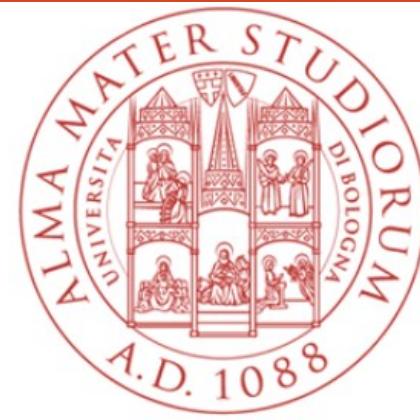


Corso di Laurea Magistrale in Ingegneria e Scienze Informatiche  
Università degli Studi di Bologna, sede di Cesena

# Open data, oggetti connessi e web semantico: dal giornalismo alla medicina di precisione

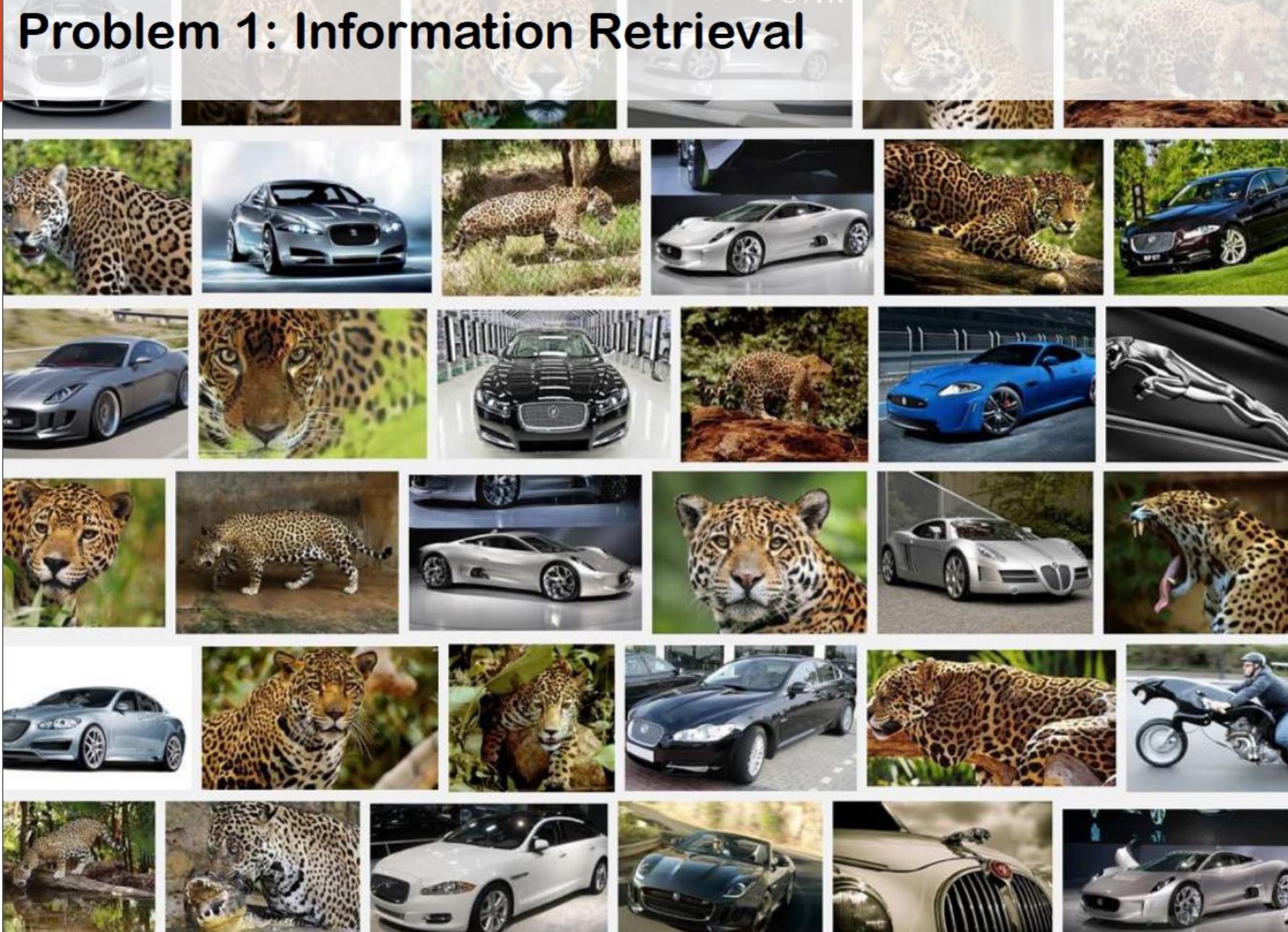
Antonella Carbonaro  
[antonella.carbonaro@unibo.it](mailto:antonella.carbonaro@unibo.it)

ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

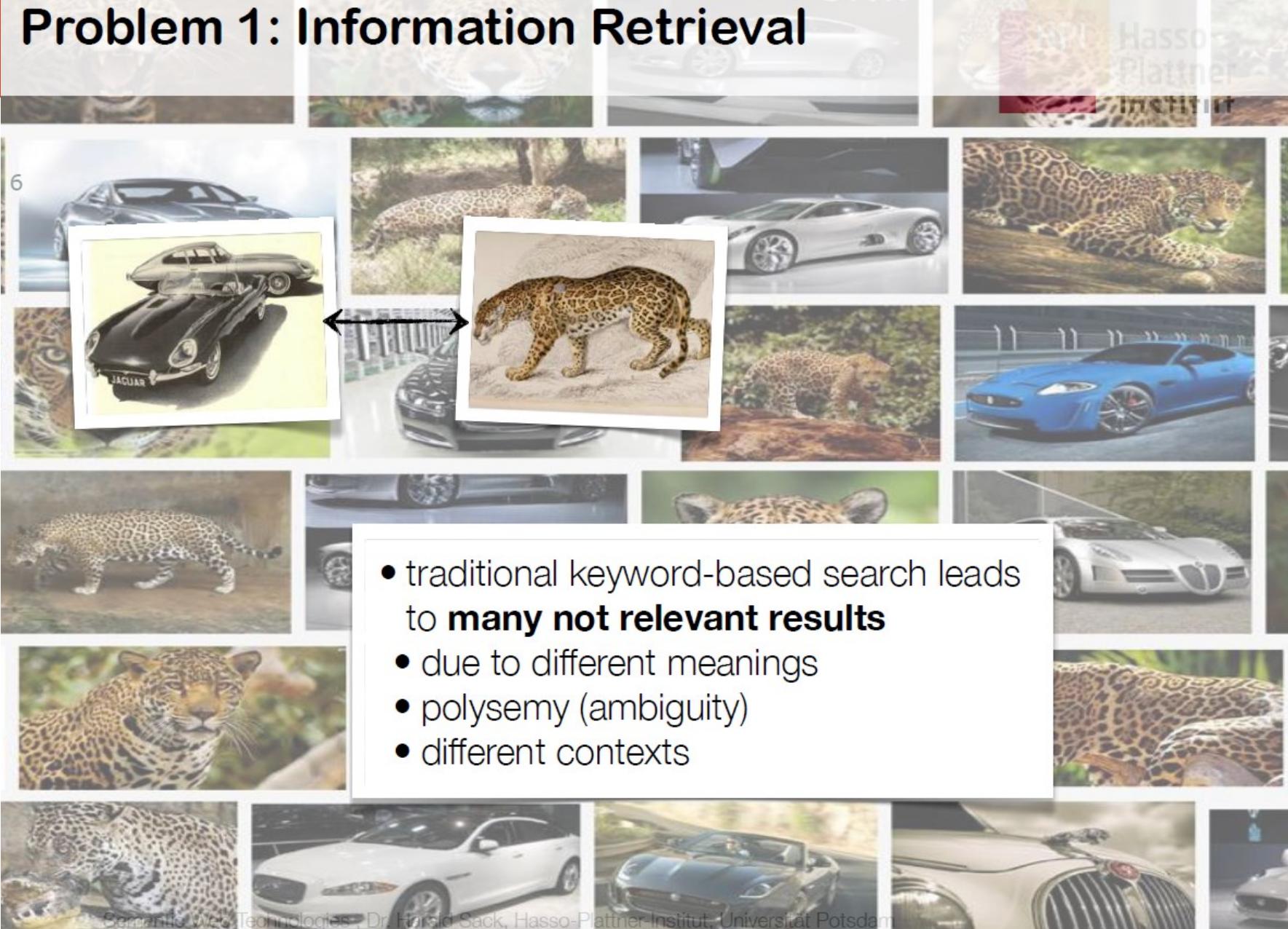


The problems are ...

