

## Google and the Algorithmic Infomediation of News

Nikos Smyrnaiois

For a decade, online news distribution has faced a growing intervention of firms that originate from the physical and logical layer of the Internet.<sup>1</sup> Web service providers, software firms, and ISPs progressively establish themselves as unavoidable passage points between publishers and Internet users. Some of these players are simple aggregators that gather content from a variety of sources on platforms with little added value. Others, such as Google News, distribute content through complex algorithms that aim to evaluate, hierarchize, and promote content, as well as organizing social interactions around it. By making supply meet demand more easily, they position themselves in the heart of the online news sector and thus become indispensable to publishers. I call this function *news infomediation*,<sup>2</sup> based on a concept used in information science<sup>3</sup> and economics<sup>4</sup> in order to describe the new, complex forms of intermediation that take place in digital networks. Infomediarities are not simple aggregators or distributors. They are situated in the very core of the internet and thus strongly influence the whole ecology of content production, distribution and consumption while, at the same time, they absorb a significant part of the surplus value produced in the internet economy.

Infomediation is a mix of aggregating, editing and distributing third party content that operates by matching between supply and demand. Infomediarities capture a significant part of online revenue and set standards about online publishing that influence journalistic practices. This paper examines the complex relations between online publishers and infomediarities from a socioeconomic point of view. First, I analyze the particular characteristics of online news that make the function of infomediation one of its major components. Second, I explain why relations of mutual dependency between publishers and infomediarities are *cooperative*, combining cooperation and competition. Finally, I look into one of the main news infomediarities, Google News, in order to explain its working principles as well as its repercussions on the online news sector. I show that Google simulates the process of the news agenda

formation through the use of quantitative criteria. The launch of Google News in Europe was litigious, and still is in places like Germany; yet the majority of publishers have gradually integrated SEO constraints imposed by Google. From a sociopolitical point of view this represents a major issue for the future of diversity and pluralism in online news.

### **How Infomediaation Operates in the Field of Online News and Journalism**

One of the trends of cultural and media industries is that of content oversupply.<sup>5</sup> Internet users have access to online news produced by traditional publishers such as newspapers, TV stations and radios, but also online-only professional media, and amateur and citizen journalism websites and blogs. This plethora of content includes traditional journalistic formats such as articles, photographs, and video reports, but also links, graphics, databases, blog posts, comments, and polls. All this material is indexed, stocked, duplicated and reproduced in a way that generates redundancy.<sup>6</sup> The result is a dazzling volume of information available online. A study of online news circulation in the United States succeeded in harvesting 170 million news items published over a period of eight months, more than half a million published per day.<sup>7</sup> This phenomenon of online information oversupply calls for filtering, selection and prioritization. This is where technology firms enter the news market through infomediaation services.

The individual pieces of content that are the basic bricks of news only make sense for the reader once they are assembled in a coherent, organized, and hierarchized ensemble. In traditional media this fundamental service is rendered by the newspaper layout and by the program grid. Therefore newspapers and television channels are not simple products but systems combining news, entertainment and publicity.<sup>8</sup> As such they are aimed at a two-sided market made of consumers and advertisers.<sup>9</sup> System markets like the newspaper industry also generate crossed subsidies. The most read pages of a local newspaper, such as the sports section, subsidize the less popular and more expensive genres such as investigative journalism. This fragile balance is dislocated on the Internet. The modular nature of online news allows technology firms to generate systems combining unique content items and links from a variety of different sources.

This transformation is also due to changes in news consumption. Today, the majority of Internet users go through deep links supplied by search engines, aggregators or social networking sites in order to access content pages directly.<sup>10</sup> Visits to news websites that start from their homepage—

an action interpreted in marketing studies as a sign of fidelity and of confidence—are diminishing as more complex navigation practices emerge.<sup>11</sup> Infomediaries build upon these new manners of accessing news by matching supply with demand through the use of algorithms.

