

# Sommario

1.	An	nalisi de	el problema	. 3			
	1.1.	Indic	cazione del contesto	. 3			
	1.2.	Desc	rizione della realtà di interesse relativa al tema	. 3			
	<b>1.3.</b> limita		Descrizione preliminare sintetica del contenuto del sito, delle eventuali assunzioni fatte ed eventuali oni				
2.	Ar	Architettura del sistema					
	2.1.	Desc	zione della struttura del sito				
	2.2.	Fogli	ogli di stile esterni4				
	2.3.	Conf	figurazione del database server				
	2.4.	Infor	rmazioni Semantiche	. 5			
	2.4	4.1.	Microdata	. 5			
	2.4	4.2.	RDFa	. 5			
	2.4	4.3.	JSON-LD	6			
3.	Impleme		ntazione e Interfaccia	7			
	3.1.	Man	utenzione del sito	7			
	3.2	I.1. Homepage		. 7			
	3.1	1.2.	Registrazione	8			
	3.2	1.3.	Assistenza	8			
	3.2	1.4.	Genere	9			
	3.1	1.5.	Videogioco	9			
4.	Te	st		12			
	4.1.	Test	su query DBpedia	12			
	4.2.	Test	dati strutturati	12			
	4.2	2.1.	Strumento di convalida del markup	12			
	4.2	2.2.	Strumento ufficiale di Google	14			
	4.2	2.3.	Test RDFa e JSON-LD	15			
	4.2	2.4.	Test intero sito dinamico in locale	16			
5.	Pip	peline a	ılternativa	17			
_	Б.			17			

## 1. Analisi del problema

#### 1.1. Indicazione del contesto

**NIGG Games** è un sito web per la recensione e descrizione di videogiochi, che offre ai visitatori non solo informazioni tecniche ma anche impressioni personali sull'esperienza di gioco.

#### 1.2. Descrizione della realtà di interesse relativa al tema

I videogiochi al giorno d'oggi sono divenuti un vero e proprio fenomeno di massa che sta conoscendo un rapido sviluppo, per tale ragione l'acquisto di uno di essi non è sempre cosa facile o scontata. I videogames hanno, inoltre, conquistato sempre più spazio nella realtà quotidiana perché hanno invaso vari dispositivi elettronici e console come cellulari, computer, Xbox e Playstation. Per tale ragione ne esistono di innumerevoli tipi e categorie. Di conseguenza è sempre maggiore il rischio di perdere qualche novità o caratteristica fondamentale.

È da qui che nasce la necessità di creare un sito di recensioni online di videogiochi che guidi il cliente ad una scelta consapevole, tenendolo continuamente aggiornato sia sui dettagli che sulle offerte riguardanti un prodotto, come già alcuni siti fanno, ad esempio Everyeye e Multiplayer, con recensioni e videorecensioni di esperti del genere.

Inoltre, per offrire un'esperienza quanto più completa possibile agli utenti, è necessario fornire loro quante più informazioni possibili: è qui che entrano in gioco le **tecnologie semantiche**, che permettono sia di ottenere dati e risorse senza la necessità di averle a disposizione in un database proprietario del sito, sia di annotare semanticamente gli elementi all'interno delle pagine web del sito in modo da fornire ai motori di ricerca dati significativi che possano rispondere in modo inequivocabile ad uno specifico intento di ricerca.

# 1.3. Descrizione preliminare sintetica del contenuto del sito, delle eventuali assunzioni fatte ed eventuali limitazioni

Il sito base è caratterizzato da diversi titoli di videogiochi, suddivisi in base alla loro categoria di appartenenza, tutti accompagnati da: descrizione, recensione, prezzo e un riferimento per l'acquisto. Inoltre, per ogni gioco, vengono mostrate informazioni aggiuntive utili prese tramite query Sparql su DBpedia. Ogni utente avrà la possibilità di effettuare o meno la registrazione al sito e, a seconda della modalità scelta, i contenuti saranno differenziati. A tal proposito, infatti, solo chi effettua il login avrà a disposizione la recensione, il prezzo e un possibile rimando all'acquisto. L'homepage del sito presenta i contenuti in evidenza, le varie categorie trattate, riferimenti ad alcune delle case produttrici di videogiochi ottenuti tramite query Sparql a DBpedia e la possibilità di effettuare il login, il quale può essere effettuato solo dopo la registrazione al sito; per tale scopo è presente una sezione dedicata che permette ciò tramite l'inserimento dei dati da parte dell'utente. Scelta una categoria d'interesse, saranno mostrati tutti i videogiochi appartenenti ad essa e, selezionandone uno, si otterranno tutte le informazioni desiderate. In tale pagina l'utente può esprimere il proprio parere valutando quanto letto tramite delle reaction. È presente, inoltre, una sezione dedicata all'assistenza degli utenti, i quali potranno contattare gli amministratori del sito.

