
GOOGLE ANSWER BOX KEYWORD-RELATED ANALYSIS A CASE STUDY

**Andrej Miklošik^{*}, Ivan Hlavatý, Ferdinand Daňo and
Peter Červenka**

*University of Economics in Bratislava, Faculty of Commerce, Dolnozemska cesta 1,
852 35 Bratislava, Slovak Republic*

(Received 31 May 2016, revised 16 August 2016)

Abstract

The opportunity to be included in the Google Answer Box is important for website owners to build brand awareness and gain extra conversions. There is an on-going debate amongst SEO experts regarding the factors that are important for a website content to be extracted to a Google Answer Box. In this article, the authors examine a Google Answer Box form and contents and the situation regarding the structure of the content on the target website. The aim is to determine important factors that affect positioning of the website in the Google Answer Box from the point of view of keyword structure and density. The case study presents results of this analysis and important findings that include the inclusion of keywords in the website's content in different locations visible and invisible to the user. Furthermore, the form of presenting content in the Google Answer Box is analysed and its relationship to structuring the content on the website is examined.

Keywords: keyword, analysis, organic search results, search engine, results page

1. Introduction

Website owners/operators are usually considering Google as a very important source of traffic. Because of this fact, search engine optimization (SEO) experts study the factors affecting the positioning of a website in the search results. As it was proven by Tomasi and Li, increasing the rankings on Search Engines Results Pages results not only into a higher number of visitors on the site, but also into the increased average time duration of users visiting the site, more user engagement and an increase of annual sales revenue [1]. Google is offering several different types of search results to its users. The main categorization includes organic results that originate as a result of application of Google's proprietary algorithm and sponsored results, which are displayed on a pay per click paid basis. Within the organic results, there are several types of

^{*}E-mail: miklosik@euba.sk

listing as well. One of the recent innovations of Google is the Google Answer Box (GAB) that tries to present a complete information to the user, so that he does not have to use a third-party website to complete his search for information. It has been introduced several years ago. However, it is still being developed and improved intensely.

There are not many studies showing the benefits of a GAB for a website. Some argue that it is even negative, because it can lower the traffic directing to the website. However, it is also expected that positioning in the GAB significantly helps building brand awareness. Then, as researched by Baye, De los Santos and Wildenbeest [2], brand awareness further increases organic clicks through both a direct and an indirect effect. In this article, the authors deal with the topic of inclusion of a website into a GAB, specifically a GAB of the type Website Extraction. The aim of this research study is to determine important factors that affect positioning of the website in the Google Answer Box from the point of view of keyword structure and density.

2. Theoretical background

Emerging technologies represent a great challenge for company owners and managers. Digital technologies are evolving rapidly and it is apparent that even keeping touch with the changes can be quite demanding. However, if companies want to benefit from them, they have to look for a way of implementing them efficiently to distinguish themselves from their competitors [3]. Google Answer Box is one example of such a technology. It is a form of content representation that is driven by the Knowledge Graph. It has evolved over the past few years aiming at providing the user the most relevant and complete information. For recipes, for example, a rich content organic search result can include steps to prepare the food, time to prepare or rating by other users. There are not many studies that have analysed the Google Answer Box selection algorithm. Most of them come from SEO experts and are published on free accessible websites including company website or blogs. No studies containing relevant resources for this topic in highly ranked journals have been identified by the authors.

Generally, the following prerequisites need to be met if a site wants to be excerpted in the GAB:

- rank as highest as possible for the respective keyword;
- include as many of the keywords from the search query in the website content as possible;
- locate these keywords in different locations, including headings, page title, URL, paragraph text, image titles, inbound or outbound links, discussion, etc.;
- structure the content to a form of list, numbered or unnumbered.

The first point is important because the higher the position of the website in organic results within the SERP, the higher chance to be included in the GAB. This relationship is shown in Figure 1 [R. Gavrilas,

<http://cognitiveseo.com/blog/6266/decoding-google-answer-box-algorithm-serp-research-10-353-keywords/>, accessed 15.12.2015]. It displays the number websites included in a GAB according to their position in organic search results. The study was based on 1 236 search results with Google Answer Boxes.