### ***Coopetition between publishers and infomediaries***

Infomediaries such as Google and Facebook are rivals, targeting different market segments and, at the same time, competing against publishers for online revenue. In the interspaces of this worldwide competition emerge innovative startups undertaking financial risks. Global Internet groups and small technological firms form an oligopolistic market structure with a competitive fringe, comparable to that of traditional cultural industries.<sup>12</sup> The effect of this is breaking up the function of infomediatio n and allowing the emergence of strategies of cooperation between different infomediaries located in different places of the chain of value and usage.<sup>13</sup>

The common characteristic of these technology firms, whether they are small or global, is the fact that they do not produce original content intended for the general public. In other words, they do not undergo the particular constraints of managing the creative components of cultural and media industries. This means that infomediatio n platforms can only function on the condition of being supplied with content produced by publishers. Even if user generated content is growing in importance, quality news produced by professionals remains necessary for those who wish to attract vast audiences. This explains the ambiguous relations between publishers and infomediaries: the latter impose technical norms and revenue sharing to the former, while trying also hard to entice them. Publishers on the other hand are often critical of infomediaries but still tend to follow their rules.

Consequently, relations between publishers and infomediaries are made of cooperative competition, also called coopetition.<sup>14</sup> The cooperative element, which is materialized through legal, technical and financial agreements that bind media and infomediaries, is based on their mutual interest to share content and traffic. Internet firms use publishers' content to attract users, while publishers benefit from the traffic that is redirected to their websites from infomediatio n platforms. The common interest of these two categories of firms engaged in a situation of coopetition is to enlarge their respective markets.<sup>15</sup> The competitive element comes from the fact that publishers and infomediaries compete for advertising, a market that is essentially dominated by infomediaries. In the UK for instance Google and Facebook are set to comprise more than

half of the eight billion pounds of online advertising revenue in 2015.<sup>16</sup> This trend drives down prices for online media and makes it difficult for them to establish viable business models.

This problem of revenue sharing is coupled with a paradigm opposition, particularly strong in the news sector: journalists and publishers share a particular professional ideology, considering themselves to be the only legitimate players in news production and distribution. Thus they are particularly fastidious about how their content is used by third parties. On the other side, infomediaries such as Google are marked by “the Californian ideology,” a heterogeneous mix of engineering culture, free market economics, and counterculture libertarianism originating in Silicon Valley.<sup>17</sup> Thus their main concern is efficiency and “user experience.” These cultural differences, added to economic disagreements, produce a potentially conflicting framework of relations between them.

The example of France is characteristic. In 2003 Google decided to launch a French version of Google News without any prior discussion with the local representatives of the firm, who had only commercial prerogatives at that time, nor with French publishers whose content was going to be aggregated. In response French publishers blamed Google for copyright infringement and user data “confiscation” and threatened to go to trial. Some of the major French publishers, including *Le Monde*, *Libération* and *Les Echos*, temporarily withdrew their content from Google News. The same movement took place in Belgium, where Google was condemned for copyright infringement in 2006 and 2007. The AFP agency also initiated legal proceedings against Google in the United States for copyright infringement in 2005. All these cases were settled by 2013, mainly by Google paying European publishers through different forms of contract, but without explicitly recognizing their copyright claims.

### **Inside the Black Box of News Infomediatioin**

News infomediaries operate as black boxes that feed on content and social interactions. They organize their content following complex algorithms that remain unknown in detail. Google News is one of the most popular, created in 2002 by Krishna Bharat and his team. It aimed to answer a problem that Google encountered during the attacks of September 11, 2001—how to respond to user news queries efficiently and in real time? As shown in Figure 1, at the time Google was not at all reactive to events that made the news.

Searched the web for **world trade center**. Results 1 - 10 of about 1,260,000. Search took 0.23 seconds.