# Problem 1: Information Retrieval



# Problem 1: Information Retrieval



- traditional keyword-based search leads to **many not relevant results**
  - due to different meanings
  - polysemy (ambiguity)
  - different contexts

# Problem 1: Information Retrieval

Plattner  
Institut

7

Ad related to "Panthera onca" ⓘ

[Lion, Tiger, Jaguar Pics - Panthera Works to Save Wild Cats](#)

[www.panthera.org/](#)

See Pictures of all the Big Cats

[Jaguar - Wikipedia, the free encyclopedia](#)

[en.wikipedia.org/wiki/Jaguar](#)

Jaguar

[www.iucnredlist.org/details/15953/0](#)

by A Caso - 2011 - Cited by 10 - Related art...

Panthera onca. Status\_ne\_off Status\_dd\_of

Status\_vu\_off Status\_en\_off Status\_cr\_off St

[ADW: Panthera onca: INFORMATION](#)

[animaldiversity.ummz.umich.edu/accounts/Panthera\\_onca/](#)

Panthera onca|jaguar. By Jonathan Nogueira. Geographic Range; Habitat; Physical Description; Reproduction; Lifespan / Longevity; Behavior; Communication ...

[Panthera | Jaguars](#)

[www.panthera.org/species/jaguar](#)

The jaguar (*Panthera onca*) is the third largest cat in the world, but it is the largest feline in the Americas.

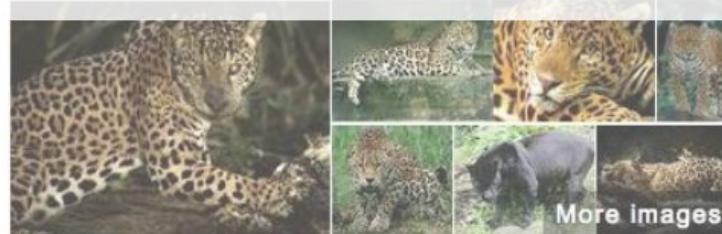
[Jaguar \(Panthera onca\) - thebigzoo.com](#)

[www.thebigzoo.com/Animals/Jaguar](#)  
The Jaguar is the largest cat in the Americas and the third largest cat in the world (after the Lion and the Tiger).

[Jaguar - Panthera onca - Nat](#)

[www.nhptv.org/natureworks/jaguar/](#)

The jaguar is the largest cat in North America and the third largest cat in the world. It stands about three feet tall from shoulders to feet, 6½ to 7½ feet long from ...



[More Images](#)

Jaguar

The jaguar is a big cat, a feline in the *Panthera* genus, and is the only extant New World member of the Panthera genus. It is the third largest feline found in the Americas. The jaguar is the third largest feline in the world, after the tiger and the lion, and the largest in the Western Hemisphere.

Scientific name: *Panthera onca*

Rank: Species

Higher classification: *Panthera*

[Feedback / More info](#)

- traditional keyword-based search
- does not find all results**
- synonyms and metaphors
- missing context definition

## Problem 2: Information Extraction

The image shows a collage of Japanese web pages from various sources, including a main page for 'Web 東奥' (Tohoku Web) and several sidebar snippets. A large red question mark is overlaid across the center of the collage.

**What does the information mean?**

Key elements visible in the collage:

- Top Left:** A sidebar for 'Web 東奥' with links to '県外在住者の方へ!!' (For people living outside the prefecture!), '東奥日報電子版' (Tohoku Mainichi Electronic Edition), and '長期訂閱あり' (Long-term subscription available).
- Top Center:** A banner for '47CLUB' with the text '50人が選んでみた! 「推し商品」はコレ!' (50 people selected! 'Pushi商品' is this!). Below it are links to categories like 天地人・社説・訃報・お悔やみ・宵宮, プロ野球・MLB・Jリーグ・大相撲・高校野球, and エンタメ・コラム・新刊本・新作映画・音楽CD・格闘技.
- Top Right:** A sidebar for 'Go ウェブ' with links to '個別店情報', '中古車ネットガーデン', 'ネット住宅展示場', 'メール申し込み', and 'Miniガイド'.
- Middle Left:** A sidebar for '東奥日報' with links to '新聞ご購読申し込み', '広告ご案内', 'ウェブ 新聞 折込', 'ドコモ・スマホサイト new', and 'ケータイ・スマホ情報'.
- Middle Center:** A news snippet: '速報 19:46 床下から2体目の遺体' (Breaking news 19:46: Two bodies found in the floorboards). Below it is a large headline for the '保育所対抗チャンバラ交流大会' (Nursery School Competition Chambala Exchange Conference) featuring a large sign that reads '第六回 保育所' (6th Annual Nursery School).
- Middle Right:** A sidebar for 'TNT 東奥NET テレビ' with a list of news items:
  - 風間浦で「同志社フェア」
  - 圓児らチャンバラ交流/板橋
  - 染工場など全焼、2人死亡
  - 津軽の食と農業まつり開幕
- Bottom Right:** A sidebar for 'ロレックス' (Rolex) labeled '正規品取扱店' (Authorized distributor) and another for 'マエバラ' (Maebara) with a link to '人材派遣なら 東奥日報人材センター アウル' (For staff派遣, Tohoku Mainichi Staffing Center, AUR).

## Problem 2: Information Extraction



9

What does the information mean?

## Problem 2: Information Extraction

The collage illustrates various scenarios where information extraction is challenging:

- Top Left:** A screenshot of a Japanese news website (東奥日報) featuring a headline about a product recommendation. The page includes links for sports (プロ野球, MLB, Jリーグ, 大相撲, 高校野球), entertainment (エンタメ), and media (コラム, 新刊本, 新作映画, 音楽CD, 格闘技).
- Top Right:** A screenshot of a website for 'Go Web' (ゴーウェブ) with a navigation bar for '中古車ネットガーデン' (Used car network), 'ネット住宅展示場' (Network home exhibition hall), 'f メール申し込み' (Email application), and 'Mini ガイド' (Mini guide).
- Middle Left:** A screenshot of the 'Information Extraction' section of the slide, listing the following points:

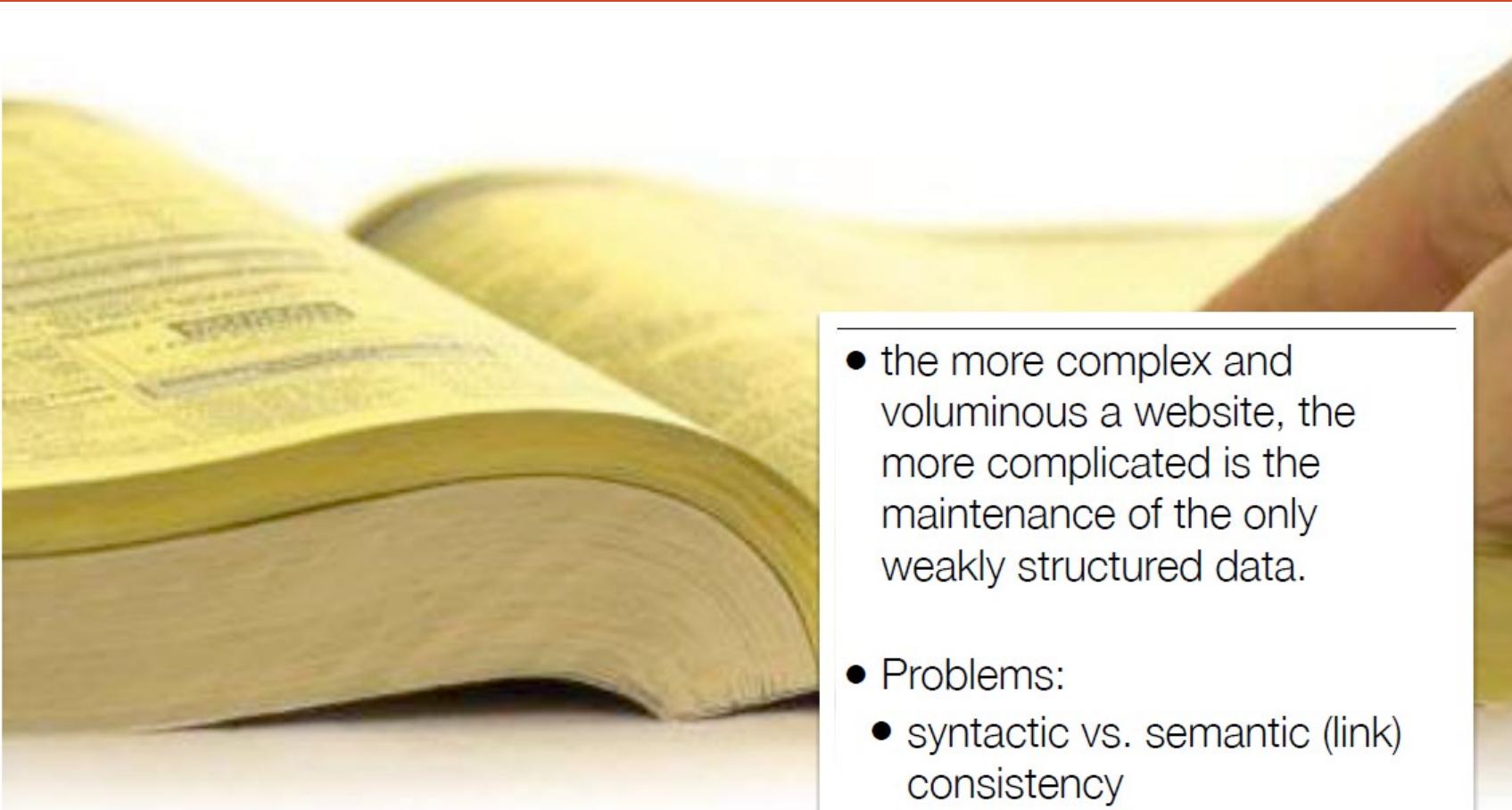
  - can only be solved ‘correctly’ by a human agent
  - heterogeneous distribution and order of information
  - a software agent does not have
    - sufficient knowledge of contexts
    - sufficient world knowledge and
    - sufficient experience

- Middle Center:** A photograph of a sign for '第六回 保育所' (6th保育所) (Nursery School) with a large black arrow pointing right.
- Middle Right:** A screenshot of a news article from 'GO WEB' with several bullet points:
  - 風間浦で「同志社フェア」
  - 園児らチャンバラ交流／板橋
  - 染工場など全焼、2人死亡
  - 津軽の食と産業まつり開幕
- Bottom Left:** A screenshot of a mobile app or website for 'おでかけ' (Local travel) with sections for 'お店検索' (Store search), 'イベント' (Events), and '展覧会' (Exhibitions).
- Bottom Right:** A screenshot of a news article from 'GO WEB' featuring a house image and the word '風家' (Kaze-no-ka).

## Problem 2: Information Extraction

- **implicit knowledge**, i.e. information does not have to be specified explicitly, but must be derived via logical deductions from available information.



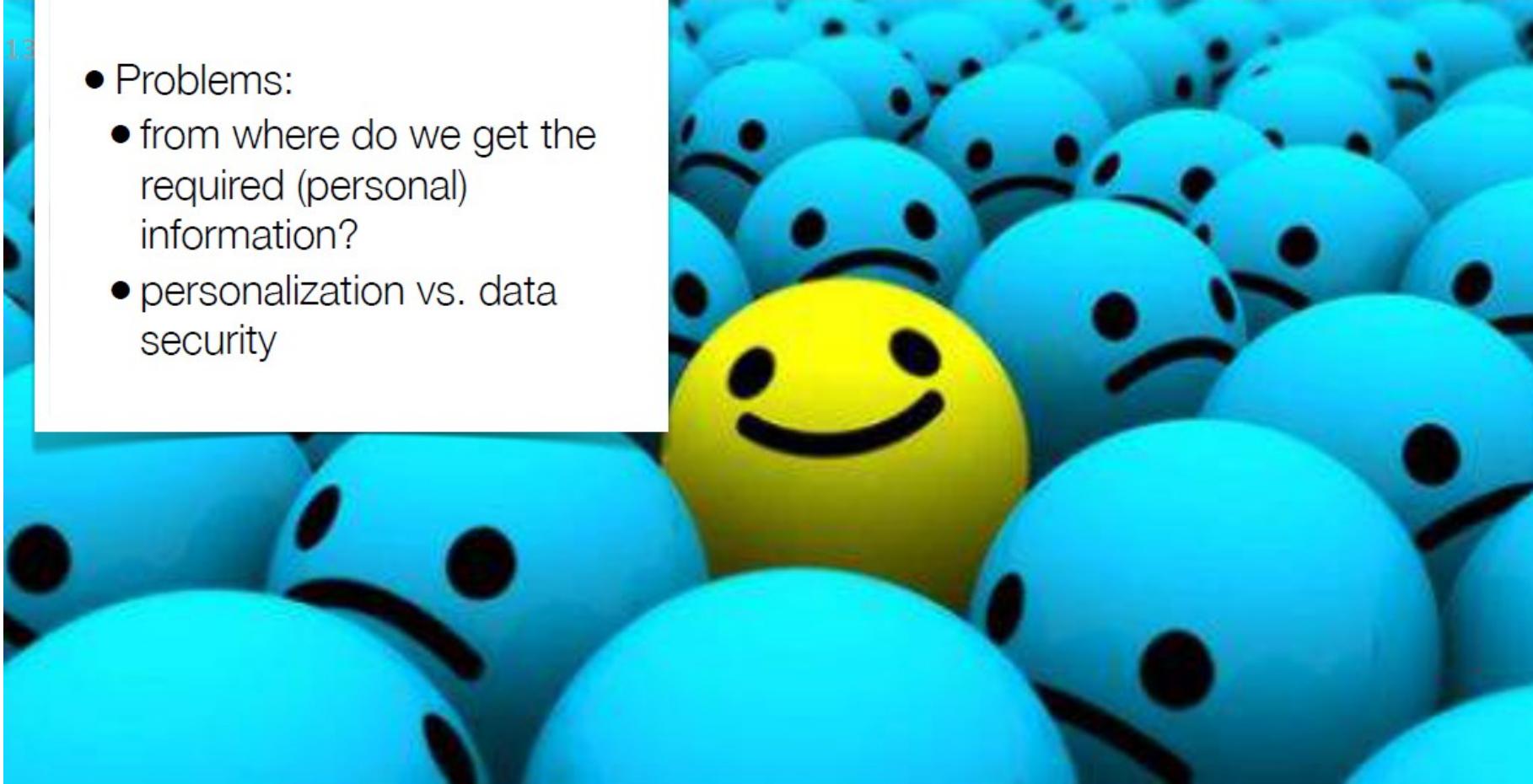


## Problem 3: Maintenance

- the more complex and voluminous a website, the more complicated is the maintenance of the only weakly structured data.
- Problems:
  - syntactic vs. semantic (link) consistency
  - correctness
  - timeliness

- Adaption of the presented information content to personal requirements
- Problems:
  - from where do we get the required (personal) information?
  - personalization vs. data security

13



## Problem 4: Personalization

**Implicit Knowledge**, i.e. information does not have to be specified explicitly, but must be derived via logical deductions from available information.



