Def}} T_1 \sigma \wedge \dots \wedge T_m \sigma.$$

$$\begin{array}{lll} (y \ childOf \ x \land \ Male(y))[x \rightarrow Alice] & =_{\mathsf{Def}} & y \ childOf \ Alice \land \ Male(y) \\ (y \ childOf \ x \land \ Male(y))[x \rightarrow Alice, z \rightarrow Bob] & =_{\mathsf{Def}} \\ (y \ childOf \ x \land \ Male(y))[x \rightarrow Alice, y \rightarrow Bob] & =_{\mathsf{Def}} \end{array}$$

L'applicazione di sostituzioni a query congiuntive si definisce come segue.

Sia
$$\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$$
 una sostituzione e siano T_1, \dots, T_m formule atomiche. Allora
$$(T_1 \wedge \dots \wedge T_m)\sigma =_{\mathsf{Def}} T_1 \sigma \wedge \dots \wedge T_m \sigma.$$

$$\begin{array}{lll} (y \ childOf \ x \land \ Male(y))[x \rightarrow Alice] & =_{\mathsf{Def}} & y \ childOf \ Alice \land \ Male(y) \\ (y \ childOf \ x \land \ Male(y))[x \rightarrow Alice, z \rightarrow Bob] & =_{\mathsf{Def}} & y \ childOf \ Alice \land \ Male(y) \\ (y \ childOf \ x \land \ Male(y))[x \rightarrow Alice, y \rightarrow Bob] & =_{\mathsf{Def}} \end{array}$$

L'applicazione di sostituzioni a query congiuntive si definisce come segue.

Sia
$$\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$$
 una sostituzione e siano T_1, \dots, T_m formule atomiche. Allora
$$(T_1 \wedge \dots \wedge T_m)\sigma =_{\mathsf{Def}} T_1 \sigma \wedge \dots \wedge T_m \sigma.$$

```
 \begin{array}{lll} (y \ childOf \ x \ \land \ Male(y))[x \to Alice] & =_{\mathsf{Def}} & y \ childOf \ Alice \ \land \ Male(y) \\ (y \ childOf \ x \ \land \ Male(y))[x \to Alice, z \to Bob] & =_{\mathsf{Def}} & y \ childOf \ Alice \ \land \ Male(y) \\ (y \ childOf \ x \ \land \ Male(y))[x \to Alice, y \to Bob] & =_{\mathsf{Def}} & Bob \ childOf \ Alice \ \land \ Male(Bob) \\ \end{array}
```

Soluzioni per una Query

Siano $\sigma=[x_1\to a_1,\ldots,x_n\to a_n]$ una sostituzione, $Q=T_1\wedge\ldots\wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad \mathcal{O} se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in \mathcal{O} .

Consideriamo l'ontologia $\mathcal O$ e la query Q definite come segue:

$$Q = y \ childOf \times \wedge \ Male(y).$$

Sia $\sigma_1 = [x \to Alice, y \to Daniel]$. σ_1 è una soluzione per Q rispetto ad \mathcal{O} ?

Soluzioni per una Query - Esempio 1

Siano $\sigma=[x_1\to a_1,\ldots,x_n\to a_n]$ una sostituzione, $Q=T_1\wedge\ldots\wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad $\mathcal O$ se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in $\mathcal O$.

Consideriamo l'ontologia $\mathcal O$ e la query Q ("Chi sono gli individui con almeno un figlio maschio?") definite come segue:

$$Q = y \ childOf \ x \land Male(y).$$

Sia $\sigma_1 = [x \to Alice, y \to Daniel]$. σ_1 è una soluzione per Q rispetto ad \mathcal{O} ? SI.

$$Q\sigma_1 = Daniel \ childOf \ Alice \land Male(Daniel).$$

Siano $\sigma=[x_1\to a_1,\ldots,x_n\to a_n]$ una sostituzione, $Q=T_1\wedge\ldots\wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad \mathcal{O} se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in \mathcal{O} .

Consideriamo l'ontologia $\mathcal O$ e la query Q definite come segue:

$$Q = y \ childOf \times \wedge \ Male(y).$$

Sia $\sigma_2 = [x \to Alice, y \to Bob]$. σ_2 è una soluzione per Q rispetto ad \mathcal{O} ?

