Big data on firms and predictive B2B marketing

François Bancilhon,
Sidetrade

Abstract:
What impact do big data have on business-to-business (B2B) marketing? After a brief recall of recent trends in big data and data science, focus is shifted to the data about firms and the situation in France. How is the big data revolution modifying marketing studies and the management of the "marketihng funnel"?

Anything new with regard to data?

Big data is the conjunction of the threesome: the availability of new data in large quantities, the availability of ever more powerful machines (Moore's law still holds), and the availability of parallel computing technology for processing masses of data. The new data mainly come from the social media, the Internet of objects, the Web and open sources. Since the first two are well known, let us examine the last two.[1]

The Web generates two types of data that contribute to the big data phenomenon: the Web data themselves and the data about users' actions on the Web. Thanks to their ability to gather Web data by crawling (sucking up all contents on a website without trying to understand them) and scraping (fetching contents from a website with knowledge of the structure so as to extract structured data), many operators have built large, detailed sets of data from the Web. This activity accounts for the fact that robots instead of humans generate 52% of Internet traffic. Given that they can gather data on users' behavior patterns, website managers have set up data management platforms. The DMP memorizes information on users' behaviors — Who has visited the website? When? Why? And to see what? — and offers this information to other websites.

Open data are the data used by public authorities to perform their duties. They are available to firms (for use) and citizens (for consultation and for the sake of transparency). The data spared this "opening" are those that are top secret (*e.g.*, the armed forces) or personal data. The open data movement, launched by the Obama administration after 2008, has had an echo, variable, in all democratic countries. We shall examine the echo in France.

To process big data, techniques have been developed that enable anyone, from the smallest startup to big industrial groups, to massively gather data, process them, build applications and develop new uses for them. These big data and the capacity for using them have proliferated in many fields: transportation, medicine, smart cities, advertising, farming, marketing, business-to-consumer (B2C) relations, etc. No field is *a priori* exempt. This article will concentrate on business-to-business (B2B) marketing.

---

[1] This article, including any quotations from French sources, has been translated from French by Noal Mellott (Omaha Beach, France).

Anything new with regard to corporate data?

Official information about firms, which used to lack visibility, has suddenly become available to everyone thanks to open data. This holds for the data produced by firms about themselves and for the data produced about them by third parties (newspapers, the Internet). Firms communicate, thus describing themselves, on social networks, mobile applications, e-business platforms and the websites where they post job offers. We shall take a look at the percentage of firms equipped with various means of electronic communication. Besides the information on companies coming from themselves, the media on the Internet regularly pass information about firms. This is the second source of business information.

What open data are available in France about firms? What do firms communicate about themselves over the Web?

The opening of data

The trend toward opening data started in France in 2010 before being strongly impelled by Etalab in 2013. Several acts of law and decrees, under various governments, have restated the principle of opening data by default — a principle that is not fully applied but is steadily advancing. After the central government, local authorities have taken up on the idea.

This trend has led to opening several sources of data on companies. The registry of identification of firms and establishments (SIRENE: Système National d'Identification et du Répertoire des Entreprises et de leurs Établissements), managed by the National Institute of Statistics and Economic Studies (INSEE), was opened in early 2017. The RNCS (Registre National des Commerces et des Sociétés), a national registry of businesses and companies that groups information from tribunals of commerce, was opened in early 2019 after a long rearguard action by Infogreffe. It fell on INPI, the national institute of industrial property, to make available these data collected by the clerks of tribunals of commerce. INPI had already, in 2016, opened its own data (trademarks, patents and industrial designs). At the same time, the Direction of Legal and Administrative Information (DILA) made available the data from *Bodacc*, the official bulletin of legal notices (from tribunals of commerce in particular). The Ministry of Research and Higher Education took part in this trend by publishing the list of public research units and of cooperative programs between research and industry. The application ScanR provides access to these data. The national registry of nonprofit organizations (RNA) is open since 2017.

Within a period of two years, all data held by the public administration on firms has been opened. Previously, a payment was required to access most of this information, and a subscriber would have paid thousands of euros monthly for access.

As a consequence of this opening of data, startups have immediate access to these data and are able to use them to offer innovative services. Another consequence is that the quality of data has improved thanks to feedback from the persons who consult them.