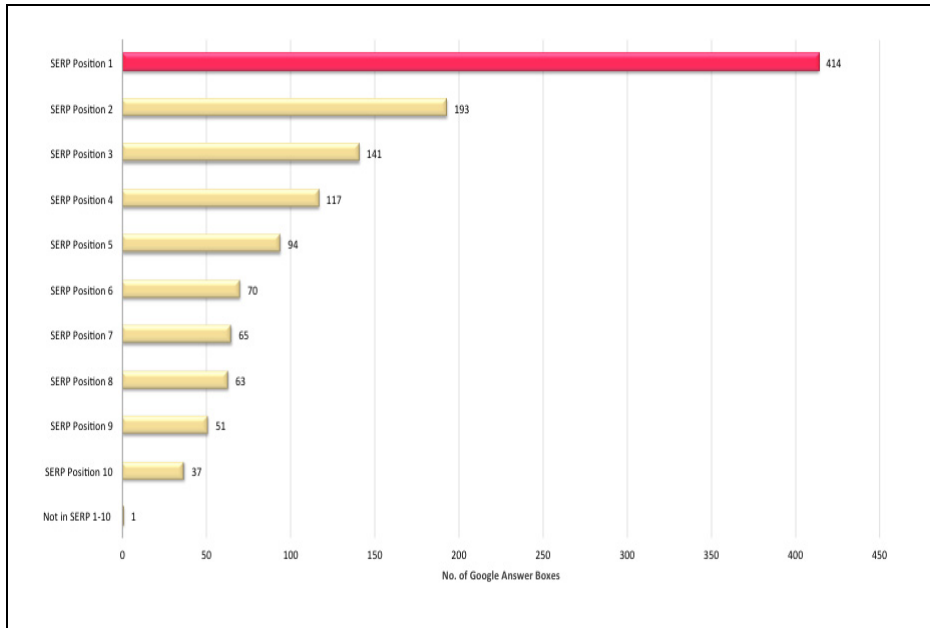


Figure 1. TOP 10 ranking position distribution within a Google Answer Box.

The available resources concentrate on content and keyword issues within the website. Gabbert indicates that the attention should be paid to these three areas [E. Gabbert, <http://www.wordstream.com/blog/ws/2015/10/13/google-answer-box>, accessed 15.01.2016]:

- content strategy,
- keyword research and selection,
- keyword optimization within your content.

To increase the chance of being included into a GAB, Kronrod [K. Kronrod, <http://www.brightedge.com/blog/google-quick-answers/>, accessed 26.12.2015] recommends the following strategy for website owners/operators:

- select a topic that is interesting for users of the website;
- create quality content relevant to the theme (use ‘buyer persona’ studies, focus groups, and engagement analytics to determine the content mix),
- structure the page with user experience in mind;
- where relevant, clearly define a how-to list, including the phrase ‘how to’ in content and bullet points clearly present answers;
- implement SEO practices like basic on-page optimization, theme-relevant cross-linking and technical SEO tactics.

Within the GAB, there are several components present. They include the title, content, image, target URL and may include a call to action button as well. It can contain a link to the website, e.g. 'see more', or a complete document e.g. 'Download'. There are several factors that can influence the decision of Google regarding the call to action box inclusion [J. Yu, <http://searchengineland.com/optimizing-google-quick-answers-box-215037>, accessed 22.12.2015]:

- Page authority plays a significant role in being featured as a call-to-action link; top landing pages usually perform best, which is closely related to authority.
- The quantity and quality of links is a tipping point for deeper pages being featured.
- The presence of a 'download' link on the page is not an influencer in some cases.

In the research results part of this article, the authors will examine the content-related issues of inclusion into a Google Answer Box in a form of a case study.

3. Research methodology

To achieve the goals set, the authors have applied both secondary and primary research methods. They reflect the current trends in marketing research, which has to adapt to new research challenges in the digital era [4]. Within the secondary research, publications of SEO experts regarding the Google Answer Box have been studied including information regarding the situation in displaying different types of GAB. These results of previous research studies have been examined, analysed, their results compared, excerpted and selected for application in the presented research. In the primary research, a method of case study has been selected. This method of qualitative research allows to examine the object of research in more detail. This enabled the authors to find interesting knowledge regarding the GAB, its functioning and the situation regarding the keyword usage on the target website. To present the results of the case study, the authors have proceeded through these five main steps:

1. Selecting the optimal search query type: The 'How to' query was selected for closer analysis, as this query has produced the most results including a GAB in the previous research [5].
2. Selecting one of the results producing a GAB containing a page extraction: The first search query 'How to bake a Christmas cake' was selected, as it contained a GAB of the most frequent type – Website Extraction.
3. Analysis of the search engine results page (further as 'SERP') – the overall structure of the SERP has been analysed including the GAB and 10 other organic search results.
4. Deep analysis of the website including visible and invisible content and its structure.
5. Formulation of results and conclusions.

