ELSEVIER

Futures of a distributed memory. A *global brain* wave measurement (1800–2000)

CrossMark

Steffen Roth [a,d,*], Carlton Clark [b], Nikolay Trofimov [c], Artur Mkrtichyan [d], Markus Heidingsfelder [e],
Laura Appignanesi [f], Miguel Pérez-Valls [g], Jan Berkel [h], Jari Kaivo-oja [i]

[a] La Rochelle Business School, France
[b] University of Wisconsin-La Crosse, USA
[c] Russian Academy of Science (ISS), Russia
[d] Yerevan State University, Armenia
[e] Habib University, Pakistan
[f] University of Macerata, Italy
[g] University of Alméria, Spain
[h] Independent, Costa Rica
[i] Turku School of Economics, Finland

ARTICLE INFO

ABSTRACT

If the *global brain* is a suitable model of the future information society, then one future of research in this global brain will be in its past, which is its distributed memory. In this paper, we draw on Francis Heylighen, Marta Lenartowicz, and Niklas Luhmann to show that future research in this global brain will have to reclaim classical theories of social differentiation in general and theories of functional differentiation in particular to develop higher resolution images of this brain's function and sub-functions. This claim is corroborated by a brain wave measurement of a considerable section of the global brain. We used the *Google Ngram Viewer*, an online graphing tool which charts annual counts of words or sentences as found in the largest available corpus of digitalized books, to analyse word frequency time-series plots of key concepts of social differentiation in the English as well as in the Spanish, French, German, Russian, and Italian sub-corpora between 1800 and 2000. The results of this *socioencephalography* suggest that the global brain's memory recalls distinct and not yet fully conscious biases to particular sub-functions, which are furthermore *not* in line with popular trend statements and self-descriptions of modern societies. We speculate that an increasingly intelligent global brain will start to critically reflect upon these biases and learn how to anticipate or even design its own desired futures.

© 2017 Elsevier Inc. All rights reserved.

1. Introduction

As researchers in technological and social change, we want to track and trace significant trends in past and future societies. One such trend is secularization, the declining importance of religion, which is so important to the self-concept of modern societies that the mere thought of a trend reversal brings back memories of the Middle Age. Another widely recognized trend is the growing influence or even dominance of the economy in our societies today. There is also discussion on further and sometimes competing trends, which include the prominent idea of an information society dominated by the mass media system. Yet another stable trend is that these and similar trends have been assumed and implied rather than studied so far, which constitutes a *third order risk* (Godet, 1986) whenever we extrapolate the trend truisms into the future, thus using the right tools to meet the wrong expectations. Most of us nonetheless rely on traditional trend knowledge, while only a few have called or tried for systematic large-scale tests (Blumler and Kavanagh, 1999; Kjaer, 2010; Roth, 2014; Roth et al., 2016), and our uncritical attitude to the facticity of some of the most significant trends in modern societies is justified to the extent that their examination presents a veritable challenge even in the plain middle of the presumed information age. The on-going proliferation of information and communication technology in general and the Internet in particular has indeed given hope that the analysis of social macro trends will be more feasible or at least more convenient, but has also shown that a network of IT-supported interactions presents more than a comprehensive search tool for big data. As much as any complex tool, the Internet is observed to have taken on a life of its own, which in the case of the World Wide Web encompasses an entire globe. Pioneers go as far as to state

that this "single information processing system (…) plays the role of a nervous system for the planet earth", thus referring to the Internet as *global brain* (Heylighen and Lenartowicz, 2016, p. 1).

In this article, we use a considerable proportion of the Internet to review macro trend hypotheses such as the *secularization*, *economization*, *mediatization*, or *politicization* of society. We draw on the global brain paradigm, first, as a constant reminder that the Internet is not one of our usual research tools, and, second, to further develop the paradigm by contributing a method we refer to as *global brain wave measurement*. Somewhat similar to the pending *planetary electroencephalography* suggested by Russell (1982), our procedure will measure certain aspects of the electromagnetic activity of the global brain. Yet, the comparably short history of the Internet also suggests that a traditional real-time electroencephalography (EEG) will not be adequate to monitor long-term social macro trends. It is due to the *Google Books* initiative, which has generated "the largest online body of human knowledge"[1] in the form of a word corpus of >25 million digitalized books, that we see that the global brain has a memory older than the Internet itself, and that we still can access this virtually pre-conscious memory using the Internet in an unprecedented way. We hence used the Google Ngram Viewer, an online graphing tool that charts annual word counts as found in the Google Book corpus, to run comparative analyses of word frequency time-series plots for the *English*, *Spanish*, *Russian*, *French*, *German*, and *Italian language areas*. The outcomes of this procedure positively resemble classical EEG recordings and indicate that the attention the global brain devoted to religion, economy, politics, the mass media and further social systems featured substantial changes in time and significant regional differences. The results also suggest that a number of popular trend statements and definitions of modern society are completely divorced from the global brain's memories between 1800 and 2000.