### 2. Architettura del sistema

#### 2.1. Descrizione della struttura del sito

Il sito dinamico realizzato, disponibile nella cartella "gruppo3" è composto da cinque pagine web:

- Homepage, accessibile tramite l'url "http://localhost/gruppo3/homepage.php", è la pagina di presentazione del sito, in cui il visitatore oltre ad avere la possibilità di visualizzare i videogiochi di maggiore rilevanza, accompagnati da una breve anteprima, può, grazie alla suddivisione per categorie, accedere alla sola categoria di interesse. Nella homepage, così come nelle altre pagine del sito, è possibile effettuare il login, raggiungere la pagina di registrazione, di assistenza, di genere, accedendovi dal menù verticale, o di videogioco, selezionandone uno fra quelli in evidenza. In questa pagina è possibile vedere i risultati di una prima query su DBpedia nel riquadro "Semantic Info" che riporta alcune delle case produttrici di videogiochi.
- **Genere**, raggiungibile tramite l'url "http://localhost/gruppo3/genere.php", ha lo scopo di presentare tutti i videogames appartenenti alla categoria selezionata, garantendo al visitatore di visualizzare solo ciò che gli interessa. Dalla suddetta è possibile effettuare il login, accedere alla pagina di registrazione, di assistenza, alla homepage e, scegliendo un videogioco, si può arrivare alla pagina ad esso dedicata.
- Videogioco, accessibile tramite l'url "http://localhost/gruppo3/videogioco.php", è la pagina in cui il visitatore potrà leggere solo la descrizione del videogioco in questione, a differenza dell'utente che ha effettuato il login, il quale avrà a disposizione anche la recensione, un riferimento ai siti d'acquisto con i relativi prezzi e la possibilità di aggiungere una reazione a quanto appena letto. Da qui è possibile effettuare il login, accedere alla pagina di registrazione, di assistenza, alla homepage e quella di genere scegliendo una categoria dalla lista. Qualsiasi tipologia di utente potrà visualizzare le informazioni sul videogioco ricavate tramite query su DBpedia nel riquadro Semantic Info.
- Assistenza, raggiungibile tramite l'url "http://localhost/gruppo3/contact.php", contiene le informazioni
  dei tre amministratori affinché l'utente, in caso di problemi o qualora volesse effettuare delle richieste
  specifiche, come l'aggiunta di un videogioco con i relativi contenuti, abbia dei contatti di riferimento.
  Permette di effettuare il login, accedere alla pagina di registrazione, di assistenza, alla homepage e quella
  di genere scegliendo una categoria dalla lista.
- **Registrazione**, accessibile tramite l'url "http://localhost/gruppo3/registrazione.php", è la pagina in cui un visitatore può iscriversi per diventare un utente, avendo così la possibilità di visualizzare maggiori contenuti sui videogiochi e di esprimere un parere tramite reazione. Tale pagina è direttamente collegata con la homepage e la pagina di assistenza.

#### 2.2. Fogli di stile esterni

Per tutte le pagine del sito web viene utilizzato il foglio di stile **style.css**, fatta eccezione per lo slideshow presente nella homepage per il quale è stato utilizzato il foglio di stile **slideshow.css**.

#### 2.3. Configurazione del database server

Versione PostgreSQL: PostgreSQL 14Versione PgAdmin: 6.1Indirizzo del server: localhostNome database: gruppo3

