

---

# Análisis de respuestas enriquecidas en Google

*Analysis of rich answers in Google*

---

Sonia SANCHEZ-CUADRADO (1), Jorge MORATO (2)

(1) Universidad Complutense, C/Santísima Trinidad, 37 28010 Madrid, España, sscuadrado@ucm.es

(2) Universidad Carlos III, Avda. Universidad, 30, 28912 Leganés, Madrid, jmorato@inf.ucm.es

## Resumen

En recuperación de información web, los motores de búsqueda como Google incluyen funcionalidades que devuelven respuestas directas a las consultas de los usuarios. Estas respuestas tratan de resolver una necesidad informativa y se conocen como *rich answers*. Para determinar cómo se presentan estos resultados y cómo afecta la optimización de los motores de búsqueda, se ha realizado un análisis de las respuestas directas destacadas que presenta Google. En este trabajo se han examinado preguntas de tipo informacional expresadas en lenguaje natural con los términos "what is". Se ha analizado el listado de los resultados para identificar las características de las respuestas directas. Además, se han explorado las estrategias SEO que puedan determinar la relevancia del fragmento respecto de la consulta. Con este trabajo se constata que la respuesta no se extrae necesariamente de forma literal de una página web y se comprueba que la solución a las preguntas puede proceder de varios recursos. Los fragmentos de respuesta directa y otros *rich answers* pueden llegar a ocupar cerca de la mitad de la página principal de resultados, cobrando un mayor protagonismo y desplazando al resto de los resultados orgánicos. Las respuestas directas proporcionan un cambio en los hábitos de búsqueda y un nuevo modo de navegar en la red basado en un sistema hiperenlazado de pregunta respuesta.

**Palabras clave:** Respuestas enriquecidas. Fragmentos enriquecidos. GAB. SEO. Motores de búsqueda. Google. Optimización de respuestas.

## 1. Introducción

Los motores de búsqueda son el principal punto de acceso para encontrar y consultar información disponible en Internet. En la actualidad, los buscadores web representan una herramienta imprescindible para los internautas y la recuperación de información. Aunque existe un amplio número de buscadores, Google figura como el buscador con mayor cuota de mercado tanto a nivel mundial (StatCounter, 2023) como en España (StatCounter, 2022). Esto supone un ingente volumen de páginas web y una elevada competencia para disputarse el tráfico. Por eso, los sitios web que desean tener visibilidad aspiran no solo a estar incluidos en el índice del buscador, sino a posicionarse entre los primeros resultados (Killo-ran, 2013). Los buscadores ofrecen el listado de

## Abstract

In Web Information Retrieval, search engines, such as Google, include features that return direct responses to user queries. These answers try to solve an information need and are known as rich answers. To find out how these results are presented and how it affects search engine optimization, an analysis of the featured direct responses presented by Google has been carried out. In this work, a collection of informational questions expressed in natural language with the terms "what is" has been examined. The ranking of results has been analysed, identified the characteristics of the direct responses. Also, SEO factors that can determine the relevance of the fragment with respect to the query have been explored. This work confirms that the answer is not necessarily taken literally from the web page and confirms that the solution to the questions can come from various resources. The direct response fragments and other rich answers can occupy close to half of the main results page, taking on a greater role and scroll down the rest of the organic results. Direct answers provide a change in search habits and a new way of browsing the web based on a hyperlinked question-answer system.

**Keywords:** Rich answers. Featured snippets. GAB. SEO. Search engine. Google. AEO.

documentos ordenados por relevancia basada en decenas de criterios. Aunque se desconoce el peso de cada factor, las páginas web tratan de optimizar el contenido para el motor de búsqueda (SEO) de algunos criterios conocidos, con el fin de conseguir valores altos de relevancia y una posición entre los primeros resultados (Morato et al., 2013; Enge et al., 2015; Harto, 2019; Ziakis, 2019).

En la página principal del motor de búsqueda, también aparecen respuestas precisas a las consultas en un lugar destacado (Miklosik, 2019). Para Google, la primera página de resultados (SERP, *search engine results page*) es esencial y se traduce en continuos cambios en su anatomía y distribución del contenido. La disposición de los elementos de la SERP se modifica en función de

las bases de conocimiento del buscador, la consulta que ejecuta el usuario y las variaciones en el algoritmo de recuperación y posicionamiento. Fruto de las estrategias de posicionamiento y de la navegación vertical aparecen en la SERP resultados procedentes de las pestañas de videos, noticias o *shopping*. Visualmente, este tipo de elementos supone un resultado atrayente para el usuario, porque ocupa un espacio considerable además de proporcionar información detallada (por ejemplo, título, enlace, miniatura). Al final de la primera página de resultados también se ubica un bloque para búsquedas relacionadas con consultas de otros usuarios. Se muestran unos ocho términos compuestos de búsqueda con las palabras clave de la consulta ubicados en dos columnas.

Otro de los cambios significativos es la incorporación de respuestas directas (*rich answers*) que proporcionan una contestación breve a una necesidad de información concreta. Estas respuestas adoptan formas distintas en función de la fuente. Sin embargo, se desconocen los criterios por los que estos recursos son seleccionados para proporcionar la respuesta.

El propósito de este estudio es analizar las respuestas directas destacadas que presenta Google en la SERP. Se ejecutan una batería de consultas informacionales y se exploran los resultados en busca de factores que contribuyan a determinar los criterios de relevancia. De acuerdo con este propósito general de la investigación, se formulan los siguientes objetivos: 1) descubrir los tipos de respuestas enriquecidas que aparecen en la SERP; 2) explorar los resultados de las respuestas enriquecidas; y 3) determinar los factores de relevancia para las respuestas enriquecidas.

## 2. Consultas y respuestas en Google

### 2.1. Tipo de consultas

Las consultas pueden ser de varios tipos según los datos requeridos (Moldovan et al., 2000; Broder, 2002; Mishra & Jain, 2016) y precisan de un proceso concreto para ser resueltas. No existe un único criterio para categorizar la necesidad del usuario y las respuestas del buscador, pero los primeros trabajos determinaron tres tipos esenciales por la intención de la consulta (navegacionales, transaccionales, informacionales) (Broder, 2002). Otras propuestas se centran en el criterio de clasificación sobre el tipo de información buscada (p.e. objetiva, de opinión, noticias) (Lewandowsky, et al., 2012) o el tipo de respuesta proporcionada, ya sea breve o extensa (Khashabi, et al., 2021).