Siano $\sigma=[x_1\to a_1,\ldots,x_n\to a_n]$ una sostituzione, $Q=T_1\wedge\ldots\wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad $\mathcal O$ se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in $\mathcal O$.

Consideriamo l'ontologia $\mathcal O$ e la query Q ("Chi sono gli individui con almeno un figlio maschio?") definite come segue:

$$Q = y \ childOf \times \wedge \ Male(y).$$

Sia $\sigma_2 = [x \to Alice, y \to Bob]$. σ_2 è una soluzione per Q rispetto ad \mathcal{O} ? **NO**.

$$Q\sigma_2 =$$
Bob childOf Alice \land *Male*(*Bob*).

Siano $\sigma=[x_1\to a_1,\ldots,x_n\to a_n]$ una sostituzione, $Q=T_1\wedge\ldots\wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad \mathcal{O} se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in \mathcal{O} .

Consideriamo l'ontologia $\mathcal O$ e la query $\mathcal Q$ definite come segue:

$$Q = y \ childOf \times \wedge \ Male(y).$$

Sia $\sigma_3 = [x \to Charlie, y \to Francis]$. σ_3 è una soluzione per Q rispetto ad \mathcal{O} ?

Siano $\sigma=[x_1\to a_1,\ldots,x_n\to a_n]$ una sostituzione, $Q=T_1\wedge\ldots\wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad \mathcal{O} se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in \mathcal{O} .

Consideriamo l'ontologia $\mathcal O$ e la query Q ("Chi sono gli individui con almeno un figlio maschio?") definite come segue:

$$Q = y \ childOf \ x \land Male(y).$$

Sia $\sigma_3 = [x \to Charlie, y \to Francis]$. σ_2 è una soluzione per Q rispetto ad \mathcal{O} ? **NO**.

$$Q\sigma_3 = Francis \ childOf \ Charlie \land Male(Francis).$$

Siano $\sigma=[x_1\to a_1,\ldots,x_n\to a_n]$ una sostituzione, $Q=T_1\wedge\ldots\wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad \mathcal{O} se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in \mathcal{O} .

Consideriamo l'ontologia $\mathcal O$ e la query Q definite come segue:

$$Q = y \ childOf \ x \land Male(y).$$

Sia $\sigma_4 = [y \to Daniel]$. σ_4 è una soluzione per Q rispetto ad \mathcal{O} ?

Siano $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ una sostituzione, $Q = T_1 \wedge \dots \wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad \mathcal{O} se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in \mathcal{O} .

Consideriamo l'ontologia $\mathcal O$ e la query Q ("Chi sono gli individui con almeno un figlio maschio?") definite come segue:

$$Q = y \ childOf \ x \land Male(y).$$

Sia $\sigma_4 = [y \to Daniel]$. σ_2 è una soluzione per Q rispetto ad \mathcal{O} ? **NO**.

$$Q\sigma_4 =$$
Daniel childOf x \land Male(Daniel).

Affinchè una sostituzione σ sia una soluzione per una query Q (a prescindere dall'ontologia) è necessario che in σ compaiano tutte le variabili di Q.

Siano $\sigma=[x_1\to a_1,\ldots,x_n\to a_n]$ una sostituzione, $Q=T_1\wedge\ldots\wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad \mathcal{O} se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in \mathcal{O} .

Consideriamo l'ontologia $\mathcal O$ e la query Q definite come segue:

$$Q = y \ childOf \times \wedge \ Male(y).$$

Sia $\sigma_5 = [x \to \textit{Elise}, y \to \textit{Charlie}, z \to \textit{Francis}]. \ \sigma_5$ è una soluzione per Q rispetto ad \mathcal{O} ?

Siano $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ una sostituzione, $Q = T_1 \wedge \dots \wedge T_m$ una query congiuntiva e $\mathcal O$ una ontologia.

 σ è detta essere una *soluzione* per Q rispetto ad \mathcal{O} se e solo se $T_1\sigma,\ldots,T_2\sigma$ compaiono in \mathcal{O} .