Username: www Password: tsw2020

#### 2.4. Informazioni Semantiche

Grazie alla semantica possiamo assegnare un significato ai nostri elementi HTML, che di per sé non portano alcuna informazione semantica. La vecchia SEO (Search Engine Optimization, la branca del Digital Marketing che si occupa di migliorare la posizione delle pagine web nei risultati organici non sponsorizzati dei motori di ricerca) era basata soltanto su keywords e backlink; invece, la SEO semantica fornisce ai motori di ricerca ulteriori dati significativi: invece di puntare a una parola chiave, è possibile costruire un contesto forte in un formato che il motore di ricerca possa liberamente leggere e interpretare. In questo modo è possibile dialogare meglio, offrire informazioni più precise e migliorare anche il traffico organico tramite i Rich Snippets. Per fare ciò gli algoritmi di ricerca fanno uso di entità: un'entità può essere tutto ciò che esiste nel mondo reale; in pratica, un modo per trasformare una pagina web in un'entità è trasformare il contenuto in dati, usando un vocabolario che i motori di ricerca possono capire. Nel nostro caso, è stato deciso di usare Schema.org, che rappresenta un modo sistematico e riconosciuto per descrivere le pagine web in modo che i motori di ricerca possano capirle senza ambiguità, facendo uso di dati strutturati, un modo per organizzare le informazioni per una migliore accessibilità.

Per marcare gli elementi della pagina html esistono diversi approcci: Microdata, RDFa e JSON-LD. La nostra scelta finale è ricaduta sull'uso dei **Microdata** in quanto, essendo il sito a nostra disposizione scritto con HTML5 e php, risultava più complesso convertirlo in JSON-LD e, per quanto riguarda RDFa, pur essendo simile a Microdata, è più vecchio e complesso. Nonostante è stato scelto di usare i Microdata per il sito dinamico, sono state create le versioni statiche di tutte le pagine sia con i Microdata, con RDFa e con JSON-LD. Queste pagine statiche si trovano nella cartella "Pagine statiche" e sono salvati nei seguenti formati "xxx-rdfa.html", "xxx-microdata.html" e "xxx-jsonld.html"

#### 2.4.1. Microdata

I microdati sono una specifica HTML utilizzata per annidare i metadati all'interno del contenuto esistente nelle pagine web. Motori di ricerca, web crawler e browser possono estrarre ed elaborare microdati da una pagina web e utilizzarli per fornire un'esperienza di navigazione più ricca per gli utenti. La sintassi dei microdati si basa su coppie di nomi e valori, chiamati "items", e si sviluppa in tre passaggi: prima si identifica un elemento e lo si evidenzia come item, che viene assegnato ad un tipo specifico tratto dal vocabolario di Schema.org. Dopo questo secondo passaggio, alla coppia item-type vengono assegnate diverse proprietà. Il markup avviene tramite attributi dell'HTML5, cioè itemscope, itemtype e itemprop:

- **itemscope:** questo attributo è usato nei tag div per definire un'area ben delineata come item, specificato meglio tramite gli attributi itemtype e itemprop;
- **itemtype:** questo attribuito si può applicare a tutti gli elementi, che sono evidenziati con l'elemento itemscope sulla base di uno schema predefinito. Così, sempre basandosi su Schema.org, è possibile ordinare gli elementi importanti di un sito secondo tipologie generali, che possono essere elaborate da tutti i motori di ricerca;
- **itemprop:** questo attributo fornisce una proprietà all'itemtype spiegato sopra. Per sapere quali proprietà assegnare ad un itemtype si rimanda al <u>sito di Schema.org</u>.

#### 2.4.2. RDFa

RDFa (Resource Description Framework in Attributes) è un formato per l'inserimento di affermazioni RDF nei file HTML, XHTML o altre varianti XML, che il W3C consiglia di usare come standard. I programmatori utilizzano RDF per specificare meglio i contenuti delle pagine di un sito grazie ai metadati, in modo da permettere ai browser e ai crawler di comprendere le relazioni semantiche tra gli elementi. RDFa introduce una serie di nuovi attributi per l'inserimento dei metadati, che ampliano il vocabolario di base dei documenti scritti nei rispettivi linguaggi di markup:

 vocab: definisce il vocabolario desiderato alla base del markup degli elementi con RDFa (ad esempio Schema.org);

- typeof: si assegnano agli elementi degli schemi precisi, chiamati "types", secondo il vocabolario scelto;
- property: vengono assegnate delle proprietà agli elementi.

#### 2.4.3. JSON-LD

JSON-LD (JavaScript Object Notation for Linked Data) è progettato intorno al concetto di "contesto", per fornire delle mappature aggiuntive da JSON verso un modello RDF. Il contesto collega le proprietà di un oggetto in un documento JSON ai concetti di un'ontologia. Per poter effettuare il mapping della sintassi JSON-LD verso RDF, JSON-LD permette di forzare il tipo di certi valori. Contrariamente agli altri formati per la strutturazione dei dati, come RDFa e i microdati, la segnalazione non avviene come un'annotazione del codice sorgente, ma i metadati vengono implementati sotto forma di un tag script separati dal contenuto del sito.