2. Global brain waves: from electrophysiological to electrosociological brain wave measurement

In our research, we used a small Internet tool to observe a big Internet database. Or put briefly, we used the Internet to monitor the Internet. This situation is different from the case of a traditional electrophysiological brain wave measurement, where the research in brains is thought to be performed from a standpoint external to the examined brains. By contrast, our research was literally *in* the global brain throughout the entire process. Our only logical starting point hence was a thorough exploration of our own research environment.

One of the most up-to-date, compact, and still comprehensive accounts of this research environment has recently been published in *Technological Forecasting and Social Change*. In their editorial to the special issue devoted to the global brain, Heylighen and Lenartowicz (2016) introduce the concept as a realistic model of the information society. They define the global brain "as the self-organizing, adaptive network formed by all people on this planet together with the information and communication technologies that connect them into a coherent system". Their idea is clearly that ICT-mediated interactions have increased interpersonal dependences up to the point where we can observe the emergence of a single superorganism, "i.e. an organism (global society) consisting of organisms (individual people)", with the Internet playing the role of the nervous system for this planetary superorganism. Next to the rapidly intensifying interdependences, the authors also stress the constantly increasing information storage and processing capacities that go along with the present Internet revolution. The authors conclude that we shall soon live to see a qualitative leap in or to the evolution of an adaptive, globally *distributed intelligence* that has a life of its own.

Among the many compelling contributions to the corresponding special issue we found co-guest editor Marta Lenartowicz' (2016) single-authored article particularly instructive as it deviates from a number of classical assumptions in the global brain literature and even in her above co-authored introduction. In "Creatures of the semiosphere. A problematic third party in the 'humans plus technology' cognitive architecture of the future global superintelligence" she argues that neither human beings nor IT-supported networks of human beings, but rather social systems can be conceived as "the most advanced intelligence currently operating on Earth" (Lenartowicz, 2016). As she draws on the work of Niklas Luhmann (1995, 2012, 2013), she defines social systems as autopoietic systems of communication, the first emergence of which she traces back to the origins of spoken language tens of thousands of years ago. This approach is remarkable in two ways: first, she proposes to change the traditional human-technology focus prevailing in the global brain literature[2] for a technology-communication focus, which, to our mind, is more suitable for the observation of complex information and communication technology systems. This proposed observational shift from networks of humans to networks of communications[3] allows access to a so far under-researched macro region of the global brain. Second, her short and appropriate recourse to the history of communication and communication media suggests that distributed intelligence might be older than the global consciousness about it (Heylighen, 2011).

If we trace these two ideas back to their systems' theoretical origins, then we find indeed that the idea of a social global brain consisting of a network of communication and technology is as plausible as is the classical idea of a bio-technological global brain made of human organisms and technology. This is true particularly because a basic form of intelligence, memory, is inherent to all forms (Luhmann, 1997, p. 364), including all forms of communication (Luhmann and Rasch, 2002, p. 160). Communication as threefold selection of information, utterance, and understanding operates in time, which implies the management of the difference between past and future, the token of which is memory (Luhmann, 2012, p. 350); and systems of communication imply memory in order to link one communication to another. Memory is hence not an isolated subfunction of a social system, but rather involved in all of its operations, and Luhmann emphasizes that "these operations are communications, and thus not neurobiological changes in the state of the [biological] brain nor what enters the awareness of a single consciousness" (id, p. 349). The more complex the social system, the more complex its memory. We consequently can image highly complex forms of collective, distributed, or simply social memory that are made of communication and nothing but communication. The main function of all these forms of memory would be the same as with all forms of memory: *forgetting*. This only prima facie counter-intuitive take on the memory function is stringent insofar as the memorization of no matter what presents a necessarily selective operation which recalls only very little information, thus filtering out numerous alternatives.

Memory works as a filter located at the interface of the past and the future, and is therefore necessarily always in the present. As a filter, the

> function of memory relates to distinctions; or, more exactly, to indications of something as opposed to something else. The memory operates with what has been successfully indicated and tends to forget the other side of the distinction. Although it can also mark distinctions as forms, for instance, the distinction between good and

---

[1] Only the bold beauty of this fittingly anthropomorphical metaphor made us quote the *Wikipedia* article on "Google Books" as accessed on July 28, 2016.

