

Study Overview

1.1 – Presentation of the Study

This study by the Fondation Descartes aims to describe the way in which the French consume news information on the Internet.^{1,2} To achieve this, we asked ResponDi, a company specializing in digital data, to record for 30 consecutive days – from September 20 to October 19, 2020³ – the complete Internet activity of a panel of individuals representative of the French population.⁴ The connection data of the 2,372 participants constituting this panel was collected based on their use of various connected devices (computers, mobile phones, tablets). We were then able to analyze these individuals' consultations of 2,946 information sources previously identified by means of a website classification algorithm designed by the company Storyzy.

The information sources identified and tracked in this study both include and far exceed those usually considered in studies of online information consumption in France. They range from the web pages of national print media and public service news channels to those of more confidential "alternative" media, regularly accused of publishing false information, as well as those of men's or women's magazines, tabloids, cultural, sports or entertainment news outlets, and even those of online news aggregators and of digital-only media outlets (such as *L'Internaute*, *Doctissimo*, *01.net*, etc.).

At the end of the study's 30-day period, we forwarded a survey to our participants in order to establish their socio-demographic profile. Furthermore, this survey inquired about their relationship to news information, the online and offline information sources they consult, and the trust they place in them. The innovative methodology employed in this study allowed us to elaborate a rich and unprecedented overview of the way in which the French inform themselves on the Internet.

1. We would like to thank the Observatoire Société et Consommation (<https://lobsoco.com>) which offered us its expertise in the initial design of this study.

2. By "news information," we mean all political, economic, social, cultural, sports, or international news and current events, or those of more specific topics – celebrity news, science, technology, etc. – as well as opinion pieces or blogs with a social or political scope.

3. This period was chosen because it lies between the first and second sanitary lockdowns of 2020. Thereby, the information behavior of the French on the Internet during this period was more representative than it would have been during a nation-wide lockdown.

4. Recorded Internet activity consists of any and all data transfers between the participants' connected devices and the rest of the global network (excluding Bluetooth transfers). Internet activities considered as information consultation include connection to URLs associated with an information source selected in this study and consultation of mobile applications associated with these information sources. It should be noted that a "time-out" system interrupted the recording of time spent on an information source if participants were inactive for more than three minutes (this does not concern the consultation of audio or video content).

Selected information sources

The sources included in this study regroup 2,295 French-language websites and 651 French-language YouTube channels identified by Storyzy's algorithm as relating to news information and having registered at least 15,000 visitors worldwide during the month preceding our study or, for YouTube channels, having at least 15,000 subscribers at the time the study began.

Some of the 2,295 websites selected also dispose of mobile applications and/or are associated with Facebook and Twitter pages and YouTube channels. We recorded the consultation of these information sources by our study participants through all of these access channels. The set of websites, mobile applications, Facebook and Twitter pages and YouTube channels tracked in this study covers a very large spectrum of the French-language information sources available on the Internet.

We note that the time potentially spent by participants to obtain information from the following sources or through the following access channels was not accounted for in this study: information sources in foreign languages, information sources consulted on connected devices other than their personal devices, information podcasts consulted on sources not included in this study, information content downloaded online but consulted offline, information shared *via* newsletters, push notifications, private messaging (such as WhatsApp) or private social media groups whose consultation does not require visiting an external web page, as well as information shared directly on participants' Facebook "wall" or Twitter "feed" and that can be read without having to visit the website of the media that published it (see Inset 3).

— INSET 1

An innovative hybrid approach

Two annual surveys seem indispensable in understanding online information consumption habits in France: the Reuters Institute's Digital News Report⁵ and the Digital Market Barometer conducted by CREDOC and ARCEP⁶. The results of these surveys are regularly interpreted as indicators of the actual consumption of online information in France, when they in fact provide us with information regarding respondents' representation of their information consumption habits on the Internet. Yet, there may well exist a marked discrepancy between the actual information consumption habits of the French on the Internet and their own representation of these habits.

One way of overcoming this significant limitation of declarative studies is to measure Internet user traffic on information-related websites, as the company Médiamétrie does in France. The behavioral – rather than declarative – data collected by Médiamétrie has the merit of providing an overall representation of the consultation of various online information sources in France.