Las consultas transaccionales (por ejemplo, escuchar música, jugar o descargar un archivo,

comprar un smartphone, o vuelos baratos a...) interesan especialmente por la rentabilidad económica. Disponen incluso de un espacio adicional (*shopping*) dirigido a la actividad comercial. Los resultados de las consultas transaccionales suelen optimizarse con SEO y marketing en motores de búsqueda (SEM).

Las consultas informacionales proporcionan resultados para satisfacer una necesidad de información del usuario, ya sea factual, procedimental o conceptual. El desarrollo tecnológico de los últimos años ha favorecido el desarrollo de preguntas completas en lenguaje natural para traducir la necesidad del usuario. Este proceso de comunicación con la computadora ha originado la ejecución de preguntas completas que incorporan los términos "*what is / qué es*" o "*how to / cómo ...?*" a la frase que define la consulta.

### 2.2. Tipo de respuestas directas

Google, en su faceta de sistema pregunta-respuesta, trata de devolver resultados en forma de respuesta enriquecidas (*rich answers*) para resolver una necesidad informativa. Entre los tipos de respuestas enriquecidas presentadas en la SERP se encuentran los basados en imágenes, elementos gráficos, texto, clasificación en pestañas, diapositivas, tablas o resultados de formularios. Para una consulta pueden aparecer distintas respuestas enriquecidas combinadas, ya sean visualizadas al mismo tiempo u organizadas en pestañas, diapositivas o carruseles de imágenes (Zhao et al., 2019). También se utiliza un espacio destacado en la SERP denominado *Google Answer Box* (GAB) para respuestas concisas a la consulta de un usuario (Miklosik et al., 2016; Sullivan, 2018; Miklosik, 2019; Khashabi et al., 2021). Este espacio destacado se ubica en la posición cero. Esta posición supone un lugar privilegiado en la parte superior de la página que asegura la visualización e incremento del tráfico para el recurso. Se materializa con los fragmentos destacados (*feature answers*) a modo de respuestas directas enriquecidas (Strzelecki & Rutecka, 2019). Estos resultados directos ofrecen al usuario una solución de consulta sin necesidad de clicar (Ishkin, 2019).

#### 2.2.1. *Featured answer*

Una respuesta destacada o *featured answer* es un fragmento o resumen extraído del texto de una página web que responden de forma directa a una consulta ejecutada en el buscador (Miklosik, 2019; Strzelecki & Rutecka, 2020a). Los fragmentos destacados tratan de proporcionar una respuesta completa ya sea simple o con varios pasos o piezas de información (por ejemplo, respuestas

a *how to?*). Los mecanismos para comprender la consulta del usuario y ofrecer la respuesta precisa presentan cierto grado de sofisticación (Nayak, 2019; Nayak, 2022). Se visualizan como párrafo, listados, tablas, gráficos, imágenes o una combinación de estos elementos. Supone un recurso para acelerar el proceso de resolver la necesidad de información, permitiendo optimizar la experiencia del usuario (Strzelecki & Rutecka, 2020b).

### 2.2.2. Panel de conocimientos

Otro tipo de respuestas destacadas son las tarjetas de conocimiento que aparecen en el lateral derecho de la SERP enmarcadas en un cuadrado. Muestra información a modo de resumen sobre la consulta. Esta información procede del propio buscador o fuentes de conocimiento como la Wikipedia y páginas de instituciones públicas o privadas. Estos paneles también están relacionados con los gráficos de conocimiento (Singhal, 2012) y con la información local de las ubicaciones de negocios para su representación en el mapa. En ocasiones aparecen *widgets* con alguna funcionalidad como realizar una traducción, buscar un término, o efectuar una conversión de unidades.

### 2.2.3. Rich snippet

Los fragmentos enriquecidos o *rich snippets* muestran una información destacada del contenido de la página web. Suelen ofrecer una respuesta concisa a la consulta de un usuario en forma de fragmento o párrafo marcado previamente. Se utilizan datos estructurados dentro del código fuente para clasificar la información o expresar una función (Chen et al., 2018). Asimismo, se proporciona metainformación sobre el recurso y el contenido de la página web. Los datos se presentan en forma de resumen, reseñas, valoraciones, eventos, etc. de acuerdo con la clasificación de schema.org (Enge et al., 2015; Zhao et al., 2019).

### 2.2.4. Otras preguntas de los usuarios

También aparece en la página de resultados un bloque denominado "Otras preguntas de los usuarios" conocidas como PAA del inglés *People Also Ask*. Este bloque recoge sugerencias de consultas en forma de pregunta autogeneradas. Las preguntas se expresan en lenguaje natural y comienzan por palabras interrogativas como qué, quién, cuál, cómo, por qué y dónde, principalmente. Tiene un formato dinámico y recursivo donde al desplegar una pregunta para obtener la respuesta se origina otro bloque con nuevas preguntas de otros usuarios relacionadas. Representan técnicas de navegación y redirección de resultados, pero también iniciativas para afinar la

experiencia del usuario al estilo de las propuestas alternativas a la consulta "Sugerencias de Google".

### 2.3. Selección de rich answers

En todos estos tipos de respuesta, el algoritmo de Google decide los criterios para mostrar una *featured answer* o una PAA (Sullivan, 2018; Nayak, 2019; Nayak, 2020). Las investigaciones, que tratan de averiguar cómo se seleccionan y extraen las respuestas directas, apuestan por criterios como la velocidad de carga de las páginas, las palabras clave, la autoridad del dominio de la página o la longitud de las consultas (Miklošik et al., 2016).

### 2.4. Impacto de la rich answer en la SERP

Los resultados enriquecidos, ya sean fragmentos directos o con formato visual, repercuten de forma positiva en la experiencia del usuario (Yu, 2014: 492), mejorando la accesibilidad y usabilidad de la SERP. Respecto al buscador, la percepción del usuario mejora con los resultados directos pues el acceso directo permite descartar cualquier otro resultado orgánico. Desde el punto de vista del recurso, las páginas web experimentan una mayor visibilidad y un aumento del tráfico (Strzelecki & Rutecka, 2020a; Duong, 2019), contribuyendo a mejorar su autoridad del dominio. Desde el punto de vista del usuario, algunos trabajos han estudiado cómo las respuestas enriquecidas condicionan a los internautas quienes tienden a sobrevalorar la confianza en un sitio web (Bernstein et al., 2012; Lurie & Mustafaraj, 2018). Con estos estudios los autores afirman que las respuestas enriquecidas alteran el comportamiento informacional de los usuarios (Bilal & Huang, 2019; Wu, et al., 2020; Bink et al., 2022).