Digitally mature firms

French firms are taking part in the general trend toward Web access. They have websites, and are present on the social media as well as the Internet via mobile applications. Table 1 lists, for four countries, the corporate equipment rates in website, e-business platforms, and certain social media (Facebook, Twitter and Linkedin). The equipment rate rises rather fast as a function of size. For instance, the equipment rate in websites of firms with sales of at least one million euros is 60% in France.

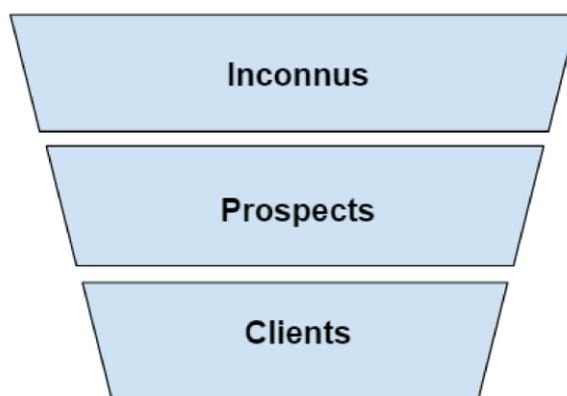| Table 1: Equipment rates: Percentage of firms having websites, e-business platforms or social media accounts						
Country	Firms	Websites	E-business platforms	Twitter	Facebook	Linkedin
UK	4,260,628	23.89%	7.59%	8.29%	7.76%	3.66%
FR	2,454,396	25.20%	9.64%	4.06%	9.22%	1.80%
BE	1,700,340	6.69%	5.87%	0.73%	2.61%	0.50%
NL	2,760,265	30.64%	6.63%	5.34%	9.36%	2.68%
NB: For France, only "commercial companies" (a category used by INSEE) have been retained.						

What impact on B2B marketing?

B2B marketing activities fall into two major categories: on the one hand, the management of the "marketing funnel", the central process in operational marketing for identifying prospects and making them customers; and, on the other hand, marketing studies, for understanding markets and changes in them.
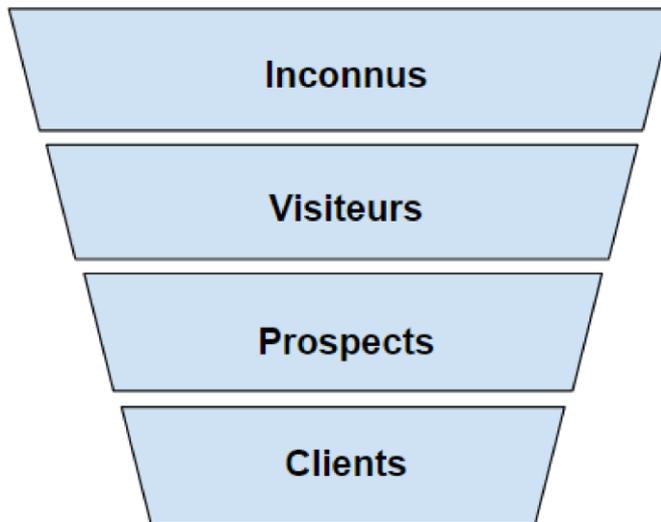
The marketing funnel

Figure 1 illustrates the classical funnel in marketing. In B2B marketing, the purpose is to identify, among firms, those that might become clients, to make them prospects with whom contacts are established, and then to change them into actual clients. The passage from being an unknown to being a prospect is the objective of traditional marketing methods: advertising, trade shows, direct mailing campaigns, telemarketing and bulk mailing using address lists acquired from vendors.

Figure 1: The marketing funneling of unknown persons into prospective and then actual customers

When business activities pass through a website, the marketing funnel is altered, as shown in Figure 2. A new category appears: the website's visitors.

Figure 2: The marketing funneling on a website: A new stage in the process (visitor)

What firms (or persons) are present in each stage of the B2B marketing funnel?
- The UNKNOWN are, in France, the 2.5 million firms that you do not know, some of which might potentially be customers of your products or services.
- Website VISITORS are of two sorts, identified or unidentified. The UNIDENTIFIED are attracted to the website by your inbound marketing. You do not know them, but see them circulating from web page to page, and you store their browsing history. The IDENTIFIED are also attracted to the website by your inbound marketing, but they have registered on the site (to access specific contents, such as a demonstration or set of specifications). So, you know them, and see them circulating from web page to page, and you store their browsing history.
- The PROSPECTS have had contacts with your sales force or interacted with your information system.
- The CUSTOMERS use your products or services. You see what they do, or do not do, with them.