# “Understanding“ Content on the Web



Text: „Heisenberg can wait?“

Entity Mapping  
Disambiguation

Benjamin Heisenberg ? a film director

Jochen Heisenberg ? a physicist

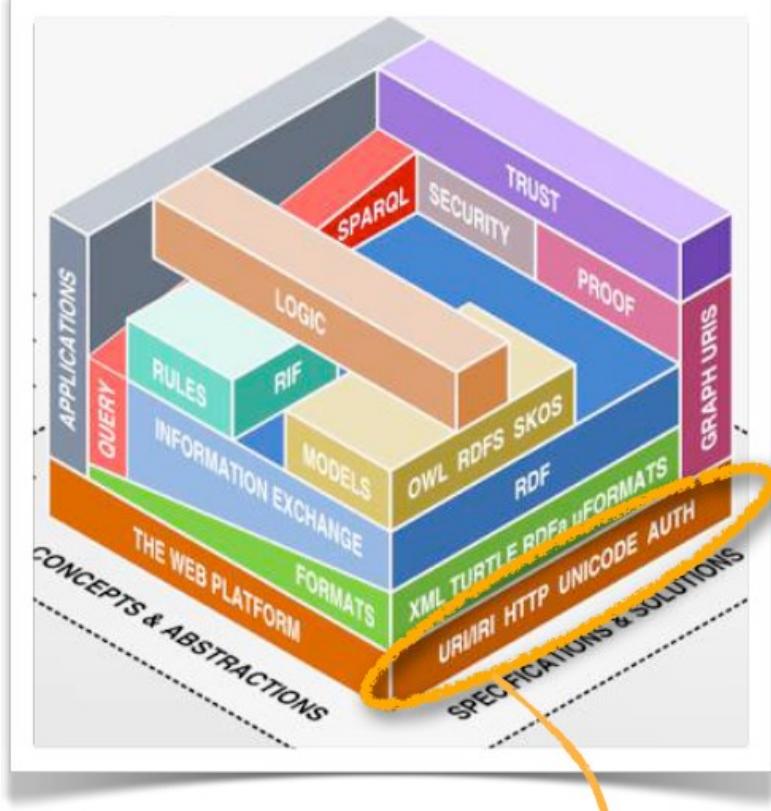
Martin Heisenberg ? a neurobiologist

Werner Heisenberg ? a physicist and Nobel Laureate

Walter White lead character of Breaking Bad

- **Disambiguation**

solution of linguistic ambiguities



Walter White

-----> [http://dbpedia.org/resource/Walter\\_White\\_\(Breaking\\_Bad\)](http://dbpedia.org/resource/Walter_White_(Breaking_Bad))

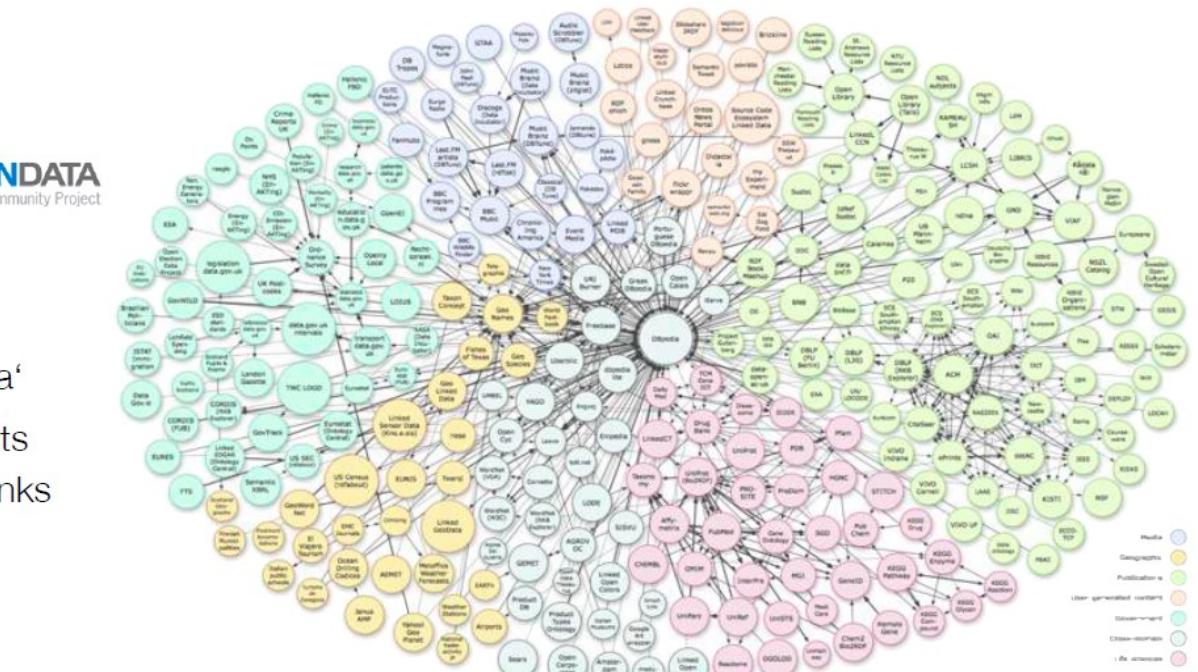
## Linked Data

- Linked Open Data (LOD) denote publicly available (RDF) Data in the Web, identified via URI and accessible via HTTP. Linked data link to other data via URI.



The 'Web of Data'

- >81 billion facts
- >800 million links  
(Mar. 2014)



[http://en.wikipedia.org/wiki/Walter\\_White\\_\(Breaking\\_Bad\)](http://en.wikipedia.org/wiki/Walter_White_(Breaking_Bad))

WIKIPEDIA  
The Free Encyclopedia

Article Talk

Walter White (*Breaking Bad*)

From Wikipedia, the free encyclopedia

Walter Hartwell "Walt" White Sr. (also known by his clandestine alias, Heisenberg) is a fictional character in the American television drama series *Breaking Bad* on AMC. He was portrayed by Bryan Cranston and was created by series creator Vince Gilligan. Once a promising chemist who was one of the founding members of the multi-billion dollar company Gray Matter Technologies, Walter left the company for personal reasons and became an unhappy and disillusioned high school chemistry teacher. After being diagnosed with Stage IIIA lung cancer, he resorts to manufacturing methamphetamine to ensure his family's financial security when he dies. As the series progresses, Walter gradually becomes darker and takes on a more villainous role.