JSON-LD aggiunge le classiche coppie nome-valore dell'annotazione JSON tramite **keyword (parole chiave)**, introdotte dal simbolo @anteposto. Sono molto importanti le keyword **@context** e **@type**. Mentre **@context** definisce il markup del vocabolario alla base, **@type** indica di quale tipo di dati (schema) si tratta.

## 3. Implementazione e Interfaccia

#### 3.1. Manutenzione del sito

Partendo dal sito progettato nel corso di Tecnologie Software per il Web, sono sia stati apportati dei miglioramenti grafici sia sono state aggiunte componenti semantiche per strutturare i dati e per mostrare più informazioni tramite delle query a DBpedia.

In questa sezione verranno mostrati tutti i cambiamenti apportati alle singole pagine e le nuove aggiunte.

#### 3.1.1. Homepage







Figura 2: nuova homepage

Nella homepage, disponibile nel file *homepage.php*, oltre al restyling grafico, sono state aggiunte due tipologie di contenuti semantici:

• Markup semantico tramite Microdata usando il vocabolario di Schema.org:



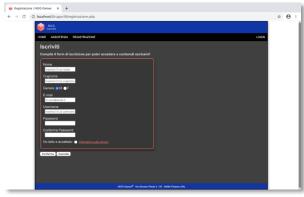
• Query Sparql su DBpedia per ottenere informazioni riguardanti alcune delle case produttrici di videogiochi più note. In particolare, la query di Sparql è stata formulata nel seguente modo:

```
$toChange = array("<",">"," ");
$changeWith = array("%3C","%3E","%20");
   $siti = array("Sega development studios", "Blizzard Entertainment", "Electronic Arts", "Capcom");
   foreach ($siti as $nome){
      $url = str_replace($toChange, $changeWith, 'http://dbpedia.org/sparql?que y=SELECT ?name ?img ?home WHERE
                                                                      {?name dbp:name|rdfs:label "
                                                                                             ''.trim($nome).'"@en
                                                                      OPTIONAL{?name dbo:thumbnail ?img}
                                                                     OPTIONAL{?name dbp:homepage|dbp:website ?home}}&format=text/csv')
      $response = file_get_contents($url);
      $row = explode("\n", $response);
$result = explode(",", $row[1]);
      $sito = str_replace("http://dbpedia.org/resource/","",$result[0]);
      $sito = str_replace("_",
                            ",$sito);
      $img = $result[1];
      $website = $result[2];
```

La query restituisce il nome dell'azienda, il link ad un'immagine del suo logo e il link al sito della stessa. Dopo una serie di manipolazioni della stringa ottenuta in risposta dalla query, i loghi vengono mostrati tramite il tag <img> e resi cliccabili tramite un collegamento ipertestuale con il tag <a> che reindirizza alla pagina dell'azienda: gli url usati per fare ciò sono quelli ottenuti dalla query.



#### 3.1.2. Registrazione



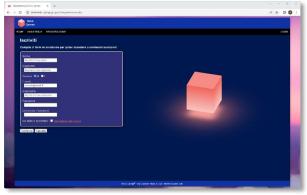
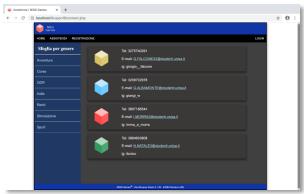


Figura 3: form precedente

Figura 4: nuovo form

In tale pagina, disponibile nel file registrazione.php, è stata effettuata la sola operazione di restyling.

#### 3.1.3. Assistenza





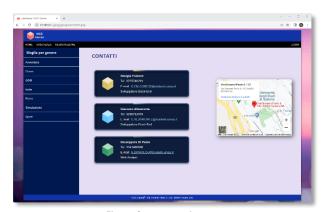


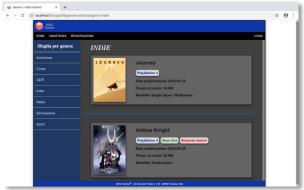
Figura 6: nuova assistenza

In questa pagina, disponibile nel file *contact.php*, oltre all'operazione di restyling sono stati aggiunti dei tag semantici tramite i Microdata usando il vocabolario di Schema.org:



La classe **Person** è stata usata per marcare semanticamente gli sviluppatori del sito, mentre **Organization** e **PostalAddress** sono impiegati nel footer per marcare le informazioni relative all'azienda.

