
Carlton Clark, Lei Zhang, and Steffen Roth

What’s trending? Google Trends tracks search trends every day, throughout the day. Social media, smartphone notifications, and unwanted pop-ups keep us abreast of current trends even when we have no interest in what is trending on a particular day or in a particular hour. But we might not realize, or might forget, that those surface trends often have very deep, ancient roots. This chapter considers trends from 1950 to 2008 in China and compares them to trends found in other parts of the world. We begin with remarks on the political theories and assumptions traceable to Periclean Athens and republican Rome. We then compare this history to that of China.

A fundamental tenet of constitutionalism, as shaped by the political writings of Aristotle, Cicero, Spinoza, Locke, Montesquieu, Rousseau, the authors of The Federalist Papers, and of course many others, is that the best way to prevent tyranny is through a separation of powers. Constitutionalism, which is evident in the political systems of Periclean Athens and republican Rome, does not depend on a written constitution and is not limited to constitutional democracy, but it recognizes that “power can only be controlled by power” (Gordon, Controlling the State, 15). As Montesquieu (1748) wrote, “there is no liberty if the judiciary power be not separated from the legislative and the executive” (Montesquieu, Complete Works, 1:199). Similarly, in Federalist 47 (1788), James Madison wrote, “The accumulation of all powers, legislative, executive, and judiciary, in the same hands, whether of one, a few, or many, and whether hereditary, self-appointed, or elective, may justly be pronounced the very definition of tyranny” (Hamilton, Madison, and Jay, Federalist Papers, 245). Constitutionalism, in other words, calls for the differentiation of political functions—in other words, functional differentiation, which is a key concept in this chapter.

Political separation of powers is just one of innumerable kinds of functional differentiation. The human brain, for instance, has through the long course of evolution
differentiated itself according to various functions. Organizations of all types also internally differentiate themselves according to various functions or responsibilities. Moving on to consider society in general, Niklas Luhmann (Reality of the Mass Media; Social Systems; Theory of Society) argues that modern globalized society has differentiated itself into several autonomous function systems. While there are debates on what “counts” as a function system, the existence of at least a core set of such systems is regarded as a key concept of modernity (Berghaller and Schinko, “Introduction”; Brier, “Construction of Knowledge in the Mass Media”; Jönhill, “Inclusion and Exclusion”; Kjaer, “Metamorphosis of the Functional Synthesis”; Leydesdorff, “Communication Turn”; Schirmer and Hadamek, “Steering as Paradox”; Vanderstraeten, “System and Environment”; Ward, “Functional Differentiation”). Roth and Schütz (“Ten Systems”) make the case for ten function systems: economy, politics, law, science, religion, education, health, art, mass media, and sport.

Contemporary globalized society protects the operational autonomy of the function systems. When we speak of the separation of church and state, freedom of the press, or academic freedom, or when we promote universal basic education or health care, we are speaking of functional differentiation. Thus, we invoke a “semantics of corruption” (Newbury, “Semantics of Corruption,” 163) when one area of society unduly influences another area, as when money corrupts politics, when religion infringes on science or education, or when politics and the mass media seem to become one. We may judge the perceived importance of functional differentiation by surveying the literature on functional de-differentiation. To cite a small sample of these “-izations,” scholars have analyzed the politicization of art and the aestheticization of politics (Benjamin, Illuminations); the politicization of education (Roper, “Politicization of Education”; Wirth, Whiddon, and Manson, What Is Wrong with Academia Today?); the economization or commercialization not only of education (Fazl-E-Haider, “Commercialization of Education”; Spring, Economization of Education) but of health care (Ewert, “Economization and Marketization”), and of culture and the liberal arts (de Valick, “Film Festivals”; Eikhof and Haunschild, “For Arts Sake!”; Fludernik, “Threatening the University”); the mediatization of politics (Esser and Strömbäck, Mediatization of Politics; Kepplinger, “Mediatization of Politics”); the politicization, economization, and mediatization of religion (Robertson, “Economization of Religion?”; Thomas, “Mediatization of Religion”); the politicization of science (Bolsen and Druckman, “Counteracting the Politicization of Science”); the mediatization of culture and society (Hjarvard, Mediatization of Culture and Society; Mazzoleni, “Mediatization of Society”; McLuhan, Gutenberg Galaxy); and, of course, the economization of everything (Lamont and McGuirk, “Introduction”).