Consideriamo l'ontologia $\mathcal O$ e la query Q ("Chi sono gli individui con almeno un figlio maschio?") definite come segue:

$$Q = y \ childOf \ x \land Male(y).$$

Sia $\sigma_5 = [x \to \textit{Elise}, y \to \textit{Charlie}, z \to \textit{Francis}]. \ \sigma_5$ è una soluzione per Q rispetto ad \mathcal{O} ? SI.

$$Q\sigma_5 = Charlie \ childOf \ Elise \land Male(Charlie).$$

Soluzioni Minimali per una Query

Siano $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ una sostituzione, Q una query congiuntiva e $\mathcal O$ una ontologia.

 σ è una soluzione minimale per Q rispetto a $\mathcal O$ se e solo se:

- lacktriangledown σ è una soluzione per Q rispetto ad $\mathcal O$ e inoltre
- ② tutte le variabili x_1, \ldots, x_n che compaiono in σ compaiono anche in Q (criterio di minimalità).

Consideriamo ad esempio

$$Q = y \ childOf \times \wedge \ Male(y).$$

 $\sigma_5 = [x \to \textit{Elise}, y \to \textit{Charlie}, z \to \textit{Francis}]$ è una soluzione minimale per Q rispetto a \mathcal{O} ?

Soluzioni Minimali per una Query

Siano $\sigma = [x_1 \to a_1, \dots, x_n \to a_n]$ una sostituzione, Q una query congiuntiva e $\mathcal O$ una ontologia.

 σ è una soluzione minimale per Q rispetto a $\mathcal O$ se e solo se:

- lacktriangledown σ è una soluzione per Q rispetto ad $\mathcal O$ e inoltre
- ② tutte le variabili x_1, \ldots, x_n che compaiono in σ compaiono anche in Q (criterio di minimalità).

Consideriamo ad esempio

$$Q = y \ childOf \times \wedge \ Male(y).$$

 $\sigma_5 = [x \to Elise, y \to Charlie, z \to Francis]$ è una soluzione minimale per Q rispetto a \mathcal{O} ? **NO**.

$$\sigma_6 = [x \to Elise, y \to Charlie]$$
 è una soluzione minimale per Q rispetto a \mathcal{O} ?

Soluzioni Minimali per una Query

Siano $\sigma=[x_1\to a_1,\dots,x_n\to a_n]$ una sostituzione, Q una query congiuntiva e $\mathcal O$ una ontologia.

 σ è una soluzione minimale per Q rispetto a $\mathcal O$ se e solo se:

- lacktriangledown σ è una soluzione per Q rispetto ad $\mathcal O$ e inoltre
- ② tutte le variabili x_1, \ldots, x_n che compaiono in σ compaiono anche in Q (criterio di minimalità).

Consideriamo ad esempio

$$Q = y \ childOf \times \wedge \ Male(y).$$

 $\sigma_5 = [x \to Elise, y \to Charlie, z \to Francis]$ è una soluzione minimale per Q rispetto a \mathcal{O} ? **NO**.

 $\sigma_6 = [x \to \textit{Elise}, y \to \textit{Charlie}]$ è una soluzione minimale per Q rispetto a \mathcal{O} ? SI.

Conjunctive Query Answering

Il problema del *Conjunctive Query Answering* consiste nel trovare tutte le soluzioni minimali di una query congiuntiva rispetto ad una ontologia.

Esse sono sempre in numero finito, infatti:

- le variabili che compaiono nelle soluzioni sono esattamente quelle che compaiono nella query,
- i nomi di individui che compaiono nelle soluzioni sono un sottoinsieme di quelli che compaiono nell'ontologia.