Las respuestas enriquecidas también han contribuido al cambio en los hábitos de la búsqueda de información a través de los asistentes de voz (Trippas, 2020). Esta opción permite formular consultas de forma verbal a través de las funciones del habla en sistemas como Google Search Assistant, Siri o Alexa. La respuesta que proporciona el asistente de voz procede de los resultados directos (*rich answers, google answer box, feature snippet...*). Este único resultado puede ser escuchado en los asistentes de voz (Epstein et al., 2022), pero se omiten los demás. El uso de los asistentes de voz para demandar información se ha extendido con rapidez. Algunos estudios estiman un incremento de más del 20 % en servicios de búsqueda por voz (Mander & Buckle, 2018, eMaker & Insider Intelligence, 2022) por lo tanto también aumenta la repercusión de las respuestas directas.

Ante tal aumento de oportunidades, cobra relevancia el desarrollo de resultados enriquecidos para entornos de pregunta-respuesta (Enge, 2017). Los gestores de contenido se afanan en elaborar posibles resultados más relevantes y pertinentes que puedan ser seleccionados por el buscador. Podemos hablar del desarrollo de una nueva materia sobre la optimización de respuestas (*Answer Engine Optimization*) que explora los criterios de relevancia de los algoritmos de recuperación de información para proporcionar respuestas directas.

### 3. Metodología

Este trabajo ha acometido un estudio empírico de las respuestas enriquecidas en la SERP de Google centrado en analizar las respuestas directas y explorar el comportamiento del buscador para consultas en lenguaje natural del tipo “*What is?*”. Este análisis está basado en la observación y exploración de los resultados de la SERP de Google. Se examinan criterios como el tipo de respuesta, la disposición, el tamaño, la extracción total o parcial del párrafo, la ubicación del fragmento en el texto original, entre otras (véase la Tabla I). El procedimiento seguido para el análisis incluye las fases de: selección de consultas, recolección de datos y análisis de los resultados.

#### 3.1. Selección de consultas

Se recopilaron 16 consultas de tipo informacional (Broder, 2002) en lenguaje natural y en inglés que incluyen las palabras “*What is?*”:

- Q1) What is NATO?
- Q2) What is Monkey Pox?
- Q3) What is Wordle?
- Q4) What is an Oligarch?
- Q5) What is an NFT?
- Q6) What is Aphasia?
- Q7) What is Hepatitis?
- Q8) What is the population of Ukraine?
- Q9) What is Salmonella?
- Q10) What is the Met Gala?
- Q11) What is a serp?
- Q12) What is a emotional breakdown?
- Q13) What is a degree in math?
- Q14) What is a handshake error?
- Q15) What is a ferromagnetic material?
- Q16) What is a career goal?

Se entiende que todas ellas tienen como propósito recuperar una respuesta específica. Para ello, se seleccionaron las preguntas de la lista de *Google Trends* del 2022 que se basan en términos de búsqueda cuyo mayor incremento de consulta se produjo en ese año en comparación con el anterior (Google, 2023; Brita-Paja, 2022). Las consultas contienen palabras clave simples y/o compuestas.

Bloque	Criterio	Descripción	
SERP	Info + idiomas	Número de idiomas utilizados en los bloques	
	Presencia <i>featured answer</i>	Presencia o no de una <i>featured answer</i> directa	
	Imagen	Presencia de imagen junto a la <i>feature answer</i>	
	1ª posición <i>feature answer</i>	Si la <i>feature answer</i> directa se encuentra en 1ª posición	
	Posición <i>featured answer</i>	Posición de la <i>feature answer</i> en la disposición en la SERP	
	Posición fuente en orgánicos	Posición del recurso fuente en los resultados orgánicos	
	N_Bloques	Número de recursos enriquecidos	
	Nº preguntas sugeridas	Nº de preguntas sugeridas por el buscador distintas	
	Tarjeta de conocimiento	Fuente de la tarjeta (Wikipedia, Google)	
	Imagen tarjeta	Presencia de imágenes	
	Clasificación entidad	Clasificado como tipo de entidad	
	GC	Gráfico de conocimiento	
	Gráfico	Respuesta en forma de gráfico	
	Calculadora	Si se muestra el <i>widget</i> de la calculadora	
	Traductor	Si se muestra el <i>widget</i> para traducir términos a otros idiomas	
	Diccionario	Si se muestra el <i>widget</i> para buscar términos	
<i>Featured answer</i>	Polisemia <i>Keyword</i>	Si en los resultados se muestran varias acepciones	
	Longitud	Número de palabras que contiene la respuesta	
	Negrita	Si se utiliza formato negrita para presentar la respuesta	
	Formato subrayado	Uso formato subrayado para presentar la respuesta	
	Respuesta en origen	Nº de párrafos	Nº de párrafos seleccionados para la respuesta
		Total/parcial	Respuesta seleccionada total o parcial del párrafo
		Posición párrafo	Posición del párrafo respecto a la página fuente
		Marcas en fuente	Marcas de formato del texto seleccionado en la página fuente
<i>Keywords</i> en encabezado	Presencia de la <i>keyword</i> en el título o encabezados		
Negrita	Presencia de la <i>keyword</i> de consulta en formato negrita		

Tabla I. Criterios para análisis de respuestas directas

### 3.2. Recolección de datos

El estudio fue desarrollado durante el periodo comprendido entre los meses de febrero y marzo de 2023. Una vez ejecutadas las consultas se procedió a examinar los bloques de respuestas y se recopilaron los datos. Para cada respuesta del buscador se recogió información acerca de cómo se mostraban los resultados, si aparecían respuestas directas y contenido enriquecido, cuál era el contenido presentado, el recurso o recursos originales de las que proceden y la posición en el listado de la SERP.

### 3.3. Análisis de los resultados

El análisis de los datos se realizó mediante la observación y la exploración de características principalmente de los fragmentos destacados y su relación con otros elementos de respuestas enriquecidas a varios niveles (Tabla I, en la página anterior). Por un lado, los indicadores definidos para este estudio se dirigieron a observar la aparición de respuestas enriquecidas en la SERP, con respecto al tipo de contenido y cómo se presentaban en función de las palabras clave de la consulta. También se recogió información sobre el texto de la respuesta en la fuente de origen, así como la posición que ocupaba en el listado ordenado de los resultados orgánicos. Además de información presentada en los distintos GAB como imágenes gráficas y *widgets*, se anotó la posición y disposición que ocupaban.

Por otro lado, se han recogido factores de posicionamiento como la presencia de palabras clave de la consulta en encabezados, presencia de palabras clave en formato negrita, marcas de formato destacado y la posición del párrafo en el recurso. Para un examen más profundo de las respuestas directas se extrajo el texto y se analizó tamaño y la estructura lingüística (Sánchez-Cuadrado et al., 2003).