In modern China, functional differentiation has unfolded differently than in other parts of the world. A thorough analysis of the factors behind this difference
is well beyond the scope of this chapter. Briefly, however, we can observe that the ancient Chinese dynasties established a pyramidal social structure—“a massive base, with higher layers of strongly diminishing size, and an apex of one single person” (Gassmann, “Through the Han-Glass Darkly,” 533)—which is still recognizable in contemporary China. China also has a philosophical and intellectual heritage based on the teachings of Confucius (or Kong Fuzi), Mozi, Mencius, Zhuangzi, and Laozi (the reputed author of the *Dao De Jing*), along with the Legalist thinkers. These “schools of thought,” which became schools of thought only retrospectively, share a few core assumptions: “They regarded disunity and chaos as a problem to be solved, they assumed that patriarchal hierarchy is natural to families and for states, and there is a natural ‘Way’ (dao) of Heaven, and that those who discover and follow the Heavenly Way will be successful” (Tanner, *China*, 67). A key principle of Daoism is that change is natural and inevitable but, ideally, orderly. Thus, good leadership, whether of the family or the state, calls for the orderly management of change—an ideal embraced by the Chinese Communist Party (CCP).

In other words, while thinkers such as James Madison saw competing interests and even a degree of disorder as the antidote to tyranny, Chinese thinkers and leaders have, for millennia, promoted orderly, hierarchically managed change. Thus while functional differentiation, which generates autonomous, self-reproducing social systems with no central steering mechanism, is happening in contemporary China, it goes against the grain of an ancient cultural heritage. This bias for orderly change and top-down social control is observable in the modern Chinese subcorpus of the Google Books Ngram Viewer, a free online graphing tool which charts annual counts of words and word sequences as found in the largest available corpus of digitized books.

Although this study uses the Google Books Ngram Viewer to analyze trends in ten function systems, it is essentially a study of just one system—the mass media system—because books (print and digitized) represent a large part of the mass media. Indeed, the mass media system arguably began with the fifteenth-century invention of the printing press (Luhmann, *Reality of the Mass Media*; McLuhan, *Gutenberg Galaxy*). According to Luhmann (2007), everything known to contemporary society derives from the mass media.

**The Google Books Ngram Viewer**

One way to measure the evolution of functional differentiation, as well as movements toward de-differentiation, is by analyzing the contents of a large corpus of digitized books. Roth et al. ("Futures of a Distributed Memory") used the Google Books Ngram Viewer to engage in a form of computational sociology known as culturomics, “a field of study that uses numerical analysis of large volumes of data to investigate culture” (Matthews, *Encyclopedia of Environmental Change*). Michel
et al. (“Quantitative Analysis of Culture”) explain their creation of the Google Books Ngram Viewer as follows:

We constructed a corpus of digitized texts containing 4% of all books ever printed. Analysis of this corpus enables us to investigate cultural trends quantitatively. . . . [T]his approach can provide insights into fields as diverse as lexicography, the evolution of grammar, collective memory, the adoption of technology, the pursuit of fame, censorship, and historical epidemiology. Culturomics extends the boundaries of rigorous quantitative inquiry to a wide array of new phenomena spanning the social sciences and the humanities.

This tool allows users to search a large corpus of books and obtain a visual representation of cultural trends, the Ngram. While Google Books contains over 15 million digitized books, the creators of the Ngram Viewer selected 5,195,769, or approximately 4 percent of all books ever published. These books contain well over 500 billion words, including English (361 billion), French (45 billion), Spanish (45 billion), German (37 billion), Chinese (13 billion), Russian (35 billion), and Hebrew (2 billion).