[2] Theories that focus on human-technology linkages, or "humans-plus-technology," and theorize the global brain as a network connecting human beings are useful but still anthropocentric. Two important texts on network society are Harrison White's *Identity and Control: A Structural Theory of Social Action* (1992) and Manuel Castells' *The Rise of Networked Society* (1996). More recently, in *Networks of Outrage and Hope: Social Movements in the Internet Age*, Castells (2012) takes up the subject of networked social movements with reference to the Arab Spring and other movements. We are more interested in autonomous social systems than in networks of human beings.

[3] For an extensive case made for a similar turn in organization studies including instructive visualizations see also Lenartowicz (2016, p. 178) and Luhmann (2012).

evil, it tends to forget what this distinction was distinguished from. The particularity of discrimination in the forgetting/remembering schema is determined not least by language and is insofar a characteristic of social systems. (Luhmann, 2012, p. 351)

As every social system requires such a social memory, every society is, in its temporal dimension, defined by the form of this filter. The key question, then, is which distinctions a given society draws in which medium to manage its own history, and the token for the particular way a given society executes this filter function is *culture* (Luhmann, 2012, p. 355). Archaic societies already had culture, i.e., a social memory concerned with the sorting of more or less tangible objects and features in the medium of oral language. Yet, Luhmann states that it was not until the Age of Enlightenment that cultures started to distinguish between culture and nature as much as between different cultures, assuming that this reflexive turn presented a necessary reorganization to align the social memory with the requirements of an increasingly complex and dynamic modern society. Today, the reflexive memory of (post-) modern societies is increasingly flexible and skilled in the use of distinctions, including those that were constitutive for earlier forms of memory. As it is our ambition to analyze social mega trends between 1800 and 2000, we shall be particularly interested in exploring this stock of distinctions available for the organization (and constant re-organization) of a modern social memory. This implies that we need to be concerned with social differentiation.

To date, we may distinguish four basic forms of social differentiation (see Table 1):

These basic forms of social differentiation may be used to tell a short history of human society (Luhmann, 1977, 1990, 2013). *Segmentation* was the dominant form of social differentiation in archaic, oral societies, which were made up of both similar and equal segments (see the top left quadrant). Yet, in the course of the Neolithic revolution, processes of *centralization* occurred that turned some segments into centers and others into periphery (top right quadrant). Although centrality does not always constitute an advantage, centralization of resources, influence, or attention often resulted in *stratification*, i.e. a process by which subsystems of society are ranked into a hierarchy of dissimilar and unequal subsystems (bottom right quadrant). In the transition to modern societies, however, stratification was replaced by functional differentiation as the primary form of social differentiation. Functional differentiation is defined as the distinction of dissimilar and equal function systems (bottom left quadrant) such as the political system, economy, science, art, religion, legal system, sport, health, education, and mass media system.[4]

It is important to note that older forms of social differentiation are not replaced but only overruled by newer ones. Thus, we still observe segmentation of families residing in private homes; however, abused or neglected children and battered spouses are now afforded protection by the legal system, and children are subject to compulsory education. Social class inequities are also still observable, and organizations (e.g., corporations, universities, governments, militaries, bureaucracies) still have hierarchical structures, but people are no longer born into fixed, unchangeable social strata with unequal legal rights. If a significant percentage of a society remains poor, we tend to blame the education system or call for reform of the economy, politics, or the healthcare system. That is to say, we don't take social inequality as a natural given. To take another example, universities and other institutions are ranked, but these rankings are changeable. We also observe centers of power (e.g., governmental or financial centers) with weaker peripheries (e.g., rural areas, rust belts, Parisian suburbs). However, the key point is that the functional differentiation of the economy, politics, law, education, healthcare, mass media, science, art, religion, etc., overrules older forms of segmentation, stratification, and center/periphery organization.

---

[4] See Roth (2015) and Roth and Schütz (2015) for a more detailed account of the process of social differentiation and a discussion on the current number of function systems.

Table 1
Social differentiation.
[Source: Roth (2015, p. 113).]

		Equal	
		+	−
Similar	+	Segmentation (Families, tribes, nations, etc.)	Centralization (Civilizations, empires, etc.)
	−	Functional differentiation (Economy, Science, Art, etc.)	Stratification (Castes, estates, classes, etc.)

In our context, this means that in modern societies all four basic forms of social differentiation are in principle available to organize the modern social memory, although modern culture may be expected to feature a certain bias to the principle of functional differentiation if it comes to the filtering or realization of relevant information.

Functional differentiation obviously is the form of social differentiation on which we need to focus in the context of our electrosociological global brain wave measurement, because the observation of trends such as secularization or economization refers to changes in the prominence of function systems such as religion or economy, and therefore implies functional differentiation.