Table 1. Six different search queries included in the research.

Case No.	'How to' Search Query	Website in the GAB
1.	bake a Christmas cake	www.taste.com.au/recipes/30325/easy+christmas+cake
2.	cook an egg	http://goop.com/how-to-cook-an-egg/
3.	unclog the toilet	www.familyhandyman.com/plumbing/toilet...to...clogged-toilet/view-all
4.	dance quickstep	https://www.youtube.com/watch?v=c-4x2m_aE7o
5.	do yoga	www.ekhartyoga.com/.../8-tips-on-how-to-do-yoga-at-home-practice-and-a...
6.	type with ten fingers	www.hongkiat.com/blog/faster-keyboard-typing/

4. Research results

In an article published in the scientific journal *Communication Today* [5], the authors presented the first part of their extensive research regarding the GAB. It included the analysis of situation regarding the inclusion of an excerpt of website's content into the GAB when searching for different types of queries. For the search query type 'How to', the GAB was displayed most frequently.

When searching for queries of the 'How to', these 6 search queries returned results that included one of the GAB forms. They are listed in the Table 1.

In this article, the authors have studied the situation regarding the first 'How to' search query: 'How to bake a Christmas cake'. A case study containing a detailed analysis of the first search results that included a GAB is presented. Authors analyse the contents and look of the GAB closely and the situation on the target website from the point of view of content structure, keywords, their placement and density. Figure 2 displays the SERP for the search query 'How to bake a Christmas cake'.

The first search result is a GAB type Website Extraction. It has excerpted a paragraph of text from the website [taste.com.au](http://www.taste.com.au). The GAB contains text, one image on the right, page title and URL. What is interesting in this case is the fact that there is a second result for the search query containing the same information (recipe) presented in a different form. It contains the same elements plus extra information regarding the rating of this recipe and time for preparation. This rich result (containing extra information compared to a regular listing) would be at the first position, if there would be no GAB presented. Thus, in this case the first result had the best chance of being included into the GAB. Within the SERP, there are 8 organic results linking to a content in form of texts and images and 2 organic results directing to a video on YouTube.

Google how to bake a christmas cake

All Videos Images News Shopping More Search tools

About 5,840,000 results (0.45 seconds)

Spoon into prepared pan and smooth the surface. Lightly tap pan on bench top to settle the mixture. Arrange the almonds over the top of the **cake**. **Bake** in preheated oven for 3 hours 15 minutes or until a skewer inserted into centre comes out clean, covering **cake** loosely with foil if over browning.

Easy Christmas Cake Recipe - Taste.com.au
www.taste.com.au/recipes/30325/easy-christmas-cake

Feedback

Easy Christmas Cake Recipe - Taste.com.au
www.taste.com.au/recipes/30325/easy-christmas-cake
★★★★★ Rating: 4.6 - 5 reviews - 3 hrs 55 min.
Spoon into prepared pan and smooth the surface. Lightly tap pan on bench top to settle the mixture. Arrange the almonds over the top of the **cake**. **Bake** in preheated oven for 3 hours 15 minutes or until a skewer inserted into centre comes out clean, covering cake loosely with foil if over browning.

Make & mature Christmas cake | BBC Good Food
www.bbcgoodfood.com/recipes/.../make-and-mature-christ...
2 hrs 35 min. - 678 cal
Remove the **cake** from the oven, poke holes in it with a skewer and spoon over 2 tbsp of your chosen alcohol. Leave the cake to cool completely in the tin. To store, peel off the baking parchment, then wrap well in cling film. Feed the cake with 1-2 tbsp alcohol every fortnight, until you ice it.

How to Make A Christmas Cake (Part 1) - YouTube
https://www.youtube.com/watch?v=A8D_9ZOBpQ
Nov 22, 2012 - Uploaded by Tillis Busy Kitchen
Want to make a traditional Christmas cake? Then look no further than Till's Busy Kitchen. Recipe at ...