The Ngram Viewer produces time series graphs, also known as time plots, which is a data visualization tool that illustrates data points at successive intervals of time. Each point on the chart corresponds to both a time and a quantity that is being measured. These graphs are used to identify, model, and forecast patterns and behaviors in data sampled over discrete time intervals. The Ngram Viewer allows us to compare the levels of communication on economic topics, as well as political, legal, educational, scientific, artistic, and religious topics, and so forth. Using this tool, Roth and colleagues analyzed key concepts of functional differentiation in the English, Spanish, French, German, Italian, and Russian subcorpora. Each of the ten function systems was represented by five keywords.

The selection of the five most important keywords per function system was a multistep mix-methods process. First, we relied on a small collection of Python scripts that generate word frequency lists based on the Google Ngram dataset. In our case, we created lists of the 10,000 most frequent words per investigated language area. We then manually scanned these lists for words that refer to one and only one of the 10 function systems, whereby each list was screened by at least two colleagues. The major challenge in this context was to identify n-grams that unambiguously refer to not more than one function system. . . . We then picked the five most frequent keywords per function system and combined them to strings such as (business + economic + money + company + cost). If entered into the Google Ngram Viewer, each such string creates one single graph that represents the combined performance of all keywords, which in this case presents the combined performance of the five strongest indicators for the economy. (Roth et al., “Futures of a Distributed Memory,” 7–8)
What's Trending in the Chinese Google Books Corpus?

For this analysis, the sample period was limited to Google Books between 1800 and 2000, because the data is most reliable after 1800 and because the research concerned macrotrends in modern societies.

This chapter covers the simplified Chinese subcorpus between the years 1950 and 2008. It offers insight into whether Chinese society between 1950 and 2008—not just in the People’s Republic of China but among all speakers of Chinese as a language—experiences significant trends in discourse related to politics, law, economy, education, science, mass media, religion, education, art, or sport—the ten function systems covered in Roth and Schütz (“Ten Systems”). Our contention is that the Google Books ngrams discussed below reveal traces of social change in modern China. We will first discuss a few of the problems we encountered in charting the Chinese-language Google Books subcorpus. Next, we will offer interpretations of the observed trends. Finally, we discuss a number of questions that arise from the ngrams and that may be explored in further research.

Methods

The first methodological problem arose from the fact that following the founding of the People’s Republic of China (PRC) in 1949, the Chinese government, building on earlier efforts, revised the traditional Chinese script to create simplified Chinese. Google has predominantly scanned books written in simplified Chinese. Therefore, rather than starting at the year 1800, as in Roth et al. (“Futures of a Distributed Memory”), this project considered only books published from 1950 to 2008. The second problem was that it is very difficult to create a reliable Chinese word-frequency list. To help people learn entry-level Chinese, attempts have been made to create such lists. For example, in 1988, the National Working Committee on Languages and Writing Systems (国家语言文字工作委员会) and the Chinese Ministry of Education created a List of Frequently Used Characters in Modern Chinese (现代汉语常 用字表), which includes 3,500 characters. These characters are divided into groups based on the number of strokes comprising each character. Yet it is not known how these characters are ranked among themselves. Another list, the Modern Chinese Character Frequency List (现代汉语单字频率列表), has been compiled by Professor Jun Da. This list contains 9,933 unique characters. These two lists, unfortunately, are not useful for our purposes because single characters can take on completely different meanings when joined to one or more other characters to form words or phrases.

Passing beyond character frequency, Chinese word-frequency information is even more difficult to compile because, unlike writing systems in Western languages, Chinese words are not segmented by white spaces. A single Chinese word or phrase may contain up to five characters. Researchers have not yet developed software that can reliably segment Chinese words and phrases. The software might misread a meaningless combination of characters as a word: a problem we encountered
when using the Google Books Ngram Viewer. Consequently, we could not use any existing word-frequency or character-frequency list.