Our planetary EEG hence is *sociological* because we analyzed recordings of global brain activities that indicate cultural fluctuations, i.e. changes in the relevance that specific forms of social differentiation have for the self-organization of the social memory; and it is *electro* because these *culturomic* recordings are produced by an Internet tool, as we shall demonstrate in the subsequent section of this article.

3. Global brain waive measurement: an operationalization using the Google Ngram Viewer

3.1. The Google Ngram Viewer as socioencephalograph

In the previous section, we supported and radicalized Marta Lenartowicz' (2016) work on semiotic forms of superintelligence and exchanged the traditional biotechnological definition for a sociotechnological definition of the global brain as the global system of communication, including information and communication technology. We also drew on Niklas Luhmann to demonstrate that, as much as any social system, this global social system features memory, which is critical as the purpose of our brain wave measurement was to verify social mega trends and hence required some form of access to the memories of the global brain. We also explained why the key indicators of our research are necessarily related to the concept of functional differentiation, the key principle behind the distinction between function systems such as religion, economy, politics, legal system, science, education, or the mass media system.

Our basic idea was to use the Internet to analyze how relevant the individual function systems have been to the global brain within the last two centuries. This approach is adequate since the ICT revolution in general and the Internet in particular considerably leveraged the cognitive capacity of the global brain. Yet, it is also problematic because the Internet is younger than the trends we intended to verify. Were therefore lucky that the Internet represents only one specific form of social memory next to older forms such as oral tradition, writing, or printing (Lenartowicz, 2016; Lenartowicz et al., 2016; Luhmann, 2012, p. 178), and we were even luckier that that the Google Books project operates at the interface of two of these forms of social memory.

Officially announced in 2004, the Google Books project has scanned and digitalized over 25 million of the estimated 130 million published titles worldwide. The research potential of this project was first recognized by a Harvard research team (Michel et al., 2011) in 2007. The team performed quality checks, created a first consolidated Google Book corpus of more than 5 million books, coined the term *culturomics* for the "the application of high-throughput data collection and analysis to the study of human culture" (ibid, 181), and developed a prototype of

what would finally become the *Google Ngram Viewer*, an online search tool that plots line charts of annual word[5] counts as found in the Google Book corpus. Today, the updated version of the Ngram Viewer scans a corpus of over 8 million books containing hundreds of billions of words in English, Spanish, Russian, French, German, Chinese, Hebrew, and now also the Italian language [see Lin et al. (2012, p. 170) for an overview of the number of volumes and ngrams for each language area]. The tool has been quickly discovered by pioneers in the digital humanities and been used predominantly to analyze issues of language, literature, history, and culture (Gibbs and Cohen, 2011; Johnson, 2010; Michel et al., 2011; Nicholson, 2012; Ophir, 2010; Sparavigna and Marazzato, 2015). There have also been attempts to establish culturomics in the social sciences, e.g., in the context of a retroactive forecasting of social movements like the Arab Spring (Leetaru, 2011) or popularity checks of sociological theories, scholars, fields, and methodologies (Chen and Yan, 2016).

Using the Google Ngram Viewer means analyzing a corpus of words as found in books that made their way to the Internet. Whereas the appearance of a word in a book is a matter of its word importance, the appearance of a book in the Google Book corpus is a matter of book representativeness. Although the designers of the Google Book corpora did their best to avoid selection biases, the corpora have been criticized for containing words from exactly one of each book, which favors merely prolific authors over possibly less prolific but more influential authors (Pechenick et al., 2015). While the latter issue can only be addressed by including – ultimately contingent – popularity indicators in the already giant dataset, the former issue is interesting because it raises performativity issues that are important in any research using interactive media. Again, we see that our research in the global brain literally takes place in this global brain, which is true as the Google Books project continues and the results of our research might enter the very memory region we screened. Our research is therefore not likely to eventually co-perform the analyzed social mega trends, which presents a methodological challenge as much as a paradoxical access cue for those who are interested in these trends "to anticipate them and to direct them towards the most desirable outcomes" (Heylighen and Lenartowicz, 2016, p. 2).