Although AMC officials hesitated to cast Cranston due to his previous comedic roles on *Malcolm in the Middle*, Gilligan cast him based on the actor's past performance in the *X-Files* episode "Drive". Cranston has contributed much of his character, including Walter's back story, physical appearance, and personality traits. Gilligan had described his goal with Walter White as turning Mr. Chips into Scarface, and has deliberately made the character less and less sympathetic.

Both the Walter White character and Bryan Cranston's performance have received critical acclaim. Cranston won three consecutive Primetime Emmy Awards for Outstanding Lead Actor in a Drama Series, becoming the second actor to do so since Bill Cosby for *I Spy* in the 1960s.<sup>[1]</sup>

In the Spanish-language remake *Mentiras y Mentiras*, his character is renamed Walter Blanco and is portrayed by Diego Trujillo.<sup>[2]</sup>

Character biography [edit]

Background and personality [edit]

Walter White was a talented chemist who graduated with a Ph.D. in chemistry from the California Institute of Technology. He, along with colleague Vivek "Vodka" Kumar, contributed to research of proton radiography that helped a team win a Nobel Prize in Chemistry.<sup>[3]</sup> After graduate school, Walter founded the firm Gray Matter Technologies with Elliott Schwartz, his former classmate and

Read Edit View history Search

Walter White  
*Breaking Bad character*

First appearance "Pilot"  
Last appearance "Felina"  
Created by Vince Gilligan  
Portrayed by Bryan Cranston

Information

Aliases Heisenberg  
Mr. Lambert  
Drug kingpin  
Chemist at Santa Fe National Laboratories  
Co-founder of Gray Matter Technologies  
High school chemistry teacher  
Car wash owner, proprietor, and manager



[http://dbpedia.org/resource/Walter\\_White\\_\(Breaking\\_Bad\)](http://dbpedia.org/resource/Walter_White_(Breaking_Bad))



7 [http://dbpedia.org/resource/Walter\\_White\\_\(Breaking\\_Bad\)](http://dbpedia.org/resource/Walter_White_(Breaking_Bad))

About: [Walter White \(Breaking Bad\)](#)

An Entity of Type : [fictional character](#), from Named Graph : <http://dbpedia.org>, within Data Space : [dbpedia.org](http://dbpedia.org)

Walter Hartwell "Walt" White (also known by his clandestine alias, Heisenberg) is a fictional character and the protagonist of the American television drama series Breaking Bad on AMC. He is portrayed by Bryan Cranston and was created by series creator Vince Gilligan.

Property Value

[dbpedia-owl:abstract](#)

- Walter Hartwell White è il protagonista della serie televisiva Breaking Bad - Reazioni collaterali, interpretato da Bryan Cranston e doppiato da Stefano De Sando.
- Walter Hartwell "Walt" White (also known by his clandestine alias, Heisenberg) is a fictional character and the protagonist of the American television drama series Breaking Bad on AMC. He is portrayed by Bryan Cranston and was created by series creator Vince Gilligan. Once a promising chemist who was one of the founding members of the fictional, now multi-billion dollar company "Gray Matter Technologies", Walter left the company for personal reasons and became an unhappy and disillusioned high school chemistry teacher. After being diagnosed with inoperable lung cancer, he resorts to manufacturing methamphetamine to ensure his family's financial security when he dies. As the series progresses, Walter gradually becomes darker and takes on a more villainous role. Although AMC officials hesitated to cast Cranston due to his previous comedic role on Malcolm in the Middle, Gilligan cast him based on the actor's past performance in an episode of The X-Files. Cranston has contributed to much of his character, including Walter's back story, physical appearance and personality traits. Gilligan had described his goal with Walter White as turning Mr. Chips into Scarface, and has deliberately made the character less and less sympathetic. Both the Walter White character and Bryan Cranston's performance have received critical acclaim. Cranston won three consecutive Primetime Emmy Awards for Outstanding Lead Actor in a Drama Series, becoming the second actor ever to do so and the first since Bill Cosby for I Spy in the 1960s.
- Walter Hartwell - Walt - White est le personnage de fiction principal de la série Breaking Bad créé par Vince Gilligan. Il est interprété par Bryan Cranston et doublé en version française par Jean-Louis Faure. L'acteur Bryan Cranston fut récompensé à plusieurs reprises : dans la catégorie Meilleur acteur dans une série dramatique pour les Satellite Awards ; Meilleur acteur principal dans une série dramatique pour les Emmy Awards 2008, 2009 et 2010.

[dbpedia-owl:alias](#)

- Heisenberg
- Mr. Lambert

[dbpedia-owl:child](#)

- [dbpedia:List\\_of\\_Breaking\\_Bad\\_characters](#)

[dbpedia-owl:creator](#)

- [dbpedia:Vince\\_Gilligan](#)

[dbpedia-owl:firstAppearance](#)

- [dbpedia:Pilot\\_\(Breaking\\_Bad\)](#)

[dbpedia-owl:occupation](#)

- [dbpedia:Methamphetamine](#)
- [dbpedia:Sandia\\_National\\_Laboratories](#)

[dbpedia-owl:portrayer](#)

- [dbpedia:Bryan\\_Cranston](#)

[dbpedia-owl:relative](#)

- [dbpedia:Hank\\_Schrader](#)
- [dbpedia:List\\_of\\_Breaking\\_Bad\\_characters](#)

[dbpedia-owl:series](#)

- [dbpedia:Breaking\\_Bad](#)

[dbpedia-owl:thumbnail](#)

- [http://upload.wikimedia.org/wikipedia/commons/thumb/6/65/Walter\\_White2.jpg/200px-Walter\\_White2.jpg](http://upload.wikimedia.org/wikipedia/commons/thumb/6/65/Walter_White2.jpg/200px-Walter_White2.jpg)

[dbpedia-owl:wikiPageExternalLink](#)

- <http://www.amctv.com/shows/breaking-bad/cast/walter-white>

[dbpedia-owl:wikiPageID](#)

- 22530352 (xsd:integer)

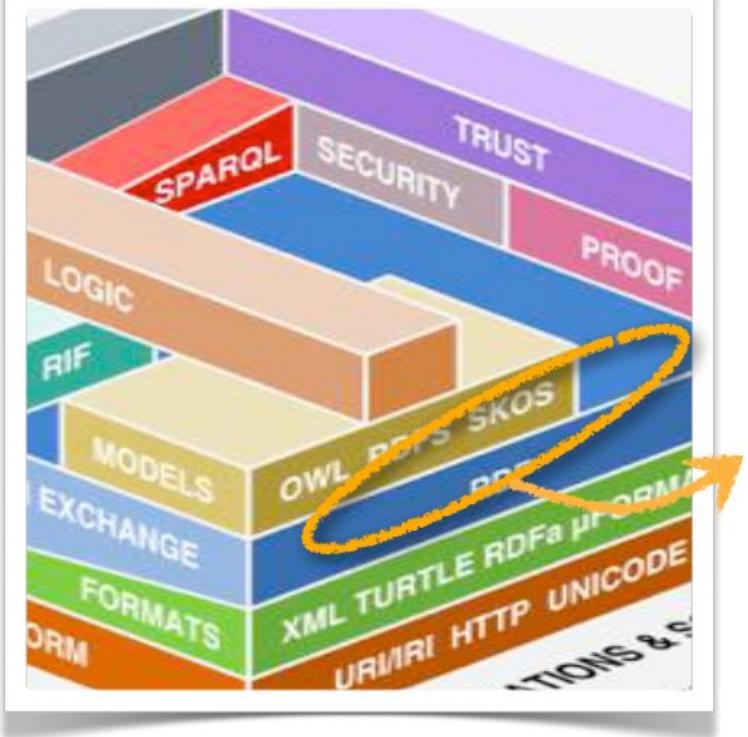
[dbpedia-owl:wikiPageRevisionID](#)