## 4. Resultados

A fin de mostrar los resultados obtenidos, se concretan a continuación los criterios en los que han sido clasificados de acuerdo con los objetivos: describir los tipos de respuestas enriquecidas, explorar los resultados, análisis de los factores de relevancia.

### 4.1. Tipo de respuestas enriquecidas y anatomía de la SERP

Entre las primeras posiciones de la página de resultados se incluyen respuestas directas alternativas como: un bloque de definición del término consultado, un bloque de tarjeta procedente de la Wikipedia, bloques y grafos de conocimiento, bloque

para la respuesta de la consulta o un bloque de preguntas (Figura 1). En cada disposición se incorporan respuestas enriquecidas y *widgets* que proporcionan funcionalidades concretas (calculadora, búsqueda de palabra en un diccionario, traducción de términos). A la hora de visualizar los resultados en la versión de escritorio, se observa un cambio en la anatomía de la SERP con respecto a versiones previas. Aunque, no todas las consultas devuelven respuestas destacadas (*featured answer*).

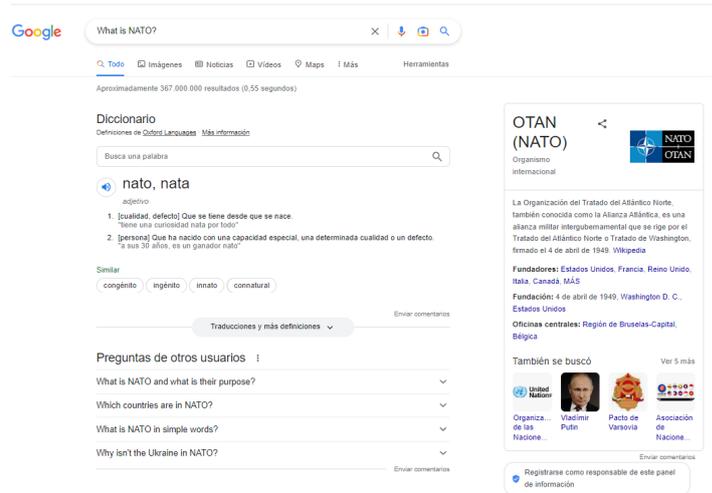


Figura 1. Resultado de consulta en Google para la pregunta 1

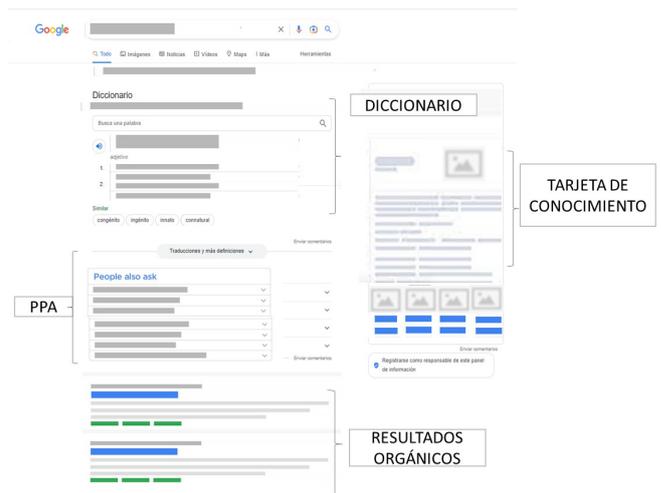


Figura 2. Anatomía de la SERP tipo 1

En cuanto a la ubicación de los resultados y las respuestas enriquecidas en la SERP se observan distintas disposiciones según las informaciones

que se vayan a ofrecer para cada consulta. Las respuestas con fragmentos enriquecidos suelen ocupar los primeros resultados, aunque también se observan algunos casos entre los resultados orgánicos comunes, y por tanto en posiciones intermedias. Las imágenes siguientes (Figuras 2, 3, 4 y 5) recogen cuatro ejemplos de disposición tipo que se han encontrado en la exploración de la SERP para los resultados a las consultas.

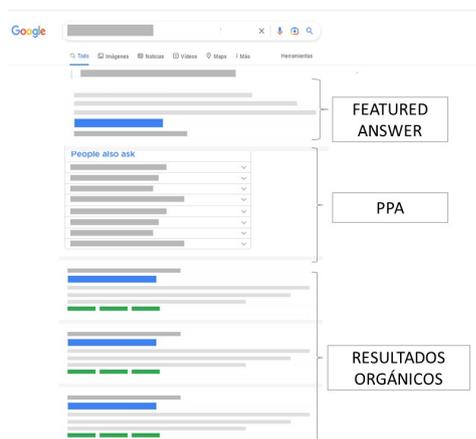


Figura 3. Anatomía de la SERP tipo 2

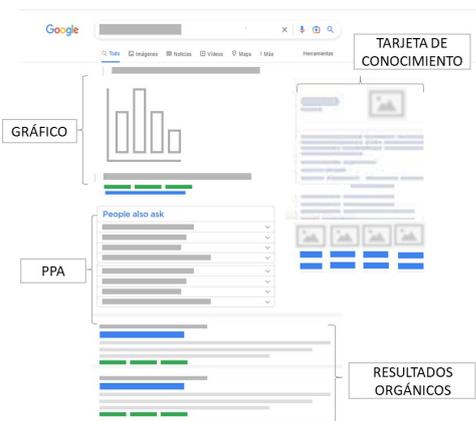


Figura 4. Anatomía de la SERP tipo 3



Figura 5. Anatomía de la SERP tipo 4

En la respuesta a la consulta sobre la población de un país se utiliza un gráfico dinámico para proporcionar un conocimiento complejo, basado en datos procedentes de diferentes fuentes (Figura 6). La incorporación de estos elementos implica que el listado de los resultados orgánicos queda relegado a la siguiente pantalla, teniendo que desplazar el *scroll* y cambiado de página para poder acceder a este tipo de resultados. Se observa como las respuestas directas mejoran las posibilidades de visualización en detrimento de los resultados posteriores.

