

USE OF SCHEMA.ORG MICRO-MARKUP IN E-COMMERCE PROJECTS

Olexandra Belz¹

Abstract. *The purpose* of the article is to identify the most effective schema.org micro-markup schemes used in e-commerce projects. *Methodology.* The research included competitive intelligence among the leading online platforms operating in Europe in general and in Ukraine in particular. The study involved TOP-8 e-commerce projects in Ukraine and TOP-9 global cross-border marketplaces operating in Europe. The service validator. schema.org was chosen as the *research tool*. *Results.* The study showed that the most popular schema.org micro-markup format is JSON-LD. In general, 82.4% of the surveyed sites use JSON-LD microdata format. Some sites use two microdata formats: JSON-LD and Microdata. But none of the top online marketplaces use the RDFa micro-markup format. Popular marketplaces operating in Ukraine and Europe often use the same types of schema.org vocabulary. However, the frequency of using micro-markup by top marketplaces operating in Ukraine is much higher than the frequency of using micro-markup by top marketplaces operating in Europe. In addition, Ukrainian marketplaces use a much wider list of schema.org micro-markup properties than marketplaces operating in Europe. However, no online store has implemented the properties of advantages and disadvantages of goods recommended by Google in the scheme. *Practical implications.* The study suggests schema.org micro-markup schemes for homepage, category page, product page, about page, payment and delivery page, warranty and returns page, contact page and blog. The proposed templates of micro-markup schemes were validated using the validator.schema.org service. The study recommends using the JSON-LD format for semantic markup of website content. *Value/originality.* Implementation of effective semantic markup of site content will allow search engines to more accurately identify the information presented on the site. This, in turn, will improve the visibility of the online marketplace in the Search Engine Results Page of Google, Bing, Yahoo! etc.

Key words: e-commerce, online marketplaces, schema.org, semantic markup of site content, SEO, micro-markup of site content, formats JSON-LD, Microdata and RDFa.

JEL Classification: O32, F23, C80

1. Introduction

Today the communication environment of the Internet creates new opportunities for business development. Promotion of business in the Internet is a necessary condition for ensuring the competitive position of the company.

The most popular platform for doing business on the Internet is a website. However, it is not enough to have a website to run a business. The website must be visible to potential customers. For this purpose, various measures are implemented to promote the site: SEO-optimization; contextual advertising; Google Shopping; dynamic and classic remarketing; targeted advertising; SMM-marketing; e-mail marketing,

etc. An integral part of successful SEO optimization is the use of web page markup data. This allows search engines to more accurately identify the information presented on the site. In addition, competent markup expands the snippet – a short description of a web resource or its individual page. Rich snippet looks more informative and attractive, increases the relevance of the page and its CTR (click-through rate). The most popular markup to web pages supported by many search engines, including Bing, Google and Yahoo!, is the schema.org. This micro-markup was developed in 2011. It involves the use of so-called "schemas". Schemas are a set of types, each of which is associated with a set of

¹ Ivan Franko National University of Lviv, Ukraine (*corresponding author*)
E-mail: oleksandra.belz@lnu.edu.ua
ORCID: <https://orcid.org/0000-0001-7462-8501>



properties. The types are arranged in a hierarchy. The vocabulary currently consists of 797 Types, 1457 Properties, 14 Datatypes, 86 Enumerations and 462 Enumeration members (Schema.org, 2022). Additionally, there are currently three schema.org micro-markup formats: Microdata, RDFa, and JSON-LD. That is why website developers have to decide which web page markup schemes to use.

The issue of using schema.org micro-markup has received much attention in scientific literature and practical projects. Since 2008, the Common Crawl project has been crawling websites to collect web page data (extracting metadata and web page text). At the time of writing, the latest scan took place from November 26 to December 10, 2022. As a result of this scan, 3.35 billion web pages were processed and 420 petabytes of content were removed (Common Crawl, 2022). Both scientists and practitioners are working with the obtained data sets of the Common Crawl project.