"Trova tutti gli individui maschi."

$$Q = Male(x)$$



"Trova tutti gli individui maschi."

$$Q = Male(x)$$



"Trova tutti gli individui maschi."

$$Q = Male(x)$$



"Trova tutti gli individui maschi."

$$Q = Male(x)$$



"Chi sono gli individui con almeno un figlio maschio?"

$$Q = y \ childOf \ x \ \land \ Male(y)$$

X	у

"Chi sono gli individui con almeno un figlio maschio?"

$$Q = y \ childOf \ x \land Male(y)$$

X	У	
Elise	Charlie	

"Chi sono gli individui con almeno un figlio maschio?"

$$Q = y \ childOf \ x \ \land \ Male(y)$$

X	у
Elise	Charlie
Alice	Daniel

"Chi sono gli individui con almeno un figlio maschio?"

$$Q = y \ childOf \ x \ \land \ Male(y)$$

X	у
Elise	Charlie
Alice	Daniel
Bob	Daniel

"Chi sono i figli di *Elise*?" Q = x childOf Elise



"Chi sono i figli di *Elise*?" Q = x childOf Elise



"Chi sono i figli di *Elise*?" Q = x childOf Elise

O = {Female(Elise), Female(Alice), Male(Bob),
 Male(Charlie), Male(Daniel),
 Alice childOf Elise, Charlie childOf Elise,
 Daniel childOf Alice, Daniel childOf Bob,
 Francis childOf Charlie}

X Alice Charlie

"Chi sono gli individui con almeno un figlio maschio ed una femmina?"

Х	у	Z

"Chi sono gli individui con almeno un figlio maschio ed una femmina?"

$$Q = y \ childOf \ x \land z \ childOf \ x \land Male(y) \land Female(z)$$

Х	у	Z

"Chi sono gli individui con almeno un figlio maschio ed una femmina?"

$$Q = y \ childOf \ x \land z \ childOf \ x \land Male(y) \land Female(z)$$

X	у	Z
Elise	Charlie	Alice
Elise	Charlie	Alice

"Chi sono gli individui maschi con almeno una figlia femmina?"



"Chi sono gli individui maschi con almeno una figlia femmina?"

$$Q = Male(x) \land y childOf x \land Female(y)$$

у

"Chi sono gli individui maschi con almeno una figlia femmina?"

$$Q = Male(x) \land y childOf x \land Female(y)$$

 $\mathcal{O} = \{Female(Elise), Female(Alice), Male(Bob), \\ Male(Charlie), Male(Daniel), \\ Alice childOf Elise, Charlie childOf Elise, \\ Daniel childOf Alice, Daniel childOf Bob, \\ Francis childOf Charlie \}$ Nessuna Soluzione

Vocabolari

Classi e proprietà vengono raggruppati in *vocabolari* che trattano specifici domini di conoscenza (eg. organizzazioni, pubblica amministrazione, biologia, commercio, etc.).

Un vocabolario può contenere anche alcuni vincoli sulle classi e le proprietà del vocabolario stesso.

Definizione di Vocabolario

Una definizione di vocabolario può essere la seguente:

$$V=(C,P,\Omega)$$

dove

- **①** C è un sottoinsieme finito di N_C ,
- ② P è un sottoinsieme finito di N_P ,
- $\ \Omega$ è un insieme finito di vincoli che coinvolgano solo nomi di classi in C e nomi di proprietà in P.

Vocabolari condivisi

L'utilizzo di vocabolari condivisi (ben noti) favorisce la scalabilità orizzontale delle applicazioni.

Ad esempio, una applicazione sviluppata sull'ontologia di un comune che utilizzi i vocabolari standard per le pubbliche amministrazioni (vedi le *Linee Guida per la Valorizzazione del Patrimonio Informativo Pubblico* dell'*Agenzia per l'Italia Digitale*) può essere estesa senza sforzi aggiuntivi per utilizzare i dati provenienti dalle ontologie di tutti i comuni.

II Vocabolario FOAF

Uno dei primi e più utilizzati vocabolari definiti nell'ambito del Web semantico è Friend OF A Friend (FOAF, vedi http://foaf-project.org).

FOAF is a project devoted to linking people and information using the Web.

In questa sede ci limiteremo solo alla parte Core.