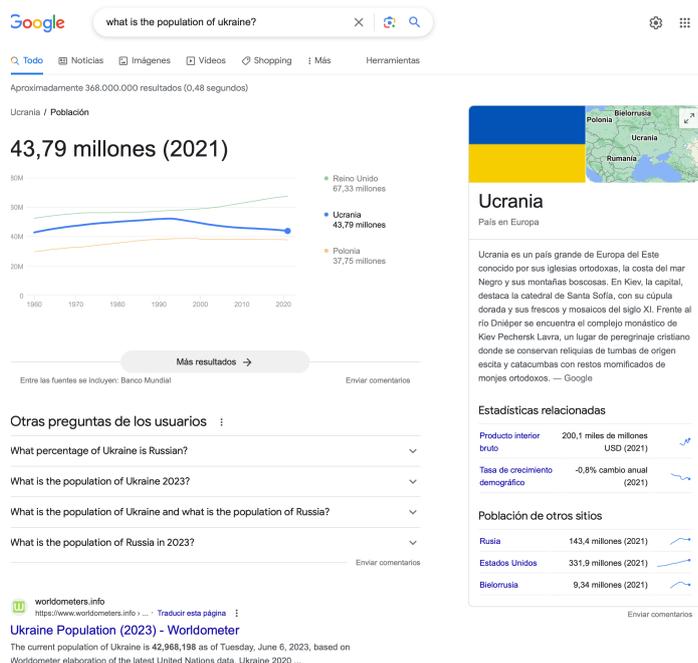


Figura 6. Resultado de la consulta 2 en Google

Explorando los resultados de las respuestas a las preguntas, se refleja control de la polisemia de las palabras. Se ha reportado el significado de un término común en el diccionario y en las tarjetas de la *Wikipedia* una entidad con nombre propio (por ejemplo, *OTAN/NATO*). Se observa además que el resultado puede ser plurilingüe basado con toda probabilidad en los datos de geolocalización de la ejecución de la consulta.

Se observa la concurrencia de recursos en idiomas diferentes en el 81 % de los casos, pudiendo mostrar los resultados orgánicos en el idioma de la consulta ejecutada (inglés) y tarjetas de información y grafos de conocimiento en otro. Por tanto, responde a procesos independientes del idioma con herramientas y mecanismos basados en el conocimiento que permiten identificar diferentes representaciones en un único concepto, así como personalizar los resultados por ubicación (Figura 7).

Google search results for "what is salmonella?". The search bar shows the query and the number of results (72,100,000). The results include:

- mayoclinic.org**: [Salmonella infection - Symptoms and causes - Mayo Clinic](https://www.mayoclinic.org/symptoms-causes/syc-2045). 29 abr 2022 — Salmonella infection is usually caused by eating raw or undercooked meat, poultry, and eggs or egg products or by drinking unpasteurized milk.
- cdc.gov**: [Questions and Answers | Salmonella | CDC](https://www.cdc.gov/general/about-cdc/questions-and-answers/salmonella). Salmonella are bacteria that make people sick. What illness do people get from Salmonella infection? Most types of Salmonella cause an illness called ...
- clevelandclinic.org**: [Salmonella: Outbreaks, Causes, Symptoms & Treatment](https://my.clevelandclinic.org/health/diseases/17281/salmonella). 24 ago 2022 — Salmonella, or salmonellosis, is an infection with Salmonella bacteria that causes diarrhea, fever and stomach pains. Salmonella usually goes ...

The right-hand side of the image shows a knowledge panel for "Salmonella" with details such as "Bacteria", "Nombre científico: Salmonella", "Clasificación inferior: Salmonella enterica, Salmonella bongori", "Clase: Gammaproteobacteria", "Dominio: Bacteria", "Familia: Enterobacteriaceae", "Filo: Proteobacteria", and "Orden: Enterobacteriales".

Figura 7. Resultado de la consulta 3 en Google

De todas las búsquedas se han recuperado respuestas enriquecidas de diversos tipos. En la mayoría de los casos se han identificado dos o tres bloques de respuesta enriquecida por consulta (Gráfico 1). En algo menos de la mitad, el buscador ha respondido con tarjetas de conocimiento procedentes del diccionario de Google y de la Wikipedia. Aunque solo en el 75 % de los casos, el buscador ha proporcionado un respuesta o fragmento destacado.

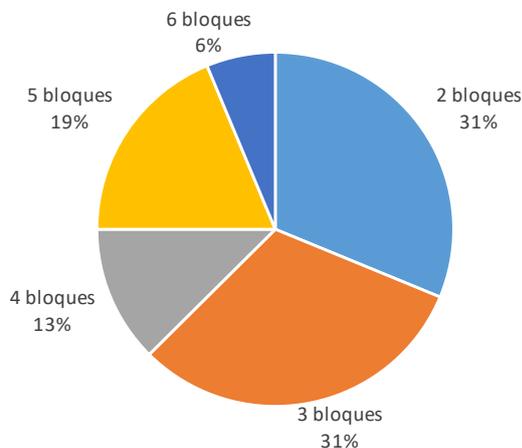


Gráfico 1. Proporción de bloques con respuestas enriquecidas por consulta

En todos los casos, el motor de búsqueda ha sugerido un bloque de PAA. Las preguntas sugeridas responden a los 5 tipos de preguntas (*what, why, who, how, when*), pudiendo mostrar más de 20 por bloque, aunque en la mayoría de los casos (43 %) se muestran 4. Cada una de estas preguntas puede ser respondida por distintas fuentes de información web.

También se detecta una correlación entre la estructura gramatical de la consulta y las otras preguntas de los usuarios que construyen frases con los mismos términos y estructuras semánticamente relacionadas. Se observa la generación de búsqueda amplia con todas las palabras de la consulta original (Tabla II) pero también se percibe un conocimiento semántico sobre el concepto. En la Tabla se muestra el resultado de la pregunta *What is a Salmonella?* Las preguntas sugeridas permiten deducir sin leer la respuesta que se trata de una cuestión relacionada con la comida y la salud. Se deduce además que tiene una sintomatología y una cura asociada. La configuración de la respuesta se expande a otros términos interrogativos frecuentes para proporcionar una visión facetada y multidimensional del concepto.

Consultas	Pregunta sugerida (PAA)
What is NATO?	What is NATO and what is their purpose? What countries are in NATO? What countries are not in NATO?
What is Wordle?	How do you play Wordle? What is Wordle and how does it work?
What is an NFT?	What exactly is an NFT in simple terms?
What is Salmonella?	What is the main cause of Salmonella? What happens if you get Salmonella? What food is Salmonella found in? What are 4 symptoms of Salmonella? What kills Salmonella in the body? Can Salmonella be cured?
What is the Met Gala?	What is the purpose of the Met Gala?

