Schema.org in Google search results

Executive Summary

The inclusion of meta-information from schema.org on websites is a major consideration of search engines in order to better understand content. Thanks to schema.org integrations, Google can instantly tell whether the content of a website concerns a movie, a person, a TV series, etc. In recent months, the search engine industry has begun to focus on this issue, with Google depending on semantic references in web pages more than ever since the Hummingbird update.

In a study of tens of thousands of keywords and more than half a million domains, Searchmetrics analyzed how often schema.org integrations are found on websites and in search results, how domains or URLs using markups are positioned in Google SERPs, and whether a schema.org integration may even have a positive effect on a domain's rankings.

Findings overview:

- Only 0.3% of the studied domains were found to include schema.org integrations.
- For only 34.4% of keywords examined, Google returned search results with neither schema.org integrations nor any other structured data involved.
- The share of websites with schema.org integrations is highest in Germany (compared to US, UK, FR and ES).
- The most common integrations are "Movies", "Offers" and "TV series".
- Review and ratings integrations from schema.org are over 60% positive.
- Pages with schema.org integrations rank better by an average of four positions compared to pages without schema.org integrations.

It is notable that usage of schema.org integrations on sites is still uncommon – fewer than one per cent of sites include them – although Google already delivers schema-derived markups in nearly 37% of search results.

search**metrics**

Moreover, sites with at least one integration seem to increase their probability of ranking higher in search results. However, this phenomenon may be due to correlation rather than causation. Hence the improved average position of pages with schema.org integrations may be related to factors above and beyond any schema.org ranking influence.

Schema.org: "Things not Strings"

When Google, Microsoft, Yahoo! and, more recently, Yandex agree on a common standard, online marketers should take note - especially if this has the potential to influence search results.

Schema.org is such a project: With the creation of a common markup scheme, the search giants want to ensure that they better understand the information contained on web pages and better represent them in search results. To this end, they have created a new vocabulary that can be found on the project's website showing how to mark certain terms or items as individual 'entities' by assigning them with a specific HTML code so that Google & Co can recognize them as such.

An entity describes a "thing" (i.e., a building, person or event) that is recognizable as such and will have certain properties.

The definition of entities via the classification from schema.org is achieved by using the "Micro Data Format" which serves to assign specific HTML tags to respective entities according to their properties. For a full technical explanation of this, you can visit the website. 'Entities' can be classified in many ways, for example:

- Events
- Books
- Recipes
- Organizations
- Persons
- Locations
- Products
- Reviews
- Movies
- Music
- Etc.

Each entity can be described by a set of specific properties.
For a recipe, for example, the following properties can be distinguished by schema.org (among many others):

- Name (text)
- Description (text)
- Image (URL source)
- URL (URL of the page)
- Ingredients (text)
- Required cooking time (duration)
- Nutrition[3]

In short, there are characteristics for each type, and these can be highlighted with HTML code. The schema.org approach was accepted by the search engines mentioned above in mid-2011 and has since become an important standard.

Since the accepted theory is that websites are analyzed according to their content, and that schema.org helps search engines better understand this content, those sites with integrations are likely to expect a better ranking in search results. This theory is what Searchmetrics has investigated.

Data and Survey Background

In order to study this, the following data from across the Web has been collected:

1. In search results for tens of thousands of keywords, we looked for entities that have been created by schema.org and/or other structured data. Here we analyzed each of the first 50 search results.

searchmetrics

2. We analyzed hundreds of thousands of domains to determine whether schema.org data can be found.

The survey was conducted in March 2014 for Google USA.

Distribution of Schema.org data

Despite the introduction of schema.org almost three years ago, only 0.3% of domains include schema integrations.