- 548300209 (xsd:integer)

[dbpprop:affiliations](#)

- [dbpedia:Jesse\\_Pinkman](#)
- [Saul\\_Goodman](#)
- [Gustavo\\_Fring](#)





[http://dbpedia.org/resource/Walter\\_White\\_\(Breaking\\_Bad\)](http://dbpedia.org/resource/Walter_White_(Breaking_Bad))

```
:Walter_White_(Breaking_Bad) rdf:type :FictionalCharacter .
:Walter_White_(Breaking_Bad) dbpp:portrayer :Bryan_Cranston .
:Bryan_Cranston dbpp:birthDate "1956-03-07" .
:Bryan_Cranston rdf:type dbpo:Actor .
:Bryan_Cranston rdf:type dbpo:Person .
...
```



*RDF Resource Description Framework*



**RDF Tripel**

:Brian\_Cranston



*RDF Subject*

rdf:type

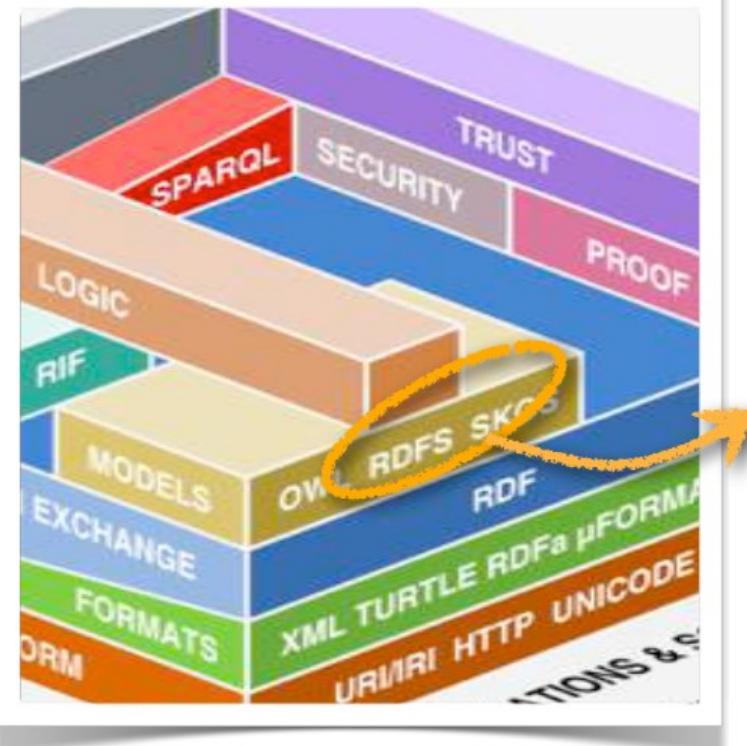


*RDF Property*

dbpo:Actor .



*RDF Object*



<http://dbpedia.org/ontology/FictionalCharacter>

```

dbpo:FictionalCharacter rdf:type owl:Class .
dbpo:FictionalCharacter rdfs:subClassOf dbpo:Person .
dbpo:FictionalCharacter rdfs:label "Fictional Character" .
dbpp:birthPlace rdf:type rdf:Property .
dbpp:birthPlace rdfs:domain dbpo:Person .
dbpp:birthPlace rdfs:range dbpo:Place .
dbpp:birthDate rdf:type rdf:Property .
dbpp:birthDate rdfs:domain :Person .
dbpp:birthDate rdfs:range xsd:date .
...

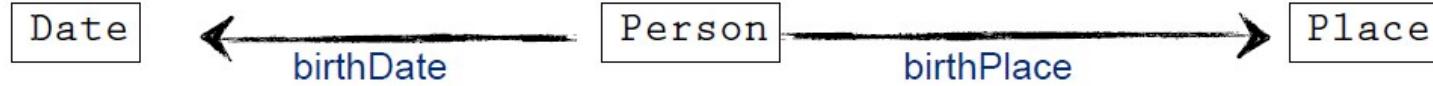
```

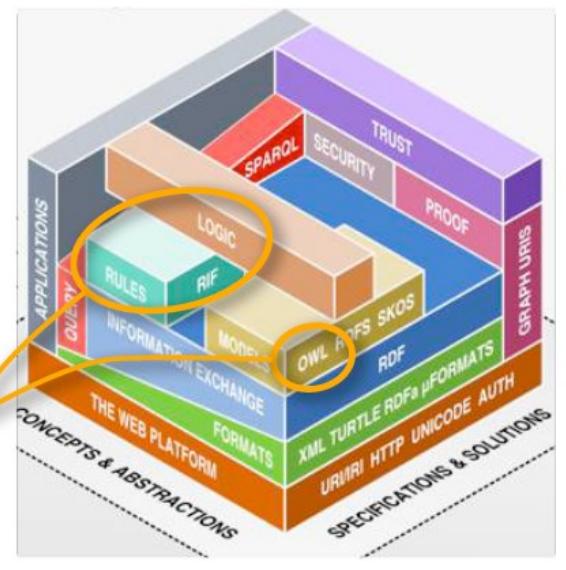
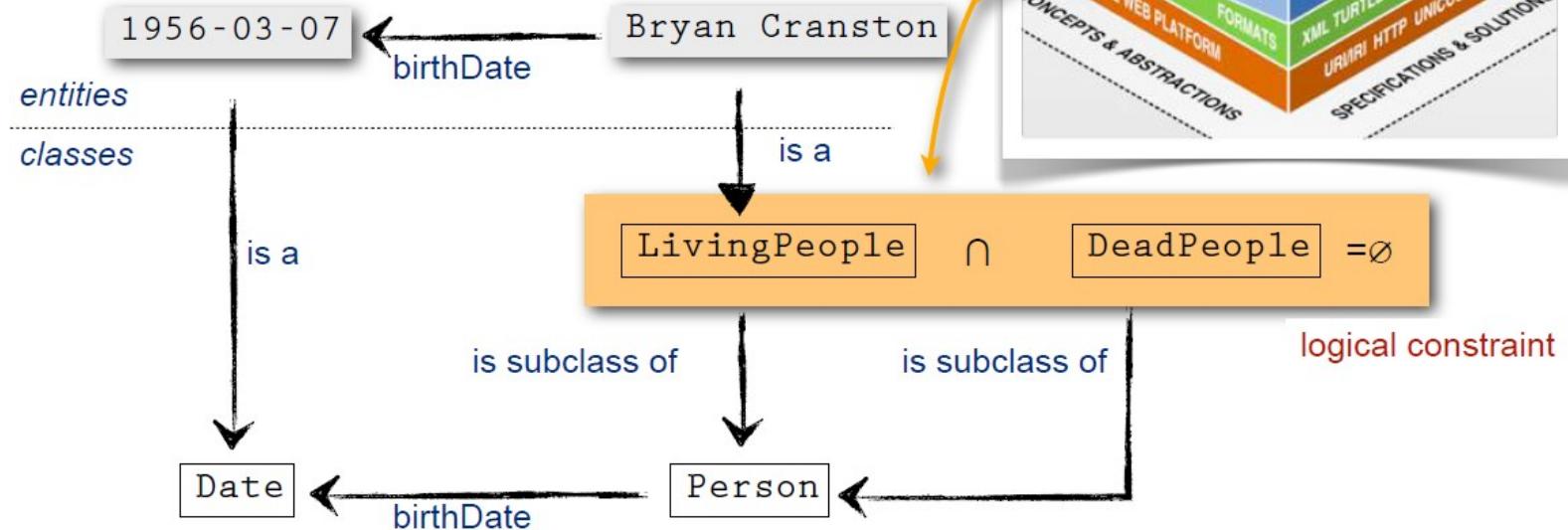