Roth et al. ("Futures of a Distributed Memory") started with word-frequency lists and then used a multistep, mixed-method process to find the top five keywords for each function system in each language area. When it came to Chinese, without a reliable word-frequency list, one option could have been to look for the best translations of the key terms used in Roth and colleagues’ work. But that method was unworkable because each language has its own keywords for each function system—in other words, the function systems are the same but the keywords are different. To be more precise, we should speak of lexemes rather than words. Although there is some debate among linguists regarding the meaning or significance of this concept (Bonami et al., *Lexeme in Descriptive and Theoretical Morphology*), a lexeme may be defined as a unit of language “intermediate between morpheme and utterance” (*Oxford English Dictionary*). It can be a single word, a part of a word, or a chain of words (e.g., *dog, stop sign, by the way, raining cats and dogs*) that forms a single basic meaningful element in the vocabulary of a language. Thus to solve our methodological problem, another option was to settle on a set of Chinese lexemes that would seem, in the judgment of well-educated native speakers, to fit the function systems; but such an effort would still be too arbitrary.

After casting about for solutions, we decided to start with the best translations for the function systems themselves: economy, politics, law, education, science, mass media, art, and so on. But a single Chinese lexeme, even when it consists of two, three, or four characters, can have many different meanings when combined with adjacent characters, and there is no white space between these adjacent lexemes, or meaningful units, of a sentence. A further difficulty is that to aid in the processing of characters, the Ngram Viewer inserts white space where it does not actually exist in the scanned Chinese texts. In fact, sometimes the user must insert artificial white space between characters of a single lexeme in order to make the Ngram Viewer accept a legitimate lexeme—a time-consuming, frustrating process.

Confirmation that single-character Chinese lexemes, or 1-grams, are not useful for Google Ngram searches comes from Yang et al. ("N-gram Statistics in English and Chinese"), researchers at Google. They compared English and Chinese ngrams and found that while 1-grams in Chinese are very different from 1-grams in English, multigrams (3-grams and up) share nearly identical frequency distributions in English and Chinese. For example, Jun Da translates the character 学 as learn/study/science/-ology. If we try to compare the frequencies of this character and the English lexeme *science*, we will not get a useful result. We get a more useful result if we compare a three- or four-character Chinese lexeme with a three- or four-word English lexeme. This finding suggests that we should look at longer strings of Chinese characters. As shown in Figure 12.7, many of the Chinese lexemes we used consisted of four characters. For example, one of the lexemes we used under “Economy” was 市场经济, which translates as market economy.
Knowing, then, that translations of single words (1-grams) like *politics*, *economy*, and *law* alone would be inadequate, we started experimenting with Google’s wildcard search function. The wildcard search works as follows: When one puts an asterisk in place of a word, the Google Ngram Viewer displays the top ten substitutions for the asterisk. To illustrate this function, we performed the following wildcard search for the English word *science* from 1900 to 2008. When we place the asterisk in front of *science*, we get the chart in Figure 12.1, and when we put the asterisk after *science*, we get the chart in Figure 12.2. We may discount the determiners, prepositions, conjunctions, and verbs because they do not produce significant 2-grams, leaving the top four two-grams from Figures 12.1 and 12.2 in Figure 12.3.

In Roth et al. (“Futures of a Distributed Memory”) we found (through the method described earlier) that in English, the top five keywords for the science system were *system*, *method*, *theory*, *research*, and *analysis*. When the user inserts plus signs between lexemes, the Ngram Viewer combines all the terms into a single time-series plotline. In Figures 12.4 and 12.5, we compare the method used in Roth et al. with the wildcard method.

Because single-word lexemes (1-grams or unigrams) will occur more frequently than two-word lexemes (2-grams or bigrams)—for instance, *theory* (one
of the science keywords used in Roth et al. occurs over forty-five times more frequently than *social science* (see Figure 12.6)—the frequency percentages are over 100 times greater in Figure 12.4 than in Figure 12.5. Consequently, we could not fit both plotlines in the same chart. But if we ignore the difference in magnitude and focus on the trends, which are more important, we see a striking similarity. Both lines rise steadily from 1800 to 1900, and then more sharply (that is, the rising trend

---

**Figure 12.3.** Top four 2-grams produced from "science" wildcard searches.

<table>
<thead>
<tr>
<th>2-gram</th>
<th>Percentage of 2-grams in Google Books English language corpus for the year 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social science</td>
<td>0.0003221600</td>
</tr>
<tr>
<td>Science fiction</td>
<td>0.0002422460</td>
</tr>
<tr>
<td>Political science</td>
<td>0.0002263571</td>
</tr>
<tr>
<td>Modern science</td>
<td>0.0001327003</td>
</tr>
</tbody>
</table>

---

**Figure 12.4.** Ngram Viewer search with plus signs between single-word lexemes, resulting in a single time-series plotline. Google Books Ngram Viewer, http://books.google.com/ngrams.

