

UNIVERSITA’ DEGLI STUDI DI TORINO

Dipartimento di Informatica

Corso di Laurea Magistrale in Informatica

Anno Accademico 2023/2024

**Modellazione Concettuale del Web Semantico**

Autore: Annalisa Sabatelli Matr. 866879

*“***OntoHorses***”*

Documentazione di progetto

1. **Motivazioni**

Il dominio modellato riguarda una descrizione delle razze più diffuse di cavalli a supporto del corretto acquisto di un esemplare.

Durante l’analisi si è cercato di porre particolare attenzione ai seguenti aspetti:

* Temperamento e aspetti generali del carattere della razza;
* Principali dati dimensionali medi in termine di altezza al garrese e peso;
* Attuali impieghi, distinguendo tra lavoro agricolo, traino e sport.

La scelta di questo particolare dominio è motivata da una conoscenza personale dello stesso e dei suoi aspetti critici. Il settore equestre, inteso in senso lato e in tutte le sue declinazioni, soffre tutt’ora di una mancanza di digitalizzazione diffusa e consistente e di un’inadeguata strutturazione della conoscenza. Tale aspetto diventa ancor più spinoso, controverso e quasi paradossale se si pensa che l’addomesticamento del cavallo risale al IV millennio a.C. mentre il primo saggio “Sull’equitazione” è stato scritto da Senofonte nel 350 a.C.

Nonostante i migliaia di trattati che hanno fatto seguito al primo di Senofonte, nella maggior parte dei casi, sia nell’impiego da lavoro che in quello sportivo, si assiste a una predominanza della trasmissione orale della conoscenza. Questa modalità mette in mostra tutti i suoi punti deboli nel momento in cui ci si appresta all’acquisto del proprio cavallo. Acquistare un cavallo è, infatti, un’attività estremamente complicata in cui bisogna aver ben chiari in mente non solo le potenzialità effettive del particolare soggetto che si sta considerando, ma anche quelle generali della razza a cui appartiene e che ne influenzeranno nella sostanza i risultati effettivamente raggiungibili con l’addestramento, le modalità relazionali con l’uomo e, ovviamente, gli impieghi possibili.

L'idea di creare un'ontologia che modelli correttamente questo dominio potrebbe quindi risultare utile per offrire uno strumento potente, in grado di facilitare l’accesso ad una conoscenza strutturata sulle razze equine con lo scopo di supportare il processo di acquisto di un esemplare.

1. **Requisiti**
   1. **Finalità**

Lo scopo dell’ontologia è quello di creare una base di conoscenza per un’ipotetica piattaforma di annunci di vendita di cavalli. Gli utenti hanno la possibilità di visualizzare gli annunci filtrandoli per età, impiego, temperamento, zona di scuderizzazione. Il reasoning, inoltre, aiuterà il sistema a suggerire agli utenti i soggetti più idonei in funzione dello scopo dell’acquisto e, per i cavalli sportivi, in funzione delle capacità del cavallo e del cavaliere in modo da faciltare la creazione del “binomio” migliore. Oltre le informazioni relative ai cavalli in vendita, l’ontologia include anche informazioni, morfologiche, comportamentali e attitudinali, delle razze più diffuse di cavalli, oltre ad una galleria di immagini. L’utente potrà partire dalla ricerca della razza più idonea rispetto ai propri scopi e di conseguenza selezionare gli annunci di vendita, oppure al contrario, una volta individuato un soggetto di interesse, potrebbe appronfondire la propria conoscenza su quella specifica razza. Rispetto agli stumenti disponibili attualmente, il sistema risulta essere innovativo. In un panorama fortemente frammentato, poco digitalizzato e poco strutturato, infatti, si propone uno standard per le informazioni caratterizzanti un annuncio di vendita e si fornisce, inoltre, un accesso diretto e immediato anche a conoscenze equine più generiche ma ugualmente importanti in questo processo di acquisto.

[eventuali immagini del mock-up]

* 1. **Task specifici e contesto**

I task principali che l'ontologia punta a rendere effettuabili sono:

* Ricerca di cavalli in vendita filtrando per zona geografica, impiego, razza;
* Ricerca di cavalli sportivi in funzione del livello di esperienza del cavaliere;
* Ricerca di cavalli sportivi in funzione delle prestazioni degli stessi;
* Ricerca delle caratteristiche principali di una data razza;
* Classificazione delle razze per temperamento, impiego, area geografica di origine, morfologia.
  1. **Utenti target**
* **Sportivi che praticano Equitazionea livello agonistico e non o amatori** che vogliono finalizzare l’acquisto di un cavallo coerentemente ai propri obiettivi e capacità;
* **Utenti** che ricercano cavalli da lavoro o da traino;
* **Proprietari di maneggi** che vogliono ricercare nuovi soggetti da aggiungere alla propria scuderia;
* **Commercianti di cavalli** che hanno bisogno di pubblicare i propri annunci di vendita;
* **Utenti** che sono interessati a consultare un database contenente informazioni sulle caratteristiche principali delle più diffuse razza di cavallo.

1. **Descrizione del dominio e Fonti**
   1. **Dominio**

Il dominio è relativo al mondo equestre ed ha lo scopo di rappresentare le principali caratteristiche delle “Razze” più diffuse di cavalli, i loro “Impieghi”, i “Cavalli” in vendita e le Persone. Dal dominio sono escluse tutte le razze non europee e quelle europee che vivono allo stato brado o sono in estinzione.

Di seguito si elencano le quattro classi principali dell’ontologia e la loro correlazione:

* Ciascuna razza è definita in termini di caratteristiche morfologiche, come per esempio l’altezza al garrese e il peso medio, comportamentali, se a sangue freddo, caldo o misto, e attitudinali, ossia gli “Impieghi” nel lavoro e nello sport per i quali è ritenuta più idonea. E’ tracciata, inoltre, sotto forma di ulteriore classe, anche l’“Area geografica” di origine della razza.
* La classe “Cavallo” è in relazione con “Razza” e da questa ne deriva le caratteristiche comportamentali di base. Ad esempio, la sottoclasse “Cavallo a sangue caldo” include tutti i soggetti la cui “Razza” è una “Razza a sangue caldo”. Ciascun esemplare di cavallo avrà poi delle caratteristiche generali proprie e dei risultati prestazionali. In particolare, nel caso di cavalli sportivi sono specificate le discipline per le quali è stato addestrato e viene impiegato e un indicatore di prestazione, come per esempio la categoria in cui gareggia, che convenzionalmente è direttamente proporzionale al tipo di “Autorizzazione a montare FISE (Federazione Italiana Sport Equestri) che permette di conseguire e alla livello di abilità richiesto al “Cavaliere” per una corretta gestione del soggetto. Questi aspetti devono essere valutati attentamente in fase di acquisto al fine di creare dei “binomi” che funzionano.
* La classe “Persona” include la sottoclasse “Proprietario” che *possiede* almeno un cavallo e la sottoclasse “Cavaliere” il cui livello di esperienza determina le patenti che può conseguire e il cavallo più idoneo.
  1. **Fonti**

Questo lavoro nasce innegabilmente da una passione personale e da una conoscenza del dominio di tipo esperienziale. In ogni caso, a completamento di questo sono state consultata varie fonti web e cartacee.

Per le informazioni sulle principali classificazioni delle razze equine è stato consultato “Il manuale di equitazione” edito dalla FISE che è considerato un testo di riferimento per l’acquisizione delle conoscenze di base sull’anatomia del cavallo e sulla sua gestione, sulla classificazione delle razze e sui principali sport equestri.