However, this method does not allow us to know who the individuals contributing to the traffic measured on these websites are. In particular, it prevents us from understanding the diversity of individual information consumption habits and from determining whether socio-demographic (gender, age, socioprofessional categories, etc.) or cognitive (interest in current affairs or level of trust in the media, for example) factors affect these habits.

This study by the Fondation Descartes is, to our knowledge, the first to combine behavioral and declarative procedures to study the information practices of the French on the Internet. Owing to this hybrid methodology, we have been able to closely describe these habits, taking into consideration their diversity and relating them to the socio-cognitive profile of individuals and to the representation that these individuals have of the way in which they inform themselves on the Internet.

— INSET 2

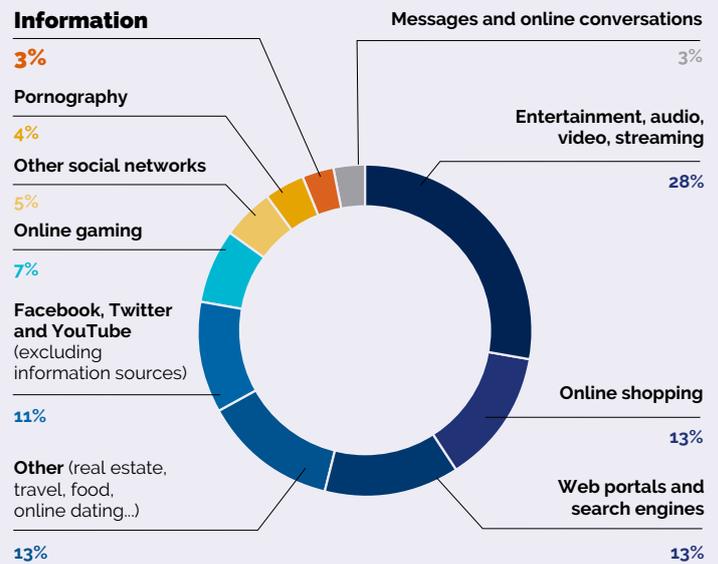
1.2 Key Findings

Our study highlights that, as a whole, the French do not devote a large portion of their connected time to the consultation of news information. On average, only 3% of the total time spent online by our study's participants consists of information consumption – which, per participant, corresponds to slightly less than 5 minutes a day. We should note that these values reflect the time our participants spent consulting information sources identified and tracked in this study on their personal devices. They therefore exclude any time spent consuming information on other connected devices (e.g., work computers) or *via* information sources excluded from our analysis, such as foreign language websites (see Inset 1).

3% of the total time spent online by our study's participants was devoted to the consumption of information.

In addition, it is important to note that while, *on average*, the French spend little time informing themselves on the Internet, online information practices can vary greatly between individuals. For instance, while 17% of our study's participants consulted no online information sources throughout the study period, 5% spent a total of over ten hours consulting such sources. We were able to identify a certain number of individual and social characteristics associated with participants' propensity to inform themselves online.

Distribution of connected time (during the study's 30-day period)



GRAPH READING: During the study's 30-day period, 3% of the total connected time of all participants was devoted to the consumption of information.

NOTE: A web portal is a website that brings together information from diverse sources – e.g., email, online forums, search engines – in a uniform way.⁷ Certain portals, such as Yahoo! and Orange, feature news or information content. In our study, time spent by participants consulting such content within web portals was counted in the "information" category.

