ORIGIN AND WEIGHT OF NEWS MEDIA OUTLETS INDEXED ON GOOGLE NEWS:

An exploration of the editions from Brazil, Colombia, and Mexico

TANIA LUCÍA COBOS
Universidad Tecnológica de Bolívar, Cartagena de Indias – Colombia
ORCID: 0000-0001-8348-3631

ABSTRACT – Google News is a controversial news aggregation service owned by Google. By analyzing 3,738,375 news stories added in the Brazil, Colombia, and Mexico editions from January to March of 2015, 2,246 news media outlets were identified. In these editions, specific to their countries, large, popular, and traditional news media such as *G1* (under *Globo.com*), *Terra Brasil, El Universal, ElTiempo.com*, and *Caracol Radio* had the highest rates of news aggregation in their respective country edition. This meant that they had greater visibility as well as a higher probability of receiving traffic from it, above the other indexed news media outlets.
Key words: Google News. Media outlets. News aggregators. Infomediation. Latin America.

ORIGEM E PESO DOS MEIOS DE COMUNICAÇÃO INDEXADOS NO GOOGLE NEWS. UMA ANÁLISE EXPLORATÓRIA DAS EDIÇÕES DO BRASIL, COLÔMBIA E MÉXICO

RESUMO – O Google News é um serviço controverso de agregação de notícias de propriedade do Google. Ao analisar 3.738.375 notícias adicionadas nas edições Brasil,

Colômbia e México, de janeiro a março de 2015, foram identificados 2.246 meios de comunicação. Nessas edições, específicas de seus países, grandes mídias, populares e tradicionais como *G1* (sob *Globo.com*), *Terra Brasil, El Universal, ElTiempo.com* e *Caracol Radio* apresentaram as maiores taxas de agregação de notícias em suas respectivas edições de cada país. Isso significava que eles tinham maior visibilidade e maior probabilidade de receber tráfego deles, acima dos outros meios de comunicação indexados.
Palavras-chave: Google News. Meios de comunicação. Agregadores de notícias. Infomediação. Latinoamérica.

ORIGEN Y PESO DE LOS MEDIOS NOTICIOSOS INDEXADOS EN GOOGLE NEWS. UN ANÁLISIS EXPLORATORIO DE LAS EDICIONES DE BRASIL, COLOMBIA Y MÉXICO

RESUMEN – Google News es un servicio de agregación de noticias controvertido propiedad de Google. Al analizar 3.738.375 noticias agregadas en las ediciones de Brasil, Colombia y México, de enero a marzo de 2015, se identificaron 2.246 medios. En estas ediciones, específicas de sus países, grandes medios, populares y tradicionales como *G1* (bajo *Globo.com*), T*erra Brasil, El Universal, ElTiempo.com* y *Caracol Radio* presentaron las tasas más altas de agregación de noticias en su respectiva edición por país. Esto significaba que eran más visibles y más propensos a recibir tráfico de éste, por encima de otros medios noticiosos indexados.
Palabras clave: Google News. Medios noticiosos. Agregadores de noticias. Infomediación. Latinoamérica.

1 Introduction

Within the ecosystem of news distribution on the web, there are several news aggregators, among them Google News, owned by the American technology company Google. It has more than 70 international editions, which includes Brazil, Colombia, and Mexico, where news from media outlets of all kinds of sizes and trajectory, including Spanish – and Portuguese – speaking respectively, are indexed.

These news items come from a news website with free access or paywall, and the aggregator may or may not have the owner's licenses to do so. Within news aggregators' scope, Google News stands out for its frequent polemics with media companies around the world (e.g.: AEDE, renamed AMI in 2017, Copiepresse, Axel Springer, Danske Medier, News Corporation, Fonds pour L'Innovation Numérique de la Presse, among others), as well as for its frequent updates (e.g.: change of design and algorithm functionalities). In the countries mentioned above within this study, there are various

positions regarding this news aggregator. In Brazil's case, hundreds of press association members decided to self-exclude from it in protest for not receiving compensation or payment for the added content (ANJ, 2012). In the case of Colombia, a very popular national newspaper refused, for some time, to mention Google in its news in protest for the aggregation (Santos, 2014). Finally, there is an absence of positions in Mexico's case.

Upon previous research on Google News in different scenarios in the last two years from the United States, we find Huyen et al. (2019) with a measurement of political personalization on Google News and Google Search; Du (2018) with the role of this news aggregator in shaping the geography of investor information, and Nechushtai and Lewis (2018) about the personalization and recommendations on this news aggregator during the 2016 US presidential campaign; from India, Bansi (2019a, 2019b) with unequal cultivation of news on the international media landscape and the current scenario of international news in this service; from Germany, Haim et al. (2018) with the effect of personalization on the diversity in this service, and from Colombia, Cobos (2018) with the perceptions and experiences about this news aggregator among Latin-American chief editors. This literature review made it possible to determine that this object of study had not been approached from the focus that this investigation makes.

Firstly, this investigation aimed to identify what news media outlets are indexed in each one of the selected editions and from which countries they come because Google does not release this information. Secondly, to determine each of these news media outlets' weight, within each edition, viewed from the perspective of the amount of news added in total or "Captured news" (which allows us to know the visibility of the news media outlets). Moreover, news stories are counted only once – after deleting duplicate captured news – or "One-off news" (which allows us to know their role as providers) to determine if Google News has favorite news media outlets to indexed and make visible. This process is relevant since this *corpus* is explored in-depth and tries to understand how the Google News algorithm or StoryRank works. Also is essential to the Latin-American chief editors since they can know how others treat their news media outlets inside this news aggregator, the visibility that has their news inside it, and their role as a provider.

2 Media outlets and news

Among the types of content produced and distributed by newspapers, television, radio, and digital media are the news. With the arrival of the internet and the consolidation of digitalization and virtualization, broadcasts/transmissions/publications of the news are received simultaneously. However, these can also be customized and are more easily storable and shareable. Similarly, through computer tools, the large, broad audiences reduce their anonymity and identify their heterogeneity – their fragmentation into different niches –, so that it is possible to interact with it and to know in detail. Also, the feedback is immediate, and the management can be done on a "face to face" basis. This identification of the audience or users and its consequent profitability, that is to say, monetization of the audience, is a cause of confrontations between tech companies (e.g., Google) and mass media companies.

In the communicative paradigm, the news corresponds to the message (Vera, 1995). In journalistic terms, it is reports of recent events or events that appear or are disseminated through the media (Demers, 2005). More precisely, it is the general interest information for the target audience; therefore, criteria such as opportunity, impact, proximity, controversy, prominence, timeliness, and strangeness, determine what the news is (Potter, 2006).

From an economic perspective, which this research emphasizes, the news is the good or product that informational industries produce and disseminate, thus transforming it into merchandise. In other words, the news is a rare "good", understanding this rarity in the sense that its use is limited. Producers do not obtain it for free; it is expensive, and, for this purpose, a specific effort is required, in short, by an organization or company. The news is by definition "useful" because it is an idea, event, or current issue that interests the public. Its life span is short, and its owner sells it twice (first to advertisers and second to audiences). As a result, the news has no use-value for its owner – the media outlet – but an exchange value for them and a use-value for others (Torres, 1985; Murdoch, 2009; Picard, 2018).

Besides, its distribution cost is also high, where the speed and efficiency of this phase or moment also determine its perishable nature. Distribution requires a sophisticated and expensive organization to meet a massive and widely dispersed demand (Torres,

1985; Picard, 2018). In this aspect, given the advent of the Internet and the WWW, new forms of news content distribution appeared, such as news aggregators. Between them, including Google News, which concentrates the news information in one place, increases their exposure to new audiences and has made these costs cheaper, but all this has more profound implications than another.

3 News aggregators

A news aggregator is a free-access digital service, although consumers pay for it sometimes, available on both web and mobile platforms. This through algorithms or with human editors constantly explores and lists and groups by topic, headlines, and, sometimes, first lines of the full text, of a news story with an active link to the original content. These news items come from a news website with free access or a paywall, and the aggregator may or may not have the owner's licenses to do so. These usually belong to multinational technology companies, and their operation is not free from controversy.