Per l’elenco di tutte le razze equine esistenti e per le informazioni di dettaglio sulle loro caratteristiche principali ed impieghi sono stati consultati principalmente i siti web di Wikipedia e Wikidata.

Infine, l’elenco di tutti gli sport equestri e delle loro categorie e l’elenco delle autorizzazioni a montare e la loro regolamentazione in temini di propedeuticità provengono dal sito ufficiale della FISE e dai suoi *Regolamenti* (https://www.fise.it).

1. **Compentency questions**

Il fine principale dell’ontologia è quello di fornire uno standard per gli annunci di vendita di cavalli sfruttando l’inferenza per suggerire il cavallo più idoneo per gli obiettivi dell’acquirente e per il suo livello di esperienza.

A questo si aggiunge un articolato database sulle caratteristiche principali delle più diffuse razze equine. Una volta selezionato un cavallo di interesse, infatti, l’utente può interrogare la base di dati per recuperare le caratteristiche principali di quella razza, supportando in modo informato e consapevole il proprio processo decisionale.

Ad esempio, si riportano di seguito delle possibili interrogazioni:

* Elenco dei cavalli in vendita con età massima 10 anni e impiegati nel Dressage;
* Elenco dei cavalli in vendita di razza “Sella Italiano”;
* Elenco dei cavalli in vendita scuderizzati in Piemonte;
* Elenco dei cavalli in vendita idonei per il conseguimento Brevetto e con categoria di salto C115 cm;
* Elenco di altezza al garrese media, peso medio, temperamento, classificazione morfologica, origine geografica e impieghi tipici della razza “Murgese”.

1. **Documentazione del dominio**
   1. **Stato dell’arte**

L’ontologia si muove essenzialmente lungo due direzioni: da un lato si pone come obiettivo quello di creare un database con le informazioni principali relative alle razze europee di cavalli più diffuse; dall’altro è una proposta di una piattaforma unica per annunci di vendita di cavalli. In pratica, si è scelto di unire una finalità commerciale ad una didascalica per favorire acquisti consapevoli in un ambito complesso come quello equestre.

Di conseguenza, la ricerca per la documentazione del dominio è stata di due tipi: informazioni sulle razze e gestione delle vendite tramite annunci on-line.

Per quanto riguarda l’elenco delle razze esistenti e le loro caratteristiche principali Wikipedia è stato utilizzato come riferimento principale. Il sito ha elencato circa 117 razze differenti ma ne esistono realmente più di 300. Si è deciso di restringere l’ampiezza del dominio per renderlo più coerente con la finalità commerciale dell’ontologia. Sono stati esclusi, perciò, i pony, tutte le razze non Europee, le razze in via di estinzione e quelle che tutt’ora vivono allo stato brado. Di queste ultime, proprio in Italia ne esistono due esempi, uno in Sicilia, il Sanfratellano, ed uno in Sardegna, il cavallino della Giara.

Di seguito, si riporta una pagina di esempio di Wikipedia per il Murgese.

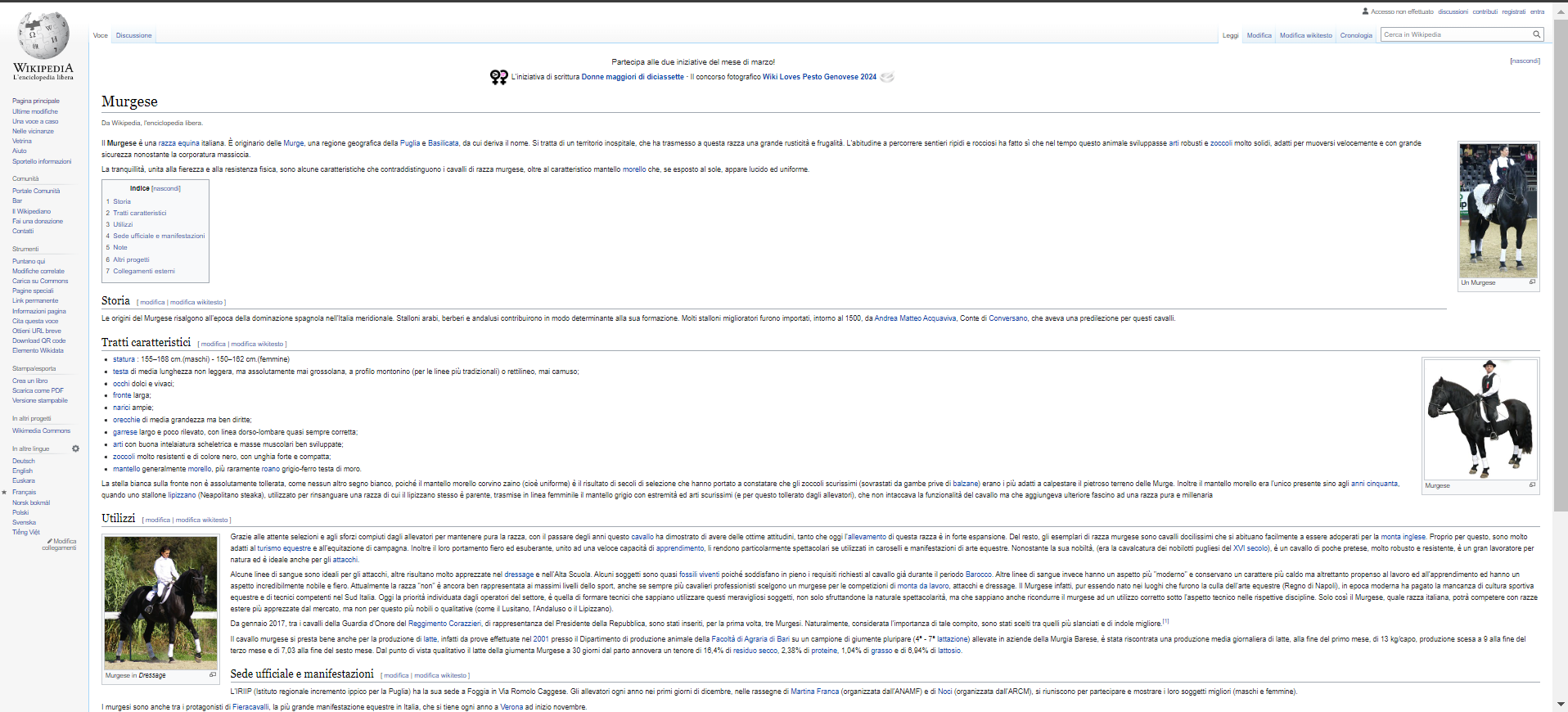


Figura 1:Esempio pagina web Wikipedia

Il sito aritcola la pagina dedicata a ciascuna razza in tre sezioni: “Storia”, “Tratti caratteristici” ed “Utilizzi”. Particolarmente utili, ai fini della creazione di questa ontologia sono state le ultime due sezioni dalle quali sono state mutuate proprietà e/o classi per l’ontologia. Per esempio, tra le proprietà inserite ci sono “l’altezza media al garrese” e il “peso medio” come suggerito da Wikipedia. Anche tutte le informazioni relative agli impieghi sono state desunte da questo sito.