Domains without Schema
integration (99.70%)
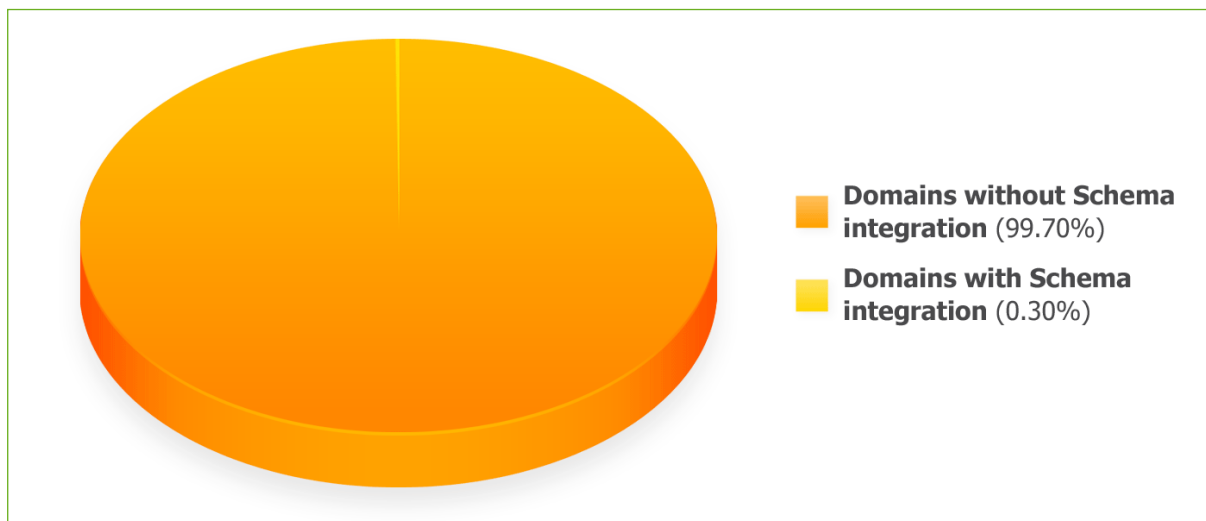
Domains with Schema
integration (0.30%)

Fig.1 – Share of domains with/ without schema.org markup

This seems very low, given the high expectations that many online marketers have for these entities. However, it should be said that the issue of 'entities' (rather than 'keywords') appeared on the radar of the search engine industry only last year, and since incorporation of new technological developments on large websites takes time, we may expect a rapid increase in this ratio from now.

Schema.org markups in Google SERPs for almost 40% of keywords investigated

The potential of Schema becomes clear when realizing that Google enhances search results with schema.org markups in more than 36% of keyword queries. Although this value has slightly decreased compared to the previous analysis, the overall proportion of search results containing snippets with additional information derived from structured data – such as results from pages like Wikipedia.org, that actually do not use schema – has actually increased.
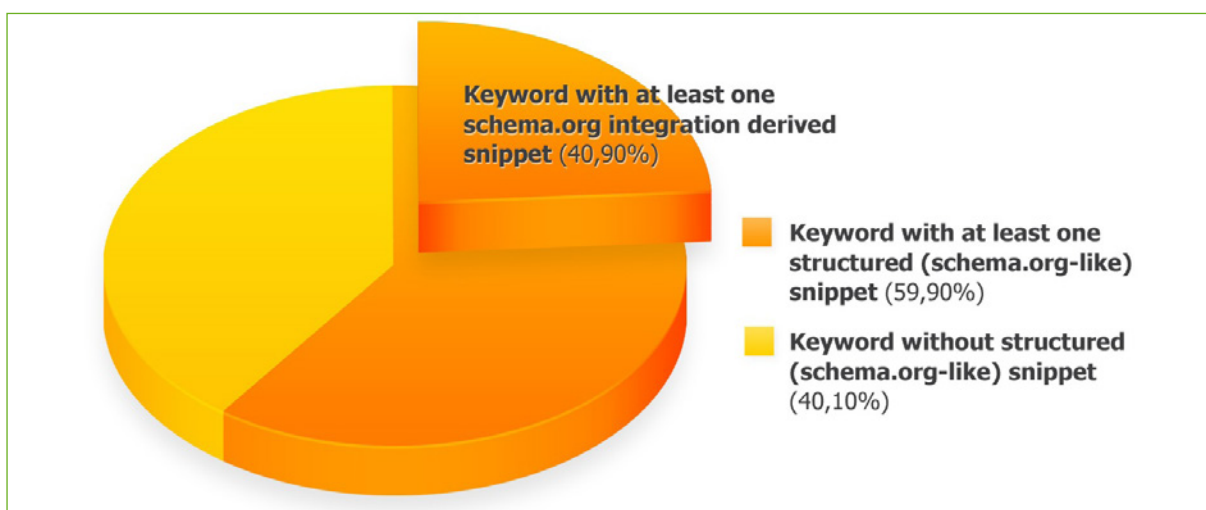

Keyword with at least one schema.org integration derived snippet (40,90%)

Keyword with at least one structured (schema.org-like) snippet (59,90%)

Keyword without structured (schema.org-like) snippet (40,10%)

Fig. 2 – Share of Google search results with/without structured data markups (incl./excl. schema.org)

searchmetrics

In contrast, the proportion of keywords for which no markups appear in snippets has declined, and is currently at about 34%.