How to Make Christmas Cake | Tesco Food - YouTube
https://www.youtube.com/watch?v=Gw_JWcbaZMw
Jul 30, 2012 - Uploaded by Tescofoodandwine
Try this simple and delicious Christmas cake recipe. Jo takes us step-by-step through the rich ingredients, right ...

How to Make a Christmas Cake (with Pictures) - wikiHow
www.wikihow.com > ... > Holiday Cooking > Christmas Cooking
Are you looking for a cake with a unique Christmas appeal? Try this recipe to make a traditional English Christmas cake that includes brandied fruit and almonds ...

How to make a Christmas cake - from Delia Online
www.deliaonline.com/how-to-cook/baking/how-to-make-a-christmas-ca...
You need to begin this cake the night before you want to bake it. These quantities will make an 8 inch (20 cm) round cake. Weigh out 1lb (450 g) of currants, 6 oz ...

Mary Berry's Christmas cake recipe - goodtoknow
www.goodtoknow.co.uk > Recipes
★★★★★ Rating: 4 - 1,273 votes - 4 hrs 30 min.
Baking queen Mary Berry shows us how to make her delicious classic Christmas cake recipe - we think it's better than Delia's!

Christmas Cake - All recipes Australia NZ
allrecipes.com.au > ... > Christmas > Christmas Baking
Look here for lots of great recipes for traditional Christmas cakes, easy fruit cakes, chocolate Christmas cakes, European Christmas cakes and the popular Yule ...

How to cook perfect Christmas cake - The Guardian
www.theguardian.com > Lifestyle > Cake
Nov 16, 2011 - As the very first real cake I ever baked, Christmas cake occupies a special place in my heart. Admittedly, the result, cruelly dubbed "the rock of ...

BBC Food - Recipes - Mary Berry's classic Christmas cake
www.bbc.co.uk/food/recipes/classic_christmas_cake_04076
2 hrs
Mary Berry's been making her Christmas cake recipe for as long as Paul Hollywood's been alive. She knows what she's talking about.

Figure 2. Search Engine Results Page for the Query ‘how to bake a Christmas cake’ [https://www.google.com.au/webhp?sourceid=chrome-instant&ion=1&espv=2&ie=UTF-8#q=how+to+bake+a+christmas+cake].

In the first part of this case study, the contents in the GAB will be examined. The text excerpted contains four words that are highlighted. Two of them are part of the search query; one is duplicated and the last fourth one is the word ‘oven’, which is not part of the search query. Thus, the text contains two of four keywords. The text looks complete; however, after the comparison with the content on the website it can be seen, that Google has selected only one small part of the original recipe. The recipe on the website contains 913 characters (without spaces), the paragraph in the GAB contains only 244 characters. The page title is not identical with the search query; however, it contains two of the four words from the search query (Christmas cake). If the URL is closer examined, it also contains these two words. Moreover, Google highlights the word ‘recipe’. Thus it is apparent, that the Google algorithm sees the direct relationship between the word ‘recipe’ and the query of the type ‘how to bake’. Similarly, there are probably other associations that Google is able to recognize automatically, such as ‘how to cook’.

Table 2. Detailed analysis of keyword occurrence.

Keyword	Page Title	URL	Headings	Lists	Paragraph	Related Content	Discussion	HTML
How to	0	0	0	0	0	0	0	9
bake	0	0	0	1	0	0	0	5
Christmas	1	1	1	0	1	4	1	32
cake	1	1	1	7	1	10	4	71

Legend:

- Page Title Page Title visible in the Internet browser tab
- URL All parts of the URL including domains of all levels and alias
- Headings All headings marked as <H> in the code (H1, H2, H3, etc.)
- Lists Numbered or unnumbered lists represented by or tags in the HTML code
- Paragraph Regular/common text represented by <p> tag in the HTML code
- Related Content Other content containing the examined keywords (e.g. recommended article, links, etc.)
- Discussion Posts in the discussion visible in the first page (in case there is pagination present)
- HTML Includes all mentions of the keyword in the HTML code of the website. This includes both visible and invisible parts of the website including meta information, image titles, image alt tags, etc.