Un altro esempio trovato in rete di classificazione delle razze equine è il seguente sito web <https://www.discoverthehorse.com/horse-breeds/profiles/#a>

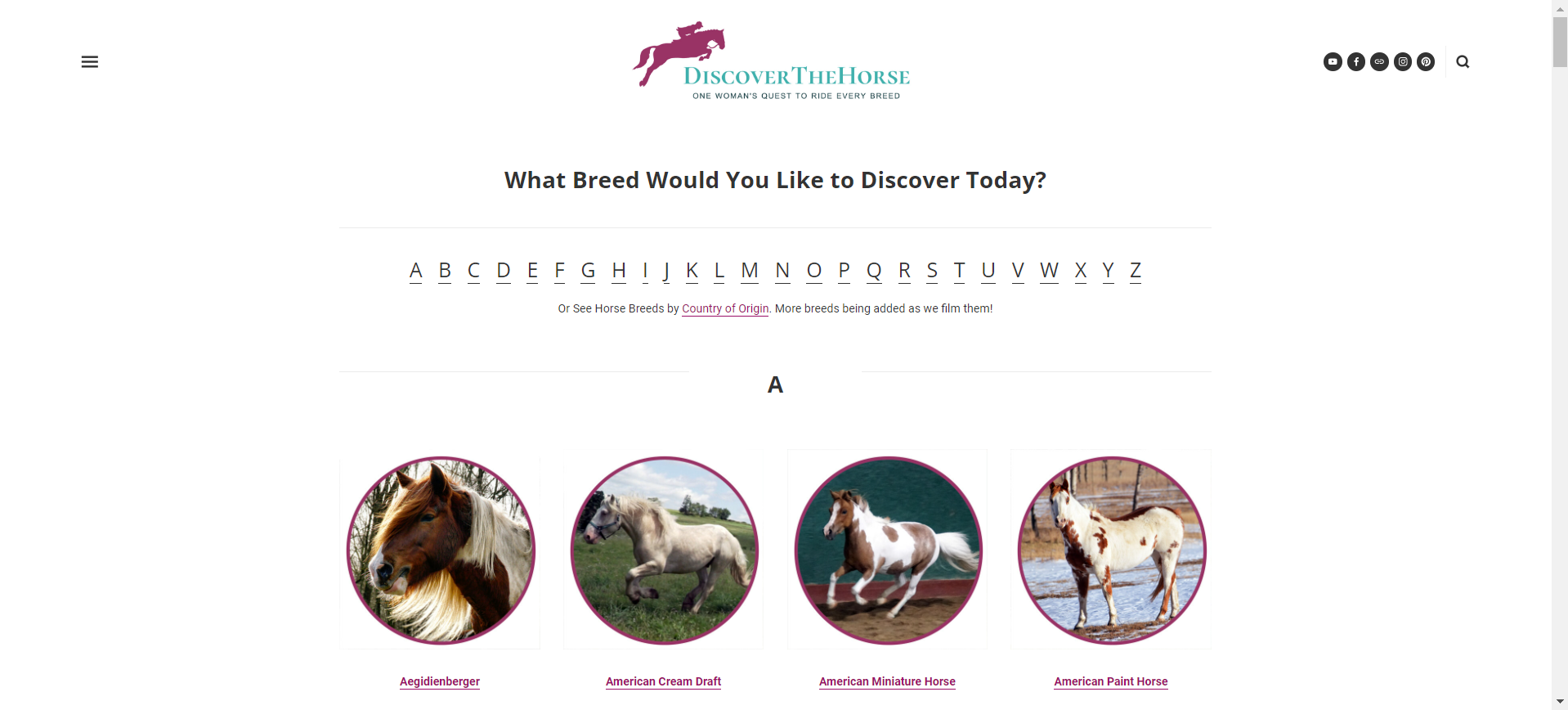


Figura 2: Home page "Discover the horse"

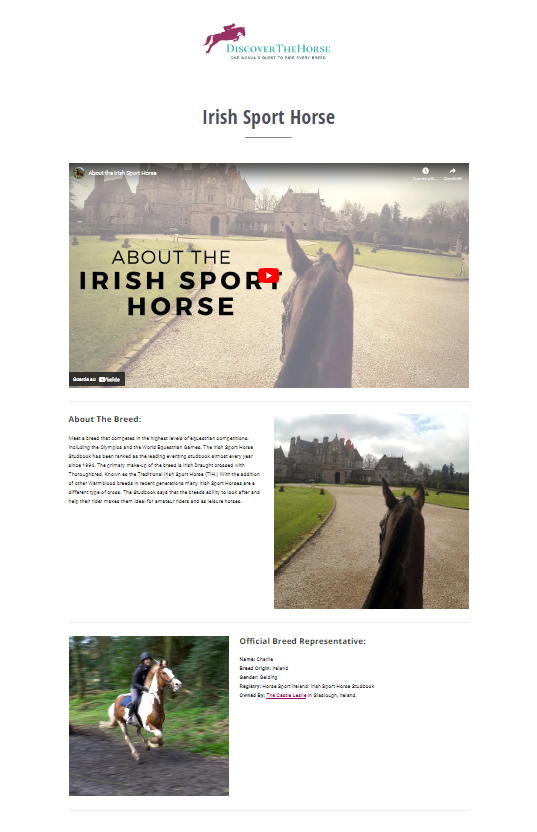
 Apparentemente questo sito sembra più user friendly di Wikipedia, ma in realtà, entrando nel dettaglio di una singola razza, si scopre che le informazioni sono più scarne e meno strutturate.

Figura 3: Informazioni di dettaglio di una razza del sito web "Discover the horse"

Entrambe queste fonti, purtroppo, non contengono informazioni sul temperamento delle razze, una caratteristica fondamentale che ne influenza il suo impiego, il suo carattere e la sua tendenziale “gentilezza” nei confronti dell’uomo. Tale informazione è stata inclusa nell’ontologia ampliando la ricerca ad altre fonti.

Il sito web <https://www.gabrielecavalli.it/2018/03/26/cavalli-a-sangue-freddo-e-cavalli-a-sangue-caldo/> è stato preso come riferimento per la classificazione delle razze in base al temperamento: sangue caldo, sangue freddo, sangue misto.

Un’altra classificazione tipicamente utilizzata e riproposta anche nel Manuale di Equitazione della FISE riguarda la corporatura e prevede tre categorie principali: brachimorfi, mesomorfi, dolicomorfi, meso-brachimorfi e meso-dolicomorfi.

Il cavallo brachimorfo è un tipo di cavallo più propenso allo sviluppo di muscolatura, con un torace ampio, forme potenti, massicce e non molto lungo. Alcuni esempi tipici: il cavallo Belga, molto utilizzato in agricoltura vista la sua forza; il cavallo Bretone, di stazza tanto grande da arrivare a pesare anche 900 kg; il cavallo Ardennese, da sempre utilizzato per il tiro della legna.

Il cavallo mesomorfo ha una struttura molto meno massiccia, che da sempre viene sfruttata a livello sportivo e per la caccia; era anche il cavallo che veniva generalmente utilizzato per gli eserciti. Alcuni esempi di cavalli mesomorfi: il Murgese, proveniente dalla Puglia, deriva da incroci di razze avvenuti durante il periodo in cui gli arabi si trovavano nel sud Italia; la sua specialità è quella di essere un cavallo adatto un po’ a tutto,e che può essere utilizzato da tutti; l’Andaluso, cavallo che viene considerato il fondatore delle altre razze nel continente americano, è un cavallo di circa 6 quintali, adatto al salto e utilizzato soprattutto per gli sport, è molto famoso per essere docile e molto elegante.

Il cavallo dolicomorfo non è adatto ai lavori di forza ma bensì a lavori dove c’è bisogno della sua agilità. Ha linee molto allungate e articolazioni assai snodate. Un esempio: il Purosangue inglese, un cavallo di grande bellezza e che ha una peculiarità, la difficoltà di montatura, essendo un cavallo molto veloce e necessita, quindi, di un cavaliere molto esperto; è ottimo sia per il salto ad ostacoli che per le corse.