### RDF Schema

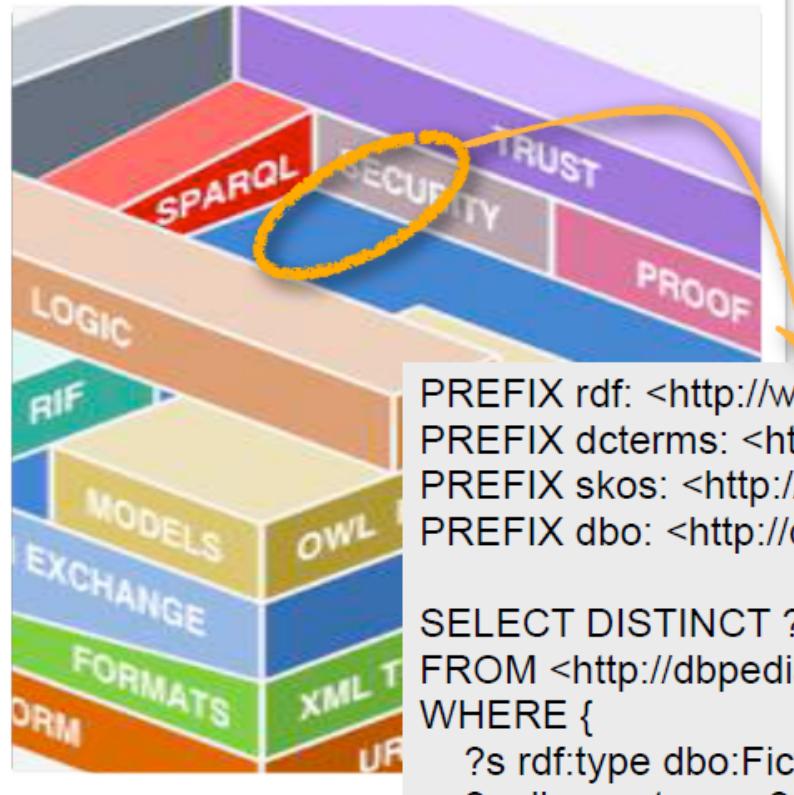
Fictional\_Character

↓  
is subclass of





$\forall x. \exists y. \text{hasDeathDate}(x,y) \wedge \text{Person}(x) \wedge \text{Date}(y) \rightarrow \text{DeadPeople}(x)$   
 + Rules  
 (Description Logics)



Look for all Fictional Characters whose portrayer have won an Oscar .

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX dbo: <http://dbpedia.org/ontology/>