On September 22, 2022, the Web Data Commons (WDC) project released the Schema.org Table Annotation Benchmark (SOTAB) for public download (Web Data Commons, 2022). SOTAB is based on data from schema.org and Common Crawl. SOTAB has two annotation tasks: Column type annotation and column property annotation. The purpose of the Column Type Annotation (CTA) task is to annotate table columns with 91 schema.org types. The purpose of the Columns Property Annotation (CPA) task is to annotate pairs of table columns with one of 176 schema.org properties. That is, SOTAB collects the most used types of web page markup with the most used properties describing them. The benchmark consists of 59,548 tables annotated for CTAs and 48,379 tables annotated for CPAs, originating from 74,215 different websites.

Scientists analyzed SOTAB data. The authors identify the top 19 most common types of schema.org. The study shows that only a small number of schema.org properties are widely used on the Internet (Bizer, Primpeli, Peeters, 2019).

Based on data from the Common Crawl project, researchers analyzed the popularity of data micro-markup formats and identified the most popular types of schema.org vocabulary (Tonon, Felder, Difallah, Cudre-Mauroux, 2016; Muhleisen, Bizer, 2012). Based on data from the Common Crawl

and Web Data Commons projects, researchers analyzed the best practices of using the Product type of the schema.org vocabulary (Selvam, Kejriwal, 2020). The vision of extending schema.org micro-markup to describe the semantic profile of a corporate web resource is presented (Zosimov, Bulgakova, Pozdeev, 2021).

As can be seen from the above review, scientists and practitioners do not focus on the use of effective schema.org micro-markup schemes.

The aim of the work is to identify the most effective schema.org micro-markup schemes used in e-commerce projects.

The research toolkit is the validator.schema.org service.

The search for effective schema.org micro-markup schemes will be conducted by analyzing the most successful e-commerce projects. The following were selected for the study:

1. TOP 8 by attendance of e-commerce projects of Ukraine: rozetka.com.ua; olx.ua; prom.ua; epicentrk.ua; tabletki.ua; allo.ua; comfy.ua and ria.com (Krotovskaia, 2022).

2. TOP 9 global cross-border marketplaces operating in Europe, namely: amazon.com; aliexpress.com; ebay.com; etsy.com; discogs.com; vinted.com; bandcamp.com; ubereats.com; zalando.com (Cross-Border Commerce Europe, 2022).

2. The schema.org micro-markup format

In the studied projects the most popular schema.org micro-markup format is JSON-LD. In general, 82.4% of the studied sites use this microdata format. However, the study found that not all website developers follow the recommendations for using this format. Only sites rozetka.com.ua, olx.ua, comfy.ua, etsy.com and discogs.com have JSON-LD code in the <head> section. Marketplaces prom.ua, epicentrk.ua, tabletki.ua, allo.ua, amazon.com, aliexpress.com, ebay.com, ubereats.com and zalando.com place JSON-LD code in the <body> section or after the <body> section. This practice violates the validity of the program codes of sites. This has a negative impact on search engine optimization in general. Some sites (rozetka.com.ua, epicentrk.ua, allo.ua, comfy.ua, ria.com and ebay.com) use two microdata formats: JSON-LD and Microdata. Marketplaces ria.com, vinted.com and bandcamp.com use only Microdata format. 57.1% of the surveyed sites use Microdata format.

None of the top online marketplaces use the RDFa micro-markup format. Search engines also do not recommend using this format for semantic markup of website content.

3. Popular schema.org microdata types in top e-commerce projects

TOP-8 Ukrainian online marketplaces use twenty-nine types of schema.org micro-markup to form extended snippets. TOP-9 European online marketplaces use twenty-eight types of schema.org micro-markup. However, there are only twenty-one common types, namely: AggregateOffer, AggregateRating, Answer, Brand, BreadcrumbList, FAQPage, ImageObject, ItemList, ListItem, Offer, Organization, Person, Place, PostalAddress, Product, Question, Rating, Review, SearchAction, WebSite and WebPage.