A slight upward trend is already apparent – last year's value was 0.27%.

Domains using integrations from schema.org have a higher SEO Visibility on average

In addition, we have examined the average SEO Visibility scores for domains with and without Schema integrations.

Fig. 3 – Average SEO Visibility of domains with/ without schema.org markup

Our results show that domains with Schema integrations have significantly higher average SEO Visibility scores than those domains that do not include them.

Domains using integrations from schema.org have a higher SEO Visibility on average

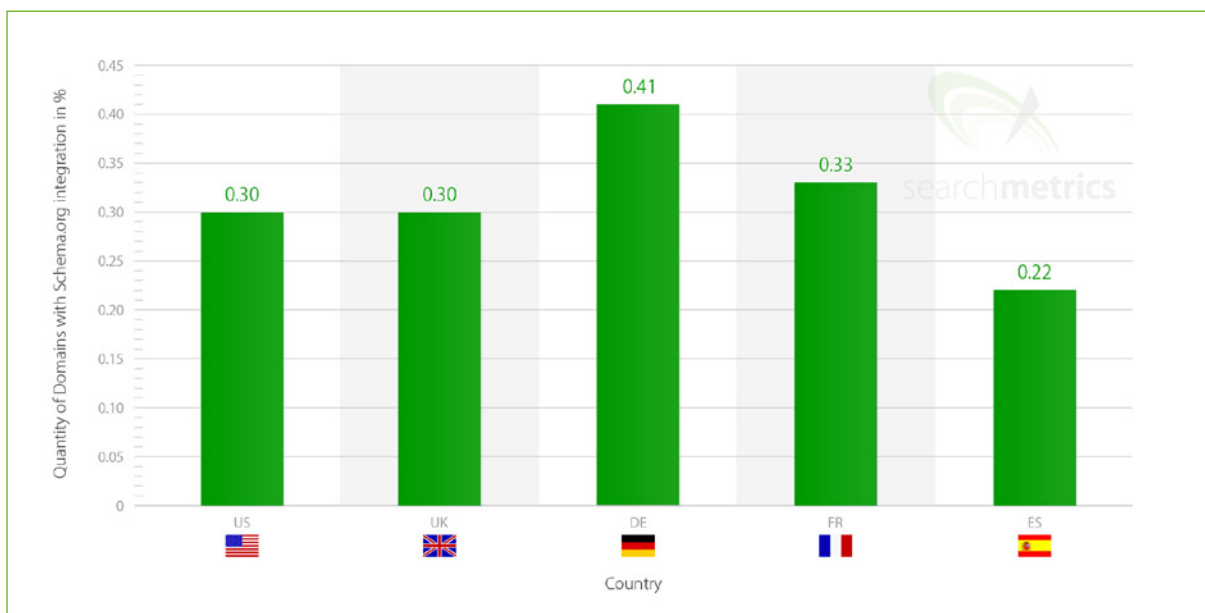The geographic distribution of integrations across the countries investigated is also interesting:

Fig. 4 – Share of domains with schema.org markups: comparison by country

Here, Germany appears to be a country of engineers, and leads the list with 0.41% of domains with schema.org integrations. The U.S. is level with the UK – just ahead of Spain.

Whether this can be interpreted as a German technological lead is doubtful. In any case, there seems to be evidence that the willingness to boost rankings with schema.org data in the U.S. is weaker than in Europe.

searchmetrics

Most common Schema.org data

Which types of information are actually used by website operators from schema.org that can be sent to search engines?