Concerning the attainment of the news information that is added, Athey and Mobius (2012) mention that the aggregator, in a first case, does not make payments of any kind, nor does it maintain a formal relationship with the original authors of the news content. However, in very few cases, the aggregator can have a direct business relationship with a provider (e.g., Google News). In a second case, part of the added news content comes from contractual partners, that is, it is licensed (e.g., Yahoo! News).

In reference to the form of access, Legerén et al. (2011) mention that:

> There are basically two modes of access to information: on the one hand, the search engine of the aggregator, where the user must enter related words with the desired theme, as it happens in the use of a search engine (querying). On the other hand, it also offers a summary that the user can explore (browsing). This summary is a list ordered by the most recent items and that can also be reviewed by large sections (national, international, sports, entertainment, etc.). On this list, entries are grouped in such a way that the same news, but from different media, appear together. (Legerén et al., 2011, p.67).

In this sense, Madsen and Andsager (2011, p.2) affirm that aggregators "represent a type of news source that compiles articles from thousands of sources to create a unique news product – the

constantly changing list of headlines and story links on its homepage – without creating original content". Regarding their service, news aggregators affirm that this operation directs traffic to the media outlets' websites at no cost to them, which provides them with the public and the possibility of increasing profits through digital advertising and subscriptions. For the user, having a wide variety of sources allows them to read a range of news on the same subject and learn quickly what the day's topics are.

On the other hand, by facilitating the news's personalization and geolocation, the user receives the thematic contents of most interest to them. However, this premise's apparent simplicity has more profound implications and complexities. In that sense, Trielli and Diakopoulos (2019, p.3) say: "Algorithmic news curation still represents a concern for source diversity, since it can concentrate societal attention on a narrow range of privileged outlets".

On the other side, news aggregators implement the "infomediation" business model to monetize. This also called "cybermediation" or digital intermediation, appeared with the advent of Web 2.0 and the search by internet companies for profitability or monetization strategies. Its agents, the "infomediaries", "cybermediaries", or digital intermediaries, also came along with this. Hagel and Rayport (1997) coined the term "infomediary", formed by the words "information" and "intermediary", which is used to refer to websites that collect and organize large amounts of data and act as intermediaries between those who want the information and those who provide it. "Infomediaries" are in the information business, and they compete in their ability to capture and manipulate it in a way that adds value to their customers. They are not holders of the products or services that they broker in, and their profits are based on-screen advertising, on the number of pages that the user views, and on sales commissions (Bayonet, 2009).

In the field of news, Foster (2012) indicates that:

> Digital intermediaries can be defined as organizations which bring news content from third-party providers to consumers using a variety of digital software, channels, and devices. This sounds initially like a neutral and entirely positive role. But intermediaries can, through the way they carry out this activity and the charges they levy, exert significant influence over their suppliers and customers. (Foster, 2012, p.25).

Thus, news aggregators are one of the types of digital intermediaries, along with search engines, social media, and mobile app stores.

Cádima (2013) references that the new "infomediaries" or digital intermediaries are the largest existing American multinational technology companies, including Google. These manage the information and automate the process through their algorithms. In that sense, Del Águila et al. (2007, p.189) mention that "Google offers a news service to their users putting together sources such as *Reuters*, *Bloomberg* or the *Washington Times*. These new types of intermediaries are named cybermediaries".

In other words, although the news media outlets, both traditional and digital natives, are the ones who continue to produce the news, the work of distributing them on the web is no longer only in their hands. As "infomediaries", multinational internet technology companies (e.g., Google) have also assumed this function through news aggregators (e.g., Google News) by selecting, automatically or humanely, the news information. In this sense, Yoo (2011) manifests:

> The news aggregators will be especially powerful in this new stage news consumption. Not only do they take news out of traditional media's package, but also reassemble the news in their own package to provide it to the users. When we expose ourselves to news through news aggregators, we are presented with lists of top news and incidentally encounter the news that we did not intend to read. Even when we are seeking news intentionally, aggregators provide us with shortcuts to, or orders of news we should read. News aggregators are gatekeepers true to its name. (Yoo, 2011, p.7).

Of course, the mentioned above implies a previous step; the news selection becomes the first instance of the selection of the news media outlets that will provide it.

Finally, Lee and Chi (2015) ask:

> Are news aggregators friends or foes to news organizations? From an audience-centric perspective, data from this study suggest that news aggregators are no enemy to most news organizations. But whether news organizations can really benefit from this delicate "friendship" with news aggregators remains unclear. For one thing, individual news organizations lose their content exclusivity while news aggregators take advantage of the economies of scale and market power. (Lee & Chi, 2015, p.18).

4 Google News

Present in more than 70 countries and 30 languages, in its corporate presentation in 2019, Google News (n.d.a) defines itself as "Comprehensive up-to-date news coverage, aggregated from sources

all over the world", and its purpose is "to help everyone understand the world by connecting people with high-quality news from a variety of perspectives." In detail, it states that:

> Google News helps users stay up to date on the news that matters to them and the world, enabling them to dive deeper into current events as well as discover diverse content from a range of different publishers. Users can subscribe to specific news providers and topics, read content online or offline, and bookmark and share articles. Google News makes it easy for readers to find relevant and interesting content by personalizing what they see in the For You tab. The app uses machine learning to get better at recommending personalized content over time, adapting to users' habits and routines. (Google News, n.d.b)

Concerning the process of selection and cataloging of the news, Cobos (2017a) affirms that this declared that:

> Articles and the multimedia content are selected and classified through a computer system, which evaluates, among other things, the frequency with which a news item appears on the Internet and the sites where it is included. We also classify informative content according to a series of characteristics, such as the present, location, relevance, and diversity. Consequently, the news is classified independently from its political point of view or its ideology, and the user can choose from a wide variety of perspectives for a specific set of news. (Cobos, 2017a, p.77).

The Google News site for all editions comprises a menu with nine standard channels. That comes to the left, which is similar to the newspaper sections and facilitates thematic exploration by the user. Aggregated news items, according to their topics, are located in one of these tabs: "Top Stories", "World", "National" (the name of the country appears), "Business", "Science", "Technology", "Sports", "Entertainment", "Health", and "For you". Additionally, it has modules of specific functionality for users and media outlets, such as Weather, Fact Check, Beyond the Headlines, and Spotlight, which are not featured in all editions. It is essential to bear in mind that at anytime some of the aforementioned structural elements could be eliminated or replaced, also a new one incorporated, and, even an updated website design. This has happened previously because of the permanent evolution of Google News.

Concerning the aggregation of news, Google News declares that its website requires the news media outlets for original and transparent content and does not allow content that responds to 1. Ads and sponsored content; 2. Personal and confidential information; 3. Copyrighted content; 4. Sexually explicit content; 5. Graphic

violent content; 6. Hateful content; 7. Medical advice; 8. Dangerous and illegal activities; 9. Harassment and cyberbullying; 10. Deceptive practices; and 11. Spam and malware.

In other words, Google News is a news site based on algorithms that aggregate news headlines from thousands of global news sources. We do not know precisely how many there are, since Google News does not make public a list that makes their names known, however, we can identify some of them when we browse the site. It does not have human editors and does not store all the articles on its servers. Instead, it presents a series of snippets that contain the name of the news media outlet, a thumbnail, and the hyperlinked headline that leads to the source of the article.

The StoryRank algorithm that operates on Google News determines the news headlines' position and visibility into the service. This algorithm is updated regularly. Its way to operate is a corporate secret; the news media outlets do not know precisely how it works. They are not sure if this benefits them a little or a lot. StoryRank considers "originality, freshness, quality, expertise of source and whether a lot of other sources around the Web are pointing to a particular article" (Machlis, 2009). A team of reviewers decides what news media outlet is tracked (Kramer, 2003) and it takes into account: 1. The volume of production from a news source; 2. Length of articles; 3. "The importance of coverage by the news source"; 4. The "Breaking News Score"; 5. Usage Patterns; 6. The "Human opinion of the news source"; 7. Audience and traffic; 8. Staff size; 9. Numbers of news bureaus; 10. The number of "original named entities"; 11. The "breadth" of the news source; 12. The global reach of the news sources; and 13. Writing style (Filloux, 2013). For Google, "the computer is unbiased and, therefore, better able to serve a mix of views because it does not recognize bias" (Carlson, 2007, p.1020). However, this is very debatable given that algorithms present voluntary and involuntary biases and influences according to their developers' and owners' interests.