Popular online marketplaces operating in Ukraine and Europe most often use the same

types of schema.org vocabulary (Table 1 and Table 2). However, the frequency of using micro-markup by top marketplaces operating in Ukraine is significantly higher than the frequency of using micro-markup by top marketplaces operating in Europe.

Top online marketplaces operating in Ukraine, on average, use 1.6 types of schema.org micro-markup on the Homepage, 5.9 types on the Category Page, 7.1 types on the Product Page, 2.1 types on the About Page, on the Payment & Delivery Page – 2.3 types, on the Guarantees & Returns Page – 1.4 types, on the Contact Page – 1.6 types, on the Blog – 1.3 types. The top online marketplaces operating in Europe use an average of 1.8 types of schema.org micro-markup on the Homepage, 2.0 types on the Category Page, 2.9 types on the Product Page, and 0.8 types on the About Page, on the Payment & Delivery Page – 0.2 types, on the Guarantees & Returns

Table 1

Schema.org micro-markup types that are most often used by the TOP 8 Ukrainian online marketplaces

Type of micro-markup	Percentage of online marketplaces that use micro-markup on a web page							
	Homepage	Category Page	Product Page	About Page	Payment & Delivery Page	Warranty and Returns Page	Contact Page	Blog
AggregateRating	0.0	0.0	50.0	0.0	0.0	0.0	0.0	0.0
BreadcrumbList	0.0	100.0	87.5	25.0	50.0	37.5	50.0	25.0
ListItem	0.0	100.0	87.5	25.0	50.0	37.5	50.0	25.0
Offer	0.0	50.0	75.0	0.0	0.0	0.0	0.0	0.0
Organization	37.5	50.0	62.5	37.5	12.5	12.5	12.5	12.5
Person	0.0	0.0	37.5	0.0	0.0	0.0	0.0	0.0
Product	0.0	37.5	87.5	0.0	0.0	0.0	0.0	0.0
SearchAction	50.0	37.5	25.0	25.0	12.5	0.0	12.5	12.5
WebSite	50.0	25.0	25.0	25.0	12.5	0.0	12.5	12.5

Source: compiled by the author based on competitive intelligence of online marketplaces operating in Ukraine (service validator.schema.org, 2022)

Table 2

Schema.org micro-markup types most often used by TOP 9 online marketplaces operating in Europe

Type of micro-markup	Percentage of online marketplaces that use micro-markup on a web page							
	Homepage	Category Page	Product Page	About Page	Payment & Delivery Page	Warranty and Returns Page	Contact Page	Blog
AggregateRating	0.0	0.0	33.3	0.0	0.0	0.0	0.0	0.0
Brand	0.0	0.0	44.4	0.0	0.0	0.0	0.0	0.0
BreadcrumbList	0.0	33.3	11.1	0.0	11.1	0.0	0.0	0.0
ListItem	0.0	33.3	11.1	0.0	11.1	0.0	0.0	0.0
Offer	0.0	11.1	44.4	0.0	0.0	0.0	0.0	0.0
Product	0.0	11.1	55.6	0.0	0.0	0.0	0.0	0.0
SearchAction	44.4	0.0	0.0	11.1	0.0	0.0	0.0	0.0
WebSite	55.6	11.1	11.1	33.3	0.0	0.0	11.1	11.1

Source: compiled by the author based on competitive intelligence of online marketplaces operating in Europe (service validator.schema.org, 2022)

Page – 0.0 types, on the Contact Page – 0.1 types, on the Blog – 0.3 types. In addition, Ukrainian marketplaces use a much wider list of schema.org micro-markup properties than marketplaces operating in Europe. However, not a single online store has implemented the properties of advantages and disadvantages of goods recommended by Google in the micro-markup schemes (Google Search Central, 2022).