Il cavallo meso-brachimorfo è a metà strada tra i cavalli da tiro e quelli da sella.

Il cavallo meso-dolicomorfo è una razza potente, forte ma allo stesso tempo agile e nobile nel portamento.

Le informazioni relative a tutti i possibili impieghi dei cavalli soprattutto in ambito sportivo derivano dalla documentazione e dai regolamenti messi a disposizione dalla FISE (Federazione Italiana Sport Equestri). Da questa fonte provengono anche le informazioni relative alle categorie di salto ad ostacoli e dressage e quelle inerenti le autorizzazioni a montare e le loro propedeuticità.

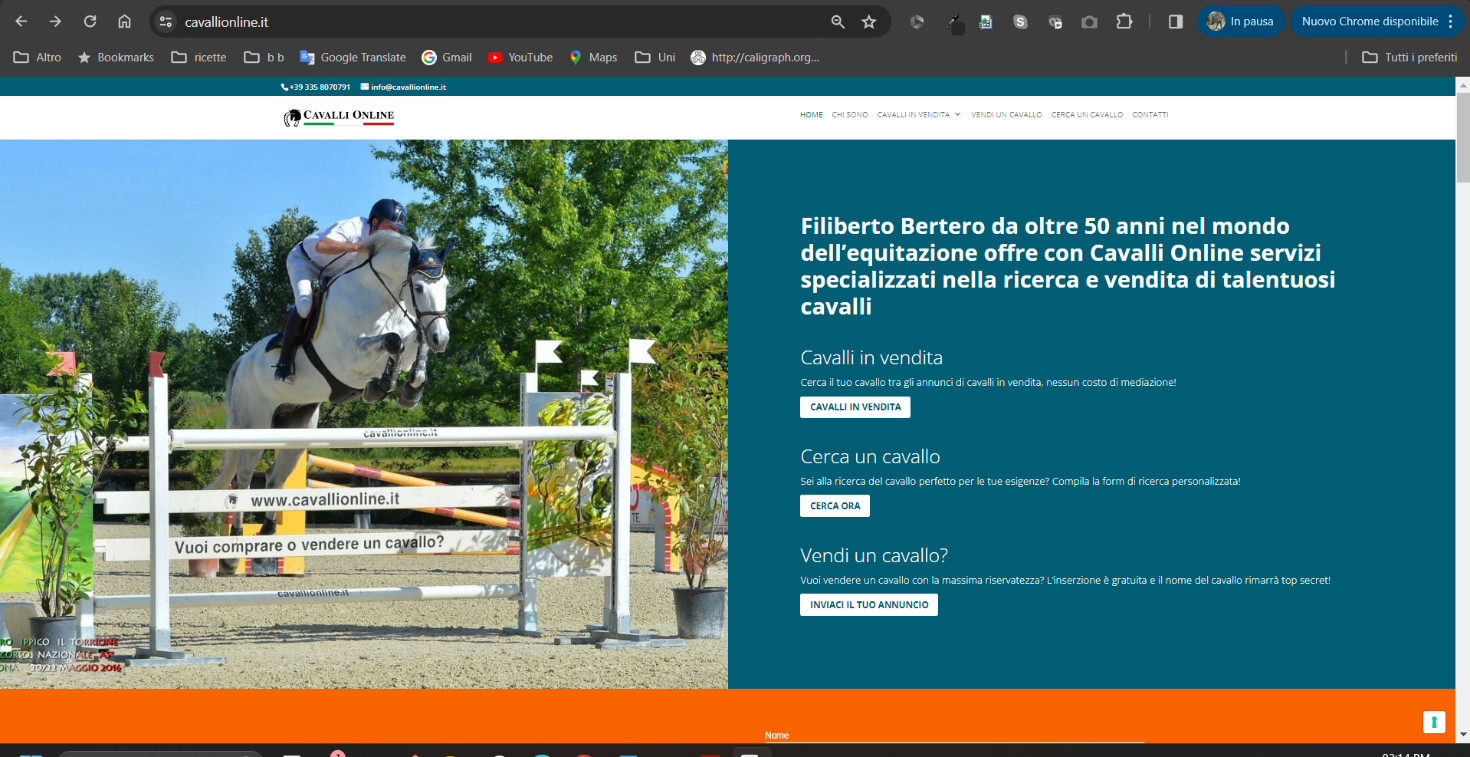
 La gestione degli annunci di vendita on line di cavalli ha richiesto uno sforzo di ricerca differente. Al momento esistono svariate modalità di pubblicazione degli annunci: allevatori e/o commercianti che hanno un sito web per la propria attività, piattaforme di vendita generiche come “*Subito*”, gruppi, pagine e post su social media come per esempio “*Facebook*”. Di seguito vengono proposti degli esempi.

Figura 4: Home page "Cavallionline"

L’immagine mostra l’home page di un commerciante di cavalli di Vinovo (TO) <https://www.cavallionline.it/>.

Cliccando su “cavalli in vendita” si accede alla sezione annunci vera e propria di cui si riporta un’immagine della pagina.



Figura 5: Sezione annunci del sito "Cavallionline"

La lista di cavalli comprende cavalli sportivi, da equiturismo e da riproduzione ma le opzioni di filtro sono molto limitate: è possibile filtrare solo per impiego. Questo rende particolarmente oneroso il tempo di ricerca. Per esempio, sarebbe utile poter filtrare per anno di nascita o età, zona di scuderizzazione, razza, prestazioni sportive.



Figura 6: Informazioni di dettaglio di un annuncio del sito "Cavallionline"

La figura sopra mostra uno degli annunci del sito (<https://www.cavallionline.it/cavallo/femmina-baia-selle-francais-del-2009-cav-498/>) che appare completo relativamente alle informazioni preliminari necessarie per valutare un acquisto. Manca però il luogo di scuderizzazione. Le informazioni fornite hanno un formato strutturato e il metodo di contatto del venditore è indicato esplicitamente ed è ben visibile. Il prezzo non è pubblicamente dichiarato. Questo è una costante in tutti gli annunci di vendita di cavalli.

Un’alternativa generalmente utilizzata da privati per la pubblicazione di annunci di vendita di cavalli sono le innumerevoli pagine *Facebook* dedicate.



Figura 7: Esempio di annuncio di vendita su gruppo Facebook

L’annuncio sopra riportato è stato estratto dalla pagina *Facebook* “Cavalli in vendita”. Sicuramente in questo caso l’informazione è meno strutturata, manca l’indicazione sulla razza e come sempre il prezzo. Le pagine facebook, inoltre, non offrono alcun tipo di filtro a supporto della selezione degli annunci.

Una menzione particolare merita la piattaforma Equiresult (<https://equiresults.com/>) che è un riferimento importante per tutti gli appassionati del settore. Equiresult è una enorme banca dati di tutti i cavalli sportivi che hanno partecipato ad almeno una competizione ufficiale FISE. E’ possibile consultare i risultati di qualsiasi concorso FISE ospitato da un maneggio italiano. Presenta, inoltre, una sezione con annunci di vendita.