5. www.digitalnewsreport.org/

6. www.arcep.fr/uploads/tx_gspublication/rapport-barometre-num-2019.pdf

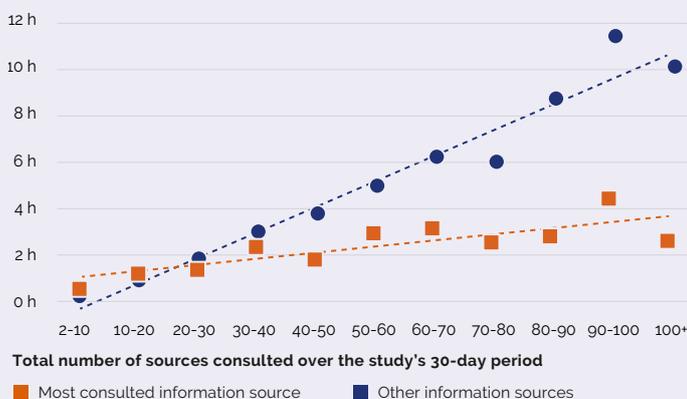
7. https://en.wikipedia.org/wiki/Web_portal

Among these, the most significant characteristic appears to be participants' age, with individuals 50 and over informing themselves more so on the Internet than others. Nevertheless, the observed correlations between participants' sociodemographic characteristics and their use of the Internet to inform themselves are all relatively weak – these characteristics alone therefore cannot explain our observation of high inter-individual variance in online information behaviors. Other unidentified factors, such as individual curiosity or time available for information consumption, necessarily contribute to this variance.

17% of participants consulted no online information sources throughout the study period, while **5%** consulted such sources for a total of over ten hours.

Meanwhile, we observe that participants who spend a lot of time informing themselves on the Internet do so less by increasing the amount of time spent consulting the same source than by diversifying the sources they consult. Indeed, the more time participants spent informing themselves online during the study period, the greater the number of different sources they consulted.

Evolution of total time spent on the most consulted source according to the number of sources consulted



GRAPH READING: The dotted line in orange (regression line) represents the increase in the total time spent on the information source most consulted by each participant as the total number of sources consulted increases. The dotted line in blue represents the corresponding increase in total time spent consulting other information sources.

This reflects a "browsing" behavior among consumers of online information, who seem to switch from one information source to the next often without consulting any one source for a significant amount of time. This is shown by the fact that time consecutively spent on an information source upon each individual consultation is, on average, less than 2 minutes.

Individuals 50 and over inform themselves more on the Internet than others.

Although this value varies significantly among participants, it equals or exceeds 4 minutes for only 6% of participants.

Another element further supports this "browsing" effect: over the 30 days of the study, participants consulted a total of 1,290 different information sources – a number that far exceeds the number of major media outlets present on the Internet, and that therefore highlights the tendency of certain participants to consult lesser-known information sources.

We observe a "browsing" behavior among consumers of online information. As such, participants who inform themselves the most online do so by multiplying the number of sources they consult.

The fact that Internet users display this "browsing" behavior in their consumption of information seems to go against the widespread, but contested,⁸ idea that the Internet encloses users in "echo chambers" in which they are only exposed to information that is in line with their expectations and beliefs.

The information sources that were most consulted by participants during the study's 30-day period were those of regional print media. Information sources dedicated to sports news ranked second, followed by information sources associated with national print media. News aggregators ranked fourth, followed by sources of other general news sites.

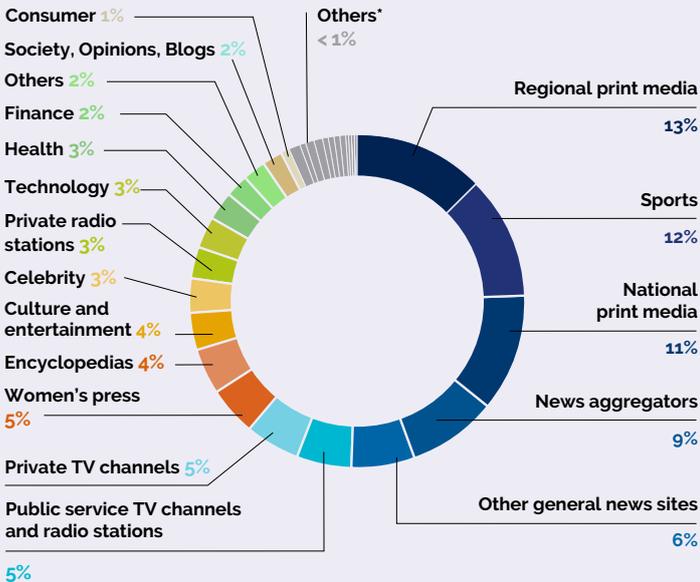
In examining the online information sources most consulted by our study's participants, we found that the top 26 sources alone accounted for 40% of the time participants dedicated to informing themselves online. This "top 26" is largely made up of online sources belonging to traditional media outlets – a noteworthy exception being Wikipedia, which was the information source consulted by the greatest number of participants over the study's 30-day period.