Tabla II. Ejemplos de preguntas sugeridas por Google para la consulta

#### 4.2. Presentación de la featured answer

Para las respuestas destacadas, Google visualiza la respuesta directa en la SERP. La respuesta que define el término cuestionado se presenta en forma de párrafo. El GAB puede contener el párrafo de la respuesta, el nombre y el enlace al recurso fuente, información sobre la fuente, el título de la url seleccionada y una fecha, correspondiente a la fecha de actualización de la página. Se mantiene el contenido del párrafo original, pero se destaca la respuesta precisa en negrita, aunque en el recurso original no aparezca esa marca. El formato destacado puede referirse a todo el párrafo, o bien a una parte de este. En otras ocasiones, el texto completo de la respuesta se muestra en negrita, y se destaca con subrayado algunas palabras clave.

En cuanto al texto se identifica una relación entre el patrón sintáctico de las frases, de la pregunta

y las respuestas. Se recogen estructuras sintácticas que denotan una correspondencia léxico-semántica por las cuales se identifican las relaciones entre los términos (por ejemplo, *means* (X,Y), *ensures\_that* (X,Y), *is\_a* (X,Y)). Esas estructuras permiten identificar patrones que revelan respuestas a consultas simples que responden al interrogante “*What is?*”.

#### 4.3. Recurso fuente de la respuesta: posición y formato

En el caso de los fragmentos destacados, los datos obtenidos indican que la respuesta suele proceder de los primeros recursos posicionado en el ranking de resultados de la SERP de Google. Como se observa en el Gráfico 2, el 46% se situó en la posición 0. Los resultados de este estudio muestran que puede estar situado en otra posición del ranking incluso alejada del top10.

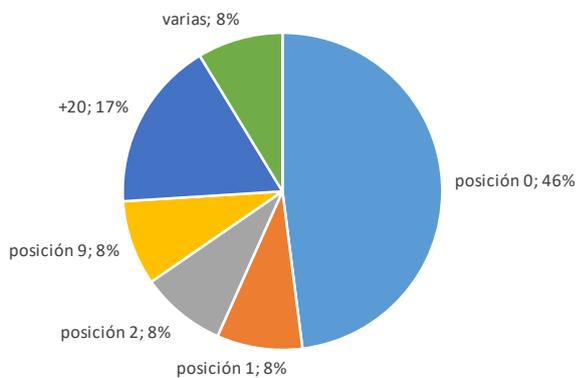


Gráfico 2. Posición del recurso fuente en el listado de resultados orgánicos

En relación con estrategias SEO *on-page*, tampoco se observa una relación directa con la posición estratégica del primer párrafo. Ni siquiera en el caso de las entradas de la Wikipedia usadas para dar respuesta a la pregunta se selecciona necesariamente el primer párrafo coincidente con la palabra clave de la consulta.

Para las respuestas destacadas se observa que la frase extraída puede estar ubicada en cualquier lugar de la página web, ya sea al inicio del párrafo, en una zona intermedia o al final. Se advierte que, en la mayoría de los casos examinados, la palabra clave de la consulta forma parte del encabezamiento del párrafo donde se encuentra la respuesta. Aunque el párrafo puede estar dentro de un bloque de contenido cuyo epígrafe no contiene los términos de la consulta, y la pregunta no es idéntica a la *query* ejecutada en el buscador. El código fuente de la página recoge el párrafo en un bloque de contenido `<div>` y contiene la palabra clave consultada en etiquetas

de encabezamiento `<h2>` y `<h4>`. En otras ocasiones aparece como un párrafo común y precedido de un encabezamiento que contiene parcialmente el término clave consultado en el caso de que sea compuesto. La longitud de la respuesta está entre 37-55 términos, aunque en ocasiones se ha observado un resultado menor ( $\pm 22$  palabras) y no coincide necesariamente con un párrafo completo, pudiendo estar formada por partes de dos párrafos, o varias fuentes en el caso de los gráficos.

## 5. Conclusiones

Este estudio ha explorado cómo se presentan las respuestas directas para preguntas de carácter informacional del tipo *What is?* en dispositivos de escritorio. A lo largo de este trabajo se ha tratado de clarificar la tipología de los bloques de contenido para proporcionar respuestas directas (respuesta corta con texto, imágenes, *widgets* para la búsqueda del término en un diccionario, traductores, respuestas destacadas y otras preguntas de los usuarios). Con respecto a los resultados devueltos por Google, se han observado diversos tipos de *rich answers* para las preguntas *What is?* que pueden aparecer simultáneamente en la SERP para la misma consulta. Los datos indican que para todas las consultas se sugieren PAA. Estas preguntas sugeridas pueden estar situadas en posición inicial, intermedia y final, o incluso con varios bloques de preguntas a la vez. Estas preguntas tienen un carácter dinámico generando nuevas consultas relacionadas, lo que constituye un nuevo modo de navegar en la red.

A partir del análisis de la página de resultados, se ha demostrado la incorporación de múltiples bloques de respuestas directas. Por consulta se han recuperado de 1 a 5 bloques de respuestas en la primera página de resultados. Aunque lo más frecuente es que aparezcan entre 2 y 3 bloques de respuestas enriquecidas. La incorporación de estos contenidos en la SERP reduce considerablemente el espacio a los resultados orgánicos comunes, y en consecuencia se produce un desplazamiento del resto de resultados. Se ha observado que la Wikipedia no aparece como primer resultado, aunque aún mantiene una presencia significativa con las tarjetas de conocimiento laterales, y por tanto una alta visibilidad. Por lo tanto, entendemos que también experimenta este desplazamiento, manifestado por la pérdida de la primera posición en los resultados orgánicos listados.

Las respuestas directas correspondientes a las *featured answers* influyen de forma positiva en el buscador al proporcionar un resultado pertinente

directo con cero clics. Estas respuestas destacadas también ayudan a que el usuario tenga confianza del sitio, pues el sistema le otorga un espacio relevante en la SERP. A menudo este espacio se refuerza además con otras preguntas de los usuarios similares en un bloque contiguo o entre los resultados orgánicos. Con respecto a las *featured answers*, los datos indican que existe relación entre la selección como respuesta más relevante y posicionamiento SEO del recurso. Sin embargo, también parecen intervenir otros factores que alteran esa relación directa y que el fragmento provenga del primer resultado posicionado.

En la mayoría de los casos, la respuesta destacada (*featured answers*) procede del resultado que ocupa la primera o segunda posición. El recurso del que ha sido tomada la respuesta puede estar situado en cualquier posición de la primera página de resultados o incluso en páginas web posteriores. Por esto, las páginas que aplican estrategias SEO para lograr una posición entre los primeros resultados tienen una mayor probabilidad de ser visualizadas como una respuesta enriquecida. Sin embargo, parecen tener algunas peculiaridades. En consecuencia, se precisa profundizar en la optimización de preguntas en motores de búsqueda (AEO) y dilucidar cómo conseguir visibilidad de una página a través de respuestas destacadas.

