Baidu, Its Services and Its Competitors

1.1. The history of Baidu

The Chinese search engine Baidu had a significant monopoly in China before the arrival of HaoSou (formerly 360) in 2012. Baidu held around 80% of the market share in August 2010 but, four years later, had no more than 56%.

HaoSou won around 20–30% of the market in 2 years, thus becoming the second most used search engine in China.

Sogou, the third search engine used in China, holds around 12% of the market share. Around 5–10% is left for other search engines, including Google.

Faced with such fierce competition, Baidu began to innovate, producing original and increasingly interesting applications. It provides services that other search engines do not have. For example, Internet users can place sponsored links on images (Search Engine Advertising (SEA), which has never been done on any other search engine.

One of Baidu’s strongest innovations is the prediction of trends relating to the Chinese economy, using queries made by users in the search engine.
Baidu may have looked like an imitation of Google in the past, but it very quickly evolved, becoming an innovative search engine. For example, in keyword generation tools, Baidu can clearly determine and separate “Desktop” and “Mobile” search volumes, which Google does not yet do (in 2016). Furthermore, in terms of the trends tool, the Baidu Index (the equivalent of Google Trends) provides more data on semantic fields, demography and current news topics, and more forums, which are linked to the searches made in this tool.

Figure 1.1. Baidu logo written in Chinese characters and pinyin

We are going to explore Baidu’s general services (its subdomains), and its tools that are important for SEO, as well as the main ways of ranking a site in the Chinese search engine.

1.2. Very “rich” Search Engine Results Page

In 2014, while Google reduced Authorship and author photos, Baidu increased the display of visuals in Search Engine Results Page (SERPs). Today, with the new update of the search engine, an image is often displayed alongside the snippet, producing very loaded search result pages. Of 10
results, half have an image alongside the meta-description. These images are extracted automatically from the sites.

Images in SERPs tend to attract users and increase the “click-through rate” in SEO. A high click-through rate has a positive effect on Chinese SEO.

On Baidu, the majority of icons and visuals must be paid for. SEO elements are supposed to be free, and are generated with JavaScript or HTML code, so that they appear in the search results (for example Rich Snippets with Schema.org or Authorship in 2013–2014 on Google).

We will see further on in this book that microdata do not work for the Chinese search engine.

The itemprop and conditions tags of the Google Schema.org site do not work on Baidu. The Chinese search engine does not take them into account.

In reality, it is the new SEA format “Baidu Shantou (百度闪投推)” that makes it possible to generate Rich Snippets by paying for the service. We only need to create an SEA campaign with Baidu Shantou to display an image or event in our snippets.

We wondered how images could be generated in the Baidu SERP, and defined the following conditions:

1) use photos of around $121 \times 75$ pixels;
2) name image files with keywords, product names, brands, etc.;
3) fill the alt attributes of images with popular keywords;
4) place keywords in the textual content around the image, relating to the alt attribute of the image.
This service was created in China in 2014, and many major brands are beginning to get involved. There is documentation of this in Chinese. However, for logos, there is a different method for generating the image in the Baidu search results. This is done in Webmaster Tools. This subject will be discussed in Chapter 5.

1.3. Baidu versus HaoSou and Sogou

As previously mentioned, the most used Chinese search engines in China are Baidu, HaoSou (formerly 360) and Sogou.

Since Google left China in 2010 for reasons relating to censorship and China’s non-transparency in information processing, the American search engine has hardly been used at all. We are going to compare the market shares in 2010 and 2014 to show what happened in 4 years.

In August 2010, Baidu clearly had the monopoly, with almost 80% of the market share. Nonetheless, Google China had 12% of the market share, outperforming Sogou (HaoSou did not yet exist).

However, between 2010 and 2014, there were many changes in search engines in China: Google China left the Middle Kingdom due to censorship problems in 2010\(^1\), and the search engine 360 (the future HaoSou) arrived on the market in 2012.

The dispute between 360 and Baidu was tense to the extent that searching “360” in Baidu produced a warning.

---

\(^1\)“Google Chine quitte la Chine”, see https://fr.wikipedia.org/wiki/Censure_d'Internet_en_R%C3%A9publique_populaire_de_Chine.
message from the number one Chinese search engine in 2014.

This warning, generated by Baidu, referred to the violation of the 20th law of the Ministry of Industry and Information Technology in China. The message revealed that 360 was preventing Internet users from downloading Baidu’s antivirus, while promoting its own.

In order to ensure the “freedom” and autonomy of Internet users in their choice of security software installation, Baidu asked those who wanted to install their antivirus to uninstall 360’s first.

In 2014, a Chinese web marketing expert told us that Baidu and 360 regularly met in court over unfair competition. This battle divided Internet users more and more, and 360 began to eat away at Baidu’s market shares.

Sogou has also progressed and provides unique services that neither Baidu nor HaoSou (formerly 360) have integrated into their search engines, such as WeChat messages (the number one social network in China, an instant messaging site), thus outperforming Weibo (Chinese equivalent of Twitter and Facebook since 2014–2015).

When we click on 微信 (WeChat), we reach the subdomain http://weixin.sogou.com, where we can read articles that have been posted.

Clicking on one of these articles redirect us to the “mobile” WeChat URL. We can also obtain WeChat articles indexed in the Sogou search engine. It is therefore possible to search for articles posted in the Chinese social network.

A notice then appears: WeChat is intended to be a closed/private space for sharing information with contacts. But if the articles come up in the Sogou search engine, can
WeChat still be considered a space that is “private between personal contacts”?