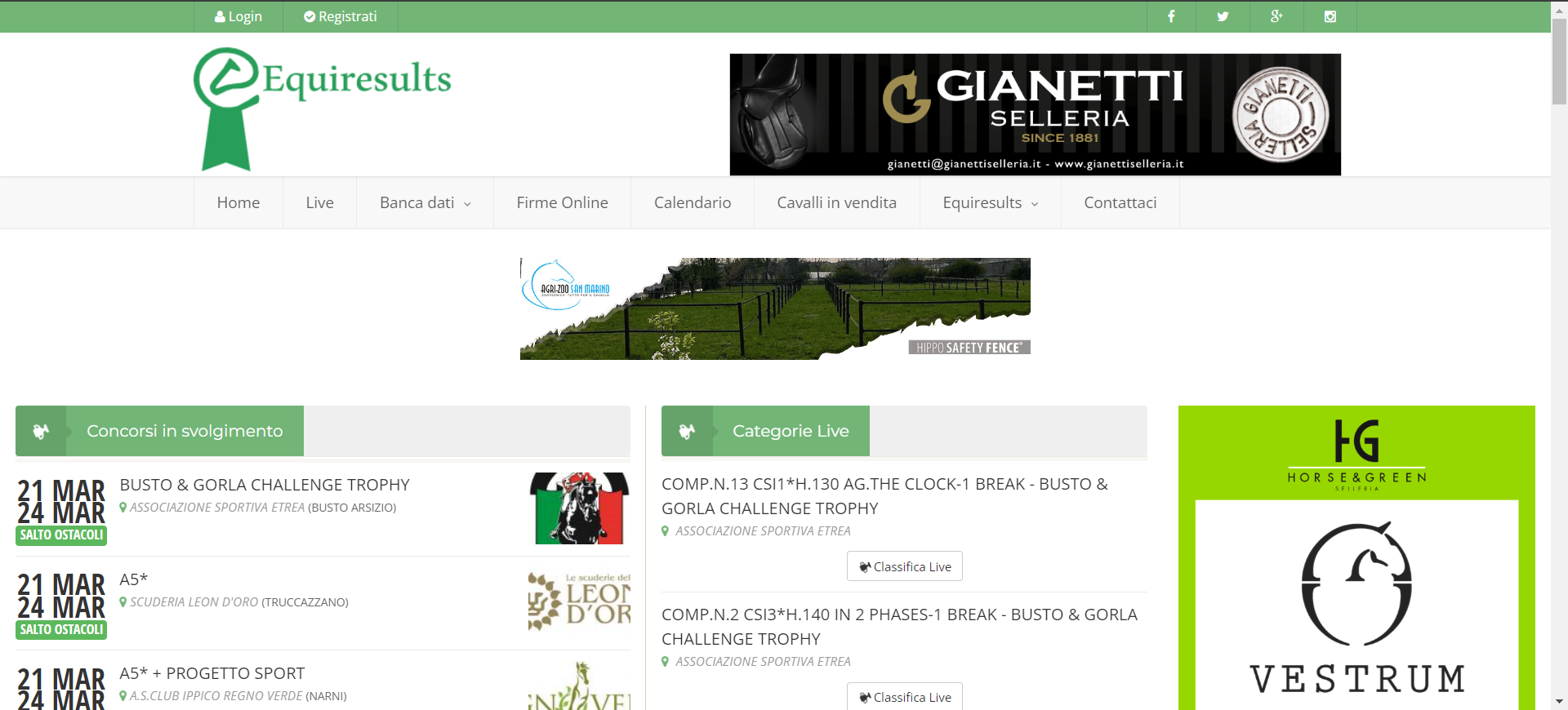


Figura 8: Home page di “Equiresults”

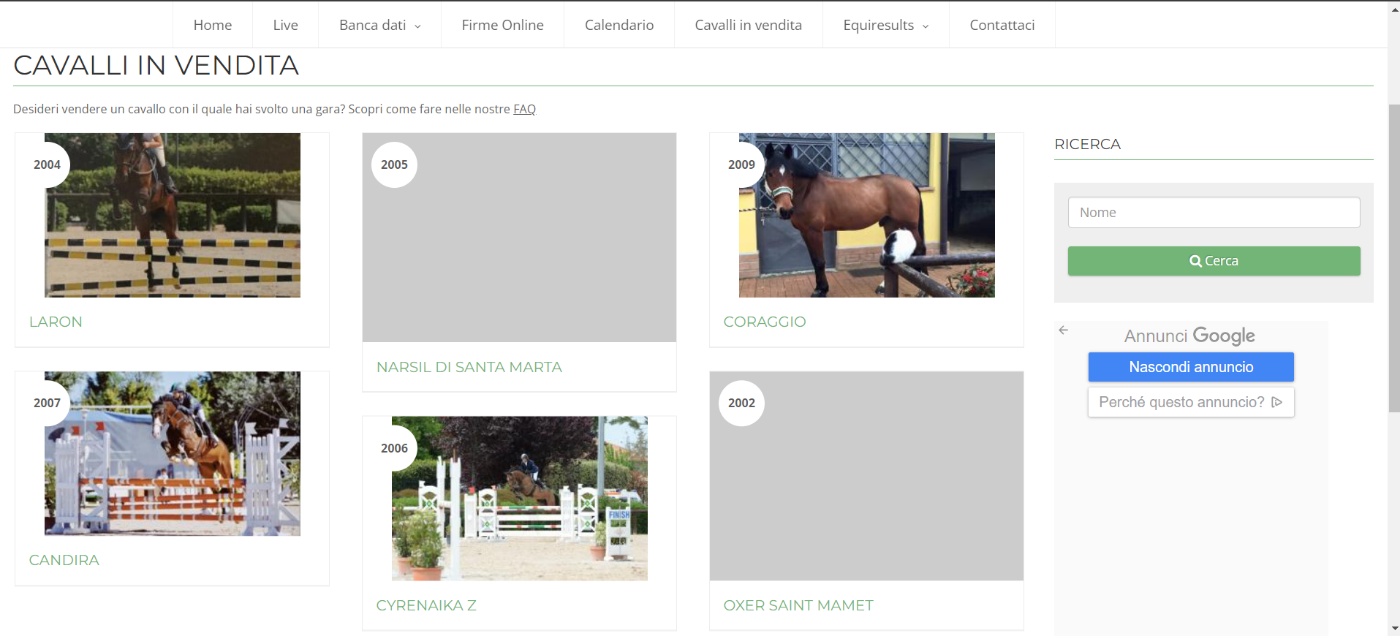


Figura 9: Sezione annunci di vendita del sito "Equiresults"

L’unico filtro possibile è per “nome del cavallo” che generalmente è l’informazione meno conosciuta nel processo di acquisto.

