

Methodology for Information Retrieval

Author: Chiang Min-yen, OTF Fellow at SecDev Group

Studying China has become increasingly challenging as traditional qualitative methods, such as interviews and fieldwork, face growing restrictions. Fortunately, the vast information available in the digital world has opened up alternative pathways, enabling continued exploration of internal social dynamics in China. However, the vast amount of online data can be overwhelming, and valuable insights are not always readily accessible. Without the structured understanding that traditional qualitative methodologies provide, researchers may struggle to organize and make sense of the intricate and often obscure information found online. This note aims to consolidate Information Retrieval methods to assist researchers in related fields, enabling them to derive insightful perspectives from China's online data by leveraging collective experience. Although it may not address all research needs, it serves as a small step toward building a body of collective expertise.

Key challenges in studying China in recent years include:

1. The worsening human rights situation in China has increased the risk of researchers inadvertently violating national security laws, thereby endangering their personal safety.
2. China's authoritarian regime has grown increasingly hostile toward democratic nations, preventing many foreign researchers who value democratic freedoms from obtaining visas to enter China.
3. China's economic environment has become more closed, decreasing the willingness of professionals across industries in democratic countries to work in China, thereby reducing external connections with Chinese society.
4. The Chinese government has intensified surveillance on both its citizens and international communities, leading many informed individuals to refrain from sharing opinions out of concern for personal safety. Furthermore, statements from insiders may not always reflect the full truth due to these pressures.

Using online data to research issues related to China aligns with the increasingly popular methodology in China and authoritarian studies: open-source intelligence (OSINT). In essence, OSINT involves leveraging publicly accessible information to advance research. This approach offers several benefits, with one of the most significant being that, by avoiding sensitive, non-public sources, research findings can be openly shared. This transparency strengthens democratic nations' understanding of China's situation and enables broader collaboration. Additionally, the circulation of open-source information and methodological discussions fosters innovative research practices and creates a foundation of collective knowledge, allowing newer researchers to build on previous findings. Other advantages include the verifiability of open-source data, enabling other researchers to authenticate findings and extract additional insights from the same sources, thus cultivating a productive, competitive research environment.

Given the potential hostility of the Chinese government toward any research effort aimed at understanding internal information, this Information Retrieval note will not necessarily cover all content exhaustively. Instead, it will focus on organizing different types of publicly accessible information sources and their corresponding methods of use, providing a structured guide for researchers navigating these resources. If there are specific sections readers would like to discuss in further detail, please feel free to contact the author (myc96@proton.me). I am a researcher specializing in the semiconductor industry in Taiwan and China, AI chips, and economic security, with a focus on examining China's digital authoritarianism. Therefore, the information I share typically falls within these research domains.

1. Corporate Annual Reports

Annual reports of publicly listed companies offer some of the most comprehensive information available, enabling researchers to systematically understand a company's development trajectory and uncover specific insights essential for critical analysis. Below are several key categories of information that are particularly valuable when organized from financial reports and serve as a foundation for observation and analysis:

a. Business Scope

The scope of operations and product types determine a semiconductor company's position within the global supply chain. Typically, semiconductor firms do not operate across diverse business types; due to resource constraints, most focus on a single or limited set of business areas. Therefore, closely examining a company's historical business scope and product offerings can reveal its place in the broader supply chain network and provide insights into potential relationships with other companies. This understanding helps identify relevant media sources for further research.

b. Revenue Sources

Revenue sources often include income from different product types and operational regions. The former allows researchers to gauge where a company primarily allocates resources, or which product-related divisions hold significant internal weight. The latter is particularly useful for analyzing the proportion of revenue generated from China, helping researchers assess the company's dependency on the Chinese market. This understanding further indicates the level of effort required to explore the company's ties with the Chinese market.

c. Customer and Supplier Contracts

In certain cases, a company's annual report may disclose its primary customers and suppliers, though this is not common. Most semiconductor firms choose not to publish customer and supplier lists to protect trade secrets. However, if a company reveals such lists to boost its credibility and brand recognition, this becomes a valuable information source. Even more valuable is when the report includes transaction amounts or proportions associated with specific customers or suppliers. This data provides a foundation for researchers to seek additional supplementary news sources. If the annual report lacks direct disclosure of customer and supplier information, researchers may still find relevant clues in sections detailing contract terms, which may indirectly reveal such connections.

d. Board of Directors and Management Team

In the unique context of Chinese corporate culture, if a semiconductor company seeks greater access to resources aligned with China's industrial policies, it may aim to publicly demonstrate its collaboration with the Chinese government. For instance, company

executives may appear alongside government officials at project announcements or mention these connections at industry forums. Consequently, the list of board members and professional managers in a company's annual report becomes an essential clue. It enables researchers to identify any potential relationships between the semiconductor company and the Chinese government through publicly available media sources. Although these connections may not be explicitly documented within the annual report, preliminary awareness of these links allows researchers to pinpoint relevant sections and assess whether they are related to official government activities.

e. Corporate Organizational Structure

Many semiconductor companies and related suppliers have established extensive international business networks. This often involves a network of proxy entities or shell companies set up in offshore financial centers, enabling them to create a complex and difficult-to-track organizational structure and global operational network. However, due to certain regulatory requirements for publicly listed companies, these entities may still need to disclose specific details at times. Consequently, the structure of information within a company's annual report may vary yearly. In some years, a semiconductor company might disclose details about its global footprint and organizational structure—valuable information for researchers tracking the company's activities and reach within the supply chain. Even if such information isn't consistently available, having an initial grasp of the structure enables researchers to cross-check subsequent developments through other sources, providing continuity in understanding the company's global operations.

f. Sources for Annual Report Information in Taiwan, Hong Kong, and China

Listed companies typically publish annual reports on their official websites, though the information may not always be complete. In such cases, researchers may need to rely on disclosures mandated by regional securities authorities:

- **Taiwan:** The primary source for listed company information in Taiwan is the [*Market Observation Post System*](#) (MOPS). While the website interface is somewhat basic, with

the English version less organized than the Chinese one, most essential data can be found under the "e-Book" section below the search bar on the homepage.