**Figure 12.5.** Ngram Viewer search with plus signs between two-word lexemes, resulting in a single time-series plotline. Google Books Ngram Viewer, http://books.google.com/ngrams.
accelerates) through most of the twentieth century. In Figure 12.5, the angle of the rising trend increases markedly in the 1950s, and nearly as much in the same years in Figure 12.4. In Figure 12.4, the trend peaks in 1978; in Figure 12.5, it peaks in 1984. Both lines then decline, for unknown reasons, until the 2008 cutoff. From the similar plotlines in Figures 12.4 and 12.5, we concluded that the wildcard search method would be a valid substitute for the method used in Roth et al., which, as explained above, would not work at all for Chinese.

Now for the Chinese-language area. As stated above, we found that there was no perfect method, but we found the most promising method was as follows. First, a native speaker determined the best Chinese translations for the ten English names for the function systems: economy, politics, law, science, education, health, sport, art, mass media, and religion. Second, because (as explained above) we knew these translations alone would be inadequate, we used the wildcard search method. Additionally, given the processing limitations of the Google Ngram Viewer, we decided that a string of four Chinese lexemes, rather than five as used in Roth et al., would be adequate.

**Results**

Figure 12.7 lists the ten function systems with their lexeme combinations in Chinese and English, ranked by the percentage of occurrence in the Google Books simplified Chinese subcorpus.

Figures 12.8, 12.9, and 12.10 are the Chinese ngrams for the years 1950–2008 for the ten function systems (broken into three sets) ranked by the frequency of occurrence in the Google Books simplified Chinese subcorpus. Below, we also compare the rankings for the year 2000 in the English, Spanish, and French languages from Roth et al. (“Futures of a Distributed Memory”) with the Chinese rankings for the same year. The reader may refer to the article by Roth et al. for the rankings of the Italian, German, and Russian language areas.
<table>
<thead>
<tr>
<th>Function system</th>
<th>Lexeme combination</th>
<th>Percentage in 2008 among all lexemes in Google Ngram Viewer Simplified Chinese Corpus (rounded to five decimal places)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economy</td>
<td>市场 经济 + 国民 经济 + 经济发展 + 经济增长 Market economy + national economy + economic development + economic growth</td>
<td>0.09763</td>
</tr>
<tr>
<td>Science</td>
<td>社会 科学 + 科学 发展 + 科学 技术 + 科学 研究 Social science + scientific development + science &amp; technology + scientific research</td>
<td>0.03240</td>
</tr>
<tr>
<td>Law</td>
<td>法律 的 + 法律 法规 + 法律 制度 + 有关 法律 Legal + legal regulations + legal system + relevant law</td>
<td>0.01891</td>
</tr>
<tr>
<td>Education</td>
<td>教育 的 + 高等教育 + 教育部 + 义务教育 Educational + higher education + Ministry of Education + compulsory education</td>
<td>0.01491</td>
</tr>
<tr>
<td>Politics</td>
<td>中央 政治 + 政治 体系 + 民主 政治 + 政治 家 Central politics + political system + democratic politics, politician</td>
<td>0.00711</td>
</tr>
<tr>
<td>Health</td>
<td>医疗 机构 + 医疗卫生 + 合作 医疗 + 基本医疗 Medical institutions + medical hygiene + cooperative medical treatment + basic medical care</td>
<td>0.00710</td>
</tr>
<tr>
<td>Mass media</td>
<td>新闻媒体 + 新媒体 + 网络媒体 + 传统媒体 News media + new media + internet media + traditional media</td>
<td>0.00326</td>
</tr>
<tr>
<td>Art</td>
<td>文化艺术 + 文学艺术 + 艺术作品 + 艺术形式 Culture and art + literature and art + art work + forms of art</td>
<td>0.00313</td>
</tr>
<tr>
<td>Religion</td>
<td>宗教 信仰 + 宗教 活动 + 民族宗教 + 宗教事务 Religion + religious activities + national religion + religious affairs</td>
<td>0.00300</td>
</tr>
<tr>
<td>Sports</td>
<td>体育 活动 + 体育 运动 + 体育事业 + 国家体育 Sports activities + sports + sports career + national sports</td>
<td>0.00113</td>
</tr>
</tbody>
</table>

Figure 12.7. Ten function systems in simplified Chinese, year 2008.