4. Recommended micro-markup schemes for use in e-commerce projects

Analyzed the experience of using schema.org micro-markup by popular online marketplaces and analyzed the recommendations of micro-markup developers (Schema.org, 2022 & Google Search Central, 2022). Therefore, the author recommends using the JSON-LD format for semantic markup of site content. Also for e-commerce projects, the author has developed recommended micro-markup templates for the main types of pages of the site.

Recommended micro-markup schemes for the Homepage:

1. Organization:

```
<script
type="application/ld+json">{"@context":
"http://schema.org", "@type": "Organization",
"brand": "My-website", "logo": "https://my-website.
com.ua/icons/icon.svg", "name": "MY-WEBSITE",
"alternateName": "MY-WEBSITE", "sameAs":
["https://www.facebook.com/mywebsite.ua",
"https://twitter.com/mywebsite_ua", "https://
www.youtube.com/channel/UCr7r1-z79TY-
fqS2IPeRR47A", "https://www.instagram.com/
mywebsiteua/", "https://t.me/mywebsite"],
"url": "https://my-website.com.ua/", "depart-
ment": {"@context": "http://schema.org",
"@type": "Organization", "image": "https://my-
website.com.ua/assets/icons/icon.svg", "telephone":
["+38(032)000-00-00", "+38(032)111-11-11", "+38
(032) 222-22-22"], "name": "Name-My-Shop", "email":
"info@my-website.com.ua", "address": {"@
type": "PostalAddress", "addressLocality": "My-City",
"streetAddress": "My-Street-Address"}}}</script>
```

2. WebSite:

```
<script
type="application/ld+json">{"@context":
"http://schema.org", "@type": "WebSite",
"url": "http://my-website.ua", "potentialAc-
tion": {"@type": "SearchAction", "target":
"http://my-website.ua/search?search_ter-
```

```
m={search_term}", "query-input": "required
name=search_term"}}</script>
```

3. FAQPage:

```
<script
type="application/ld+json">{"@context":
"http://schema.org", "@type": "-
FAQPage", "mainEntity": [{"@type": "Ques-
tion", "name": "My-Question 1", "acceptedAn-
swer": {"@type": "Answer", "text": "My-Answer
1"}}, {"@type": "Question", "name": "My-Ques-
tion 2", "acceptedAnswer": {"@type": "Answer",
"text": "My-Answer 2"}}]}</script>
```

Recommended micro-markup schemes for the Category Page:

1. Organization, FAQPage – similar to the corresponding schemes of the Homepage.

2. Action:

```
<script
type="application/ld+json">{"@context":
"http://schema.org", "@type": "Action",
"potentialAction": {"@type": "SearchAction",
"target": "https://my-website.com.ua \
/search/?text={search_term_string}", "query-
input": "required name=search_term_string"}
}</script>
```

3. BreadcrumbList:

```
<script type="application/ld+json">{"@con-
text": "http://schema.org", "@type": "Bread-
crumbList", "ItemListElement": [{"@type":
"ListItem", "position": 1, "item": {"@id":
"https://my-website.com.ua", "name": "Name-
My-Shop"}}, {"@type": "ListItem", "position": 2,
"item": {"@id": "https://my-website.com.ua/
", "name": "My-Category"}}]}</script>
```

4. Product (if a line of products of the same type is sold on the site):

```
<script
type="application/ld+json"> {"@context":
"http://schema.org/", "@type": "Product",
"name": "Name-Product", "offers": {"@type":
"AggregateOffer", "priceCurrency": "UAH",
"lowPrice": "4.44", "highPrice": "9.99", "offer-
Count": 3, "availability": "InStock"}} </script>
```

5. WebPage:

```
<script
type="application/ld+json"> {"@context":
"http://schema.org", "@type": "WebPage",
"name": "My-Title", "description": "My-Descrip-
tion", "url": "http://my-website.ua/my-page"}
</script>
```

Recommended micro-markup schemes for the Product Page:

1. Organization, FAQPage – similar to the corresponding schemes of the Homepage.

2. Action, BreadcrumbList, WebPage – similar to the corresponding schemes of the Category Page.

3. Product:

```
<script
type="application/ld+json">{"@context":
"http://schema.org", "@type": "Product", "sku":
"000000000", "url": "https://my-website.
com.ua/product1", "name": "Name-Product 1",
"Image": ["https://my-website.com.ua/
goods/images/product1.jpg", "https://
my-website.com.ua/goods/images/prod-
uct2.jpg"], "description": "Description-Product
1", "itemCondition": "https://schema.org/
NewCondition", "offers": {"@type": "Offer",
"availability": "http://schema.org/InStock",
"url": "https://my-site.com.ua/product1", "-
price": "999.99", "priceCurrency": "UAH", "-
priceValidUntil": "2023-11-16T17:52:22"},
"brand": {"@type": "Brand", "name": "My-
Brand", "url": "https://my-site.com.ua/
producer/brand1/"}, "review": [{"@type":
"Review", "author": {"@type": "Person",
"name": "Name-Author 1"}, "datePublished":
"2022.12.01", "description": "My-Description
1", "positiveNotes": {"@type": "ItemList", "item-
ListElement": [{"@type": "ListItem", "position": 1,
"name": "Positive 1"}, {"@type": "ListItem",
"position": 2, "name": "Positive 2"}]}, "negative
Notes": {"@type": "ItemList", "itemListElement":
[{"@type": "ListItem", "position": 1, "name":
"Negative 1"}, {"@type": "ListItem", "position": 2,
"name": "Negative 2"}]}], {"@type": "Review",
"author": {"@type": "Person", "name":
"Name-Author 2"}, "datePublished": "2022.10.30",
"description": "My-Description 2"}], "aggregat-
eRating": {"@type": "AggregateRating", "rating-
Value": 4.7, "ratingCount": 12, "bestRating": 5,
"worstRating": 1}}</script>
```

Recommended micro-markup schemes for the About Page, the Payment & Delivery Page, the Warranty & Returns Page, the Contact Page:

1. Organization, FAQPage – similar to the corresponding schemes of the Homepage.

2. Action, BreadcrumbList, WebPage – similar to the corresponding schemes of the Category Page.

Recommended micro-markup schemes for the Blog:

1. Organization, FAQPage – similar to the corresponding schemes of the Homepage.

2. Action, BreadcrumbList, WebPage – similar to the corresponding schemes of the Category Page.

3. Blog:

```
<script
type="application/ld+json">{"@context":
"https://schema.org/", "@type": "Blog", "@id":
"https://my-website.com.ua/blog/",
"mainEntityOfPage": "https://my-website.
com.ua/blog", "name": "My-Title", "descrip-
tion": "My-Description", "publisher": {"@
type": "Organization", "@id": "https://my-
website.com.ua", "name": "Name-My-Shop",
"logo": "https://my-website.com.ua/icons/
icon.svg"}}</script>
```

4. BlogPosting:

```
<script
type="application/ld+json">{"@context":
"https://schema.org/", "@type": "BlogPost-
ing", "mainEntityOfPage": "https://my-website.
com.ua/blog/my-post", "name": "My-Post-Ti-
tle", "description": "My-Post-Description",
"datePublished": "2022-01-14", "dateModi-
fied": "2022-12-14", "author": {"@type": "Per-
son", "@id": "https://my-website.com.ua/
author/my-name", "name": "My-Name"},
"image": {"@type": "ImageObject", "@id":
"https://my-website.com.ua/image/image.jpg",
"url": "https://my-website.com.ua/image/
image.jpg"}, "url": "https://my-website.com.ua/
blog/my-post"}</script>
```

The proposed micro-markup schema templates were validated using the validator.schema.org service. The validation result is as follows: zero errors and zero warnings.

5. Conclusions

In this work, a competitive intelligence of the top online stores operating in the Ukrainian and European markets was conducted in order to use schema.org micro-markup. In the process of work, the shortcomings of using the schema.org micro-markup scheme by top online stores were revealed. It was also found that Ukrainian online marketplaces use more informative schema markup schemes.