Core - These classes and properties form the core of FOAF. They describe characteristics of people and social groups that are independent of time and technology; as such they can be used to describe basic information about people in present day, historical, cultural heritage and digital library contexts. In addition to various characteristics of people, FOAF defines classes for Project, Organization and Group as other kinds of agent. Related work:

(tratto da FOAF Vocabulary Specification 0.99, Namespace Document 14 January 2014, Paddington Edition, http://xmlns.com/foaf/spec/)

FOAF Core

Il vocabolario Foaf Core è definito come segue:

FOAFCore
$$=_{\mathsf{Def}}$$
 $(C_{foaf}, P_{foaf}, \Omega_{foaf})$

$$C_{foaf} =_{\mathsf{Def}} \{Agent, Person, Project, Organization, Group, Document, Image\}$$

$$P_{foaf} =_{\mathsf{Def}} \{name, title, img, depiction, depicts, familyName, givenName, based_near, age, made, maker, primaryTopic, primaryTopicOf, member}$$

$$\Omega_{foaf} =_{\mathsf{Def}} \{Person \sqsubseteq Agent, Group \sqsubseteq Agent, Organization \sqsubseteq Agent, Image \sqsubseteq Document, dom(title) \sqsubseteq Document, range(depiction) \sqsubseteq Image, img \sqsubseteq depiction, dom(img) \sqsubseteq Person, dom(knows) \sqsubseteq Person, range(knows) \sqsubseteq Person, \ldots\}$$

Descrizioni Intuitive degli Elementi dei Vocabolari

Le classi e le proprietà di un vocabolario vengono spesso fornite di una descrizione intuitiva nel documento che descrive il vocabolario. Ad esempio, le classi *Agent* e *Person* vengono descritte come segue in http://xmlns.com/foaf/spec/

Agent - The Agent class is the class of agents; things that do stuff. A well known sub-class is Person, representing people. Other kinds of agents include Organization and Group.

The Agent class is useful in a few places in FOAF where Person would have been overly specific. For example, the IM chat ID properties such as jabberID are typically associated with people, but sometimes belong to software bots.

Person - The Person class represents people. Something is a Person if it is a person. We don't nitpic about whether they're alive, dead, real, or imaginary. The Person class is a sub-class of the Agent class, since all people are considered 'agents' in FOAF.

Descrizioni Rigorose degli Elementi dei Vocabolari

Tuttavia, già nei vincoli di un vocabolario si trovano indicazioni importanti sulla semantica dei nomi di classe e di proprietà del vocabolario stesso.

```
\begin{array}{l} \textit{Person} \sqsubseteq \textit{Agent} \\ \textit{range}(\textit{depiction}) \sqsubseteq \textit{Image} \\ \textit{img} \sqsubseteq \textit{depiction}, \\ \textit{dom}(\textit{img}) \sqsubseteq \textit{Person}, \\ \textit{dom}(\textit{knows}) \sqsubseteq \textit{Person}, \\ \textit{range}(\textit{knows}) \sqsubseteq \textit{Person}, \\ \end{array}
```

Vocabolari Compositi

È possibile costruire un vocabolario estendendone un'altro. Ad esempio, il vocabolario Organization Ontology (in breve ORG, vedi http://www.w3.org/TR/vocab-org/) estende FOAF con classi e proprietà, per modellare in dettaglio le strutture organizzative di aziende, associazioni e tutte le forme di organizzazioni.

Formalmente, dati due vocabolari

$$egin{array}{lll} V & =_{\mathsf{Def}} & (C,P,\Omega) \ V' & =_{\mathsf{Def}} & (C',P',\Omega') \end{array}$$

si dice che V' importa (o estende) V se e solo se

$$\begin{array}{ccc} C & \subseteq & C' \\ P & \subseteq & P' \\ \Omega & \subseteq & \Omega' \end{array}$$

Organization Ontology, in breve ORG, 2 estende FOAF con classi e proprietà, per modellare in dettaglio le strutture organizzative di aziende, associazioni e tutte le forme di organizzazioni.

Il vocabolario è disponibile alle URL

```
http://www.w3.org/ns/org.rdf
http://www.w3.org/ns/org.ttl
```

II namespace del vocabolario ORG è

org: http://www.w3.org/ns/org#

```
\begin{array}{ll} \text{ORG} & =_{\mathsf{Def}} & \left( C_{org}, P_{org}, \Omega_{org} \right) \\ \\ C_{org} & =_{\mathsf{Def}} & C_{foaf} \cup \{FormalOrganization, OrganizationalUnit, Site \\ & Post, Role, \ldots \} \\ \\ P_{org} & =_{\mathsf{Def}} & P_{foaf} \cup \{basedAt, classification, hasPost, hasPrimarySite, \\ & hasRegisteredSite, hasSite, hasSubOrganization, \\ & hasUnit, headOf, heldBy, holds, location, postIn, \\ & purpose, role, siteAddress, siteOf, unitOf, \\ & subOrganizationOf, transitiveSubOrganizationOf, \ldots \} \end{array}
```

Organization Ontology, in breve ORG, ² estende FOAF con classi e proprietà, per modellare in dettaglio le strutture organizzative di aziende, associazioni e tutte le forme di organizzazioni.