**Discussion**

In Figure 12.8, Economy sits at the top in 1950 and 2008; however, it fell to second position between 1954 and 1969. The graph indicates a sustained rise of economically (or commercially) oriented communication from 1970 to 1995. The rise in the Economy line accelerates from about 1987 to 1995. It drops after 1995, but still remains well above Science and Law in 2008. In trying to interpret this trend, we note that from 1978 until his retirement in 1992, Deng Xiaoping, as Paramount
Leader, oversaw major economic policy reforms, which remained in place under the PRC presidency of Jiang Zemin (1993–2003) and Hu Jintao (2003–2013). The decline from 1995 to 2008 is difficult to interpret, but it suggests that the economic reforms were no longer news. This decline, however, warrants further research. Science and Law, the other two lines, are essentially tied in 1950. Science immediately rises but then declines from 1963 to 1971, concurrent with the Cultural Revolution.6
Following 1971, the Science line rises and then plateaus during the 1980s. This trend is consistent with the historical record. One of the lexemes for Science is *science and technology*. Clearly, with the exception of most of the 1960s, the CCP saw science and technology as essential to China’s emergence as a world power. Law does not begin to rise until 1977 but maintains a steady rise before plateauing in 2004. The legal system has gained in importance as China’s economy has developed, as commercial activity relies on the stability provided by law. Following the Cultural Revolution, Deng Xiaoping and other reformists recognized that a stronger legal system was necessary for continued economic growth and social stability. As Tanner writes,

> One of their first priorities in 1979 was to have the National People’s Congress approve China’s first post-1949 Criminal Law and Criminal Procedure Law. . . . A number of other laws followed during the 1980s, many of them designed to deal with the emerging market economy: the Inheritance Law (1985), Civil Law (1986), and Bankruptcy Law (1986), laws concerning private enterprises, joint enterprises, foreign investment, business taxes, and so on. (Tanner, *China*, 551)

We now move on to Figure 12.9, where Education sits at the top from 1950 to 2008. The Education line fell between about 1957 and 1975, which may again be associated with the Cultural Revolution. Education then rose steadily after 1975, peaking in about 2001. With the exception of the Cultural Revolution years, education has always been a high priority in China. The second interesting line in Figure 12.9 is Health, which rose from near the bottom in 1950 to second position in 2008, just above Politics. The rise in the Health line from 1965 to 1973 may be related to the Barefoot Doctors initiative, which brought affordable health care to the countryside. Until the mid-1960s, Chinese health care resources had been directed mostly to urban areas. But according to Zhang and Unschuld,

> Mao Zedong criticised the urban bias of medical services and pointed out the stress placed on rural areas in 1965, [and] mobile teams of doctors from urban hospitals were sent to deliver health care and train indigenous paramedics. . . . [T]he barefoot doctor programme effectively reduced costs and provided timely treatment to the rural people. . . . Reforms in the health-care system in the early 1980s . . . resulted in the collapse of the cooperative medical system to a payment-based system of medical care in rural areas. The percentage of villages with a cooperative medical system fell from 90% in the 1960s to 5% by 1985. (Zhang and Unschuld, “China’s Barefoot Doctor,” 1865–66)

The decline between 1973 and the early 1980s may reflect the collapse of the cooperative medical system. This line rose sharply from the late 1980s to 2008. It is
What's Trending in the Chinese Google Books Corpus?