Based on the results of the work, templates to schema.org micro-markup schemes are proposed for the Homepage, the Category Page, the Product Page, the About Page, the Payment & Delivery Page, the Warranty & Returns Page, the Contact Page and Blog.

Since the schema.org dictionary is dynamically updated, SEO specialists need to follow the recommendations of search engines

regarding the use of micro-markup and, if necessary, make changes to the program code of the site pages.

References:

- Bizer, C., Primpeli, A., & Peeters, R. (2019, April 12). *Using the Semantic Web as a Source of Training Data*. Available at: http://www.webdatacommons.org/largescaleproductcorpus/papers/Bizer2019_Article_UsingTheSemanticWebAsASourceOf.pdf
- Common Crawl (2022, November/December). *November/December 2022 crawl archive now available*. Available at: <https://commoncrawl.org/2022/12/nov-dec-2022-crawl-archive-now-available>
- Cross-Border Commerce Europe (2022). *TOP 100 Marketplaces in Europe: Annual Ranking 2022*. Available at: <https://www.cbcommerce.eu/press-releases/top-100-marketplaces-in-europe-annual-ranking-2022-out-now>
- Google Search Central (2022). *Product (Product, Review, Offer) structured data*. Available at: <https://developers.google.com/search/docs/appearance/structured-data/product>
- Google Search Central (2022, August 5). *New in structured data: Pros and cons*. Available at: <https://developers.google.com/search/blog/2022/08/pros-and-cons-structured-data>
- Krotovskaia, O. (2022, August 15). *Ukrainci pochaly chastishe tsikavytysia tsinamy na haz i dyvytysia filmy: top-25 saitiv za lypen 2022 roku [Ukrainians became more interested in gas prices and watching movies: top 25 sites for July 2022]*. Available at: <https://thepage.ua/ua/news/top-25-najpopulyarnishih-sajtiv-v-ukrayini-za-lipen-2022-roku> (in Ukrainian)
- Muhleisen, H., & Bizer, C. (2012). *Web Data Commons – Extracting Structured Data from Two Large Web Corpora. WWW2012 workshop: Linked Data on the Web (LDOW2012)*. Lyon, France, 16 April, 2012. Available at: <https://ceur-ws.org/Vol-937/ldow2012-inv-paper-2.pdf>
- Schema.org (2022). *Full Hierarchy*. Available at: <https://schema.org/docs/full.htm>
- Schema.org (2022). *Organization of Schemas*. Available at: <https://schema.org/docs/schemas.html>
- Selvam, R. K., & Kejriwal, M. (2020, July 27). *On using Product-Specific Schema.org from Web Data Commons: An Empirical Set of Best Practices. KDD '20, August, 2020, San Diego, California*. Available at: <https://arxiv.org/pdf/2007.13829.pdf>
- Tonon, A., Felder, V., Difallah, D. E., & Cudre-Mauroux, P. (2016). *VoldemortKG: Mapping schema.org and Web Entities to Linked Open Data. In ISWC. Springer International Publishing, 2016*. Available at: <https://exascale.info/assets/pdf/voldemort.pdf>
- Web Data Commons (2022, September 22). *Web Data Commons – Schema.org Table Annotation Benchmark*. Available at: <http://webdatacommons.org/structureddata/sotab>
- Zosimov, V., Bulgakova, O., & Pozdeev, V. (2021). *Semantic Profile of Corporate Web Resources. Selected Papers of the VIII International Scientific Conference 'Information Technology and Implementation' (IT&I-2021). Workshop Proceedings. Kyiv, Ukraine, December 1-3, 2021, 389–397*. Available at: https://ceur-ws.org/Vol-3179/Short_13.pdf

Received on: 2th of November, 2022

Accepted on: 10th of December, 2022

Published on: 30th of December, 2022