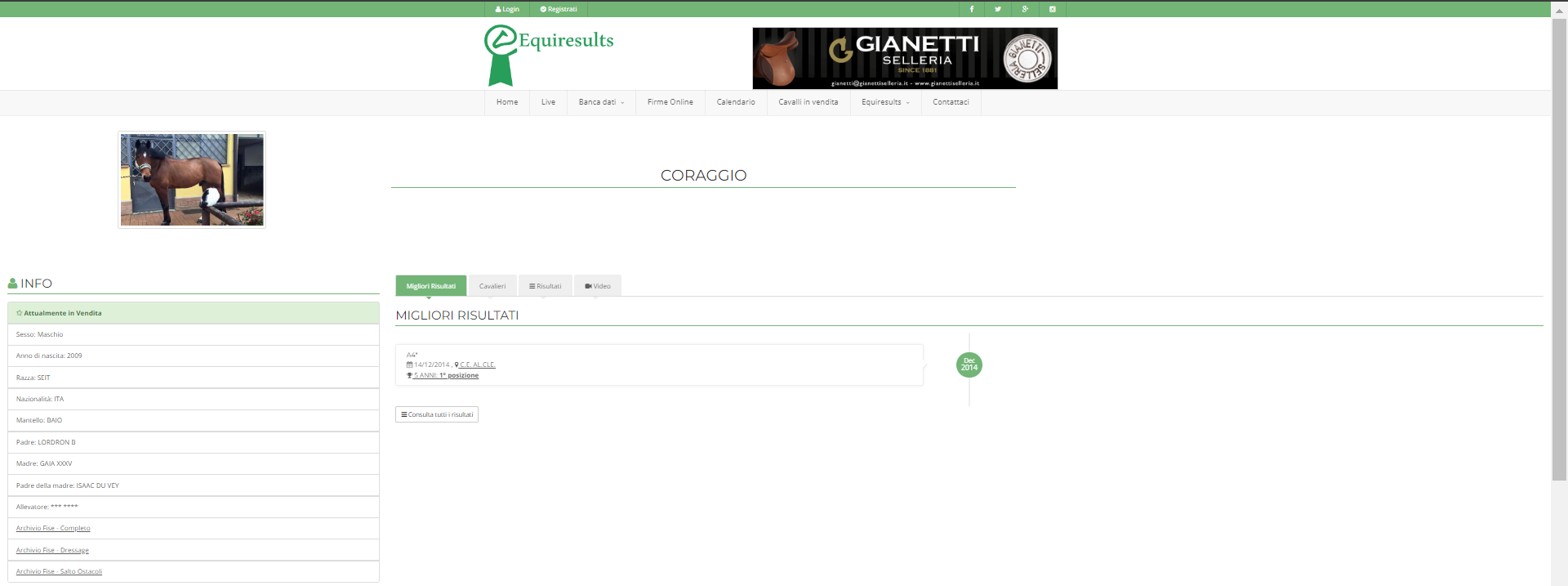


Figura 10: Dettaglio annuncio di vendita del sito "Equiresults"

La figura sopra mostra il dettaglio di un annuncio. Le informazioni non presenti nel riguadro a sinistra sono ottenibili navigando il sito. Manca il prezzo di vendita e non è chiaro come poter contattare il venditore.

Sulla base degli esempi illustrati che sono ben rappresentativi di quello che attualmente è lo stato dell’arte del settore, si comprende che non esiste di fatto una piattaforma unica che gestisca sia le informazioni generali di una razza che quelle di un determinato soggetto in vendita. L’ontologia che viene proposta con questo lavoro ha come fine quello di coniugare la conoscenza e l’opportunità.

* 1. **Allineamento con ontologie esistenti**

Particolarmente laboriosa è stata la ricerca di ontologie esistenti alle quali poter riferire classi, individui e proprietà dell’ontologia oggetto di questo lavoro.

Per quanto riguarda le razze è stato realizzato un allineamento con Wikidata, sebbene esistano anche risorse ontologiche come Animal breed Ontology e Livestock Breed Ontology.

E’ stato effettuato anche un allineamento con l’ontologia “*Friend of a Friend*” per la classe “Persona”, con l’ontologia “*dbpedia*” per le classi “Cavallo” e “Cavaliere” e con “*schema.org*” per la classe “area geografica.

La tabella seguente sintetizza l’elenco delle ontologie con il loro namespace e prefisso.

|  |  |  |
| --- | --- | --- |
|  | **NAMESPACE PREFIX** | **PREFIX** |
| **Friend of a Friend** | http://xmlns.com/foaf/0.1/ | foaf: |
| **Dbpedia** | http://dbpedia.org/ontology/ | dbo: |
| **Schema.org** | http://schema.org/ | schema: |

Tabella 1:Elenco ontologie utilizzate per l'allineamento

L’allineamento con risorse non ontologiche come Wikidata è stato realizzato per mezzo di skos.

1. **Tool LODE**

Il progetto LODE (https://essepuntato.it/lode), sviluppato da Silvio Peroni, è un servizio fornito da un sito web che genera una documentazione completa di una ontologia a partire dal suo file .ttl. In realtà, poichè tale servizio on line presenta da tempo alcuni problemi tecnici che lo rendono irraggiungibile o inutilizzabile, per poter ottenere una documentazione completa dell'ontologia oggetto di questo lavoro è stato utilizzato pyLODE, un'implementazione in python del medesimo servizio fornito dal sito LODE utilizzabile in locale. Una copia aggiornata della documentazione è disponibile sul repository github.

1. **Visualizzazione dell’ontologia**
   1. **Tassonomia delle classi**

Le figure seguenti mostrano a livelli progressivi di dettaglio la tassonomia delle classi dell’ontologia evidenziando unicamente la relazione di sussunzione tra queste. Tale rappresentazione grafica è stata prodotta per mezzo del tool OntoGraf di Protegè.



Figura 11:Tassonomia delle classi di primo livello

Si espandono, ora, le classi evidenziate per rappresentare la loro ulteriore articolazione in sottoclassi fino ad arrivare ai livelli più bassi della gerarchia.

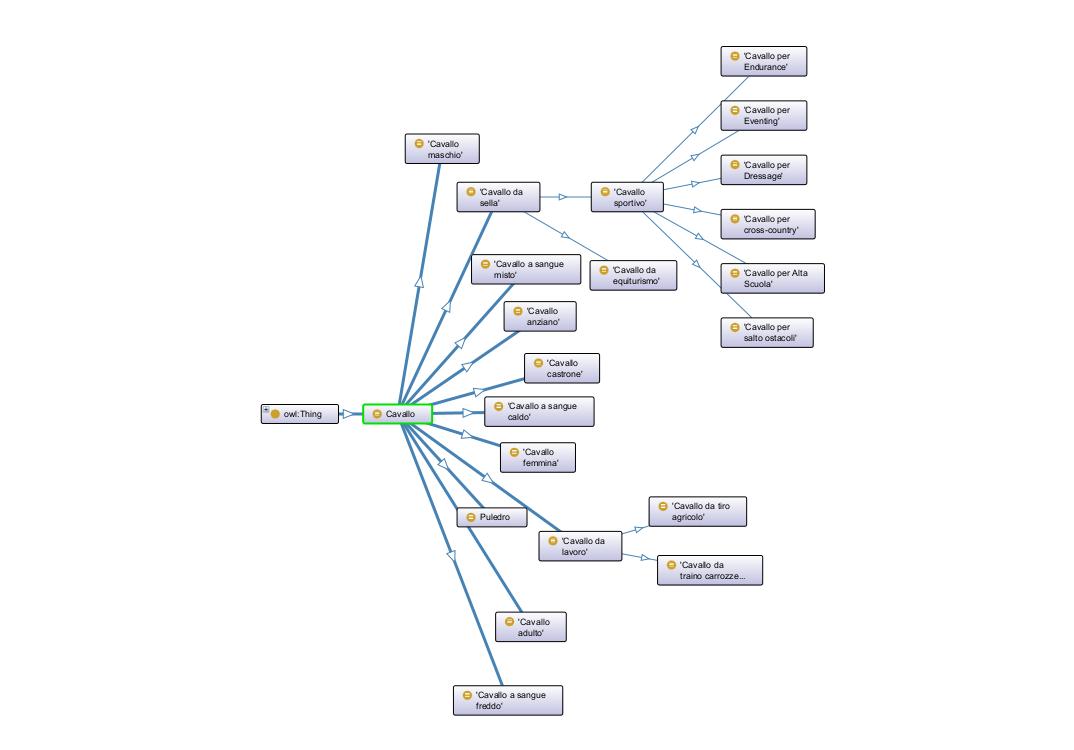


Figura 12: Sottoclassi della classe "Cavallo"