Posicionar preguntas es una estrategia de posicionamiento muy acertada para el que consigue la posición 0, porque presenta una respuesta que resuelve de forma directa la necesidad del usuario, atrae tráfico web y puede mejorar la tasa de conversión de productos relacionados.

Las respuestas destacadas directas suponen un nuevo cambio en la búsqueda web. Con el análisis se advierte que la respuesta seleccionada no coincide con un párrafo completo. En ocasiones se seleccionan varios fragmentos de forma total o parcial, lo que inevitablemente plantea la cuestión de cuáles son los criterios para estimar que la respuesta está completa y como establecer los límites de la respuesta. Estos interrogantes evidencian que nos encontramos ante un sistema algorítmico más complejo.

Los avances en el procesamiento del lenguaje natural y la extracción de información están contribuyendo al desarrollo de criterios que determinen la relevancia informacional en los sistemas de recuperación de información. Los datos relativos al uso incremental de estrategias dirigidas a los asistentes de voz demuestran una clara apuesta de los motores de búsqueda por proporcionar respuestas directas. Los principales efectos que provoca son la mejora de la experiencia

del usuario y el suministro de respuestas a los asistentes de voz conversacionales.

En el caso de Google, las respuestas enriquecidas no solo proporcionan una valoración positiva sobre el buscador, sino que también se percibe una permanencia más duradera en la SERP. El usuario trata de explorar las distintas opciones ofrecidas por las *rich answers*, *widgets*, grafos de conocimiento y sugerencias de búsqueda, tratando de conceptualizar y contextualizar las consultas mediante relaciones y facetas esenciales.

Con la presente investigación se han obtenido resultados que confirman un cambio en la forma de determinar la relevancia de la información y de mostrar los resultados. Aunque el análisis de este estudio preliminar ha revelado datos representativos para establecer cambios en el algoritmo, entendemos que aumentar la muestra podría facilitar relaciones estadísticas más significativas. También sería de interés analizar las variables relativas a enlaces entrantes y dominio de autoridad que permitan observar el impacto de factores externos. Google hace cambios en el algoritmo y el aumento de nuevos recursos y preguntas puede alterar la disposición de los elementos en su página de resultados. Con todo, los resultados son válidos por cuanto muestran una situación en una determinada franja temporal y aunque haya modificaciones seguirá siendo una fase transitoria hacia otro modelo de recuperación en el que entrará en juego el posicionamiento de bloques de respuestas enriquecidas. Para futuros trabajos nos proponemos un estudio en profundidad sobre los criterios de selección de las respuestas destacadas y las otras preguntas de los usuarios. Otro tipo de búsquedas como *“How to?”* también merecen un análisis particular, pues presentan características diferenciadas en forma de listas para reflejar los pasos del procedimiento.

## Referencias

- Bernstein, Michael S.; Teevan, Jaime; Dumais, Susan; Liebling, Daniel; Horvitz, Eric (2012). Direct answers for search queries in the long tail. // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Austin Texas USA: ACM 237-246. ISBN 9781450310154. <https://doi.org/10.1145/2207676.2207710>
- Bilal, Dania; Huang, Li-Min (2019). Readability and word complexity of SERPs snippets and web pages on children's search queries: Google vs Bing. // *Aslib Journal of Information Management*. 71:2, 241–259. <https://doi.org/10.1108/AJIM-05-2018-0124>
- Bink, Markus; Zimmerman, Steven; Elsweller, David (2022). Featured Snippets and their Influence on Users' Credibility Judgements. // *ACM SIGIR Conference on Human Information Interaction and Retrieval*, 113–122. <https://doi.org/10.1145/3498366.3505766>
- Brita-Paja Nuñez, P. (2022). Elementos de posicionamiento de los sistemas de pregunta respuesta de la web. *Getafe:*