Managing the funnel means making firms go down the funnel, the aim being to have as high as possible passage rates between stages in the funnel. This means for:
- the UNKNOWN: detecting those who are potential clients (who, therefore, need your products or services) and attract them to your website by inbound marketing (based on contents or referrals) or outbound marketing (bulk mail).
- UNIDENTIFIED VISITORS: identifying some of them (via IP-tracking, which links the visitor's IP address to a firm). You might obtain additional information on them via a DMP, which stores their browsing history on other websites, and you might convince them to identify themselves to benefit from additional information.
- IDENTIFIED VISITORS: turning them into prospective customers.
- PROSPECTS: making them customers.
- CUSTOMERS: following their behavior patterns, persuading them to buy more (upselling) or something else in your product line (cross-selling) and convincing them not to leave the website (retaining them).

Let us see what sort of data is available in each stage of the funnel:
- the UNKNOWN: any data available on firms.
- UNIDENTIFIED VISITORS: data about their comportment on your website and possibly on other websites.
- IDENTIFIED VISITORS: data about their behavior on your website and possibly on other websites, as well as data on the firm.
- PROSPECTS: data about their behavior on your website and on other websites, as well as data on their firm and the history of business relations with them.
- CUSTOMERS: data on their behavior on your website and possibly on other websites, as well as data on their firm, on the history of business relations with them, and on their uses of your product.

Which techniques can be used with these data in order to reach our objective of moving visitors down through the funnel?
- the UNKNOWN: tools for filtering, searching and ratings (via machine learning).
- UNIDENTIFIED VISITORS: IP tracking, ratings (via machine learning), marketing automation.
- IDENTIFIED VISITORS: ratings (via machine learning), marketing automation.
- PROSPECTS: ratings (via machine learning).
- CUSTOMERS: ratings (via machine learning).

What does machine learning entail during each stage in the funnel? By using accumulated data about what has happened in the past in the funnel, a more or less sophisticated predictive model is built for identifying the businesses most likely to become customers. Table 2 summarizes what has just been described.

Table 2:

TARGETED GROUP	Objective	Data	Techniques and tools
THE UNKNOWN	Identify and make contact	External data on firms	Machine learning
UNIDENTIFIED VISITORS	Identify and make contact	Behavior on the website	IP tracking, machine learning & marketing automation
IDENTIFIED VISITORS	Change into a prospect	External data on firms & behavior on the website	Machine learning & marketing automation
PROSPECTS	Change into a customer	External data on firms & behavior on the website plus business interactions	Machine learning & customer relationship management
CUSTOMERS	Retain on the website, upselling, cross-selling	External data on firms, behavior on the website & business interactions plus uses of the product or service	Machine learning & customer relationship management

Marketing studies

In this category are the studies for assessing the size of a market and trends in it, identifying an ecosystem, and segmenting the market by type of activity. This sort of study helps, for example, to identify innovative firms or to understand a technological tipping point (such as the shift from diesel to gasoline for internal combustion engines). The intent is to describe or quantify a phenomenon. Studies of this sort can be ordered from a marketing agency. After identifying from twenty to a hundred firms considered representative of the business sector, the agency will make a questionnaire, administer it to a panel, produce a qualitative and quantitative report, and make recommendations. The report will be printed and distributed to customers. Typical of this usual approach to marketing studies are: that the report is on paper (an electronic version in pdf will be available, but that does not change the problem), that it provides a picture at a given instant but it might be outdated six months later, and that it is based on a sample.

Corporate data are now available in real time; data can be obtained on a list of firms instead of a sample and updated automatically over time; and interactive tools can be developed for visualizing data. Given all this, a completely different approach can now be adopted with a tool:

● that operates on a full scale. We no longer use a sample but the firms actually concerned.

● that operates in real time. We do not make a snapshot but follow changes in the data as they occur.

● that is interactive. This refers not to relations but to an interactive application that provides various viewpoints.

These dynamic, real-time, interactive and exhaustive dashboards do not rival other, more qualitative methods; instead, they tend to complete them.

Big data is changing most, even all, business and social activities. It has penetrated the field of B2B marketing somewhat later, but is now disrupting marketing approaches there. New data and new techniques provide new tools that significantly improve the results offered to firms. They help firms acquire new customers and increase their income from existing customers. They provide firms with a detailed vision of their markets in real time.