When we click on one of these accounts, we are redirected to a page with an introduction to the account and recent articles posted online by this account.

Thus, to effectively rank a brand or a site on the Sogou search engine, we must consider moderating our WeChat account and optimizing it as much as possible so that it appears prominently in the search results. Chapter 6 will explain how to optimize a WeChat account.

The popularity and update frequency of the account may have an effect on its position in the results. Baidu and HaoSou may provide this function in the future.

Moving away from Baidu, HaoSou and Sogou seem for the moment totally unknown to the European and Francophone market.

However, if we wish to improve the visibility of sites in China, these search engines must be taken into account. The aim of this book, though, is to reveal the tricks of the trade for indexing and ranking effectively on Baidu.

1.4. Baidu’s services

As well as the sponsored links on Baidu mentioned previously (section 1.1), the Chinese search engine frequently promotes its own subdomains in its search results. For generic searches and branding searches alike, some sites belonging to Baidu will always be found in the results.

The Chinese Wikipedia, Baike Baidu, always appears in the search results and is often ranked very high – between the first and fifth positions on the first page of the SERP.
Zhidao Baidu (a sort of Chinese Yahoo! with questions/answers), Tieba Baidu (forums), Wengou Baidu (shopping) and Fanyi Baidu also appear on the first results page.

This is why, between Baidu’s paid-for search results and the fact that the search engine promotes its own sites in the first results page, it is very difficult to appear in the first SERP of the Chinese search engine. However, because of the techniques that will be presented in the course of this book, it is not impossible.

### 1.5. Eye tracking on Baidu versus Google

When a Chinese SEO service is to be performed, it is useful to know where Internet users look first in the search results in order to identify and target the best position in the SERPs:

– So which are the “strategic” areas that Chinese people look at in Baidu search results?

– What is the difference compared with Google?

On eye tracking screens, green clouds and red dots show that Chinese Internet users look at the Baidu SERP as a whole, whereas Internet users on Google look only at the beginning (Figure 1.2).

However, one similarity is observed: Internet users look first of all at the first two search results (yellow).

This is partly because the first results come from Baidu subdomain sites and are not necessarily the sites Internet users are looking for. Baidu subdomains are often forums and questions/answers sites like Yahoo!, as we have seen previously. Official information and corporate Websites are not provided in these areas.
For Baidu, we therefore recommend targeting the first 10 spots on the first results page, as visibility can be gained. The rest of the book will explain the techniques required to achieve this.

Figure 1.2. Comparative eye tracking of Chinese Internet users (source: Autoveille). For a color version of this figure, see www.iste.co.uk/duong/SEO.zip

1.6. How does “BaiduSpider” work?

BaiduSpider is a robot (or bots) that sorts and indexes data provided by Internet users. In an official report on the functioning of its crawler, Baidu reveals how BaiduSpider works. The article is very detailed, and it is clear from
reading it that Baidu primarily uses keywords to rank pages. Here is the title of the report in Chinese: SEO的系统化策略.

It is observed that the functioning of BaiduSpider, shown in the shape of a funnel (Figure 1.3), distinguishes between many elements and places them in a hierarchy during the indexing process:

– the database of keywords (red);
– the content and the webpage (blue);
– the crawl by the crawler (green);
– the indexing (violet);
– the positioning and the distributed traffic (yellow).

![BaiduSpider, sorting and hierarchization of data. For a color version of this figure, see www.iste.co.uk/duong/SEO.zip](https://zhanzhang.baidu.com/college/articleinfo?id=837)

Baidu believes that SEO can be summarized in the following three points:

1) the Internet user searches for a piece of information on Baidu, which gives him/her a list of responses corresponding to his/her search;

2) SEO specialists try to attract Internet users to their sites using various methods and SEO strategies;

3) Baidu users become users of the websites they have visited.

Baidu also separates the elements that do or do not carry weight in terms of SEO and influence the crawling and navigation speed.

Those which carry weight are as follows:
– internal network;
– mobile agent;
– on-site optimized elements adapted according to the Baidu guidelines;
– external links.

Those which do not carry weight, but are, nonetheless, significant in a Chinese SEO strategy:
– sitemap;
– Ping;
– manual indexing.

Internal and external links are important for Baidu, because an external link corresponds to a “vote” for the site in question. The more “votes” a site has, the more popular it is for search engines. We will return to this aspect throughout this book.

NOTE.– Since the launch of BaiduSpider version 3.0 in 2016, the indexing rates of sites with a low page volume have fallen by 62%.
1.7. Understanding the difference between crawl and indexing on Baidu

When discussing SEO, individuals often confuse crawl and “indexing”. Baidu has explained, in the form of an infographic, the difference between these two terms.

The objective of the crawl is as follows:
– accessibility of the site for indexing pages in a database;
– compatibility with mobile devices;
– 301 redirects correctly set up.

Baidu’s tips for benefiting from an optimal crawl of the site include having a stable server and a well-structured robots.txt.

The objective of indexation is as follows:
– attract traffic to the site;
– continually refresh the pages and content of the site.

Baidu’s tips for a correct page indexing rate include designing an easily accessible site, setting up a sitemap.xml, having a good internal network and submitting site links to the search engine.

To monitor the crawl, the Baidu logs must be checked every day to ensure that the site remains accessible by the search engine.

To monitor indexing, the Baidu Webmaster Tools rubric can be used to check that all site pages are properly indexed.

3 See http://zhanzhang.baidu.com/indexs.