The practice of Google News carried out through automation without asking for permission, in other words, its "infomediary" role, has been strongly criticized by the news industry in different geographies around the planet. The media industry referred to Google News as a "parasite", a "plagiarist", a "content kleptomaniac", a "digital vampire" sucking "newspapers' blood" (Lee & Chyi, 2015, p.4). All this because the news aggregator "stealing" their content

and unfairly profiting on their labor" (Chyi et al., 2016, p.2).

For copyright disputes, some editions have been closed (e.g., Spain) and others have had to be modified to adapt to local laws (e.g., Germany) (Athey et al., 2017). In others, they signed agreements to finance initiatives related to the press (e.g., France); in others have never been launched (e.g., Denmark), and in others, some news media outlets chose self-exclusion (e.g., Brazil) (Cobos, 2014). It has also had well-known disputes with media magnates, such as Rupert Murdoch of News Corporation or big media companies as Grupo PRISA.

This news aggregator and the news media outlets maintain a "coopetition" (competition and cooperation) relationship, where they need each other, but also, they compete against each other for audience (Lee & Chyi, 2015). In some countries, the prominent chief editors have resorted to their government to legislate in their favor as a way to equilibrate this "coopetition" that appears to benefit more to Google News and by extension to Google.

On the other hand, also Google News implements the practice of deep linking. It consists of linking to a specific piece of content, generally searchable or indexable, within a website. For example, link to http://www.media.com/section/news1 instead of http://www.media.com. In other words, it means linking to specific web pages of a website avoiding the passage through the home page where the most expensive advertising is usually located. That reduces this space's value, which may also entail another problem, the devaluation of the news media outlet brand and the brand appreciation of the news aggregator. The user may pay little or no attention to who the provider of the news is, or consider himself/herself satisfied with the headlines he/she has read and not click to read the complete news, producing a substitution effect and establishing unfair competition (Edo et al., 2018). In this direction Athey et al. (2017) state:

> Our findings also highlight that while large publishers may not see an effect in overall page views as a result of aggregators, they may lose traffic to their home pages, as well as their role in curating news, as readers read articles referred by Google News at the expense of articles referred by their own home pages (where newspapers monetize the home pages much better than articles). If readers do not pay attention to the identity of the publisher when they read articles on Google News, then the large publishers may lose their incentives to maintain a reputation for quality, and consumers may be less willing to subscribe to the publisher or use the publisher's mobile application. (Athey et al.,2017, p.27).

5 Methods

An investigation of a quantitative with an exploratory and descriptive nature was carried out. Exploratory scope studies "are carried out when the objective is to examine a little-studied research issue or problem, of which there are many doubts or which none has addressed before" (Hernández et al., 2010, p.79). Descriptive studies seek "to specify properties, characteristics and important features of any phenomenon that is analyzed" (Hernández et al., 2010, p.80). The research did not have a comparative approach since what was to make the first diagnosis of each edition's news ecosystem's structure, that is, to identify what was in each one.

In this research's specific case, the aim was to determine which news media outlets were indexed and where they came from since this is a piece of unknown information. As well as, the frequency of news aggregation of these or its weight within the news ecosystem (their visibility and provider role) to determine if Google News has a favorite news media outlet to index. For this, we took the editions from Google News Brazil, Colombia, and Mexico. The sample was built for convenience, choosing these countries for cultural proximity reasons.

About the editions selected and their languages, the edition from Mexico, in Spanish and under the subdomain news.google.com.mx, was launched in December 2004 (Krantz, 2004). The edition from Colombia, also in Spanish, was launched in January 2006 under the subdomain news.google.com.co (Van Dijk, 2006). We should note that news media's indexing in these countries' editions had already been done since September 2003 when Google opened Google News Spain under the subdomain news.google.es, which covered more than 700 news sources (El Mundo, 2003). In Brazil's case, it was launched in Portuguese in November 2005 under the subdomain news.google.com.br and claimed to index more than 1.500 news sources in such a language (Fonseca, 2005).

Also, this research used digital methods. This refers to the methodology that uses software belong to the field of computing for the capture, visualization, and analysis of data on phenomena that occur on the internet. All of this from the perspective of Social Sciences, and in this specific study, from the communication sciences and journalism. In this sense, Rieder (2013) affirms:

Research methods using software to capture, produce, or repurpose digital data in order to investigate different aspects of the Internet have been used for well over a decade. Datasets can be exploited to analyze complex social and cultural phenomena and digital methods have a number of advantages compared to traditional ones: advantages regarding cost, speed, exhaustiveness, detail, and so forth, but also related to the rich contextualization afforded by the close association between data and the properties of the media (technologies, platforms, tools, websites, etc.) they are connected with; data crawling necessarily engages these media through the specifics of their technical and functional structure and then produces data that can provide detailed views of the systems and the use practices they host. (Rieder, 2013, p.1).

One of these digital methods is web scraping. This is a data extraction technique or a computing technique for obtaining data by "scraping" web pages, which focuses on the transformation of data without structure, such as HTML format, which can be analyzed in a spreadsheet such as Microsoft Excel. To do it, a scraper bot was developed in PHP in collaborative work between a developer and the researcher. It tracked the editions of Google News Brazil, Colombia, and Mexico from January 1, 2015, to March 31, 2015. This period was chosen for convenience. During this, the scraper bot visited the different channels (e.g., Sports, Health, Business…) of these three editions every hour and captured the variables: edition, channel, headline, URL, source, and date of capture, from every news item that it found and stored the data collected on an external database in MySQL. It is noteworthy that the tracking was limited only to the text; photographs and videos were discarded.

We exported captured data to spreadsheets in .xlsx format, and we then proceeded first to make a manual review of it. This process detected some anomalies in the function of the StoryRank algorithm. These were related to the URL of the source and the name of the source. This probably due to human error at the initial moment of the insertion of the identifying data (e.g., the URL does not match with the source, the name of the source is the same for different news media outlets; as well as spelling mistakes, among others). Once such anomalies were corrected, the news media outlets or sources indexed were identified. From these news stories, it was possible to identify what news media outlets they came from and how many news items corresponded to each. We carried out this analysis on Microsoft Excel by implementing functionalities such as pivot tables, duplicate removal, sort, and filter data depending on what we want to see. Data is available at Dipòsit Digital de Documents de la UAB, see Cobos (2017b).

6 Results

The scraper bot determined the capture of 3,738,375 news stories, of which 1,230,539 came from Brazil, 1,222,320 from Colombia, and 1,285,516 from Mexico over the ninety days. From the analyses of this data, 839 media outlets were counted and identified in the Brazil edition, 1,216 in the Colombia edition, and 1,259 in the Mexico edition. Keep in mind that Colombia and Mexico's editions share some news media outlets for language reasons. Such values in these editions could be significantly affected by the closure of Google News Spain in December 2014 (Gingras, 2014) and the elimination (or, at least, almost all) of the Spaniard news media outlets from the service.

In Brazil's case, this number also could have been affected by the self-exclusion since June 2011 of many Brazilian newspapers that are members of the *Associação Nacional de Jornais* (ANJ) for payment disputes or compensations demanded by them from Google for aggregation services (ANJ, 2012). Moreover, one must note that the number of news media outlets and their news stories added in the three cases during the period of data extraction do not correspond, at any time, to absolute and constant data; that is, these can change with the pass of the time.

The scraper bot did not discriminate against the news. That is, if the news was published several hours in one or more channels, even from one day to another, it was captured and identified as "Captured news"; therefore, one could duplicate a news story one or several times. After removing duplicated news from "Captured news", those that remained, that is, the news counted only once, were identified as "One-off news". The term "Captured news" alludes to greater visibility. A more considerable amount of captured news implied greater visibility due to more prolonged exposure and, thus, a higher likelihood of receiving traffic from Google News; therefore, it also implied a greater degree of exposure within the aggregator for the news media outlet. The term "One-off news", on the other hand, alludes to the amount of news provided by the media outlet to the aggregator or the actual number of news items from the media outlet that Google News added – the greater the amount of one-off news, the larger the media outlet was as a provider, and vice versa, the lesser the amount of one-off news, the smaller the media outlet was as a provider.