The online information sources most consulted are those of regional print media, followed by those dedicated to sports news and those of the national print media.

8. www.fondationdescartes.org/en/2020/07/filter-bubbles-and-echo-chambers/



Types of information sources consulted



*Automobile, Lifestyle, Economy, TV program, Foreign francophone media, Science, Vocational news, Men's press, Cooking, Local news, Environment, Animal, Parenting, Spiritual, Aerospace, Fact-checking

GRAPH READING: Overall, participants spent 13% of their connected time dedicated to information consumption on online sources belonging to regional print media.

NOTE: This graph does not account for online weather forecasts.

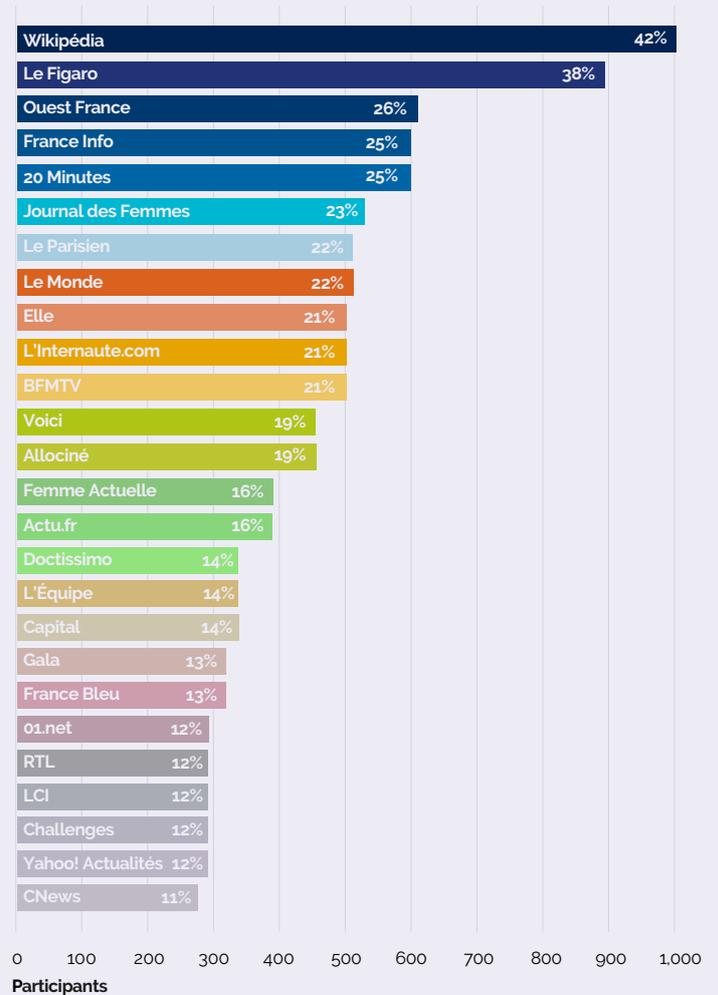
Major traditional media outlets have therefore succeeded in exploiting their offline reputation to assert themselves as major players in the online information market. It should be noted that the digital-only media present in this "top 26" (*L'Internaute*, *Doctissimo*, etc.) all belong to major media groups, with the exception of Wikipedia.

Another one of our findings that must be emphasized is that participants' reported use of the Internet to inform themselves – whether in terms of the nature of the sources they report consulting, or in terms of the frequency with which they report doing so – only weakly corresponds to their effective online information behavior, as measured by our study. In other words, the French seem to have a relatively fuzzy perception of the way in which they inform themselves on the Internet.