- Universidad Carlos III de Madrid. Trabajo Fin de Grado. <http://hdl.handle.net/10016/36335>
- Broder, Andrei (2002). A taxonomy of web search. // ACM SIGIR Forum, 36:2, 3–10. <https://doi.org/10.1145/792550.792552>
- Chen, Wei-Fan; Hagen, Matthias; Stein, Benno; Potthast, Martin (2018) A User Study on Snippet Generation: Text Reuse vs. Paraphrases. // The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. 1033–1036. <https://doi.org/10.1145/3209978.3210149>
- Duong, Véronique (2019). State of the Art of SEO. // SEO Management: Methods and Techniques to Achieve Success (1.a ed., pp. 1-9). Wiley Data and Cybersecurity. <https://doi.org/10.1002/9781119681427>
- eMarketer, & Insider Intelligence. (April 13, 2022). Number of voice assistant users in the United States from 2017 to 2022 (in millions) [Gráfico]. // Statista. <https://www-statista.com/statistics/1029573/us-voice-assistant-users/>
- Enge, Eric; Spencer, Stephan; Stricchiola, Jessie C. (2015). The Art of SEO: mastering search engine optimization. 3ª ed.. O'Reilly Media, Inc, USA.
- Enge, Eric (2017). Featured Snippets: New Insights, New Opportunities. Stone Temple Company Blog. <https://www.stonetemple.com/featured-snippets-new-insights-new-opportunities/>
- Epstein, Robert; Lee, Vivian; Mohr, Roger; Zankich, Vanessa R. (2022). The Answer Bot Effect (ABE): A powerful new form of influence made possible by intelligent personal assistants and search engines. // PLOS ONE. 17:6, e0268081. <https://doi.org/10.1371/journal.pone.0268081>
- Google (2023). Google Trend. <https://trends.google.es/>
- Harto, Agus Budi (2019). Implementing Website Design Based on Search Engine Optimization (SEO) Checklist to Increase Web Popularity. Journal of Applied Information, Communication and Technology, 6:2, 87–97. <https://doi.org/10.33555/ejaict.v6i2.67>
- Ishkin Rand. (2019 Agosto 13). Less than half of Google searches now result in a click. // SparkToro. <https://sparktoro.com/blog/less-than-half-of-google-searches-now-result-in-a-click/>
- Khatabi, Daniel; Ng, Amos; Khot, Tushar; Sabharwal, Ashish; Hajishirzi, Hannaneh; Callison-Burch, Chris (2021). GooAQ: Open Question Answering with Diverse Answer Types. // Findings of the Association for Computational Linguistics: EMNLP 2021, 421–433. <https://doi.org/10.18653/v1/2021.findings-emnlp.38>
- Killoran, John B. (2013). How to Use Search Engine Optimization Techniques to Increase Website Visibility. // IEEE Transactions on Professional Communication, 56:1, 50–66. <https://doi.org/10.1109/TPC.2012.2237255>
- Lewandowski, Dirk; Drechsler, Jessica; Mach, Sonja (2012). Deriving query intents from web search engine queries. // Journal of the American Society for Information Science and Technology, 63:9, 1773–1788. <https://doi.org/10.1002/asi.22706>
- Lurie, Emma; Mustafaraj, Eni (2018). Investigating the Effects of Google's Search Engine Result Page in Evaluating the Credibility of Online News Sources. // Proceedings of the 10th ACM Conference on Web Science, 107–116. <https://doi.org/10.1145/3201064.3201095>
- Mander, Jason; Buckle, Chase (2018). Voice Search. Insight Report 2018. // Global Web Index. <https://www.insidemarketing.eu/cdn/wp-content/uploads/2019/01/Voice-Search-global-web-indexreport.pdf>
- Miklošič, Andrej; Hlavatý, Ivan; Daňo, Ferdinand; Červenka, Peter (2016). Google Answer Box Keyword-related analysis. A case study. // European Journal of Science and Theology. 12:5,185-194. [http://www.ejst.tuiasi.ro/Files/60/19\\_Miklosik%20et%20al.pdf](http://www.ejst.tuiasi.ro/Files/60/19_Miklosik%20et%20al.pdf)
- Miklosik, Andrej. (2019). Search Engine Marketing Strategies: Google Answer Box-Related Search Visibility Factors. // Handbook of Research on Entrepreneurship and Marketing for Global Reach in the Digital Economy, edited by Luísa Cagica Carvalho y Pedro Isaias, Eds. Hershey, PA, USA: IGI Global, 2019. 463-485. <https://doi.org/10.4018/978-1-5225-6307-5.ch020>
- Mishra, Amit; Jain, Sanjay Kumar (2016). A survey on question answering systems with classification. // Journal of King Saud University - Computer and Information Sciences. 28:3, 345-361.
- Moldovan, Dan; Harabagiu, Sanda; Pasca, Marius; Mihalcea, Rada; Girju, Roxana; Goodrum, Richard; Rus, Vasile (2000). The structure and performance of an open-domain question answering system. // Proceedings of the 38th Annual Meeting on Association for Computational Linguistics - ACL '00, 563–570. <https://doi.org/10.3115/1075218.1075289>
- Morato, Jorge; Sánchez-Cuadrado, Sonia; Moreno, Valentin; Moreira, José. A. (2013). Evolución de los factores de posicionamiento web y adaptación de las herramientas de optimización. // Revista española de Documentación Científica. 36:3, e018. <https://doi.org/10.3989/redc.2013.3.956>
- Nayak, Pandu (2019). Understanding Searches Better Than Ever Before. // Google. <https://blog.google/products/search/search-language-understanding-bert>
- Nayak, Pandu (2022). How AI powers great research results. // Google. <https://blog.google/products/search/how-ai-powers-great-search-results/>
- Sánchez-Cuadrado, Sonia; Lloréns, Juan; Morato, Jorge; Hurtado, José. A. (2003). Extracción Automática de Relaciones Semánticas. 2da. Conferencia Iberoamericana // Sistemas, Cibernética e Informática. CISC 2003. Orlando (Florida) 31 - 2 de agosto de 2003. 265-268.
- Singhal, Amit (2012). Introducing the knowledge graph: things, not strings. <https://googleblog.blogspot.co.at/2012/05/introducing-knowledge-graph-things-not.html>.
- StatCounter. (abril 26, 2023). Worldwide desktop market share of leading search engines from January 2015 to March 2023 [Gráfica]. // Statista. <https://www-statista.com/statistics/216573/worldwide-market-share-of-search-engines/>
- StatCounter. (mayo 10, 2022). Cuota de mercado de los principales motores de búsqueda online usados // España // 2021 [Gráfica]. // Statista. <https://es-statista.com/estadisticas/670092/cuota-de-mercado-de-los-motores-de-busqueda-por-buscador-espana/>
- Strzelecki, Artur; Rutecka, Paulina (2019). The Snippets Taxonomy in Web Search Engines. // M. Pańkowska & K. Sandkuhl (Eds.). Perspectives in Business Informatics Research. 365, 177-188. Springer International Publishing. [https://doi.org/10.1007/978-3-030-31143-8\\_13](https://doi.org/10.1007/978-3-030-31143-8_13)
- Strzelecki, Artur; Rutecka, Paulina (2020a). Direct Answers in Google Search Results. // IEEE Access. 8, 103642-103654. <https://doi.org/10.1109/ACCESS.2020.2999160>
- Strzelecki, Artur; Rutecka, Paulina (2020b). Featured Snippets Results in Google Web Search: An Exploratory Study. In Á. Rocha, J. L. Reis, M. K. Peter, & Z. Bogdanović (Eds.), Marketing and Smart Technologies. 9-18. Springer. [https://doi.org/10.1007/978-981-15-1564-4\\_2](https://doi.org/10.1007/978-981-15-1564-4_2)
- Sullivan, Danny (2018, enero 30). A reintroduction to Google's featured snippets. Blog Google. <https://blog.google/products/search/reintroduction-googles-featured-snippets/>
- Trippas, Johanne R.; Spina, Damiano; Thomas, Paul; Sanderson, Mark; Joho, Hideo; Cavedon, Lawrence (2020).

- Towards a model for spoken conversational search. // *Information Processing & Management*. 57:2, 102162. <https://doi.org/10.1016/j.ipm.2019.102162>
- Wu, Zhijing; Sanderson, Mark; Cambazoglu, B. Barla; Croft, W. Bruce; Scholer, Falk (2020). Providing Direct Answers in Search Results: A Study of User Behavior. // *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. 1635-1644. <https://doi.org/10.1145/3340531.3412017>
- Yu, Liyang (2014). *A developer's guide to the semantic web (Second)*. Springer Berlin, Heidelberg. <https://doi.org/10.1007/978-3-662-43796-4>
- Zhao, Yiming; Zhang, Jin; Xia, Xue; Le, Taowen (2019). Evaluation of Google question-answering quality. // *Library Hi Tech*. 37:2, 308–324. <https://doi.org/10.1108/LHT-10-2017-0218>
- Ziakis, Christos; Vlachopoulou, Maro; Kyrkoudis, Theodosios; Karagkiozidou, Makrina (2019). Important factors for improving Google search rank. // *Future Internet*. 11:2, 32. <https://doi.org/10.3390/fi11020032>

---

Enviado: 2023-03-31. Segunda versión: 2023-05-26.  
Aceptado: 2023-05-06.

---