- **Hong Kong:** Information on listed companies is available through [*The Stock Exchange of Hong Kong Limited \(HKEX\)*](#), where the website is relatively user-friendly, and English information is more accessible.
- **China:** Annual reports for Chinese listed companies are accessible via the [*Shanghai Stock Exchange \(SSE\)*](#) or the [*Shenzhen Stock Exchange \(SZSE\)*](#). Searching in simplified Chinese generally yields better results.

2. Government Registration Data for Non-Public Companies

When researching supply chain topics, researchers frequently encounter non-public companies, which lack the disclosure requirements of publicly listed entities. However, existing company registration systems can still provide useful insights, allowing researchers to explore further through other publicly available information.

a. Taiwan:

Taiwan's Ministry of Economic Affairs offers a free [*Business Registration Inquiry System*](#), where users can access basic company information using the official registered name or company identification number. This includes capital amount, board members, company address, and the names of responsible parties. The board members' list often shows shareholdings and the names of shareholders they represent, which can help infer any potential Chinese affiliations. In some cases, board members may be listed under a non-Chinese foreign status, despite having connections to China, as their names may be recorded in Pinyin. The system also includes a historical data section that documents change in shareholding, enabling researchers to track significant company changes over time and, using this timeline, identify other sources of information to uncover additional details.

b. Hong Kong:

The Hong Kong Companies Registry provides [access](#) to documents submitted during the registration and modification processes of companies registered in Hong Kong. While the government does not offer an open data retrieval system, researchers can obtain documents through a detailed application process, which requires a fee. The information contained in these documents often exceeds the scope provided by Taiwan's government sources. For example, researchers can access a company's articles of association, potentially uncovering details about its decision-making and shareholding structures. Historical documents related to changes in the company may reveal both past and present shareholders and board members.

Since the registry provides original documents rather than consolidated summaries, researchers can sometimes find unintended clues, such as the names of document handlers or timestamps that suggest company activities. However, it is worth noting that Hong Kong's system requires applicants to provide personal information, which may pose certain data security considerations for researchers.

c. China:

The Chinese government does not provide an official channel for researchers to access company registration information. However, numerous online public information platforms (e.g., [Qichacha](#), [Aiqicha](#)) offer extensive company data, surpassing what is available in Taiwan and Hong Kong. These platforms often include comprehensive details, such as full board member lists, historical changes, and associated shareholder and equity transformations, allowing researchers to dissect a company's ownership structure and trace the ultimate sources of capital.

Given the state capitalist nature of China's economy, the government frequently leverages industrial policies or state-owned enterprises to control private sector operations, particularly in emerging technology sectors. Through these platforms, researchers can assess the extent of Chinese government involvement in various semiconductor companies and may even pinpoint the specific policy networks exerting influence.

However, with China's business environment becoming increasingly restricted, these platforms now face growing limitations. These include requirements for user registration,

restrictions on access for users outside China, and other measures that have made it progressively harder to research Chinese companies, further obscuring local industry information from external observation.

3. Non-Mainstream Media and Other Online Information Sources

In the fields of semiconductor industry and economic security research, mainstream media outlets like [Bloomberg News](#), [Reuters](#), and [Nikkei Asia](#) provide reliable reports, including exclusive insights, especially concerning China's tech industry. In addition, industry-focused outlets such as [The Information](#), though not mainstream, offer valuable coverage and occasionally exclusive reports on niche industry developments. These two types of news sources give researchers access to key details that might not be immediately evident in more public information sources. Such insights often serve as inspiration for researchers to explore the larger, more comprehensive picture behind these events.

Mainstream media typically verify exclusive stories through multiple sources before publication, but the immediate nature of news limits journalists' capacity to delve deeply into the underlying narratives. This gap presents an opportunity for researchers to build a fuller story based on available details.

In Taiwan, researchers can leverage respected local industry intelligence agencies to obtain the latest updates on the semiconductor sectors in Taiwan and China or use their archives to analyze historical trends. Key agencies include [TrendForce](#), [TechNews](#), and [DigiTimes](#), which provide extensive news archives and industry intelligence to help researchers identify long-term trends and potential research topics.

Moreover, in the age of widespread social media use, semiconductor engineers and industry professionals frequently engage on these platforms, leaving digital footprints that can offer researchers insights into the industry's public sentiment. While social media content may not always be fully credible, it allows researchers to stay attuned to trends within the tech industry and engineer communities, serving as a guide to meaningful information across other public sources. Common platforms in China include [Zhihu](#), and in Taiwan, [PTT forums](#),

both of which are useful for gauging opinions and sentiment within specific professional circles.