It is noteworthy that three of the four lexemes for the Health function system are *medical hygiene, cooperative medical treatment,* and *basic medical care,* terms clearly associated with the rural healthcare initiative. Continuing with Figure 12.9, Politics ranked second in 1950 but fell to third, behind Health, in 2008. Thus, Politics ranks sixth among the ten function systems. This finding contrasts with the prominent position of Politics in previous Google Ngram studies. For instance, in Roth et al. (*Futures of a Distributed Memory*), the political function system has ranked first in the English, Spanish, French, German, Italian, and Russian language areas since at least the early twentieth century: in the Spanish language area since 1840, and in the English area since 1880. The difference between first and sixth is significant. But given that the CCP, like Chinese rulers in the past, does not encourage the general Chinese public to take an active interest in politics (Tanner, *China*), the comparatively low position of the Politics line makes sense. Nonetheless, the stark contrast between the positions of Politics in the modern Chinese Google Books subcorpus and the English, Spanish, French, German, Italian, and Russian language areas warrants further study. At the bottom of Figure 12.9, the Mass Media line is interesting because although the function system does register very slightly in the early 1950s, the line only really appears in 1993 and then rises steadily up to 2008. The rise of mass-media-oriented communication may be related to a loosening of the CCP’s oversight, allowing a relatively independent print news industry which accompanied economic reforms (Clark and Zhang, “Grass Mud Horse”; Tanner, *China*). But the rise of the Mass Media line is also clearly related to the emergence of the internet, as indicated by three of the lexemes under Mass Media: *new media, Internet media,* and *traditional media.* On April 20, 1994, a 64K dedicated international line for the internet, with full functional accessibility, was set up in China with strong backing from the government (Weishan, Hongjun, and Zhangmin, “Who’s in Charge of Regulating the Internet?”). In July 2008, China surpassed the United States in the total number of internet users (Barboza, “China Surpasses U.S.”).

We now move to Figure 12.10, which features the bottom three function systems—Art, Religion, and Sport. Here the Art system rises from second in the early 1950s to first in 2008, just slightly above Religion. The Art line rises dramatically between 1953 and 1962, it then falls just as sharply between 1964 and 1971, during the Cultural Revolution. Art then rises again in 1974, peaking in 1981. During the Cultural Revolution, propaganda pieces were created anonymously and collectively, and most of it was destroyed or thrown away after Mao’s death in September 1976 (Zheng, “Brushes Are Weapons”). However, some of the art was preserved, and in 2002 an exhibition titled “Art of the Great Proletarian Cultural Revolution, 1966–1976” was held in Vancouver, Canada (King et al., *Art in Turmoil*). As for Religion, this trend line sits in second place in 1950 and remains there in 2008. Religion dropped between 1962 and 1975, but it has risen consistently since 1975. This rise probably relates to the general opening up of Chinese society in the post-Mao
decades. In tenth position is Sport, which fluctuates quite a bit over the years. This system also ranks last among the ten function systems in each of the language areas covered in Roth et al. (“Futures of a Distributed Memory”).

We close this section with a few comments on Figure 12.11. The comparison of the English, Spanish, French, and Chinese language areas, as presented by Google Books Ngram Viewer, shows that Politics and Economy essentially exchange positions. In 2000, Economy and Politics rank first and fifth, respectively, in the Chinese ngrams. In contrast, in English, Spanish, and French, Politics ranks first and Economy no higher than fourth. To save space, we have not included German, Italian, and Russian here; but in Roth et al. (“Futures of a Distributed Memory”), Politics ranks first in each of those languages and Economy no higher than third.

These findings indicate that, at least between 1950 and 2008, economic issues were discussed far more frequently in books published in Chinese than in books published in English, Spanish, French, German, Italian, and Russian; whereas political issues were discussed far less frequently in Chinese books. Another area of contrast is Education. In Figure 12.11, Education ranks fourth on the Chinese list but no higher than seventh in English, Spanish, and French. The rankings for Education are similar in German, Italian, and Russian, as shown in Roth et al. We do not know, however, how much of the education-related discourse in China is patriotic education and how much is conventional academic education. Health also ranks higher on the Chinese list; it ranks sixth in the year 2000 (rising to fifth in 2008) on the Chinese list and eighth, ninth, or tenth in other languages.