The percentage of captured news and the percentage of one-off news corresponding to each indexed news source in each edition they obtain the percentage change rate in each case and with this determine the classification of the news media outlets in different visibility groups (very high visibility, high visibility, intermediate visibility, low visibility, and very low visibility) and their supplier role (significant supplier, medium supplier, small supplier, and micro supplier).

Table 1 – Region of origin of news media outlets in the selected Google News editions

By edition	News media outlets	From Ibero-American countries	From Non-Ibero-American countries	Indeterminate origin
Brazil	839 (100%)	807 (96.1%)	32 (3.8%)	0 (0%)
Colombia	1,216 (100%)	1,083 (89%)	128 (10.5%)	5 (0.4%)
Mexico	1,259 (100%)	1,138 (90.3%)	116 (9.2%)	5 (0.3%)

Source: own elaboration

Table 1 shows the distribution of news media outlets' origin in the three editions. It is noteworthy that between 89% and 96% of them are of Ibero-American origin. They are from Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Spain, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Portugal, Puerto Rico, Dominican Republic, Uruguay, and Venezuela.

Table 2 – Country of origin of the news media outlets on Google News Brazil

Position	Country of origin	News media outlets	Captured news	One-off news
1	Brazil	675 (80.4%)	1,095,017 (88.9%)	191,670 (88.4%)
2	Portugal	132 (15.7%)	106,618 (8.6%)	20,685 (9.5%)
3	USA	7 (0.8%)	600 (0.05%)	113 (0.05%)
4	Angola	5 (0.6%)	1,155 (0.09%)	206 (0.09%)

5	Russia	3 (0.3%)	272 (0.02%)	75 (0.03%)
6	France	2 (0.2%)	693 (0.06%)	178 (0.08%)
7	UK	2 (0.2%)	20,687 (1.6%)	2,609 (1.2%)
8	Switzerland	2 (0.2%)	494 (0.04%)	158 (0.07%)
9	Germany	1 (0.1%)	2,099 (0.1%)	577 (0.2%)
10	China	1 (0.1%)	2,197 (0.1%)	295 (0.1%)
11	Ireland	1 (0.1%)	1 (0.0008%)	1 (0.0004%)
12	Italy	1 (0.1%)	34 (0.002%)	9 (0.004%)
13	Japan	1 (0.1%)	26 (0.002%)	7 (0.003%)
14	Luxembourg	1 (0.1%)	63 (0.01%)	18 (0.008%)
15	Macau	1 (0.1%)	23 (0.001%)	7 (0.003%)
16	Mozambique	1 (0.1%)	418 (0.03%)	99 (0.04%)
17	Netherlands	1 (0.1%)	36 (0%)	3 (0.001%)
18	São Tomé and Príncipe	1 (0.1%)	23 (0.002%)	5 (0.002%)
19	Vatican	1 (0.1%)	83 (0.006%)	22 (0.01%)
	TOTAL	**839 (100%)**	**1,230,539 (100%)**	**216,737 (100%)**

Source: own elaboration

In the case of Google News Brazil in Table 2, we found that 80.4% of the indexed news media outlets came from Brazil, followed by Portugal with 15.7%. The remaining 3.8% came from 17 different countries, including other Portuguese-speaking nations such as Angola, Macau, Mozambique, and São Tomé and Príncipe, and other 13 foreign countries producing news in the Portuguese language. We also observed that 88.9% of the captured news in this edition corresponded to 80.4% from Brazilian indexed news media outlets.

In turn, they obtained 88.4% of the news items counted only once or one-off news in their role as providers.

Table 3 – Country of origin of the news media outlets on Google News Colombia

Position	Country of origin	News media outlets	Captured news	One-off news
1	Mexico	340 (27.9%)	108,446 (8.8%)	23,884 (13.6%)
2	Argentina	244 (20%)	63,175 (5.1%)	14,461 (8.2%)
3	Colombia	111 (9.1%)	770,663 (63%)	74,240 (42.3%)
4	USA	80 (6.5%)	34,136 (2.7%)	7,205 (4.1%)
5	Peru	60 (4.9%)	54,842 (4.4%)	12,609 (7.1%)
6	Chile	54 (4.4%)	9,388 (0.7%)	2,571 (1.4%)
7	Venezuela	54 (4.4%)	52,943 (4.3%)	11,741 (6.6%)
8	Cuba	35 (2.8%)	18,237 (1.4%)	3,405 (1.9%)
9	Spain	30 (2.4%)	5,762 (0.4%)	1,693 (0.9%)
10	Bolivia	24 (1.9%)	11,532 (0.9%)	2,587 (1.4%)
11	Dominican Republic	21 (1.7%)	1,510 (0.1%)	272 (0.1%)
12	Uruguay	20 (1.6%)	6,029 (0.4%)	1,536 (0.8%)
13	Ecuador	19 (1.5%)	12,326 (1%)	2,563 (1.4%)
14	Panama	13 (1%)	3,689 (0.3%)	789 (0.4%)
15	Paraguay	13 (1%)	6,005 (0.4%)	1,409 (0.8%)
16	Costa Rica	11 (0.9%)	8,328 (0.6%)	1,629 (0.9%)
17	El Salvador	9 (0.7%)	2,930 (0.2%)	724 (0.4%)
18	Guatemala	8 (0.6%)	2,444 (0.1%)	701 (0.3%)

19	UK	8 (0.6%)	6,754 (0.5%)	1,283 (0.7%)
20	Honduras	7 (0.5%)	12,076 (0.9%)	3,139 (1.7%)
21	Italy	6 (0.4%)	1,656 (0.1%)	416 (0.2%)
22	France	5 (0.4%)	2,728 (0.2%)	707 (0.4%)
23	Puerto Rico	5 (0.4%)	1,210 (0.09%)	239 (0.1%)
24	China	4 (0.03%)	8,215 (0.6%)	1,530 (0.8%)
25	Nicaragua	4 (0.03%)	3,427 (0.2%)	922 (0.5%)
26	India	3 (0.02%)	5 (0.0004%)	3 (0.001%)
27	Switzerland	3 (0.02%)	2,745 (0.2%)	759 (0.4%)
28	Australia	2 (0.01%)	3 (0.0002%)	2 (0.001%)
29	Canada	2 (0.01%)	56 (0.004%)	11 (0.006%)
30	Pakistan	2 (0.01%)	2 (0.0001%)	2 (0.001%)
31	Singapore	2 (0.01%)	2 (0.0001%)	2 (0.001%)
32	Germany	1 (0.08%)	511 (0.04%)	125 (0.07%)
33	Belgium	1 (0.08%)	16 (0.001%)	2 (0.001%)
34	Brazil	1 (0.08%)	22 (0.001%)	2 (0.001%)
35	South Korea	1 (0.08%)	63 (0.005%)	22 (0.01%)
36	Greece	1 (0.08%)	9 (0.0007%)	3 (0.001%)
37	Iran	1 (0.08%)	604 (0.04%)	173 (0.09%)
38	Israel	1 (0.08%)	40 (0.003%)	12 (0.006%)
39	Japan	1 (0.08%)	50 (0.004%)	2 (0.001%)
40	New Zealand	1 (0.08%)	3 (0.0002%)	1 (0.0005%)
41	Czech Republic	1 (0.08%)	12 (0.0009%)	4 (0.002%)
42	Russia	1 (0.08%)	410 (0.03%)	115 (0.06%)

43	Vietnam	1 (0.08%)	121 (0.009%)	29 (0.01%)
	Indeterminate origin	5 (0.4%)	9,195 (0.7%)	1,978 (1.1%)
	TOTAL	**1,216 (100%)**	**1,222,320 (100%)**	**175,502 (100%)**

Source: own elaboration

In the case of Google News Colombia, Table 3 shows that only 9.1% of the indexed news media outlets came from Colombia, ranked third, far from Mexico's first position with 27.9%, and the second position grabbed by Argentina with 20%. The remaining 42.8% came from 40 different countries, 18 of them Spanish-speaking nations and 22 foreign countries producing news in the Spanish language. We can also observe that 63% of the captured news in this edition corresponded to 9.1% of Colombian indexed news media outlets. In turn, in their role as providers, they obtained 42.3% of the news items counted only once, that is, one-off news, with the Mexican and Argentine media lagging far behind.