If you look at the results of the keywords analyzed first, the following schema.org integration-types appear most frequently in Google SERPs:
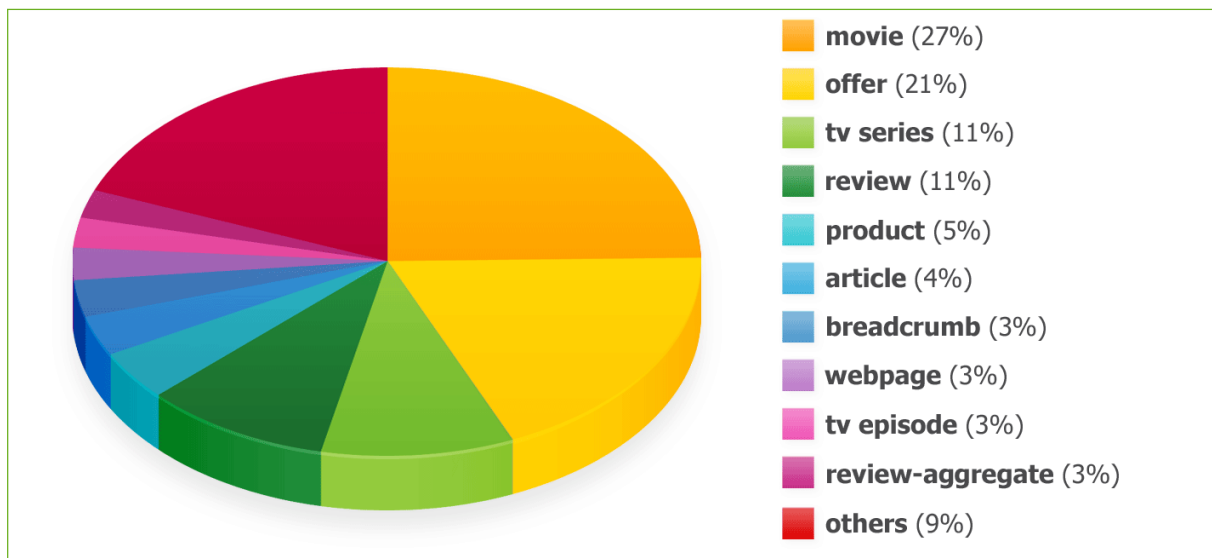

- movie (27%)
- offer (21%)
- tv series (11%)
- review (11%)
- product (5%)
- article (4%)
- breadcrumb (3%)
- webpage (3%)
- tv episode (3%)
- review-aggregate (3%)
- others (9%)

Fig. 5 – Most popular schema.org types in Google SERPs

The film industry is clearly at the top, followed by 'Offers' – i.e. e-commerce sites. It is of note that "Reviews" – the evaluation of products or articles on websites – are not at the top, even though they could be included on any web page, whereas film integrations are only suited to those sites that relate to movies.

There is a case to be made for "Breadcrumbs", i.e. the inclusion of breadcrumb navigation on Web pages, to be more popular, as this is possible on any site as well.

Quantity of structured data: domains vs Google snippets

A further graph shows that webmasters who have used some kind of schema.org integration once, often use them more and more:
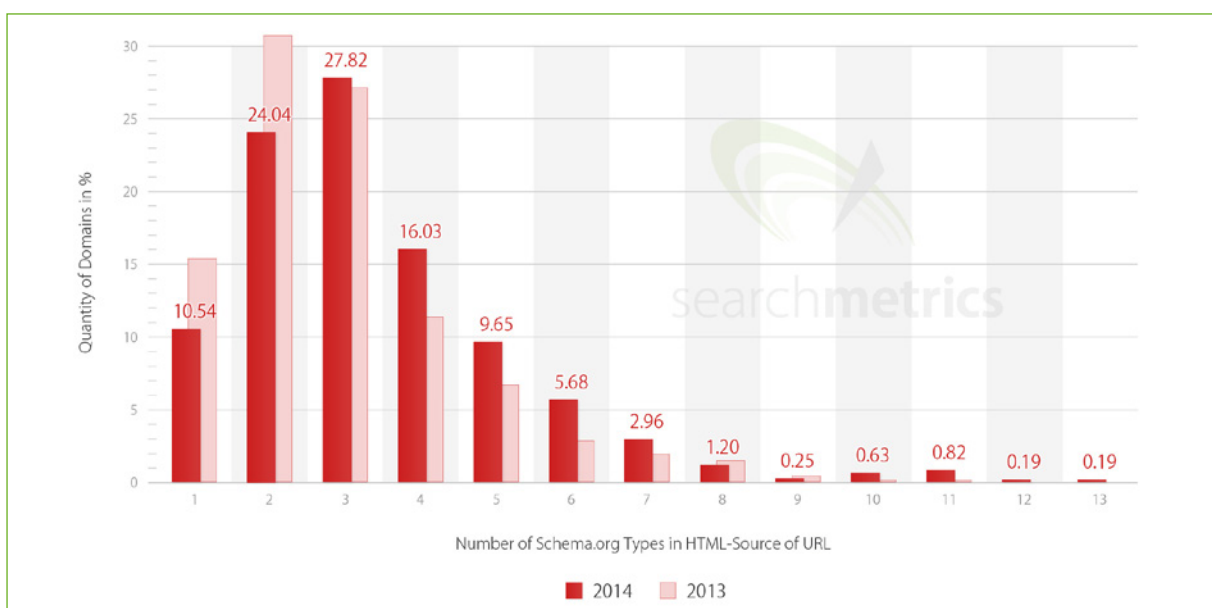