In our research, we considered the words to be forms of communication in a communicative medium (written language) and translated into another communicative medium (computer language). We further assumed that the frequency with which these forms appear in the respective medium as indicated by the Google Ngram plots be an appropriate approximation to their importance; in fact, word frequency is deemed the "simplest and most impartial gauge of word importance" (Kloumann et al. 2012, p. 1) or the popularity of objects, concepts, or persons (Bohannon, 2011; Ophir, 2010). Moreover, our research builds on earlier applications of the Google Ngram Viewer to social mega trend verification (Roth, 2014; Roth et al., 2016), which we complement and further develop in the following three dimensions: first, our reference to the global brain concept makes our approach more intuitive, concrete, and literally more reflexive. Second, by adding Spanish, Russian, and Italian, our research covered more language areas in order to check for inter-language diversity and test the generalizability of the earlier conclusions.[6] Third, we systematically used recently introduced new features of the Google Ngram Viewer such as the option of combining several words into one graph. In this sense, we scrutinize the results of earlier works applying a more reflexive and robust methodology to a broader scope of samples.

3.2. Semi-automated search term selection

To fully deploy the new options provided by the enhanced version of the Google Ngram Viewer, we furthermore had to reappraise the selection of the search terms to be entered into the Viewer's search field. So far, authors have mainly focused on how the importance of function system designations – i.e. terms such as economy, religion, or art – fluctuated in time, and only gave limited examples of how the performance of pertinent keyword chunks could be systematically analyzed. To address this limitation, our challenge was to identify the most pertinent keywords per function system. As the Google Ngram Viewer only allows for a relatively small number of keywords to be entered into the search field, we limited the number of desired keywords to five per function system. We hence decided to select the five most frequent keywords per function system and to combine them into one graph per function system so as to produce comparative time series plots of fluctuations of the importance of each function system between 1800 and 2000.

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 (see Annex). 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. For example, the n-gram *university* clearly refers to education, however, not unambiguously so, as it also refers to science because universities are research institutions, too[7]; the term does therefore not qualify as a function system indicator, whereas the n-grams *money* or *theory* can be relatively safely assumed to refer to economy or science, respectively.[8] 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.[9] As we decided to track the performance of ten systems – namely political system, economy, science, art, religion, legal system, sport, health, education, and mass media system (Roth and Schütz, 2015) – we needed to produce two plots of five function systems each. We repeated the entire procedure for each language and then compared the results against the subsequent set of hypotheses.

---

[5] The basic units of the corpus are not words but n-grams, sequences of $n \geq 1$ letters, figures, or signs, including misspellings and apparently meaningless expressions; thus the name Google Ngram Viewer. We shall nonetheless use the term *word* for the sake of readability.

[6] We understand some readers might object that we are studying the "Western brain" rather than the global brain; however, we excluded Chinese because of data quality and linguistic issues that justify being addressed in a separate article, and we did not include Hebrew because no team member is proficient in this language. We opted for 1800–2000 as the sample period because the data is most reliable for these two centuries and because this period corresponds well to our ambition to observe macro trends in modern societies.

[7] By contrast, we kept terms such as church or school. Technically speaking, churches or schools are not exactly mono-functional as we may easily observe power struggles in churches or school fees. Yet, we found that, unlike the inherently bi-functional universities, churches and schools are relatively strictly coupled to only one dominant function system.

[8] In some cases, we used the Google Ngram Viewer to estimate the degree of word ambiguity. The ngram *company*, for example, may also have non-economic meanings such as in "in good company". Yet, the ratio of "good company" to "company", which can be checked using the string (*good company/company*), is never exceeding 1.7% and declining to <0.3% in 2000. *Good* is furthermore not among the most common determiners of company (string: **_DET company*). Similarly, "electric power" is almost non-existent in the 19th century, with the (*electric power/power*) ratio peaking at hardly >1.0% in the 1950s and declining to <0.5% in the 1990s. *Power plant* or *power station* also account for <0.5% or 0.2% in 2000. Again, none of the aforementioned determiners is among the most common (strings: **_DET power* and *power *_DET*). Wildcard searches such as ** company* or *power ** further corroborated our interpretations; they also proved helpful in contexts such as "pp.", where we checked that the abbreviation actually refers to the referencing of book pages and, thus, to the mass media system.

[9] It is worth noting that the chunking cancels the recently upgraded feature that unlocks the decision between case-sensitive and insensitive searches. All our searches were hence case-sensitive searches, which implies that, for example, the n-grams *church* and *Church* be treated as independent search terms.