Table 4 – Country of origin of the news media outlets on Google News Mexico

Position	Country of origin	News media outlets	Captured news	One-off news
1	Mexico	460 (36.5%)	1,007,605 (78.3%)	192,381 (70.7%)
2	Argentina	245 (19.4%)	46,729 (3.6%)	12,585 (4.6%)
3	USA	83 (6.5%)	43,399 (3.3%)	10,485 (3.8%)
4	Colombia	68 (5.4%)	28,851 (2.2%)	10,237 (3.7%)
5	Chile	53 (4.2%)	6,345 (0.4%)	2,185 (0.8%)
6	Peru	50 (3.9%)	34,829 (2.7%)	10,229 (3.7%)
7	Venezuela	47 (3.7%)	30,436 (2.3%)	9,055 (3.3%)
8	Cuba	34 (2.7%)	11,928 (0.9%)	2,979 (1%)

9	Spain	29 (2.3%)	3,822 (0.2%)	1,177 (0.4%)
10	Bolivia	22 (1.7%)	7,926 (0.6%)	2,054 (0.7%)
11	Uruguay	21 (1.6%)	5,243 (0.4%)	1,484 (0.5%)
12	Ecuador	18 (1.4%)	6,234 (0.4%)	1,766 (0.6%)
13	Dominican Republic	18 (1.4%)	523 (0.04%)	187 (0.06%)
14	Panama	15 (1.1%)	1,160 (0.09%)	504 (0.1%)
15	Paraguay	12 (0.9%)	3,849 (0.2%)	1,159 (0.4%)
16	Costa Rica	11 (0.8%)	6,584 (0.5%)	1,612 (0.5%)
17	El Salvador	10 (0.7%)	2,075 (0.1%)	681 (0.2%)
18	Guatemala	7 (0.5%)	1,914 (0.1%)	637 (0.2%)
19	Honduras	7 (0.5%)	10,746 (0.8%)	3,417 (1.2%)
20	Italy	6 (0.4%)	1,456 (0.1%)	433 (0.1%)
21	Nicaragua	5 (0.3%)	1,897 (0.1%)	609 (0.2%)
22	Puerto Rico	5 (0.3%)	821 (0.06%)	291 (0.1%)
23	UK	5 (0.3%)	2,433 (0.1%)	847 (0.3%)
24	China	4 (0.3%)	6,366 (0.4%)	1,474 (0.5%)
25	France	4 (0.3%)	2,319 (0.1%)	645 (0.2%)
26	Switzerland	3 (0.2%)	562 (0.04%)	261 (0.09%)
27	Canada	2 (0.1%)	55 (0.004%)	13 (0.004%)
28	Germany	1 (0.07%)	336 (0.02%)	115 (0.04%)
29	Brazil	1 (0.07%)	1 (0.00007%)	1 (0.0003%)
30	South Korea	1 (0.07%)	51 (0.003%)	17 (0.006%)

31	Iran	1 (0.07%)	535 (0.04%)	173 (0.06%)
32	Israel	1 (0.07%)	23 (0.001%)	10 (0.003%)
33	Japan	1 (0.07%)	23 (0.001%)	1 (0.0003%)
34	Czech Republic	1 (0.07%)	5 (0.0003%)	2 (0.0007%)
35	Russia	1 (0.07%)	403 (0.03%)	130 (0.04%)
36	Vietnam	1 (0.07%)	82 (0.006%)	29 (0.01%)
37	Luxembourg	1 (0.07%)	17 (0.001%)	1 (0.0003%)
	Indeterminate origin	5 (0.3%)	7,933 (0.6%)	2,224 (0.8%)
	TOTAL	**1,259 (100%)**	**1,285,516 (100%)**	**272,090 (100%)**

Source: own elaboration

Finally, in the case of Google News Mexico, as seen in Table 4, we found that 36.5% of the indexed news media outlets came from Mexico, followed in second place by Argentina with 19.4%. The remaining 44% came from 35 different countries, including 18 Spanish-speaking nations and 17 foreign countries producing news in the Spanish language. We can also observe that 78.3% of the captured news in this edition corresponded to 36.5% of Mexican indexed news media outlets. In turn, they achieved 70.7% of news items counted only once or one-off news in their role as providers.

Table 5 – Media outlets and captured news items on Google News Brazil

Range of captured news	Percentages of captured news	Number of news media outlets	Percentage change rate	Group	News media outlets in the range
More than 100,001	8.32%	1	8.32	Very high visibility	G1*
70,001 – 75,000	5.88%	1	5.88	High visibility	Terra Brasil
35,001 – 40,000	2.84%	1	2.84	Intermediate visibility	Jornal do Brasil, EXAME.com, Diário do Grande ABC
25,001 – 30,000	4.36%	2	2.18		
15,001 – 20,000	10.97%	8	1.37	Low visibility	veja.com, Paraná-Online, Boa Informação, R7, globoesporte.com, Correio da Bahia, Correio do Estado, Público.pt
10,001 – 15,000	17.06%	18	0.94	Very low visibility	Others
5,001 – 10,000	11.84%	22	0.53		
**0 – 5,000	38.69%	786	0.04		
1,230,540	**100%**	**839**			

* Under *Globo.com* | ** The 0 value corresponds to *Diário Catarinense*, *Diário Gaúcho* and *Zero Hora*

Source: own elaboration

Table 5 gives us a detailed view of the news media outlets' visibility in the Google News Brazil edition based on the amount of captured news. We found that the news website *G1* (under *Globo.com*) was the one with very high visibility, 8.3%, as more than 100,001 of the total number of captured news items came from it. Secondly, the *Terra Brasil* website also ranked as having high visibility, 5.8%, as between 70,001 and 75,000 of the total number of captured news items came from it.

Table 6 – Media outlets and captured news on Google News Colombia

Range of captured news	Percentages of captured news	Number of news media outlets	Percentage change rate	Group	News media outlets within the range
85,001 – 90,000	7.3%	1	7.34	Very high visibility	ElTiempo.com
75,001 – 80,000	12.7%	2	6.35	High visibility	Caracol Radio, ElEspectador.com
55,001 – 60,000	4.7%	1	4.75	Intermediate visibility	El Colombiano, Radio Santa Fe, Vanguardia Liberal
50,001 – 55,000	4.1%	1	4.10		
40,001 – 45,000	3.2%	1	3.29		
25,001 – 30,000	11.6%	5	2.33	Low visibility	Pulzo, El Universal – Colombia, I Heraldo (Colombia), Terra Colombia, Semana.com, RCN Radio (Comunicado de prensa) (blog), W Radio, El Universal, Entorno-Inteligente, LaRepública.com.co, Terra Perú, El País – Cali Colombia, Portafolio.co, Blu Radio, Prensa Latina
20,001 – 25,000	1.6%	1	1.67		
15,001 – 20,000	12.4%	9	1.38		
10,001 – 15,000	4.6%	5	0.93	Very low visibility	Others
5,001 –10,000	11.9%	22	0.54		
*0 – 5,000	25.3%	1,168	0.02		
1,222,320	**100%**	**1,216**			

* The value 0 corresponds to *elEconomista.es*

Source: own elaboration

Table 6 offers us a detailed vision regarding the news media outlets' visibility in the Google News Colombia edition based on the amount of captured news. We found that newspaper *ElTiempo.com* came up as having very high visibility, 7.3%, since between 85,001 and 90,000 of the total number of the captured news stemmed from there. Secondly, radio station *Caracol Radio* and newspaper *ElEspectador. com* came up with a high visibility rate, 12.7%, as between 75,001 and 80,000 of the total number of captured news originated from these.