Fig. 6 – Share of domains with X schema.org markups in source code

searchmetrics

The graph shows the average number of different schema.org types used by webmasters for the respective URL in 2014 compared to 2013. Last year, two different schema.org types per URL was the most common application – for example, breadcrumbs and a movie, or a product and a review. The second most common case was three integrations (e.g. breadcrumbs, review and film). Fewer than 16 percent of sites had only one schema.org snippet.

An annual comparison, however, shows that there is a tendency for webmasters to include more schema types on their websites. The average integration of three, four, five, six and even seven different schema markup types per URL has increased, while fewer sites had only one or two distinct schema markups in their HTML than in the previous year.

This has an effect on search results, too, as there are commonly two, but increasingly more, integrations displayed in the snippets for keyword search results:
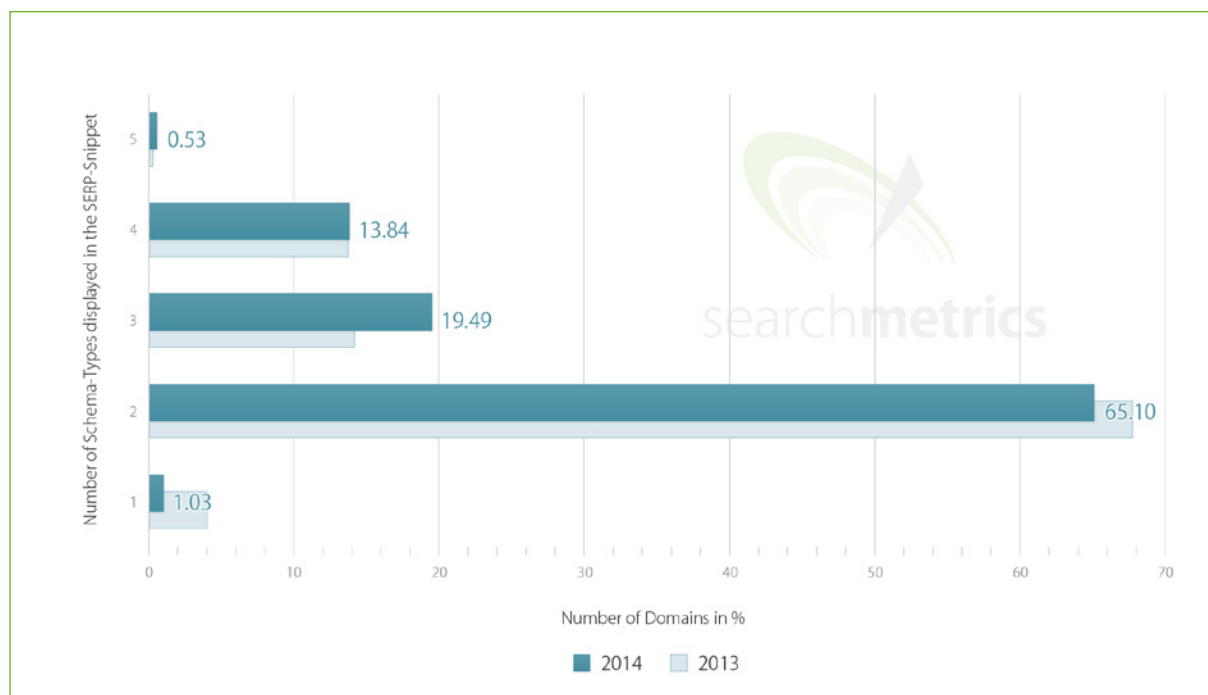

Fig. 7 – Number of markups in Google snippets to share of domains ratio

This has not changed significantly year on year, however, the tendency on the part of Google to show more rather than fewer markups per snippet can be seen.