3.3. Hypotheses

As function systems are defined as both dissimilar and equal systems (see Table 1), their inherent incommensurability makes an excellent case for our null hypothesis. In fact, they represent coequal nominal data and therefore can be assumed equally relevant to given social systems. Our null hypothesis, therefore, reads as follows:

H0. *The global brain's memory recalls that all function systems have been equally important throughout the last centuries. Our global brain wave measurement therefore shows a uniform distribution of the combined performances of the five most frequent keywords per function system from 1800 to 2000.*

Yet, prominent trend statements and self-definitions of modern societies, such as the ideas of secularized or capitalist societies, suggest that the global brain recalls unequal distributions and significant fluctuations of the significance of the individual function systems. Our alternative hypothesis was as follows:

H1. *The global brain's memory recalls that all function systems have not been equally important throughout the last centuries. The combined performances of the five most frequent keywords per function system therefore exhibit an unequal distribution both in the course of time (H1.1) and across the language areas (H1.2).*

As we intended to pursue the alternative hypothesis and link it to the verification of social macro trend statements and self-descriptions of modern societies, our second alternative hypothesis suggested that the global brain's memory is in line with the most popular common senses on trends in modern societies:

H2. *All linguistic regions of the global brain recall significant trends in functional differentiation, including (H2.1) the secularization, (H2.2) the economization, (H2.3) the politicization, and (H2.4) the mediatization of society as indicated by the combined performances of the five most frequent keywords per function system.*

4. Results

With the exception of sport,[10] the vast majority of the top 5 keywords of all function systems were found within the top 2000 of all language areas. The comprehensive list of keywords is available in Tables A1.1–A1.6 (see Annex).

Entering these keywords, we found that the combined occurrence frequencies of these five most frequent keywords per function system exhibit unequal distributions both in time and within as well as across all language areas.

Due to the above word or ngram limits to the Google Ngram Viewer input mask, our charts present five chunks of five keywords. The figures in the running text present the combined occurrence frequencies of the five most frequent keywords for the function systems most relevant to our hypotheses (political system, economy, religion, and mass media) complemented by the best-performing out of the remaining function systems (Figs. 1–6). Figures showing the performances of all function systems are available in the Annex (Figs. A2.1–6).

---

[10] It was not possible to identify unambiguous sport keywords within the first 3000 ngrams of all language areas, which might be due the relatively short (or better: interrupted) history of sport and its consideration as a function system. Moreover, we did not find even popular sports such as soccer, tennis, or chess among the top 10,000. Except for the term sport that actually appeared in the German language area, the ngrams we entered to trace the performance of sport might be approximations rather than solid indicators until further theory work on sport as a function system is done. We therefore did not include it in our presentation in the Results and Discussion sections; however, we kept the sport graph in the charts to stimulate feedback and opinions. In any case, sport clearly presents the lowest importance of all function systems.

4.1. English language area

In the English language area, religion is the most dominant function system in the 19th century and the political system the most dominant one of the 20th century. Starting soon after 1840, the decline of religion is dramatic and stopped not before World War I. The political system is the most important function system in the English language area since about 1880; the two World Wars seem to have had a significant influence on the importance of the political system. Another smaller peak for politics may be observed in the 1960s. Science became increasingly important in the 20th century and particularly during the Cold War period; in 2000, science was the second most important system in the English language area. Originally more important than science, economy became more important particularly during and between the two World Wars, but was outperformed by science at about 1950. A small rise of the information age may be traced back to 1920 with the curve getting steeper since the end of the 1960s. Another significant trend is the considerable uptrend of education since the early 20th century (see Fig. A2.1, Annex). In 2000, education enters the top 5 after political system, science, mass media, and economy. There is also a smaller rise in the importance of health since the 1960s. The health system is seventh after the legal system in 2000. Art and sport are consistently unimportant throughout the entire observation period.

4.2. Spanish language area

The most important function systems in the Spanish corpus are the political system, religion and the legal system throughout the entire observation period. Religion started as the dominant system in 1800; however, it soon displays turbulent interactions with the political and the legal system, at the end of which religion remains third at about 1870. Yet, there is no dramatic fall of religion, which again overtakes the legal system in the interwar period and remains second until the mid-1970s, and is second just again in 2000. The legal system shows a sharp decline after 1900, before levelling out at about 1970. Uninterruptedly dominant since about 1840, the peak of the political system is at about 1870, with the systems reaching an almost similar importance in the 1990s subsequent to a decline that reversed since World War II. Science (see Fig. A2.2, Annex) and economy feature little fluctuation throughout most of the observation period, both featuring a moderate take-off after 1940. After the political system, religion, legal system, science and economy are fourth and fifth in 2000. A flat growth curve of the mass media system can be observed to start as early as in the 1860s. Originally a fairly prominent function system, health features a significant decline between 1820 and 1880, and is the second least important function system in 2000. Education features a considerable increase between 1880 and 1910, and art a less pronounced uptrend between 1900 and 1950. Both systems nonetheless belong to the less important function systems in the Spanish language area.