#### 3.1.4. Genere





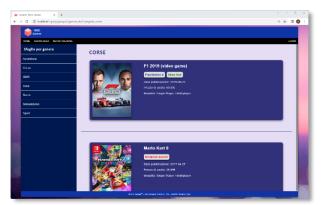


Figura 8: nuovo genere

In questa pagina, disponibile nel file *genere.php*, oltre all'operazione di restyling sono stati aggiunti dei tag semantici tramite i Microdata usando il vocabolario di Schema.org:



La classe **Videogame** è stata usata per marcare semanticamente i videogiochi appartenenti al genere della pagina, **Offer** per marcare il prezzo e **Organization** e **PostalAddress** sono impiegati nel footer per marcare le informazioni relative all'azienda.

#### 3.1.5. Videogioco



Figura 9: videogioco precedente



Figura 10: nuovo videogioco

Nella pagina videogioco, disponibile nel file *videogioco.php*, oltre al restyling grafico, sono state aggiunte due tipologie di contenuti semantici:

Markup semantico tramite Microdata usando il vocabolario di Schema.org:

Videogame	AggregateRating	Review	Offer	Person	Organization	PostalAddress
• trailer • name	<ul><li>ratingCount</li><li>itemReviewed</li></ul>	• name • itemReviewed	• price	• name	• legalName • address	• streetAddress • postalCode
• review • dataPublished	<ul><li>ratingValue</li><li>bestRating</li></ul>	• author • reviewBody			· dddic33	<ul><li>addressLocality</li><li>addressRegion</li></ul>
• playMode	• worstRating	Teviewbody				• sameAs

La classe **Videogame** è stata usata per marcare semanticamente il videogioco nella pagina apposita, **Offer** per marcare il prezzo sui vari siti, **Review** per marcare la recensione, **Person** per individuare l'autore della recensione, **AggregateRating** per il riscontro lasciato dall'utente e **Organization** e **PostalAddress** sono impiegati nel footer per marcare le informazioni relative all'azienda.

• Due Query Sparql su DBpedia per ottenere alcune informazioni aggiuntive riguardanti il videogioco in questione. In particolare, la prima query è stata formulata usando <u>EasyRDF</u>, una libreria php per formulare automaticamente query su DBpedia; questa query restituisce il link alla pagina del videogioco. La seconda query, invece, non fa uso di EasyRDF poiché era più complessa l'estrazione dei singoli dati, per cui è stata usata anche qui, come nella query nell'homepage, la funzione file\_get\_contents() sul cui output sono poi state effettuate delle manipolazioni per prende solamente i dati di interesse.

La seconda query restituisce la prima data di pubblicazione del videogioco, la casa produttrice, il direttore musicale e il genere. Dopo una serie di manipolazioni della stringa ottenuta in risposta dalla query, tutte queste informazioni vengono stampate a video in una lista e, qualora fosse disponibile il link a quella risorsa, sarà possibile accedervi tramite il collegamento ipertestuale inserito nel tag <a>>.

Nell'immagine successiva viene infine riportata la schermata della restante pagina "videogioco.php" dopo che un utente ha effettuato il login.

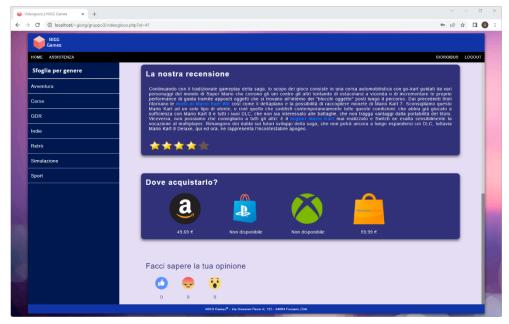


Figura 11: informazioni dopo il login

## 4. Test

#### **4.1.** Test su query DBpedia

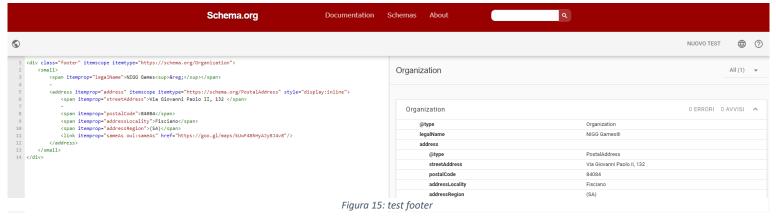
Per formulare correttamente le query Sparql su DBpedia è stato utilizzato lo <u>SPARQL query editor</u> messo a disposizione da DBpedia stesso: una volta validato il formato della query e ottenuti i risultati desiderati, si è proceduto alla formulazione delle query definitive e alla gestione di eventuali eccezioni all'interno del codice php (come, ad esempio, l'assenza di uno dei campi richiesti dalla query).