Ratings and Reviews: Purely positive?

Ratings and Reviews occupy a special position within schema.org integrations, because these relate to not just a product or a person – but show the opinion of site visitors about the product or web page evaluated.

Therefore, we have also analyzed the 'Star' ratings (i.e. the expression of assessment) - with the following results:
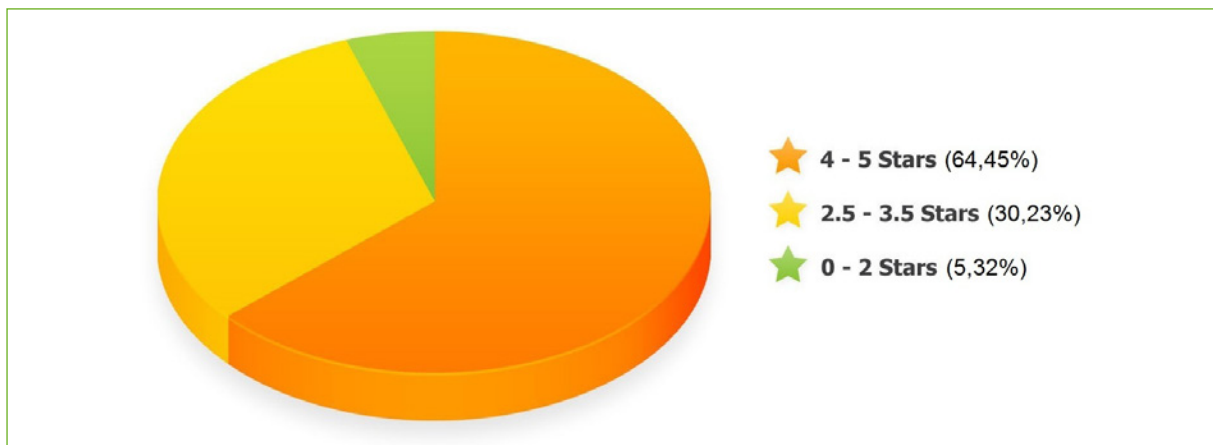
Fig. 8 – Share of positive, neutral and negative "Ratings"

It's clear to see that providing the ability to rate content is rewarded positively. Only 6% of the reviews were bad (0-2 stars) but more than 60% were positive (4-5 stars).

Higher Google rankings for pages using schema.org markups

Finally, we wanted to test the hypothesis that the integration of schema.org data in Google might be rewarded with a better ranking. So we have displayed the positions of pages with schema.org integrations compared to positions of those pages without schema.org data in Google SERPs.

We analyzed the top 50 search results, taking into account all sites both with and without Schema integrations. Given the small number of URLs using schema markups (0.3%), the average ranking of domains without schema markups (99.7%) is 25.

For 2013, sites with at least one integration ranked on average three positions higher than those without. A 2014 comparison underlines this pattern and shows sites with schema integrations ranking on average four positions better than those without, at 21.

It must be noted be noted that this is not necessarily a causal relationship. It may not be the case that pages are actually preferred by Google just because they provide schema integrations, and maybe the higher rankings can be explained by the fact that webmasters who use schema.org integrations are one step ahead of the competition due to other factors that affect their rankings in a positive way.

Conclusion: markups based on structured data are used very rarely by webmasters, but are common in Google SERPs

Schema.org integrations are still used by very few webmasters for structuring and categorization of data on Web pages. Based on our data set, the value was less than one percent.

Google SERPs, however, often appear with snippets using structured data. In our current analysis, the proportion of search results with snippets enhanced by schema integrations amounted to nearly 37 percent – the share of search results with structured data in general (including schema.org) amounted to almost 60 percent.

Whether the use of HTML markups does genuinely affect the ranking of a domain, as it would appear from the data, is hard to say. In fact, it may be due in part to interaction with other factors.

However, it can be said that the inclusion of HTML tags from schema.org to clearly flag up the content of the page will greatly help search engines better understand the Web's content – and that is an essential consideration for a good ranking.

searchmetrics
Make it visible.