Categories: [Business > International Business and Trade > Services > Information](#)  
[News > Current Events > Economy and Business > Globalization and Free Trade](#)

**WTCA On-Line**  
 The **World Trade Centers Association** is an organization of over 300 **World Trade Centers** in almost 100 countries, connected to expand your global business. ...  
 Description: One-stop **trade** information hub for posting and reviewing **trade** opportunities, accessing valuable database...  
 Category: [Business > Industries > Import and Export > Portals > Trade Boards](#)  
[www.wtca.org/](http://www.wtca.org/) - 48k - [Cached](#) - [Similar pages](#)

[View from WTCA Headquarters](#)  
 ... Bienvenue! | Willkommen! | WTC Events | **Trade** Opportunities | On-Line Catalogs | **World Trade Library** | WTC Viewcam | Customer Service | About WTCA | Board of ...  
[www.wtca.org/view.html](http://www.wtca.org/view.html) - 13k - [Cached](#) - [Similar pages](#)

**World Trade Center Boston**  
 ... Located on Boston's historic waterfront, the **World Trade Center Boston** (WTCB) is New England's premier choice for exhibitions, conferences, and special events ...  
[www.wtcb.com/](http://www.wtcb.com/) - 7k - [Cached](#) - [Similar pages](#)

**The Top of the World Trade Center**  
 ... The Top of the **World Trade Center** web site is hosted, and designed by NYCPORT Networks all comments and inquiries to: [webmaster@nycport.com](mailto:webmaster@nycport.com). ">.  
[www.wtc-top.com/](http://www.wtc-top.com/) - 10k - [Cached](#) - [Similar pages](#)

Sponsored Links  
**Be Connected With Family**  
 You never know when tragedy will strike. Stay connected at our site  
[www.PhoneHQ.net](http://www.PhoneHQ.net)  
 Interest: \_\_\_\_\_  
[See your message here...](#)

Figure 1. Screenshot of Google on September 11th, 2001 during a search with key words “world trade center.” Source: Searchengineland.com

Google News was therefore conceived as an answer to this problem. The solution implemented by Google engineers was to produce a real-time prioritization process of news items using algorithms. This complex process of news infomediation consists of a mix of editing and distributing aggregated content from a multitude of third parties. Google indexes thousands of preselected news sources in real time and republishes a fraction of their content in its pages, typically publishing a headline, a summary, and an image for each article. Unlike at other news aggregators such as *Huffington Post* and *Gawker*, the only human intervention in Google News is in preselecting the initial sample of sources that are crawled and “tuning” the site’s algorithm. All the other operations of Google News are automated. News has slowly pervaded the results of the search engine so as to become their most visible element when it comes to important events such as the death of Osama bin Laden in May 2011, as shown in Figure 2.

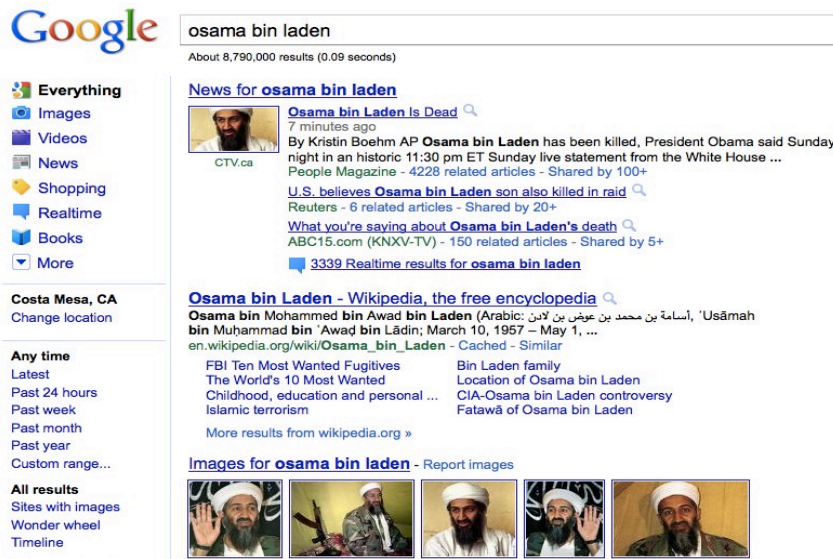


Figure 2. Screenshot of Google on May 1st, 2011 during a search with keywords “osama bin laden”. Source: Searchengineland.com

According to the patent deposits filed by Google, the prioritization of Google News is different from the ranking system of the search engine.<sup>18</sup> Google News actually tries to translate the social logics that govern journalism into quantitative variables in order to seize the media agenda in real time and with minimal human intervention. It gives each news website a rank based on a variety of criteria: page rank, productiveness, reactivity, and popularity, for example. Google News then crawls these websites and extracts new articles’ headlines, summaries and images that it bundles into clusters. Each cluster refers to an “event” or a “news topic.” The bigger a cluster is, and the higher the rank of the websites that compose it, the higher priority it receives on the Google News homepage and in the search engine’s results. The precise way in which the algorithm of Google News functions is a trade secret, but we can assume from statements by Google executives that other criteria are also used in prioritization such as novelty, originality, click-through rate and mentions in social media.<sup>19</sup> Even if Google News operates through algorithms, there is also an important social component to it.