#### **4.2.** Test dati strutturati

#### 4.2.1. Strumento di convalida del markup

Per la convalida generica dello schema, è stato utilizzato lo <u>Strumento di convalida del markup schema</u> per testare tutti i tipi di markup Schema.org, senza la convalida specifica di Google. In questo modo è stata verificata la correttezza degli attributi semantici aggiunti sulle pagine del sito rese statiche (disponibili nella cartella "Pagine statiche") al fine di poter vedere l'effettivo funzionamento. Vengono riportati di seguito i risultati ottenuti su ogni pagina dopo questo test di validazione.



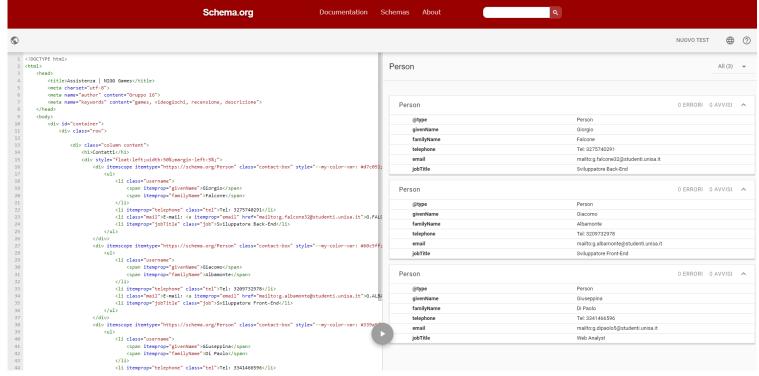


Figura 16: test contatti

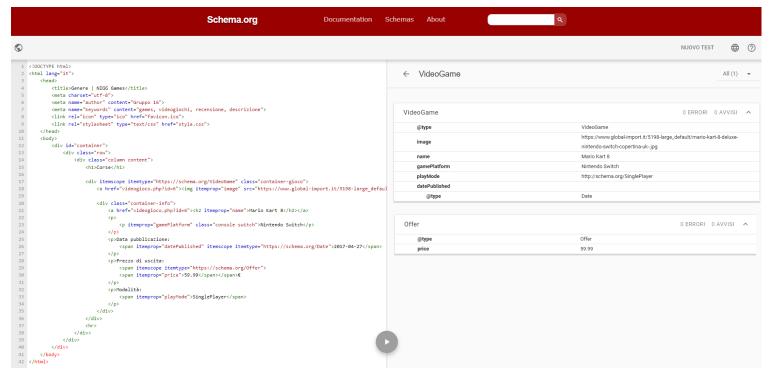


Figura 17: test genere

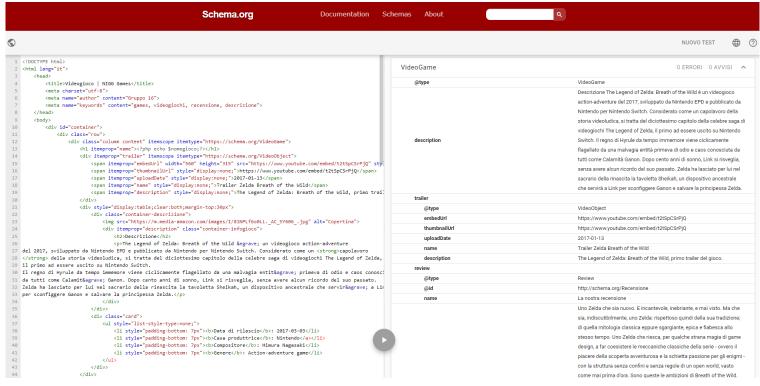


Figura 18: test videogioco

#### 4.2.2. Strumento ufficiale di Google

È stato poi usato lo <u>strumento ufficiale di Google</u> per testare i dati strutturati e verificare quali risultati multimediali di Google possono essere generati dai dati strutturati sulla pagina; inoltre, lo strumento è stato utile anche per capire quali proprietà Google ritenesse necessarie per le varie entità. Tramite questo tool si può anche visualizzare l'anteprima dell'aspetto dei risultati multimediali nella Ricerca Google: nonostante tutti i test siano andati a buon fine, e quindi tutti i microdata fossero inseriti correttamente nei tag, non ci è stato possibile ottenere i rich snippets in quanto solo alcuni tipi di microdata sono raccolti in pochi <u>modelli proposti da Google</u>. Qui di seguito vengono riportati alcuni dei risultati ottenuti con questo tool.