Il vocabolario è disponibile alle URL

```
http://www.w3.org/ns/org.rdf
http://www.w3.org/ns/org.ttl
```

Il namespace del vocabolario ORG è

org: http://www.w3.org/ns/org#

```
 \begin{array}{ll} \text{ORG} & =_{\mathsf{Def}} & (C_{org}, P_{org}, \Omega_{org}) \\ \\ C_{org} & =_{\mathsf{Def}} & C_{foaf} \cup \{FormalOrganization, OrganizationalUnit, Site \\ Post, Role, \ldots\} \\ \\ P_{org} & =_{\mathsf{Def}} & P_{foaf} \cup \{basedAt, classification, hasPost, hasPrimarySite, \\ & hasRegisteredSite, hasSite, hasSubOrganization, \\ & hasUnit, headOf, heldBy, holds, location, postIn, \\ & purpose, role, siteAddress, siteOf, unitOf, \\ & subOrganizationOf, transitiveSubOrganizationOf, \ldots\} \\ \end{array}
```

Organization Ontology, in breve ORG, ² estende FOAF con classi e proprietà, per modellare in dettaglio le strutture organizzative di aziende, associazioni e tutte le forme di organizzazioni.

Il vocabolario è disponibile alle URL

```
http://www.w3.org/ns/org.rdf
http://www.w3.org/ns/org.ttl
```

Il namespace del vocabolario ORG è

 $=_{\mathsf{Def}} (C_{\mathsf{org}}, P_{\mathsf{org}}, \Omega_{\mathsf{org}})$

ORG

$$\mathcal{C}_{org} =_{\mathsf{Def}} \mathcal{C}_{foaf} \cup \{Formal Organization, Organizational Unit, Site Post, Role, \}$$

$$P_{org} = =_{\mathsf{Def}} P_{foaf} \cup \{basedAt, classification, hasPost, hasPrimarySite, \\ hasRegisteredSite, hasSite, hasSubOrganization, \\ hasUnit, headOf, heldBy, holds, location, postIn, \\ purpose, role, siteAddress, siteOf, unitOf, \\ subOrganizationOf, transitiveSubOrganizationOf, \ldots \}$$

 $=_{\mathsf{Def}} (C_{\mathsf{org}}, P_{\mathsf{org}}, \Omega_{\mathsf{org}})$

ORG

Ad esempio, il vocabolario *Organization Ontology* (in breve *ORG*, vedi http://www.w3.org/TR/vocab-org/) estende FOAF con classi e proprietà, per modellare in dettaglio le strutture organizzative di aziende, associazioni e tutte le forme di organizzazioni.

$$C_{org} = =_{\mathsf{Def}} \quad C_{foaf} \cup \{FormalOrganization, OrganizationalUnit, Site \\ Post, Role, \ldots\}$$

$$P_{org} = =_{\mathsf{Def}} \quad P_{foaf} \cup \{basedAt, classification, hasPost, hasPrimarySite, \\ hasRegisteredSite, hasSite, hasSubOrganization, \\ hasUnit, headOf, heldBy, holds, location, postIn, \\ purpose, role, siteAddress, siteOf, unitOf, \\ subOrganizationOf, transitiveSubOrganizationOf, \ldots\}$$

$$\begin{array}{ll} \Omega_{\textit{org}} & =_{\mathsf{Def}} & \Omega_{\textit{foaf}} \cup \{ \textbf{FormalOrganization} \sqsubseteq \textbf{Organization}, \\ & \textbf{OrganizationalUnit} \sqsubseteq \textbf{Organization}, \\ & \textit{headOf} \sqsubseteq \textit{memberOf}, \texttt{dom}(\textit{hasUnit}) \sqsubseteq \textit{Organization}, \\ & \text{range}(\textit{hasUnit}) \sqsubseteq \textit{OrganizationalUnit}, \ldots \} \end{array}$$

NOTA: I vincoli semantici esplicitano la relazione tra i due vocabolari.