After its United States launch in 2002, Google News was gradually translated into different languages establishing publishers’ dependence on Google on a global level. The portion of their traffic directly streaming from the different Google services in Europe as well as in the United States varies between 20 and 50 percent, and in some cases can go up to 80 percent.<sup>20</sup> These numbers clearly show how essential Google is for the online news economy. Due to the success of Google News, European and American publishers and regulatory authorities have become aware of

the stakes of news infomediation. A double movement has followed. First, media organizations have invested in human resources and technologies in order to improve their SEO and better exploit these new channels of distribution, perhaps to the point of “enslaving themselves to Google.”<sup>21</sup>

This cooperation has had a great impact on journalistic practices inside online newsrooms.<sup>22</sup> There has been a proliferation of “Keyword:...” style headlines because Google is particularly fond of them. Also the imperatives of productivity and reactivity, central parameters when it comes to prioritizing sources in Google News, push online newsrooms towards “shovelware”: re-writing press agency and PR material and publishing it as soon as possible.<sup>23</sup> Other practices such as the use of Google Insights and other similar tools in order to know which topics generate the most queries at a given moment significantly influence editorial choices. At the same time Google encourages publishers to use tools such as Analytics, Sitemaps, or AdSense by insinuating that they have a positive impact in source ranking. In this way Google interferes in the whole chain of production, publishing, measurement and monetization of news.<sup>24</sup> Of course these are optional practices and their degree of implementation depends on each particular media outlet. But they are definitely on the rise.

Second, publishers have used their lobbying power to influence decision-makers in order to obtain regulations that limit the Californian firm’s market power. The results run from frontal litigation, as in the case of Germany,<sup>25</sup> to full cooperation as in France where the main publishers struck a deal with Google.<sup>26</sup> In this respect the case of Spain illustrates perfectly the Cornelian dilemma to which publishers are confronted when it comes to dealing with Google. In 2014 The Spanish Association of Daily Newspaper Publishers (AEDE) obtained from the government a new law coming into effect January 1, 2015 requiring Google to pay to display snippets of their content. Google’s reaction was to shut down the Spanish version of Google News, causing a dramatic drop in traffic to Spanish news websites.<sup>27</sup> Today Google News Spain is shut off. Google still references Spanish online news sites in its search results without paying publishers and without them complaining about it.