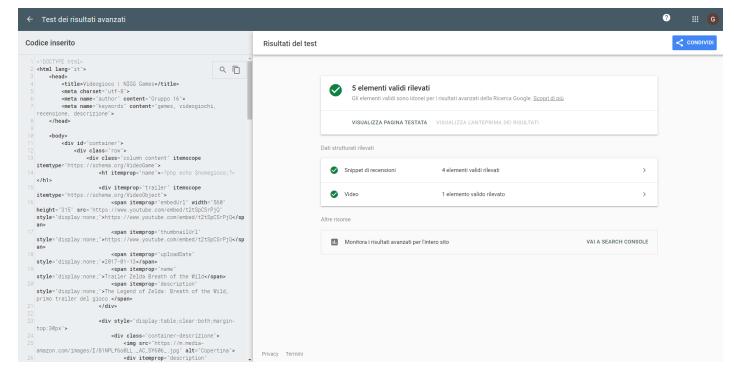


Figura 19: correttezza1 videogioco

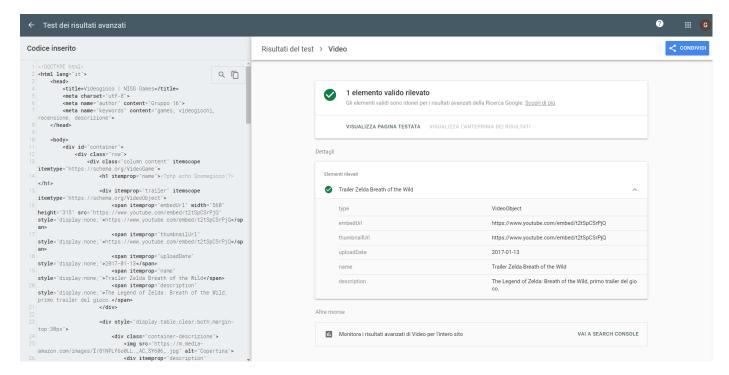


Figura 20: correttezza2 videogioco

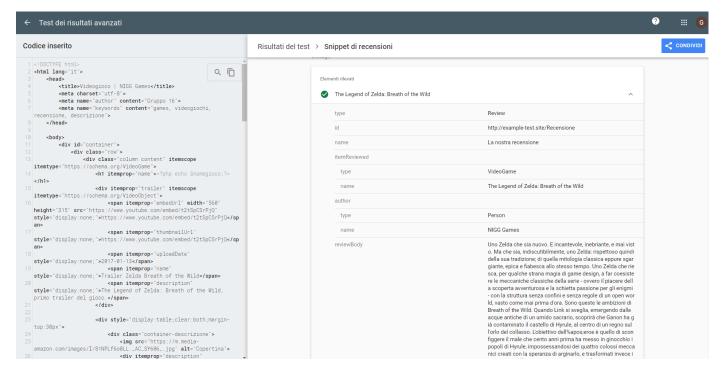


Figura 21: correttezza3 videogioco

#### 4.2.3. Test RDFa e JSON-LD

Gli stessi test effettuati con le pagine statiche con i microdata sono stati effettuati sulle pagine statiche con RDFa e JSON-LD. Queste pagine si trovano nella cartella "Pagine statiche" descritta nel paragrafo 2.4.

#### 4.2.4. Test intero sito dinamico in locale

Siccome il nostro sito non è pubblico, per poter testare l'intero sito in locale (anche con le componenti dinamiche) e renderlo momentaneamente accessibile a chiunque avesse il link, è stato utilizzato Ngrok, un tool che ci permette di esporre il localhost su Internet. Nell'immagine successiva è riportato come esempio uno dei risultati.