This chapter has analyzed trends in communication, specifically through the medium of books, in post-1949 China, relating to ten function systems identified in social systems theory: economy, politics, law, science, religion, education, health,
art, mass media, and sport. By charting these trends with the aid of the Google Books Ngram Viewer, we have been able to track social change in modern China. This study builds on previous research on the English, Spanish, French, German, Italian, and Russian subcorpora of Google Books, but makes a unique contribution by using the wildcard search function to get around the difficulties presented by the Chinese writing system.

The Google Books Ngram tool offers suggestive results that invite further, more fine-grained research. That is to say, rather than being an end in itself, this kind of big-data research generates new research questions that may be explored through other methods. For example, why does political communication (the Politics trendline) rank significantly lower on the list of function systems for the modern Chinese language area, compared to its position in English, Spanish, French, German, Italian, and Russian, as found in Roth et al. (“Futures of a Distributed Memory”)? This study offers evidence that books published in Chinese from 1950 to 2008 discuss political topics less frequently than do books published in other major world languages. But books, of course, are only one medium among many, and a very small percentage of people actually publish books. Are political topics also discussed less frequently outside of books among Chinese speakers, in comparison with speakers of these other languages?

Researchers might investigate the following questions as well. Is commercially oriented communication (the Economy function system) actually more common or more valued among Chinese speakers as compared to speakers of other major world languages? Put more simply, are the Chinese people more interested in economics than in politics? Further, the Education function system ranks high in the Chinese language area, but to what degree does this trend reflect traditional educational priorities versus Communist Party–directed patriotic education. Regarding Law, what factors lie behind the increasing importance of a less politicized legal system in contemporary China? In terms of the Health system, did the resources devoted to affordable, universal health care actually rise sharply during 1963–1973 (when other areas of Chinese society suffered under the Cultural Revolution) and then decline significantly between 1974 and 1983? And if so, why? These are just a few of the many questions provoked by this kind of big-data research.

Notes

One author gratefully acknowledges financial support from the Research Council of Lithuania and the European Regional Development Fund-ERDF/FEDER (National RandD Project 01.2.2-LMT- K-718–02-0019 “Platforms of Big Data Foresight [PLATBIDAFO]”).

1. The German Funktionssystem is translatable as “function system” or “functional system.” Both terms appear in English translations of Luhmann’s work.

2. Arguably, in this context a better word than “economy” is “commerce,” a word that avoids associations with thrift or the wise management of limited resources.
“Commercialization” is also preferable to “economization” because the latter already has an established definition, “the action or process of economizing” (*Oxford English Dictionary*). Besides “commerce”, another option is “market economy,” a term adopted by researchers such as Valentinov. But in the publications cited in this section, “economization” signifies the economic takeover of traditionally noncommercial areas of society.

3. Walter Benjamin declared, “All efforts to render politics aesthetic culminates in one thing: war” (*Benjamin, Illuminations*, 243).

4. This attitude toward change is captured in this passage from the *Zhuangzi*, the second most important text in Daoism. Explaining his craft, the cook Ding says, “A good cook changes his knife once a year—because he cuts. A mediocre cook changes his knife once a month—because he hacks. I’ve had this knife of mine for nineteen years and I’ve cut up thousands of oxen with it, and yet the blade is as good as though it had just come from the grindstone” (*Chuang Tzu, Complete Works*, 51).

5. The reader may notice that in both charts, we set the smoothing number at ten rather than the default of three. Smoothing can be useful because if a lexeme occurs in a single book in one year but not in the preceding or following years, that creates a taller spike than it would with more sustained occurrence: in other words, tall spikes would present a distorted view of the actual trend. No smoothing, or zero smoothing, would reflect raw data—in other words, no visual patterns at all (Michel et al., “Quantitative Analysis of Culture”).

6. Some histories date the Chinese Cultural Revolution to 1966–1976; however, it formally ended in 1969 when “the Ninth Party Congress declared that the Cultural Revolution had come to a victorious conclusion” (*Tanner, China*, 530). The year 1976 is often cited in this connection, as it was the year of Mao’s death and the end of ten years of factional struggle within the CCP leadership.

**Bibliography**