Table 7 – Media outlets and captured news on Google News Mexico

Range of captured news	Percentage of captured news	Number of news media outlets	Percentage change rate	Group	News media outlets within the range
80,001 – 85,000	6.5%	1	6.55	Very high visibility	El Universal
45,001 –50,000	3.8%	1	3.82	High visibility	Vanguardia.com.mx, Milenio.com, Diario Digital Juárez, RadioFórmula
40,001 – 45,000	7.0%	2	3.50		
35,001 – 40,000	3.0%	1	3.04		
30,001 – 35,000	7.3%	3	2.44	Intermediate visibility	El Economista, El Diario de Yucatán, Terra México, El Financiero, La Crónica de Hoy
25,001 – 30,000	4.2%	2	2.12		
20,001 – 25,000	5.4%	3	1.82	Low visibility	Excélsior, Noticieros Televisa, Informador.com.mx, El Siglo de Torreón, Univisión, La Jornada en linea
15,001 – 20,000	3.9%	3	1.31		

10,001 – 15,000	2.5%	3	0.84	Very low visibility	Others
5,001 – 10,000	21.3%	38	0.56		
*0 – 5,000	34.7%	1,202	0.02		
1,285,516	**100%**	**1,259**			

* The 0 value corresponds to *elEconomista.es*
Source: Own elaboration

Table 7 offers us a detailed vision regarding the visibility that the news media outlets accomplished in the Google News Mexico edition based on the amount of captured news. We found that newspaper *El Universal* grabbed very high visibility, 6.5%, since between 80,001 to 85,000 of the total number of captured news came from this one. Newspapers *Vanguardia.com.mx*, *Milenio.com*, *Diario Digital Juárez*, and radio station *RadioFórmula* came in second in terms of high visibility, 13.8%, as between 35,001 to 50,000 (covering three ranges) of the total amount of captured news came from the outlets mentioned above.

Table 8 – Media outlets and one-off news on Google News Brazil

Range of one-off news	Percentages of one-off news	Number of one-off news items	Percentage change rate	Group	News media outlets within the range
More than 10,001	12.1%	2	6.07	Big suppliers	Terra Brasil, G1*
				Medium suppliers	Jornal do Brasil, Diário do Grande ABC
6,001 –7,000	6.0%	2	3.02		
5,001 – 6,000	2.4%	1	2.46	Small suppliers	R7, Paraná-Online, EXAME.com, globoesporte.com, DCI, Reuters Brasil
4,001 – 5,000	4.4%	2	2.20		
3,001 – 4,000	4.9%	3	1.64		

2,001 – 3,000	13.7%	12	1.14		
1,001 – 2,000	15.9%	25	0.63	Micro suppliers	Others
**0 – 1,000	40.2%	792	0.05		
216,737	**100%**	**839**			

* Under *Globo.com* | ** The 0 value corresponds to *Diário Catarinense, Diário Gaúcho* and *Zero Hora*

Source: own elaboration

Table 8 provides us with a detailed view regarding the role of suppliers of the news media outlets in the edition of Google News Brazil. Here we find that the outlets mentioned above *Terra Brasil* and *G1* (under *Globo.com*), both Brazilian, were also the most significant news providers for this edition. They stood at the rank of more than 10,001 one-off news items, with 12.1% of the total.

Table 9 – Media outlets and one-off news on Google News Colombia

Range of one-off news	Percen-tages of one-off news	Number of one-off news items	Perce-ntage change rate	Group	News media outlets within the range
9,001 – 10,000	5.2%	1	5.29	Big suppliers	Caracol Radio, ElTiempo.com
8,001 – 9,000	4.8%	1	4.88		
6,001 – 7,000	3.6%	1	3.65	Medium suppliers	ElEspectador.com, El Colombiano, Terra Colombia
5,001 – 6,000	6.2%	2	3.11		
4,001 – 5,000	2.4%	1	2.41	Small suppliers	Terra Perú, EntornoInteligente, Vanguardia Liberal, El Universal
3,001 – 4,000	5.8%	3	1.96		
2,001 – 3,000	11.9%	8	1.49	Micro suppliers	Others
1,001 – 2,000	19.2%	23	0.83		
*0 – 1,000	40.3%	1,176	0.03		
175,502	**100%**	**1,216**			

* The 0 value corresponds to *elEconomista.es*

Source: own elaboration

Table 9 offers us a detailed view concerning the role as suppliers that the news media outlets had in the edition of Google News Colombia. Here we find that the outlets mentioned above, *Caracol Radio* and *ElTiempo.com*, both Colombian, were also the most significant news providers for this edition. The case of *ElEspectador.com* was ranked as a medium-sized provider as it ranged from 9,001 to 10,000 one-off news items, or 5.2% of their total number.

Table 10 – Media outlets and one-off news on Google News Mexico

Range of one-off news	Percentage of one-off news	Number of news media outlets	Percentage change rate	Group	News media outlets within the range
More than 10,001	4.8%	1	4.89	Big suppliers	El Universal
9,001 – 10,000	3.3%	1	3.33	Medium suppliers	Radio Fórmula, Diario Digital Juárez, Milenio.com, Terra México, Vanguardia.com.mx
8,001 – 9,000	6.1%	2	3.08		
7,001 – 8,000	2.5%	1	2.57		
6,001 – 7,000	2.4%	1	2.45		
5,001 – 6,000	3.7%	2	1.87	Small suppliers	El Economista, El Diario de Yucatán, Excélsior, Informador.com.mx, La Crónica de Hoy, El Financiero, Noticieros Televisa, El Siglo de Torreón, Terra Perú, La Jornada en linea
4,001 – 5,000	6.4%	4	1.61		
3,001 – 4,000	4.9%	4	1.24		
2,001 – 3,000	6.1%	7	0.87	Micro suppliers	Others
1,001 – 2,000	21.3%	40	0.53		
*0 – 1,000	37.9%	1,196	0.03		
272,090	**100%**	**1,259**			

* The 0 value corresponds to *elEconomista.es*

Source: own elaboration

Table 10 provides us with a detailed view regarding the news media outlets' providers in the edition of Google News Mexico. Here we find that the newspaper *El Universal* mentioned above, of Mexican origin, was also a great news provider for this edition since it stood at a rank of over 10,001 one-off news items, or 4.8% of the total amount.

As we see, the web portals *G1* (under *Globo.com*) and *Terra Brasil* on Google News Brazil, newspaper *ElTiempo.com* and radio station *Caracol Radio* on Google News Colombia and newspaper *El Universal* in Google News Mexico, all of them originating in each edition's countries, were the news media outlets that achieved a significant weight in the news ecosystem of Google News, at least in the three editions studied, standing out above the more than 2,000 news media outlets identified in total.

7 Concluding discussion

The research aimed to identify what news media outlets are indexed in each one of the selected editions and from which countries they come. As previously mentioned, we identified more than 2,000 news media outlets, and all almost of them come from Ibero-American countries. This research shows that Google News has a heterogeneous news ecosystem. We can find traditional and native digital media, large and popular or small and little known, especially press (newspapers and magazines) and also television, radio, and news agencies, in addition to web portals, blogs, and theme-based websites. In other words, this variety of news media outlets, and, at first sight, diversity of news, is essential for Google because it allows for a large and varied inventory of news for this information service.

This research evidenced the technical relationship between diversity and variety existing in Google News. The small, little known or new news media outlets are more abundant than the large, traditional and popular news media outlets. However, their economic conditions, reflected in their volume of news production, are lower compared to the second and even, with greater possibilities of ceasing such production permanently or temporarily at any time. This situation affects the aggregator that always needs news. In conclusion, all news media outlets

have a place – independent of their ideological orientations. However, there is a tendency to add news from certain media outlets over others because of their news production capacity to guarantee an inventory and update rate and thus, the service's optimal functioning.

This wide variety of news media outlets available has allowed for the extraction of other perceptions about the "infomediary" work of Google News. Indeed, it does not receive a direct income from advertising, nor does it charge for positioning, its "infomediation" business model is more subtle. Its role as a news distribution system and sending traffic to indexed news media outlets is secondary. Its main role is to be an effective audience concentrator interested only in news for the utility of all interests of this technology company (e.g.: know the user). Moreover, given the disinformation and misinformation that we live in the XXI century's second decade, its importance could be strengthened. Even its consumption could be increase since it is a service where users know that the aggregate information is trustworthy and reliable or at least a large part.