È buona norma definire ontologie facendo riferimento a vocabolari ben noti.

Definiamo ad esempio una ontologia per descrivere la struttura organizzativa del comune di Catania (vedi http://www.dmi.unict.it/~longo/comunect/) a partire dal vocabolario ORG.

Il sindaco del comune di Catania è Enzo Bianco

$$\mathcal{O}_{CT} = \{ \textit{EnzoBianco headOf ComuneCT} \}$$

È buona norma definire ontologie facendo riferimento a vocabolari ben noti.

Definiamo ad esempio una ontologia per descrivere la struttura organizzativa del comune di Catania (vedi http://www.dmi.unict.it/~longo/comunect/) a partire dal vocabolario ORG.

Fanno direttamente capo al comune il gabinetto del sindaco, la polizia municipale e la direzione affari legali.

È buona norma definire ontologie facendo riferimento a vocabolari ben noti.

Definiamo ad esempio una ontologia per descrivere la struttura organizzativa del comune di Catania (vedi http://www.dmi.unict.it/~longo/comunect/) a partire dal vocabolario ORG.

Il comune comprende inoltre diverse direzioni, raggruppate in due macro-aree.

$$\mathcal{O}_{CT} = \Sigma_{org} \cup \\ \{ \textit{EnzoBianco headOf ComuneCT}, \\ \textit{ComuneCT hasUnit GabinettoDelSindaco}, \\ \textit{ComuneCT hasUnit PoliziaMunicipale}, \\ \textit{ComuneCT hasUnit AffariLegali}, \\ \textit{ComuneCT hasUnit Area1}, \\ \textit{ComuneCT hasUnit Area2}, \\ \end{cases}$$

È buona norma definire ontologie facendo riferimento a vocabolari ben noti.

Definiamo ad esempio una ontologia per descrivere la struttura organizzativa del comune di Catania (vedi http://www.dmi.unict.it/~longo/comunect/) a partire dal vocabolario ORG.

Il comune comprende inoltre diverse direzioni, raggruppate in due macro-aree.

$$\mathcal{O}_{CT} = \Sigma_{org} \cup \\ \{Enzo Bianco\ head Of\ Comune CT, \\ Comune CT\ has Unit\ Gabinet to Del Sindaco, \\ Comune CT\ has Unit\ Polizia Municipale, \\ Comune CT\ has Unit\ Affari Legali, \\ Comune CT\ has Unit\ Area1, \\ Comune CT\ has Unit\ Area2, \\ Area1\ has Unit\ Risorse Umane, \\ Area1\ has Unit\ Famiglia, \\ Area1\ has Unit\ Ecologia, \\ Area1\ has Unit\ Lavori Pubblici. \\$$

È buona norma definire ontologie facendo riferimento a vocabolari ben noti.

Definiamo ad esempio una ontologia per descrivere la struttura organizzativa del comune di Catania (vedi http://www.dmi.unict.it/~longo/comunect/) a partire dal vocabolario ORG.

Il comune comprende inoltre diverse direzioni, raggruppate in due macro-aree.

 $\mathcal{O}_{CT} = \Sigma_{org} \cup \{ \text{EnzoBianco headOf ComuneCT}, \}$ ComuneCT hasUnit GabinettoDelSindaco. ComuneCT hasUnit PoliziaMunicipale. ComuneCT hasUnit AffariLegali, ComuneCT hasUnit Area1. ComuneCT hasUnit Area2. Area1 hasUnit RisorseUmane, Area1 hasUnit Istruzione. Area1 hasUnit Famiglia, Area1 hasUnit Ecologia, Area1 hasUnit Manutenzione. Area1 hasUnit LavoriPubblici, Area2 hasUnit Patrimonio. Area2 hasUnit ServiziDemografici. Area2 hasUnit Turismo, Area2 hasUnit Urbanistica. Area2 hasUnit AttivitaProduttive. Area2 hasUnit Economato }