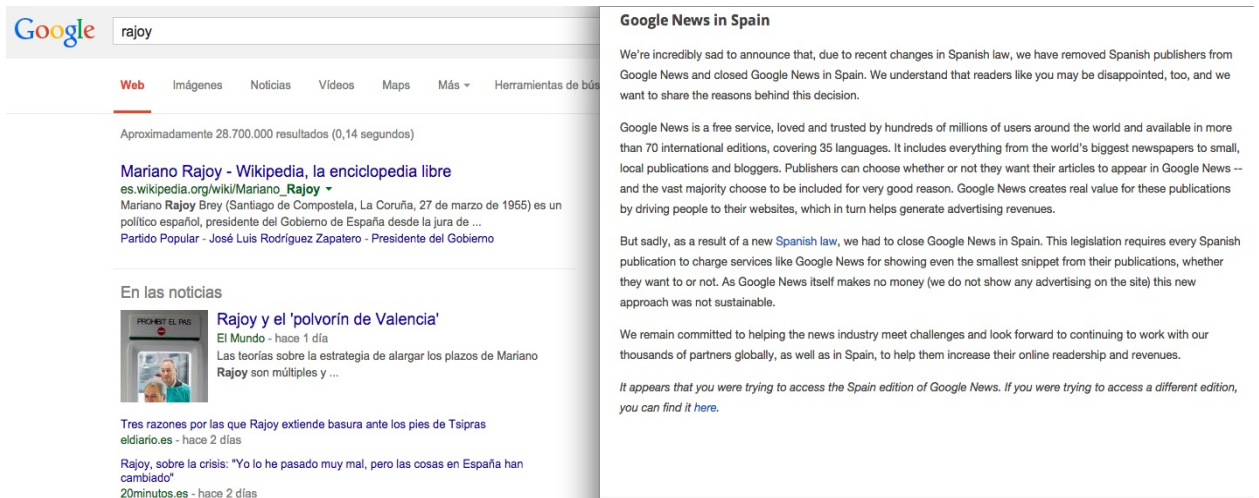


Figure 3. Screenshots of Google on February 16, 2015. On the left, results of a search with keyword “rajoy” where Google News extracts appear. On the right, Google explains why it closed the Google News portal in Spanish.

## Some Provisory Conclusions

The rise of online news has meant a growing dependency of publishers upon infomediaries such as Google. From the publishers’ point of view this trend is both a problem and a necessity. It is a necessity because news infomediation is a major source of traffic. It is a problem because dependency on technology firms diminishes publishers’ editorial autonomy and forces them to share revenue generated by their own content.

From an economic perspective, this trend poses a risk of extreme concentration in the sector of online news. In a time of global economic crisis, few worldwide media groups have the necessary resources to produce their own infomediation platforms. Fewer are those that can directly negotiate with Internet giants like Google, Facebook, and Apple. The vast majority of publishers in small or middle markets can only undergo the strategies of these multinationals trying not to be completely wiped out.<sup>28</sup> A major issue that emerges from a sociopolitical point of view is that of diversity and pluralism. If the media audience is concentrated on a small number of news websites that obey the rules of the infomediaries, then plurality of opinion and news diversity are negatively affected. The growing influence of Silicon Valley firms on the news agenda through infomediation is a manifestation of their rising political power and as such it should be put under further critical analysis.

## Notes



- 1 Yochai Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (New Haven, CT: Yale University Press, 2006).
- 2 Nikos Smyrniotis and Bernhard Rieder, "Social infomediaation of news on Twitter: A French Case Study," *NECSUS European Journal of Media Studies* 4, no. 2 (2013): 359–81, [necsus-ejms.org/social-infomediaation-of-news-on-twitter-a-french-case-study/](http://necsus-ejms.org/social-infomediaation-of-news-on-twitter-a-french-case-study/).
- 3 Audrey Knauf and Amos David, "Vers une meilleure caractérisation des rôles et compétences de l'infomédiaire dans le processus d'intelligence économique," paper presented at the conference Veille Stratégique, Scientifique et Technologique, Toulouse, 25 October 2004.
- 4 John Hagel III and Marc Singer, *Net Worth: Shaping Markets when Customers Make the Rules* (Boston: Harvard Business School Press, 1999).
- 5 David Hesmondhalgh, *The Cultural Industries* (London: Sage, 2007).
- 6 Nikos Smyrniotis, Emmanuel Marty and Franck Rebillard, "Does the Long Tail Apply to Online News? A Quantitative Study of French-Speaking News Websites," *New Media & Society* 12, no. 8 (2010): 1244–61.
- 7 Jaewon Yang and Jure Leskovec, "Patterns of Temporal Variation in Online Media," paper presented at the fourth ACM International Conference on Web Search and Data Mining, Hong Kong, 9 February 2011.
- 8 Michael L. Katz and Carl Shapiro, "Systems Competition and Network Effects," *Journal of Economic Perspectives* 8, no. 2 (1994): 93–115.
- 9 Simon P. Anderson and Jean Gabszewicz, "The Media and Advertising: A Tale of Two-Sided Markets," in *Handbook on the Economics of Art and Culture*, Volume 1, ed. Victor A. Ginsburg and David Throsby (London: Elsevier, 2006), 567–614.
- 10 AT Internet, "Online marketing trends," 13 March 2014, [atinternet.com/en/documents/news-websites-7-traffic-comes-facebook/](http://atinternet.com/en/documents/news-websites-7-traffic-comes-facebook/).
- 11 Pew Research Center, "Social, Search & Direct: Pathways to Digital News," March 2014, [journalism.org/files/2014/03/SocialSearchandDirect\\_PathwaystoDigitalNews.pdf](http://journalism.org/files/2014/03/SocialSearchandDirect_PathwaystoDigitalNews.pdf).
- 12 Françoise Benhamou, *L'Économie de la culture* (Paris: La Découverte, 2004).
- 13 For example, Flipboard developed an application offering readers a customized magazine made up of content that "friends" have shared on Twitter. In this case there is a double layer infomediaation that takes place: first inside Twitter's platform, then when the data from the Twitter API are filtered by Flipboard's algorithms.
- 14 Eric Brousseau, "e-Economie: Qu'y a-t-il de nouveau?" in *Annuaire français des Relations Internationales* (Brussels: Emile Bruylant, 2001), 813–33.
- 15 Paul Belleflamme and Nicolas Neysen, "Coopetition in Infomediaation: General Analysis and Application to E-Tourism," in *Advances in Tourism Economics*, ed. Álvaro Matias, Peter Nijkamp and Manuela Sarmento (Berlin: Springer, 2009), 217–34.
- 16 Mark Sweney, "Google and Facebook Will Have More than Half of UK Digital Ad Market," *Guardian*, 3 December, 2014, [theguardian.com/technology/2014/dec/03/google-facebook-uk-digital-advertising-market](http://theguardian.com/technology/2014/dec/03/google-facebook-uk-digital-advertising-market).
- 17 Richard Barbrook and Andy Cameron, "The Californian Ideology," *Science as Culture* 6, no. 1 (1996): 44–72.
- 18 Nikos Smyrniotis and Franck Rebillard, "L'actualité selon Google. L'emprise du principal moteur de recherche sur l'information en ligne," *Communication et Langages*, no. 160 (2009): 95–109.
- 19 Megan Garber, "Krishna Bharat on the Evolution of Google News and the Many Virtues of 'Trusting in the Algorithm,'" *Nieman Lab*, 6 May 2011,