4.3. Russian language area

Religion is the most dominant function system in the Russian language area from soon after 1800 to round about the 1905 Russian Revolution. After a short period of interaction with the political system in the inter-revolution period, religion declines and is the third least important system in 2000 despite a small revival since 1990. The political system is the most dominant function system since the 1917 Russian Revolution, with a steep rise during the Stalin era to an all-time peak during World War II and two smaller peaks around 1960 and 1980. Science is the second most important system since 1920, and remains in second position even after a considerable decline between 1980 and the late 1990s. Between about 1850 and the 1917 Revolution, the legal system was the most important system (see Fig. A2.3, Annex). During a short period in the 1820s the mass media system was second after religion. Except for this small peak, the mass media system features a
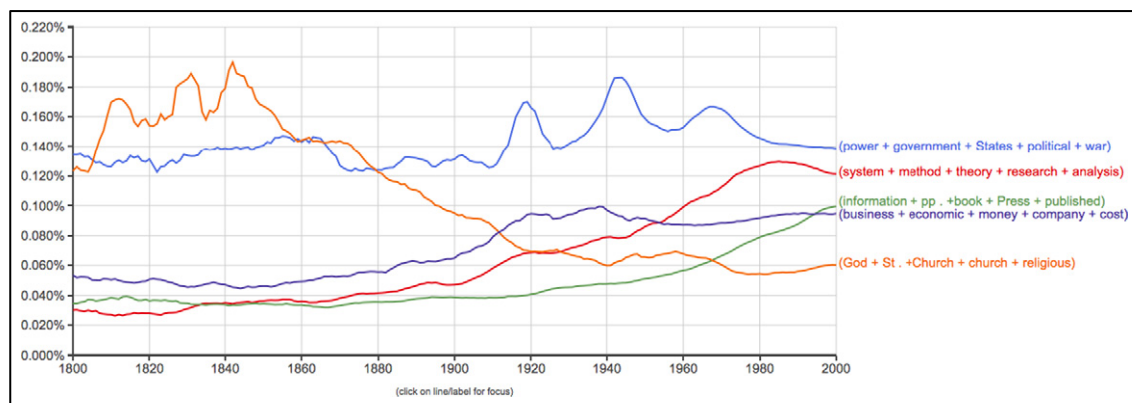
Fig. 1. Combined occurrence frequencies of the five most frequent keywords for political system (blue), economy (violet), religion (orange), mass media (green), and science (red) in the English language Google Books corpus (1800–2000).

stable performance until a period of modest growth starting after 1950. Since the legal system regained importance in the context and aftermath of Perestroika, it is the third most important system in 2000, follow by the mass media system and the economy. The latter was virtually absent before the inter-revolution period, and rose to a temporary third position between the mid 1950s and the early 1990s. Art, education, religion, and health follow on ranks six to nine.

4.4. French language area

The French language area is characterized by an intensive interaction of the legal system, religion, and the political system until the eve of World War I, when the political system booms to an all-time high that abruptly skips to a steep decline, after which the system nonetheless remains in the lead until the end of the observation period. The second peak for the political system around World War II is considerably. After a short period of dominance between about 1850 and 1870, religion declines to a fifth rank in 2000 (despite a modest revival since 1980), superseded by the legal system, which was dominant until the political boom. Subsequent to a fairly steep growth curve between 1910 and 1970, the economy is the second most important system in the French language area in 2000, followed by science (see Fig. A2.4, Annex), the legal system and the mass media system. Although quite popular around 1800, art is ranked sixth at the end of the observation period, followed by education, whose most significant change was a considerable growth between 1830 and 1880. Health is the least important function system in the French language area.

4.5. German language area

The 19th century in the German language area sees an intensive interaction of religion, legal system (see Fig. A2.5, Annex), political system and science. Religion is clearly dominant and the legal system second until 1860, a point in time when the latter starts to dominate until the eve of World War I leads to a take-off of the political system, which is further fueled during the Cold War period until a peak at around 1970. Despite a constant yet somewhat moderate decline, the political system is by far the most dominant system in 2000, followed by science, which had its 20th century peak around 1970, too. The legal system is fourth and religion fifth (after having been second between 1940 and 1955). Due to a moderate growth between 1910 and 1930 as well as the meanwhile stopped declines of religion and the legal system, the economy remains third in 2000. The German-language art system seemed to be significantly influenced by the two World Wars, surpassing the economy in both postwar periods, however, not sustainably so. Education is sixth both after a moderate decline since 1980, tightly followed by the mass media system. Once quite prominent, health is the least popular function system in the German language area in 2000.