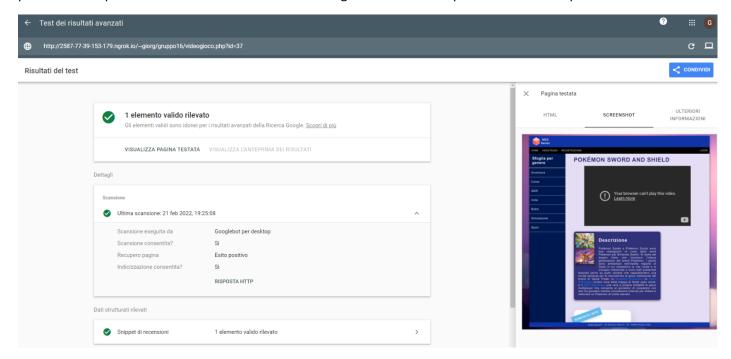


Figura 22: test Ngrok

Abbiamo, inoltre, utilizzato il software <u>Visual SEO Studio</u> che permette di simulare la visualizzazione di una pagina sulle **SERP** (Search Engine Results Pages, ovvero pagine dei risultati del motore di ricerca) **di Google**, partendo dal **tag title** e dalla **meta description** della pagina stessa. Tramite l'emulatore ci si può fare un'idea se i testi contenuti nel tag title e nella meta description risultano idonei, coerenti e rilevanti rispetto ai contenuti della pagina. Lo strumento è presente nelle varie Viste di Visual SEO Studio, bisogna selezionare l'URL della pagina da analizzare e cliccare sul tab *Anteprima SERP*.

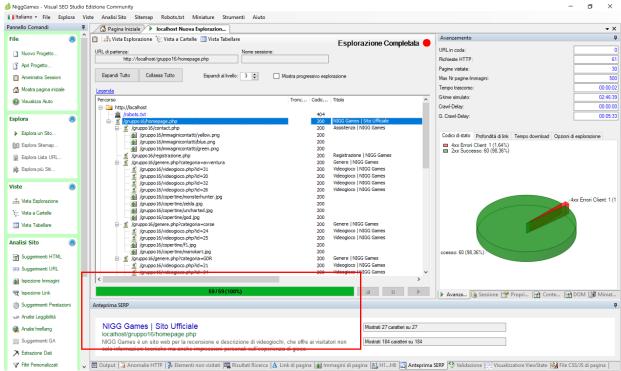


Figura 23: Anteprima SERP

# 5. Pipeline alternativa

È stato pensato un modo alternativo di integrazione dell'ontologia già presente su Schema.org, con l'intenzione di aggiungere nuove proprietà che specializzassero il nostro dominio di interesse.

Schema.org fornisce diversi meccanismi di estensione delle ontologie già esistenti. Alcuni di questi sono andati in disuso e attualmente Schema.org consiglia di condividere con la community estensioni o nuovi schemi che non siano legati a quelli già esistenti. Ovviamente, lo schema proposto deve guadagnare riscontri positivi e qualora si rivelasse utile e funzionale può essere spostato nel vocabolario principale di schema.org, in modo che i motori di ricerca possano fornire più funzionalità basate su una migliore comprensione dei dati strutturati.

Quindi la principale problematica riscontrata è stata non poter rendere disponibile l'ontologia online: pur caricando l'ontologia su GitHub come suggeritoci nella <u>documentazione</u>, sia il validator di Schema.org che il tester di Google generavano errori in quanto non riconoscevano le nuove entità e le nuove proprietà aggiunte. Quindi, pur non creando errori visibili sulle pagine web, abbiamo preferito presentare un prodotto privo di errori.

Vengono qui riportate le aggiunte pensate all'ontologia preesistente:

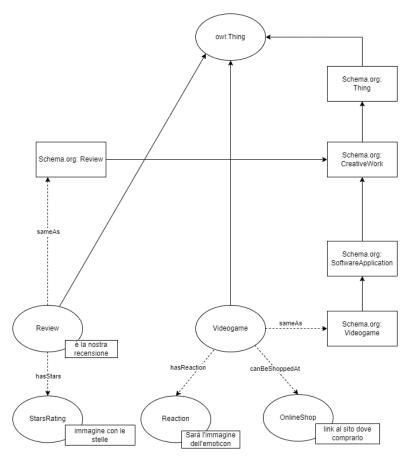


Figura 24: estensione ontologia

#### 6. Demo

Grazie al software Ngrok è possibile creare un indirizzo web che punta al nostro localhost. Sarà possibile quindi avere a disposizione un'url tramite il quale è possibile interagire con il sito web completo di informazioni semantiche direttamente dal proprio computer.

Si tratta di un link usa e getta, in quanto funziona quando il computer host è acceso. Quindi, qualora si volesse provare ad effettuare questa demo live, prima del giorno dell'esame, è necessario contattarci per ottenere il link.