Furthermore, we observed only a weak relationship between the consultation of a given information source and participants' reported level of trust in this source. It is possible that this is a consequence of the "browsing" behavior of the French on the

The **26** online information sources that were most consulted constitute **40%** of the time spent consuming information on the Internet. These mainly consist of online sources belonging to traditional media outlets.

Ranking of leading information sources by number of participants having consulted them



GRAPH READING: 42% of all participants (i.e., 990 participants) consulted Wikipedia at least once throughout the study's 30-day period.

NOTE: This graph does not account for online sources of TV programs or for online weather forecasts.

Internet. Individuals may be switching from one source to another or clicking on a headline appearing on their Twitter "feed" or Facebook "wall," less in an attempt to search for specific information than out of entertainment or intellectual curiosity.

Lastly, we used in this study an algorithm developed by Storyzy in order to sort the selected information sources into two categories: "reliable" and "unreliable." We then compared the results obtained by this procedure to the classification proposed by the newspaper *Le Monde*. This comparison

The French seem to have a relatively fuzzy perception of the way in which they inform themselves on the Internet.

showed that the two methods of evaluating the reliability of information sources produce largely compatible results. This categorization of sources allowed us to evaluate the exposure of the French to disinformation on the Internet, whether it be generic disinformation (websites providing social and political information and news content considered to be unreliable), health-related disinformation, pseudo-science, satire or "click-bait" (websites whose content is not intended to inform the reader, but solely to attract their attention in order to generate traffic on their page).

We observed only a very weak relationship between reported trust towards information sources and the consultation of these sources.

We found that, over the study's 30-day period, 39% of participants consulted information sources considered to be unreliable. On average, these participants spent 11% of their daily time dedicated to online information consumption consulting these sources, which equates to 0.4% of their total connected time. Among all participants, 5% of connected time dedicated to information was spent consulting information sources considered to be unreliable, which equates to 0.16% of total connected time. The sources of disinformation that were most consulted were sources of generic disinformation, followed by sources of "click-bait" and sources of health-related disinformation.

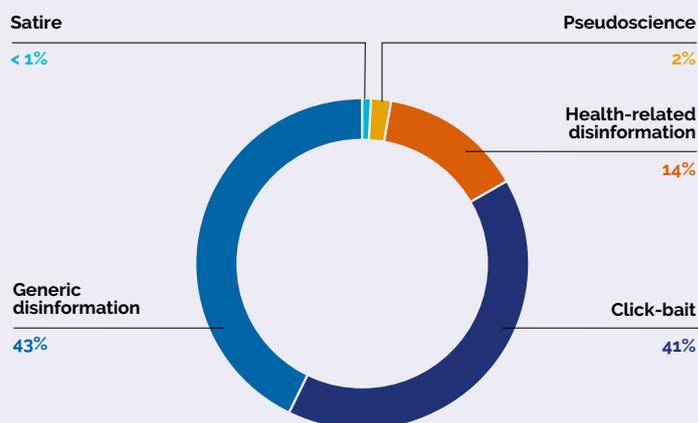
Over the study's 30-day period, 39% of participants consulted information sources considered to be unreliable.

The statistical analyses we subsequently conducted show that men are over-represented among participants who consulted sources of generic disinformation. Within this subsection of our study's participants, we also find proportionately more individuals living alone and more individuals expressing support for or belonging to the Yellow Vests movement. Furthermore, these individuals on average have a low level of trust in institutions, the government and the media, but express on average more interest than other participants in social and political news.

On the contrary, women are over-represented among participants who consulted sources of health-related disinformation, and we find among this particular subgroup of participants a disproportionate number of inactive individuals, including retirees. Additionally, participants who consulted sources of health-related disinformation scored higher than others on a standardized scale intended to evaluate individual inclinations towards belief in conspiracy theories.

Among all participants, **5%** of connected time dedicated to information was spent consulting information sources considered to be unreliable, which equates to **0.16%** of total connected time.

Distribution of time spent on unreliable information sources



GRAPH READING: 43% of total time spent consulting unreliable information sources was on sources of generic disinformation.

NOTE: N = 921 participants. These are the participants that consulted unreliable information sources throughout the study's 30-day period.