- niemanlab.org/2011/05/google-news-founder-krishna-bharat-we-see-ourselves-as-the-yellow-pages/.
- 20 Pew Research Center, "Navigating News Online: Where People Go, How They Get There and What Lures Them Away," May 2011, journalism.org/files/legacy/NIELSEN%20STUDY%20-%20Copy.pdf.
  - 21 Expression used by an executive of a French media group. Smyrniaios and Rebillard, "L'actualité selon Google."
  - 22 Murray Dick, "Search Engine Optimization in UK News Production," *Journalism Practice*, 5, no. 4 (2011): 462–77.
  - 23 Chris Patterson, David Domingo, eds. *Making Online News* (New York: Peter Lang, 2008).
  - 24 Guillaume Sire and Bernhard Rieder, "Dans les ramifications de l'arbre hypertexte," *French Journal for Media Research*, no. 3 (2015), frenchjournalformediaresearch.com/lodel/index.php?id=468.
  - 25 Sam Schechner, "Google to Stop Publishing German Newspaper Extracts," *Wall Street Journal*, 2 October 2014, online.wsj.com/articles/google-to-stop-publishing-german-newspaper-extracts-1412260073.
  - 26 Nikos Smyrniaios, "Google's Deal With the French Press: a Major Event for the Future of News," *InaGlobal*, March 2013, inaglobal.fr/en/press/article/googles-deal-french-press-major-event-future-news.
  - 27 James Vincent, "Spanish News Sites Suffer After Dropping Google News," *The Verge*, 17 December 2014, theverge.com/2014/12/17/7407735/spain-drops-google-news-loses-traffic.
  - 28 Charles Arthur, "How Google's 'Panda' Update Put Some Websites on Endangered Species List," *The Guardian*, 5 December 2011, guardian.co.uk/technology/2011/dec/05/google-panda-update-endangered-species.

**Nikos Smyrniaios** is an Associate Professor at the University of Toulouse, France. His research focuses on the socioeconomic stakes and political issues of the internet. His main fields of work are online journalism and media as well the political use of social networking sites.