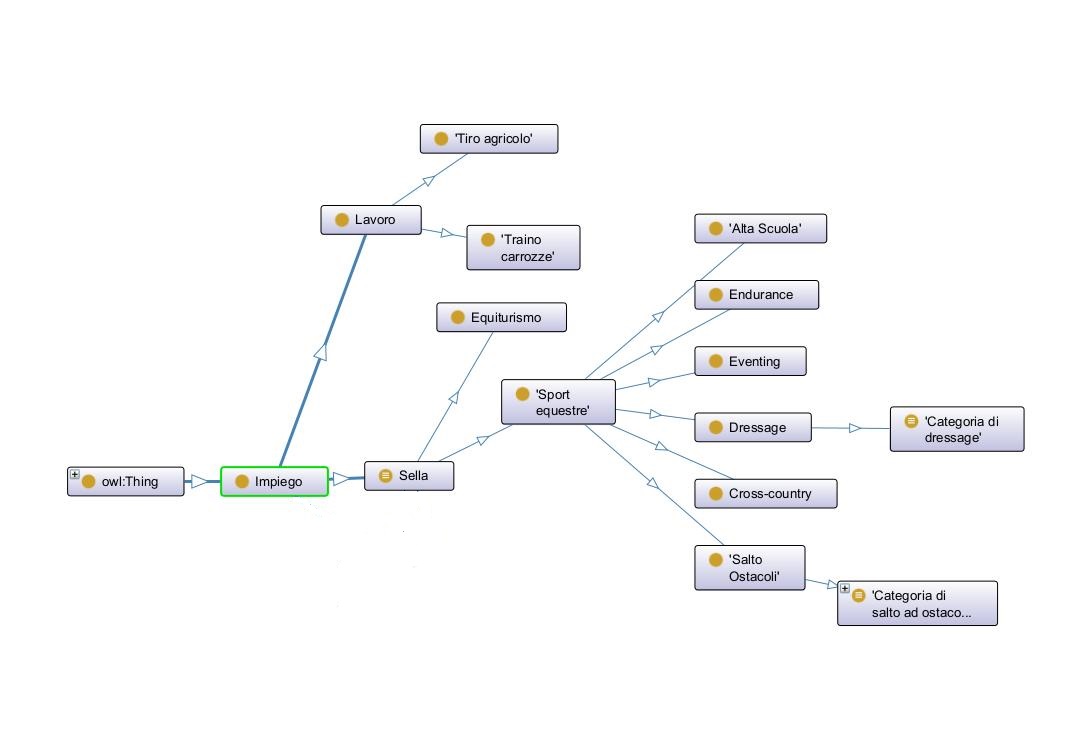


Figura 13: Sottoclassi della classe "Impiego"

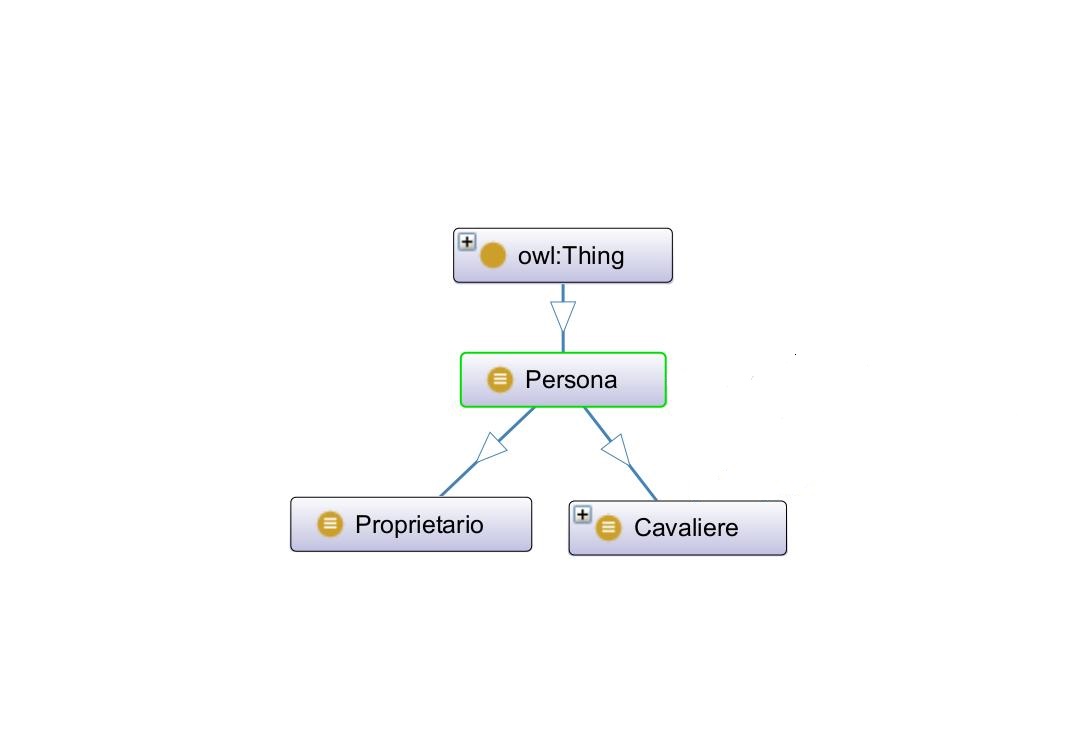


Figura 14: Sottoclassi della classe "Persona"

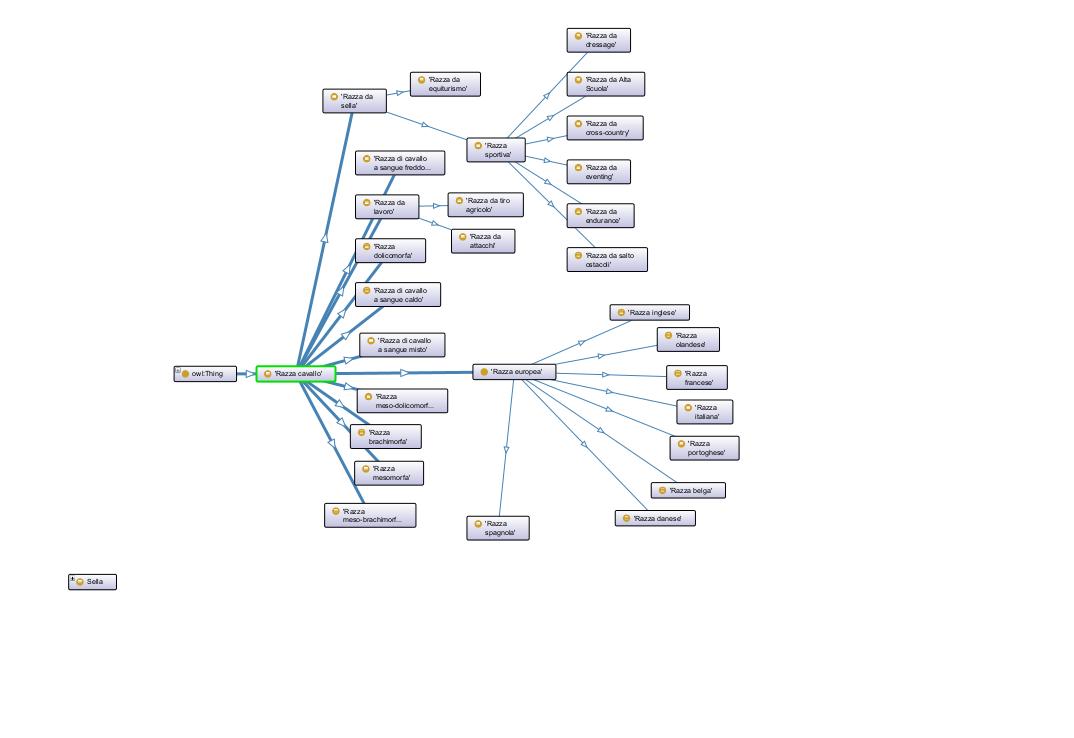


Figura 15: Sottoclassi della classe "Razza cavallo"