Participants having consulted sources of "click-bait" are characterized first and foremost by being on average older than the average age of all participants. Individuals aged 65 and over are particularly over-represented here, as are inactive individuals and retirees. We further observe that the subgroup of participants who consulted sources of "click-bait" possess, on average, a lower level of education and household income than the rest of participants.

In general, the share of connected time that our participants spent on online information sources considered to be unreliable appears to be relatively low. This seemingly encouraging results may in actuality indicate that disinformation on the Internet today is disseminated more so *via* social networks than *via* established online sources (e.g., websites). The sheer quantity of fake news surrounding the Covid-19 pandemic that have been massively disseminated on Facebook and Twitter serve as an example of this phenomenon. Misleading messages are circulated from person to person by being shared by social network users on their personal "walls" or "feeds." Very often, these pieces of disinformation take on the form of a short text or a commented picture, without referring to an external online source. As a result, we were not able to account for participants' exposure to this type of disinformation in our study.

Social networks in this study

Social networks are, in many respects, “black boxes” for researchers. It is indeed difficult, both for technical reasons and data privacy concerns, to study the nature of the content appearing in private Facebook groups or on users’ Facebook “walls” or Twitter “feeds.”

Unable to escape this limitation, we did not have access to participants’ “walls” or “feeds,” nor to the private groups they potentially visited on Facebook. But this in no way means that our study does not account for participants’ consultation of information through their social media accounts. Firstly, we recorded the time that participants spent on Facebook and Twitter pages and YouTube channels associated with some of the 2,295 information sources we selected. The time participants spent consulting the 651 “independent” YouTube channels tracked in this study was also recorded (let us recall that YouTube is a social network).

Secondly, upon clicking on a video or article appearing on their “wall” or “feed” or in a private group, social network users are often redirected towards an external webpage which hosts the content in question. Thus, participants in our study who clicked on a news article featured on their Facebook “wall,” for instance, were redirected to the website of the media outlet that published that article. The time that these participants spent reading this article on the website of the media outlet in question was therefore accounted for in our study.

However, our study’s methodology did not allow us to account for the information that participants may have consulted directly on their Twitter “feeds” or Facebook “walls” or in the private groups they visit, given that in such instances participants were not required to visit an external website. This information can generally be broken down into two groups. First, the headlines of news articles that participants may have seen but did not click on – of these articles, they therefore read at most its title or its summary, in cases where the latter was displayed on social networks. Second, the short videos produced by certain media (e.g., Brut) and specifically designed to be viewed directly on social networks.

We must however note that this limitation to our study probably only had a marginal impact on our results. Indeed, based on the scarce data available on this issue, it would seem that news information only represents a small portion of the content circulating on social networks. For example, on Facebook – which is, alongside YouTube, the most widely used social network in France⁹ – only an average of 5% of the content featured on users’ “walls” consists of news information (understood in a broad sense).¹⁰ While this value undermines the individual disparities resulting from the personalization of content carried out by Facebook’s algorithm,¹¹ it does indicate that news information remains relatively marginal on France’s foremost social network. This phenomenon stems in part from Facebook’s policy to actively limit the amount of news information visible on its pages.¹²

— INSET 3

9. Nearly two-thirds of French Internet users use Facebook (<https://fr.wikipedia.org/wiki/Facebook>). Over the study’s 30-day period, 67% of participants used Facebook, and 24% used Twitter.

10. <https://www.buzzfeednews.com/article/alexkantrowitz/facebook-wont-release-its-5-of-news-feed-content-is-news>. This estimate includes political, social, economic and international news information, as well as entertainment or lifestyle news.

11. According to the only study that, to our knowledge, directly investigates the composition of Facebook content, 50% of the “walls” of a sample of users featured no information-related content among the first ten pieces of content displayed, 23% featured only one information-related piece of content, 16% featured two information-related pieces of content and the remaining 11% featured more than two. <https://www.niemanlab.org/2017/12/how-much-news-makes-it-into-peoples-facebook-feeds-our-experiment-suggests-not-much/>

12. <https://www.buzzfeednews.com/article/alexkantrowitz/facebook-wont-release-its-5-of-news-feed-content-is-news>.