SELECT DISTINCT ?s ?p
FROM <http://dbpedia.org>
WHERE {
    ?s rdf:type dbo:FictionalCharacter .
    ?s dbo:portrayer ?p .
    ?p rdf:type dbo:Person .
    ?p dcterms:subject/skos:broader*
        <http://dbpedia.org/resource/Category:Academy_Award_winners> .
}
ORDER BY ?p
```



[Link to public DBpedia  
SPARQL endpoint  
including query](#)

s	p
<a href="http://dbpedia.org/resource/Tony_Montana">http://dbpedia.org/resource/Tony_Montana</a>	<a href="http://dbpedia.org/resource/Al_Pacino">http://dbpedia.org/resource/Al_Pacino</a>
<a href="http://dbpedia.org/resource/Michael_Corleone">http://dbpedia.org/resource/Michael_Corleone</a>	<a href="http://dbpedia.org/resource/Al_Pacino">http://dbpedia.org/resource/Al_Pacino</a>
<a href="http://dbpedia.org/resource/Carlito_Brigante">http://dbpedia.org/resource/Carlito_Brigante</a>	<a href="http://dbpedia.org/resource/Al_Pacino">http://dbpedia.org/resource/Al_Pacino</a>
<a href="http://dbpedia.org/resource/Inspector_Clouseau">http://dbpedia.org/resource/Inspector_Clouseau</a>	<a href="http://dbpedia.org/resource/Alan_Arkin">http://dbpedia.org/resource/Alan_Arkin</a>
<a href="http://dbpedia.org/resource/Fagin">http://dbpedia.org/resource/Fagin</a>	<a href="http://dbpedia.org/resource/Alec_Guinness">http://dbpedia.org/resource/Alec_Guinness</a>
<a href="http://dbpedia.org/resource/George_Smiley">http://dbpedia.org/resource/George_Smiley</a>	<a href="http://dbpedia.org/resource/Alec_Guinness">http://dbpedia.org/resource/Alec_Guinness</a>
<a href="http://dbpedia.org/resource/Maleficent">http://dbpedia.org/resource/Maleficent</a>	<a href="http://dbpedia.org/resource/Angelina_Jolie">http://dbpedia.org/resource/Angelina_Jolie</a>
<a href="http://dbpedia.org/resource/Sookie_Stackhouse">http://dbpedia.org/resource/Sookie_Stackhouse</a>	<a href="http://dbpedia.org/resource/Anna_Paquin">http://dbpedia.org/resource/Anna_Paquin</a>
<a href="http://dbpedia.org/resource/Mia_Thermopolis">http://dbpedia.org/resource/Mia_Thermopolis</a>	<a href="http://dbpedia.org/resource/Anne_Hathaway">http://dbpedia.org/resource/Anne_Hathaway</a>
<a href="http://dbpedia.org/resource/White_Queen_(Through_the_Looking-Glass)">http://dbpedia.org/resource/White_Queen_(Through_the_Looking-Glass)</a>	<a href="http://dbpedia.org/resource/Anne_Hathaway">http://dbpedia.org/resource/Anne_Hathaway</a>
<a href="http://dbpedia.org/resource/Hannibal_Lecter">http://dbpedia.org/resource/Hannibal_Lecter</a>	<a href="http://dbpedia.org/resource/Anthony_Hopkins">http://dbpedia.org/resource/Anthony_Hopkins</a>
<a href="http://dbpedia.org/resource/Abraham_Van_Helsing">http://dbpedia.org/resource/Abraham_Van_Helsing</a>	<a href="http://dbpedia.org/resource/Anthony_Hopkins">http://dbpedia.org/resource/Anthony_Hopkins</a>
<a href="http://dbpedia.org/resource/Stanley_Kowalski">http://dbpedia.org/resource/Stanley_Kowalski</a>	<a href="http://dbpedia.org/resource/Anthony_Quinn">http://dbpedia.org/resource/Anthony_Quinn</a>
<a href="http://dbpedia.org/resource/Judith_Traherne">http://dbpedia.org/resource/Judith_Traherne</a>	<a href="http://dbpedia.org/resource/Barbara_Stanwyck">http://dbpedia.org/resource/Barbara_Stanwyck</a>
<a href="http://dbpedia.org/resource/List_of_View_Askewiverse_characters">http://dbpedia.org/resource/List_of_View_Askewiverse_characters</a>	<a href="http://dbpedia.org/resource/Ben_Affleck">http://dbpedia.org/resource/Ben_Affleck</a>
<a href="http://dbpedia.org/resource/Sweeney_Todd">http://dbpedia.org/resource/Sweeney_Todd</a>	<a href="http://dbpedia.org/resource/Ben_Kingsley">http://dbpedia.org/resource/Ben_Kingsley</a>
<a href="http://dbpedia.org/resource/Fagin">http://dbpedia.org/resource/Fagin</a>	<a href="http://dbpedia.org/resource/Ben_Kingsley">http://dbpedia.org/resource/Ben_Kingsley</a>
<a href="http://dbpedia.org/resource/Hood_(Thunderbirds)">http://dbpedia.org/resource/Hood_(Thunderbirds)</a>	<a href="http://dbpedia.org/resource/Ben_Kingsley">http://dbpedia.org/resource/Ben_Kingsley</a>
<a href="http://dbpedia.org/resource/Jack_Rafferty">http://dbpedia.org/resource/Jack_Rafferty</a>	<a href="http://dbpedia.org/resource/Benicio_del_Toro">http://dbpedia.org/resource/Benicio_del_Toro</a>
<a href="http://dbpedia.org/resource/Larry_Talbot">http://dbpedia.org/resource/Larry_Talbot</a>	<a href="http://dbpedia.org/resource/Benicio_del_Toro">http://dbpedia.org/resource/Benicio_del_Toro</a>
<a href="http://dbpedia.org/resource/Baby_Jane_Hudson">http://dbpedia.org/resource/Baby_Jane_Hudson</a>	<a href="http://dbpedia.org/resource/Bette_Davis">http://dbpedia.org/resource/Bette_Davis</a>
<a href="http://dbpedia.org/resource/Judith_Traherne">http://dbpedia.org/resource/Judith_Traherne</a>	<a href="http://dbpedia.org/resource/Bette_Davis">http://dbpedia.org/resource/Bette_Davis</a>
<a href="http://dbpedia.org/resource/List_of_past_Casualty_characters">http://dbpedia.org/resource/List_of_past_Casualty_characters</a>	<a href="http://dbpedia.org/resource/Brenda_Fricke">http://dbpedia.org/resource/Brenda_Fricke</a>
<a href="http://dbpedia.org/resource/Blanche_DuBois">http://dbpedia.org/resource/Blanche_DuBois</a>	<a href="http://dbpedia.org/resource/Cate_Blanchett">http://dbpedia.org/resource/Cate_Blanchett</a>
<a href="http://dbpedia.org/resource/Scheherazade">http://dbpedia.org/resource/Scheherazade</a>	<a href="http://dbpedia.org/resource/Catherine_Zeta-Jones">http://dbpedia.org/resource/Catherine_Zeta-Jones</a>



# Giornalismo

Molti sono i settori nei quali i media stanno scommettendo su algoritmi e automazione per abbattere tempi e costi dell'informazione.

**Thompson Reuters**, specializzata in informazione finanziaria, sfrutta l'analisi semantica per automatizzare la produzione di grafici interattivi per accompagnare le sue analisi di mercato. Lo stesso fa **Associated Press** – l'agenzia americana sostenuta da 1400 testate statunitensi e migliaia di broadcaster in tutto il mondo - per accelerare la pubblicazione di analisi di borsa e risultati delle trimestrali (ne produce 3700 a trimestre, 12 volte più di prima) e sta pensando di applicare lo stesso approccio ai risultati sportivi, es. cronaca social delle partite di rugby e calcio.

**L'automatizzazione della produzione di video**, sempre più richiesti dai canali online, è una delle frontiere più interessanti dell'automazione editoriale e molti guardano a soluzioni in grado di raccogliere in maniera automatica contenuti pubblici e geotaggati da Instagram, Twitter, Facebook, GooglePlus, Vkontakte e altri, per produrre in pochi secondi video di un evento di attualità.

La diffusione della rete e dei sistemi cloud sta moltiplicando le fonti di dati e le possibilità di addestrare i software su terreni diversi.



Entro il 2025 molte delle notizie lette dal grande pubblico sarà prodotto da un computer con poco o nessun contributo umano (e il **nostro ruolo?**)

# Pubblica amministrazione

Stiamo assistendo ad un **accumulo esponenziale di informazioni**. Pensiamo ad un (apparente) innocuo "tap" sul nostro smartphone, un gesto semplice e rapido che in realtà, ogni volta che lo facciamo, genera delle conseguenze, sull'app o sul portale che stiamo utilizzando, e quindi dati. **I dati sono il vero nuovo petrolio del XXI secolo e data la loro gigantesca mole abbiamo il dovere di ordinarli, di incasellarli, di "taggarli" con etichette che ci permetteranno la loro comprensione, il loro recupero e la loro lettura in futuro, quando finiranno nell'immensità della rete.**

Le ontologie servono essenzialmente per attribuire ai dati una **semantica**, un senso. Danno ai dati una voce, la possibilità cioè di far dire a loro di che informazione si fanno carico, di che contenuto si fanno ospiti.

La **Pubblica Amministrazione** detiene un patrimonio informativo che, se opportunamente condiviso, è in grado di generare valore primariamente all'interno della PA stessa, ma è in grado anche di generare **impatti rilevanti sulla collettività**, dal momento che contribuisce al miglioramento della qualità di vita dei cittadini che possono utilizzare i dati, condividerli, aggiornarli secondo un meccanismo **partecipativo** e volto a sostenere la fiducia nell'amministrazione, secondo i principi dell'open-government.

Esempio i monitoraggi e i controlli in ambito energetico, di sicurezza o ai feedback derivanti dai social media.

# Medicina



L'aumento delle cronicità insieme con l'invecchiamento della popolazione, in un contesto nazionale ed internazionale sempre più concentrato sul trovare il giusto equilibrio tra la qualità e quantità dei servizi ed il pesante vincolo dato dalla ristrettezza di risorse impongono al sistema sanitario nel suo complesso forti **revisioni dei modelli organizzativi**, che dovranno tendere sia alla razionalizzazione dei tanti punti di erogazione di servizi sanitari ai cittadini (sempre meno ospedalizzazione a favore di prese in carico sul territorio, a domicilio, presso i centri di cura primari, presso la medicina di associazione, strutture polifunzionali, multi-specialistiche ...), sia ad enfatizzare la centralità del cittadino – paziente rispetto all'insieme dei soggetti e delle strutture deputati alla cura del suo benessere.

Gli **Open Data in campo medico** possono aiutare ad esempio a combattere la diffusione di epidemie. In Africa, per esempio, i dati di localizzazione degli smartphone vengono impiegati per tracciare i movimenti della popolazione e quindi prevedere la propagazione del virus Ebola, dando inoltre evidenze su dove realizzare centri per le cure.

# Medicina



## L'analisi semantica applicata alla medicina

è uno strumento adatto a ricavare **significato** dai dati non strutturati e trasferire elementi della storia medica del paziente in **modelli di natura predittiva**. Rappresenta la tecnologia abilitante per identificare ed estrarre informazioni, pattern e trend costruendo conoscenza ricavata da dati non strutturati e big data in ambito medico (articoli, diagnosi cliniche, ecc.), garantendo benefici sia per il sistema sanitario che per il paziente:

- **Miglioramento della “patient experience”;**
- **Riduzione dei ricoveri e degli esami in fase diagnostica;**
- **Quadro più completo dei pazienti e dei trattamenti** per ogni paziente;
- **Risparmio di tempo** da parte del personale nel riempire formulari;
- **Miglioramento della qualità dei dati raccolti;**
- **Accesso alle informazioni più efficiente.**

# Per approfondimenti

- [40 BRILLIANT OPEN DATA PROJECTS PREPARING SMART CITIES FOR 2018](#)
- <https://www.healthdata.gov/>
- <http://wiki.dbpedia.org/>
- <https://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>
- <http://ec.europa.eu/programmes/horizon2020/en/h2020-section/information-and-communication-technologies>
- <http://forumpa2017.eventifpa.it/it/event-details/?id=5347>

# Open data, oggetti connessi e web semantico: dal giornalismo alla medicina di precisione

Antonella Carbonaro  
[antonella.carbonaro@unibo.it](mailto:antonella.carbonaro@unibo.it)

ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA



Corso di Laurea Magistrale in Ingegneria e  
Scienze Informatiche  
  
Università degli Studi di Bologna,  
sede di Cesena