* 1. **Knowledge graph**

In questa sezione sono presentati i knowledge graph relativi ad alcuni esempi istanziati nell’ontologia. I grafici sono stati realizzati tramite la Linked Data Platform GraphDB che è un sistema di gestione di database a grafo progettato per memorizzare, gestire e interrogare dati basati su grafi in modo efficiente. GraphDB utilizza il linguaggio di interrogazione RDF per interagire con i dati, ed è compatibile con gli standard RDF come SPARQL per le query e OWL per l'ontologia. E’ anche dotato di motori di inferenza per supportare il ragionamento ontologico e la deduzione di nuove informazioni dai dati esistenti.

L’ontologia oggetto di questo lavoro contiene due tipologie di dati e di individui per le motivazione illustrate nei praragrafi precedenti. Da un lato individui e dati legati al concetto di razza, dall’altro individui e informazioni connesse con uno specifico cavallo in vendita. Si presentano, perciò, separatamente le due tipologie di template istanziati su un esempio.

Per i dati relativi alla razza si è scelto come esempio il knowledge ghaph relativo all’entità “Murgese”.



Figura 16: Knowledge graph di una razza di cavallo: il Murgese

I nodi del grafico sono stati posizionati intorno al nodo centrale di esempio in modo tale che nella parte alta ricadano i nodi relativi alle classi di appartenenza dell’entità considerata inferite sulla base delle proprietà rappresentate da archi e nodi presenti nella parte bassa. La triangolazione tra le proprietà “ha regione di origine”, “is-Part-of” e “has nationality” è un esempio di property chain in cui la nazionalità italiana è inferita partendo dall’informazione che la regione di origine del Murgese sono le Murge e che questa è una parte del territorio italiano.

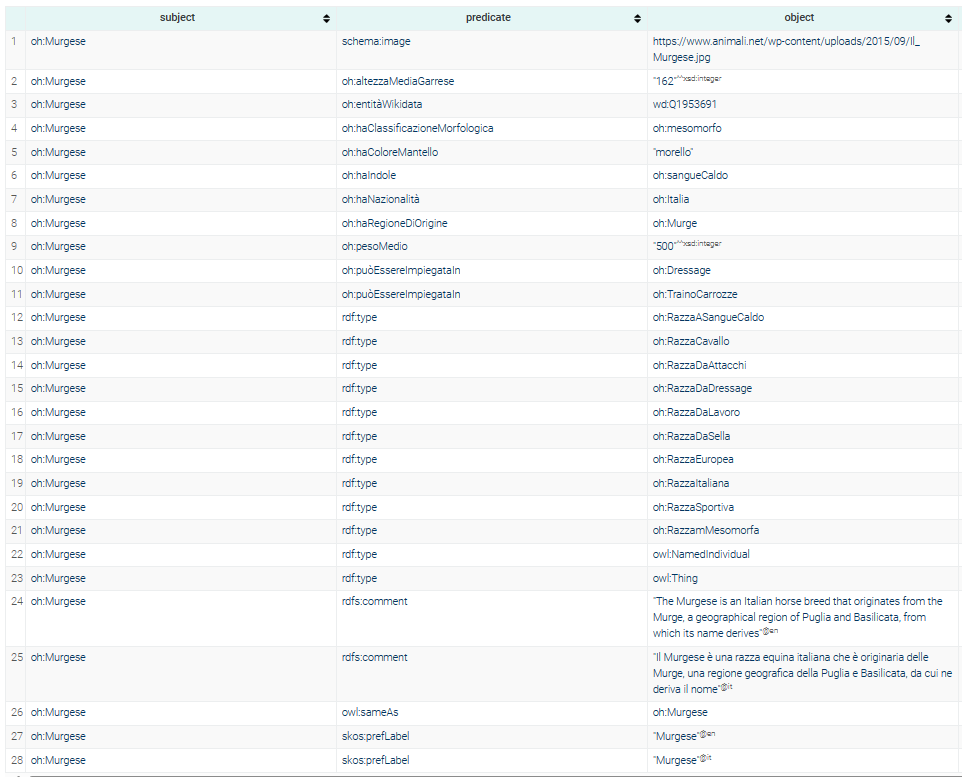
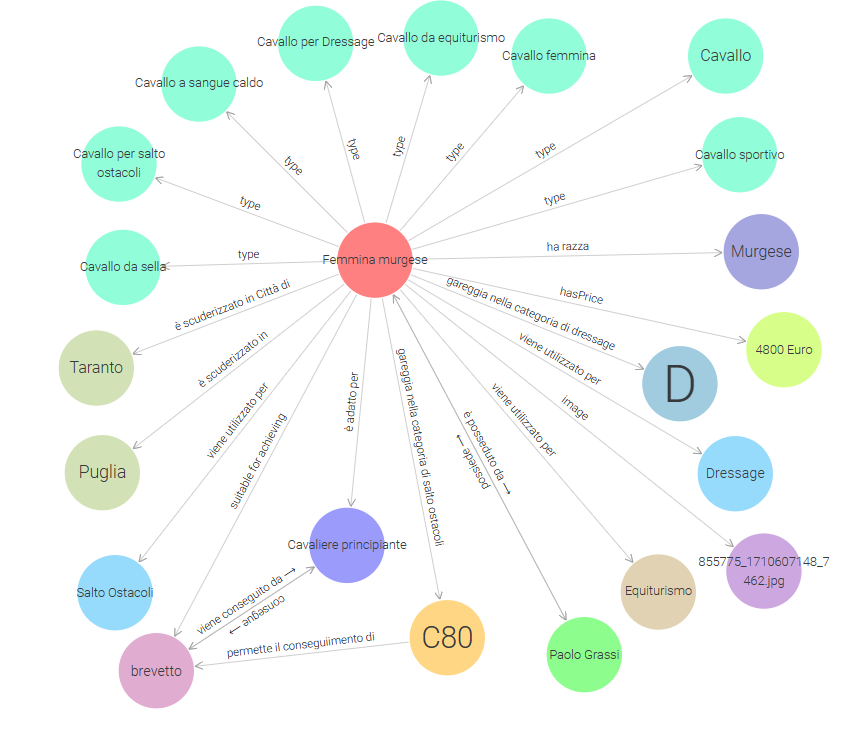


Figura 17: Triple dell’entità “Murgese” in forma tabellare

Si considera, ora, il knowledge graph di un cavallo in vendita, in particolare dell’entità “Femmina Murgese”.



Come nell’esempio precedente, nella parte alta del grafo sono posti i nodi che rappresentano le classi di appartenenza dell’entità considerata inferite sulla base delle proprietà ad essa associate. Le triangolazioni tra nodi e archi rappresentano come prima delle property chain.



Figura 18:Triple dell'entità "Femmina Murgese" in forma tabellare

1. **Queries SPARQL**

Il set di query SPARQL che si propone riflette la duplice natura di questa ontologia.