During the period studied, the predominant indexation of news media outlets in the three editions came from Brazil (680), Mexico (464), Argentina (276), Portugal (240), and Colombia (111). However, keep in mind that it is impossible to determine how much this Top 5 was affected by the Google News Spain's events in December 2014 and with the *Associação Nacional de Jornais* (ANJ) in June 2011 on Google News Brazil. Despite more than 2,000 news media outlets identified, only four were very relevant: the web portals *G1* (under *Globo.com*) and *Terra Brasil* on Google News Brazil, newspaper *ElTiempo.com* and radio station *Caracol Radio* on Google News Colombia and newspaper *El Universal* in Google News Mexico, all of them originating in the countries of each edition. This study shows a marked preference of StoryRank for adding news from traditional and popular news media outlets. Moreover, large news producers, probably, as previously mentioned, contribute due to the need to have an inventory that is updated continuously. Nevertheless, also causing a bias because it reinforces the dominance of these news media outlets and, therefore, gives more weight to their ideological lines within the ecosystem by massively aggregating their news, which also implies a new gatekeeping level.

The research also allowed us to determine that Google News Brazil is a local edition, very typical of its country. The indexation of Brazilian news media outlets was predominant, and the predominance of news aggregation thereof mainly with *G1* (under *Globo.com*) and *Terra Brasil*. The above happened equally in the case of Google News Mexico, also a local edition, with the predominance of Mexican news media outlets' indexation, and the predominant aggregation of news from such outlets, which mainly highlights *El Universal*. Although it is also local, the edition of Google News Colombia is the least local of the three studied, since Mexican and Argentine news media's indexation predominated over Colombian ones. An explanation for this could lie in the territorial extension, as Colombia is a smaller country than Mexico and Argentina; even so, it was the Colombian news media outlets that had a predominant aggregation of news, whereby chiefly *ElTiempo.com* and *Caracol Radio* stood out. This study deduced that each edition is designed as a very local product to attract and concentrate local audiences.

In the line with the second aim, the frequency of news aggregation or its weight within the news ecosystem (their visibility and provider role). These four were the news media outlets that achieved a significant weight in Google News's news ecosystem since they reached the highest news aggregation rates for both "Captured news" and "One-off news". Reviewing who they are, *G1* (under *Globo.com*) is a news website owned by Grupo Globo, a Brazilian media conglomerate founded in 1925 and considered one of the largest in Latin America and the world. *Terra Brasil* is a web portal owned by Telefónica SA, a Spanish multinational founded in 1924 and widely spread in Latin America. Meanwhile, *El Universal* is a generalist newspaper owned by Compañía Periodística Nacional S.A. de C.V, a Mexican conglomerate of print and digital media (newspapers and magazines) founded in 1916. *ElTiempo.com* is a general newspaper owned by the media conglomerate Casa Editorial El Tiempo (owned by the Luis Carlos Sarmiento Angulo Organization, a Colombian business group with banking, telecommunications, and real estate businesses present in Central America, the Caribbean, and the United States); and *Caracol Radio*, a radio station owned by Grupo PRISA, a Spanish multinational media company founded in 1972, one of the most powerful in the Ibero-American sphere.

In these four media outlets, we identified three common characteristics: 1. They are large (not only because of their size as measured by print, audience, and coverage but also because they belong to consolidated conglomerates and multinational companies). 2. They are traditional (In Colombia and Mexico, they emerged more than six decades ago in other media. They then migrated to the web, where their sites have been operating for at least twenty years. In Brazil's case, they are digital native media whose trajectory exceeds ten years). 3. They are popular (In the case of Colombia and Mexico, in the offline world, they are national media, and, in the online world, their websites have a high rate of traffic both globally and within their countries. In Brazil's case, its websites have high traffic both globally and domestically).

Of the media companies previously mentioned, Grupo PRISA stands out, owner of *Caracol Radio* in Colombia, which has had a publicized relationship of friends/enemies. They have been a strong detractor of Google News, but also a significant beneficiary of Google programs such as the Digital News Initiative, becoming a clear example, as was previously mentioned, of a coopetition relationship with this tech company (Cobos, 2017a). There is also Casa Editorial El Tiempo, owner of *El Tiempo*, who also had disagreements with Google News because of its operating way. Although it made an act of protest, it finally withdrew (Santos, 2014). This is interesting and leads to elaborating the hypothesis that the greater the amount of traffic, the news media outlet increases the importance of the website that is sending it and will seek other additional compensation for the aggregation. The lower referred traffic, the news media outlet considers the aggregation irrelevant and settles for just a few clicks. This hypothesis may explain why the large, traditional and popular news media outlets, which saw in the previous conclusions, enjoy greater visibility in the news aggregator and, therefore, a greater probability of receiving traffic from this. Also, they are to a large extent, the that more have complained against Google News and demand payments for their news.

Finally, this shows that the StoryRank algorithm showed, at least in the editions of Brazil, Colombia, and Mexico, as previously mentioned, a very marked preference for adding news from large, traditional, and popular news media outlets, which are characterized by their very high volume of news production, by

matching this with similar findings in other Google News editions (Foster, 2012; Legerén, 2014) – even Google Search (Diakopoulos, 2019). This system allows Google News to guarantee a permanent inventory of news and updates, as previously mentioned. However, it also shows a bias caused by technical factors that make the neutrality that this aggregator proclaims questionable and may have deeper social implications. Despite that wide variety at first glance, audiences are consuming news from specific news media outlets that are popular, large and traditional and thus may have obtained more traffic, for instance, than the small, unknown and new news media outlets.

REFERENCES

Associação Nacional de Jornais – ANJ. (2012, September 30). *Jornais brasileiros fora do Google News – Esclarecimento da ANJ*. Retrieved from www.anj.org.br/2013/09/30/jornais-brasileiros-fora-do-google-news-esclarecimento-da-anj-associacao-nacional-de-jornais/

Athey, S., & Mobius, M. (2012). The impact of news aggregators on internet news consumption: the case of localization. Workshop on the Economics of Web Search and Social Network. Proceedings of the *Sixth ACM International Conference on Web Search and Data Mining 2013*. Sapienza Università di Roma, Rome.

Athey, S., Mobius, M., & Pal, J. (2017). *The impact of aggregators on internet news consumption*. Stanford Institute for Economic Policy Research (SIEPR). Retrieved from https://siepr.stanford.edu/sites/default/files/publications/17-034_0.pdf

Bansi, L. (2019a). "In the world system, unequal cultivation of news on international media landscape": A content analysis of international news on Google News. *International Journal of Research in Social Sciences*, 9(3), 791–80. Retrieved from www.ijmra.us/project%20doc/2019/IJRSS_MARCH2019/IJMRA-15319.pdf

Bansi, L. (2019b). News and news agencies: An analysis of international news on Google News. *International Journal of Research in Social Sciences*, 9(4), 826–839. Retrieved from www.indianjournals.com/ijor.aspx?target=ijor:ijrss&volume=9&issue=4&article=052

Bayonet, L. (2009, May 18). *Infomediarios: un modelo de negocio en internet*. Universidad Pontificia de Salamanca, campus Madrid, España. Retrieved from www.slideshare.net/ebayonet/infomediarios-un-modelo-de-negocio-en-internet

Cádima, F. (2013). A Google, o sistema de media e a agregação de informação. *Intercom – Revista Brasileira de Ciências da Comunicação*, 36(1), 19–37. DOI: 10.1590/S1809-58442013000100002.

Carlson, M. (2007). "Order versus access: news search engines and the challenge to traditional journalistic roles". *Media, Culture & Society*, 29(6), 1014–1030. DOI: 10.1177/0163443707084346.

Chyi, H., Lewis, S., & Zheng, N. (2016). Parasite or partner? Coverage of Google News in an era of news aggregation. *Journalism & Mass Communication Quarterly*, 93(4), 789–815. DOI: 10.1177/1077699016629370

Cobos, T.L. (2014). Diarios en Brasil: Google News vs la Associação Nacional de Jornais. Proceedings of the *XX Congreso Internacional de la Sociedad Española de Periodística*. Barcelona.