Secondly, the website contents needs to be examined to determine the structure of the content, frequency of keywords and other related factors that played any role in the decision of Google selecting this resource as the optimal for the inclusion into the GAB. The website is quite rich in content. It contains this recipe, structured into heading, description, basic information presented visually, main image, ingredients, recipe in 5 steps, discussion and related recipes. Thus, there are many content types present and many of them contain one or more of the keywords. It can be assumed that this plays a major role in

the final Google's decision. One important part of this analysis needs to be determining the frequency of keywords appearing on the website in various forms. The authors have selected 7 different content forms visible to the website user. Moreover, the total frequency of keywords will be measured using the HTML code of the website as well. The results are presented in Table 2.

Table 3. Summary of findings from the keyword analysis.

Parameter	Value
GAB: Different keywords in text	2 of 4
GAB: Different keywords in page title	2 of 4
GAB: Different keywords in URL	2 of 4
GAB: Different keywords in total	3 of 4
Website: Different keywords in headings (<H> tag)	2 of 4
Website: Different keywords in text (<p> tag)	2 of 4
Website: Visible keywords in total	3 of 4
Website: Keywords including HTML code in total	4 of 4
Website: Keywords in total (including duplications)	35
Website HTML code: Keywords in total (including duplications)	117
Content structured on the website	Yes
Type of structure	Steps (5)
Content structured in the GAB	No
Type of content	Text, 1 paragraph

As it can be seen from the Table 2, the monitored keywords are frequently present in all 8 locations. Even the first keyword 'How to' can be found in the HTML code, despite not being present in the visible part of the page. In total, there are 35 occurrences of all keywords in the visible part of the page and 117 occurrences in the HTML code. These and other relevant findings from the analysis are summarised in the Table 3.

5. Summary of findings

This case study has revealed several interesting findings regarding the GAB. They can be summed up as following:

- Google is able to convert numbered list, or process represented by several steps, into a single text paragraph in case it cannot display the whole content due to its length. The content was selected for a GAB despite its length and presented without numbering despite having it.
- Google duplicates content in the search results displaying the same content in two different ways.
- The content containing the three 'major' keywords related to the core of the content was excerpted despite the fourth keyword designating the query type 'how to' was not included in the website.

- The content displayed in the GAB is not useful to the user without going to the source and reading the whole recipe. It is just one part of the recipe; however, it can generate a significant traffic to the website.
- Google prefers website that is content rich and the relevant keywords are present many times in various forms on the website.
- User generated content can be one of the factors determining the inclusion of a website into the GAB; however, this presumption needs to be verified in the future research.

6. Conclusions

In this case examined, the first organic search result was also selected to be excerpted in the GAB. However, it surely makes great sense for other websites in the listing to try optimizing the structure of their content to increase their chance of being excerpted. The analysis has shown that the situation on the website appearing in the GAB is not ideal and thus, its competitors could beat it by optimizing those keyword related factors that were shown in the case study. To closely determine the situation in the presented search engine results page, it is a good idea to have a closer look at the other websites listed in the search engine results page. There are 10 regular results except of the GAB, one of them is the same website (taste.com.au), which has been included in the GAB. Thus, there are 9 other websites to be analysed. As the deep analysis of each of them would be impossible due to limits for the scope of this article, the authors point out this fact as a subject for next steps in their research. They or other experts could examine the situation on the other websites using the presented methodology for measuring keyword structure and density and moreover, other factors determining the position within the SERP can be added, the most important of them is apparently Google PageRank. This shall put more light on the factors determining the ranking of each website in the SERP and the selection process of Google for its GAB within these results.

Acknowledgement

This paper originated as the result of working on the grant scheme 2015-PSD-PAV-01 ‘Optimization of websites for search engines as an integral component of marketing’.

References

- [1] S. Tomasi, S. and X. Li, *Journal of Electronic Commerce in Organizations*, **13(1)** (2015) 27-49.
- [2] M.R. Baye, B. De los Santos and M. R. Wildenbeest, *J. Econ. Manage. Strat.*, **25(1)** (2016) 6-31.
- [3] D. Benešová and I. Némethová, *Studia Commercialia Bratislavensia*, **5(20)** (2013) 524-535.
- [4] S. Žák, *Studia Commercialia Bratislavensia*, **8(29)** (2015) 139-152.

- [5] A. Miklošik and F. Daňo, *Communication Today*, **7(1)** (2016) 82-91.