Cobos, T.L. (2017a). *Medios de comunicación iberoamericanos y agregadores de noticias: análisis a las ediciones de Google News Brasil, Colombia, España, México y Portugal* [doctoral dissertation, Universitat Autònoma de Barcelona]. Retrieved from ddd.uab.cat/record/188096

Cobos, T.L. (2017b). *Dataset: Noticias de Google News ediciones Brasil, Colombia, México y Portugal, 2015*. Dipòsit Digital de Documents de la UAB. Retrieved from ddd.uab.cat/record/182292

Cobos, T.L (2018). Las percepciones y experiencias sobre Google News de los editores de medios noticiosos latinoamericanos indexados en las ediciones de Colombia y México. *Estudios Sobre el Mensaje Periodistico*, 24(2), 1183–1198. DOI: 10.5209/ESMP.62208

Del Águila, A., Padilla, A., & Serarols, C. (2007). Value creation and news intermediaries on Internet. An exploratory analysis of the online news industry and the web content aggregators. *International Journal of Information Management*, 27(3), 187–199. DOI: 10.1016/j.ijinfomgt.2006.12.003

Demers, D. (2005). *Dictionary of mass communication & media research: a guide for students, scholars & professionals*. Washington: Marquette Books.

Diakopoulos, N. (2019, May 10). Audit suggests Google favors a small number of major outlets. *Columbia Journalism Review*. Retrieved from www.cjr.org/tow_center/google-news-algorithm.php

Du, K. (2018). The impact of news aggregators on the geography of financial information production: Evidence from Google News in the U.S. and Spain. SSRN Electronic Journal. DOI: 10.2139/ssrn.3311235

Edo, C., Yunquera, J., & Bastos, H. (2018). La sindicación de

contenidos en los agregadores de noticias: hacia la devaluación de los criterios profesionales periodísticos. *Comunicar*, 59(27), 29–38. DOI: 10.3916/C59-2019-03

El Mundo. (2003, September 30). Nace el servicio 'Google News' en español. *El Mundo*. Retrieved from www.elmundo.es/navegante/2003/09/30/empresas/1064923609.html

Filloux, F. (2013, February 25). Google News: the secret sauce. *The Guardian*. Retrieved from www.theguardian.com/technology/2013/feb/25/1

Fonseca, B. (2005, November 17). Get news in Portuguese. *Google Official Blog*. Retrieved from www.googleblog.blogspot.com.es/2005/11/get-news-in-portuguese.html

Foster, R. (2012). *News plurality in a Digital World*. Reuters Institute for the Study of Journalism. University of Oxford. Oxford: UK.

Gingras, R. (2014, December 11). Novedades acerca de Google Noticias en España. *Blog Oficial de Google España*. Retrieved from http://googleespana.blogspot.com.es/2014/12/novedades-acerca-de-google- noticias-en.html

Google News. (n.d.a). *Appear in Google News*. Publisher Center Help. Retrieved from https://support.google.com/news/producer/answer/6016113?hl=en

Google News. (n.d.b). *Google News policies*. Publisher Center Help. Retrieved from https://support.google.com/news/publisher-center/answer/6204050

Haim, M., Andreas, G., & Hans-Bernd, B. (2018). Burst of the Filter Bubble? Effects of personalization on the diversity of Google News. *Digital Journalism*, 6(3), 330–343. DOI: 10.1080/21670811.2017.1338145

Hagel, J., & Rayport, J. (1997, n.d.). The coming battle for customer information. *Harvard Business Review*. Retrieved from https://hbr.org/1997/01/the-coming-battle-for-customer-information

Huyen, L., Raven, M., Ekdale, B., High, A., Havens, T., & Shafiq, Z. (2019). Measuring political personalization of Google News search. Proceedings of *The Web Conference 2019*. San Francisco, CA. DOI: 10.1145/3308558.3312504

Hernández, R, Fernández, C., & Baptista, P. (2010). *Metodología de la investigación*. México DF: McGraw-Hill.

Kramer, S. (2003, September 25). Google News creator watches portal quiet critics with "Best News" webby. *Online Journalism Review*. Retrieved from www.ojr.org/ojr/kramer/1064449044.php

Krantz, M. (2004, December 9). News happens everywhere. *Google Official Blog*. Retrieved from http://googleblog.blogspot.com.es/2004/12/news-happens-everywhere.html

Lee, A, & Chyi, H. (2015). The rise of online news aggregators: consumption and competition. *International Journal on Media Management*, 17(1), 3–24. DOI: 10.1080/14241277.2014.997383.

Legerén, E. (2014). *Visibilidad de las universidades andaluzas en los medios de comunicación online a través de Google News* [doctoral dissertation, Universidad de Granada]. Retrieved from www.academia.edu/11342212/Visibilidad_de_las_universidades_andaluzas_en_los_medios_de_comunicaci%C3%B3n_online_a_trav%C3%A9s_de_Google_News?auto=download

Legerén, E., Herrero, V., & Arboledas, L. (2011). "*Comunicación institucional y "agregadores de noticias": el caso de la Universidad de Granada*". In J. Vargas, J. Sierra, Y. Rodríguez & J. Durán (Eds.), *El EEES y la divulgación de los resultados investigadores sobre innovación docente* (pp.65–80). Madrid: Visión Libros.

Machlis, S. (2009, October 5). Inside the Google News algorithm. *Computerworld*. Retrieved from www.computerworld.com/article/2467854/e-commerce/inside-the-google-news-algorithm.html

Madsen, P., & Andsager, J. (2011). Aggregating agendas: online news aggregators as agenda setters. *Midwest Association for Public Opinion Research Annual Conference*. Chicago.

Murdoch, R. (2009, December 1). From town crier to bloggers: How will journalism survive the internet age? *Federal Trade Commission's Workshop*. Retrieved from www.ftc.gov/sites/default/files/documents/public_events/how-will-journalism-survive-internet-age/murdoch.pdf

Nechushtai, E., & Lewis, S. (2018). Personalization and diversity in recommendations on Google News. Proceedings of the *12th International AAAI Conference on Web and Social Media*. Stanford, CA.

Picard, R. (2018). *The economics of journalism and news provision*. In T. Vos (Ed.), *Handbook of Communication Science: Journalism* (pp.281–296). De Gruyter Mouton: Berlin. DOI: 10.1515/9781501500084-014

Potter, D. (2006). *Manual de periodismo independiente*. Programas de Información Internacional del Departamento de Estado de Estados Unidos. Retrieved from https://photos.state.gov/libraries/amgov/30145/publications-spanish/handbook_journalism-sp.pdf

Rieder, B. (2013). Studying Facebook via Data Extraction: The

Netvizz Application. Proceedings of the 5*th Annual ACM Web Science Conference*. Paris, France.

Santos, L. (2014, December 15). Salen medios españoles de Google News, que ya había rechazado a pagar a diarios colombianos. *Pulzo*. Retrieved from www.pulzo.com/opinion/salen-medios-espanoles-de-google-news-que-ya-habia-rechazado-pagar-diarios-colombianos/258741

Torres, J. (1985). *Economía de la comunicación de masas*. Madrid: Grupo Cultural Zero S.A.

Trielli, D., & Diakopoulos, N. (2019). Search as news curator: The role of Google in shaping attention to news information. Proceedings of the *2019 CHI Conference on Human Factors in Computing Systems*. Glasgow, UK.

Van Dijk, L. (2006, January 31). No Nieuws is bad news. *Google Official Blog*. Retrieved from http://googleblog.blogspot.com.es/2006/01/no-nieuws-is-bad-news.html

Vera, M. (1995). *La teoría de la comunicación: perspectivas para un debate interdisciplinario*. Málada: Universidad de Málaga.

Yoo, S. (2011). Google and Yahoo as the next gatekeeper: examining agenda setting effects of news aggregators. Proceedings of the *International Communication Association Annual Meeting*. Boston, USA.

TANIA LUCÍA COBOS. Lecturer and researcher at the Social Communication program of the Universidad Tecnológica de Bolívar, Cartagena de Indias (Colombia). She is a PhD in Communication and Journalism from the Universitat Autònoma de Barcelona (Spain) and MsC in Sciences Communication from the Tecnológico de Monterrey (Mexico). She researches about digital journalism, digital communication, audience reception, and Japanese animation. E-mail: tcobos@utb.edu.co, blog: http://tanialu.co and *Twitter*: @TaniaLu.