4.6. Italian language area

The first half of the 19th century is characterized by an intensive interaction of a number of function systems in the Italian language area, too. The dominance of the political system starts early, however, not uninterruptedly so: religion is dominant for a shorter period around 1840,
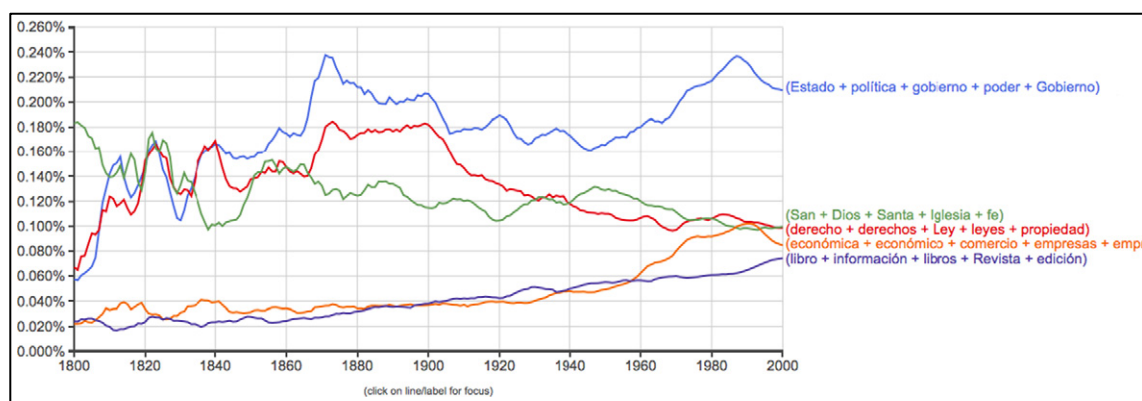

Fig. 2. Combined occurrence frequencies of the five most frequent keywords for political system (blue), economy (orange), religion (green), mass media (violet), and legal system (red) in the Spanish language Google Books corpus (1800–2000).
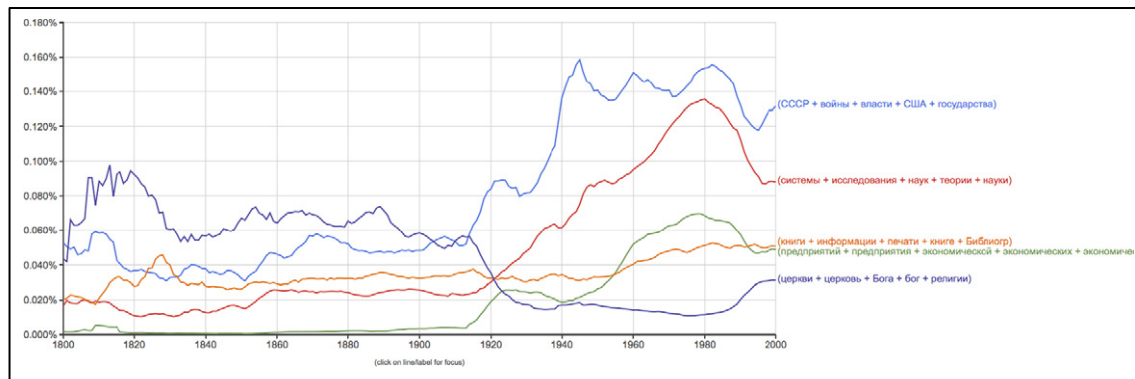
Fig. 3. Combined occurrence frequencies of the five most frequent keywords for political system (blue), economy (green), religion (violet), mass media (orange), and science (red) in the Russian language Google Books corpus (1800–2000).

and the legal system for a longer period between 1870 and again until the eve of World War I. The political system peaks during the first and has an interim peak during (and after) the second World War; a third peak is visible in the 1970s. Science is second in 2000 after an undramatic history of modest growth (see Fig. A2.6), followed by religion, whose decline since 1840 stopped as early as 1890 and was partly compensated by its post-1980 growth. Fourth is the mass media system, which featured its most significant growth trend between 1850 and 1890 and a second smaller one between 1960 and 1990. The legal system is fifth after two waves of decline, the first of which started even before the rise of the political system and the second around 1960. Next is art, whose peaks again correspond to the two World Wars. Economy is seventh in the Italian language area, followed by education, which featured continuous growth as of 1800 that ended around 1910. The initially relatively high importance of health soon declined after 1810 and followed a flat degrowth curve displaying two small dents during the two World Wars.

4.7. Interregional results

The clearest interregional trend is the dominant position of the political system in the 20th century. This trend applies to all language areas without any reservation other than that it started already as early as 1880 in some cases, whereas in others the take-off of politics was not before around World War I.

Most language areas display a 19th-century bounce of religion followed by a significant decline, which is most pronounced in the English and German case and least in the Spanish and Italian. All language areas except the Spanish feature an at least moderate revival of religion starting around 1980.

There are intensive fluctuations and interactions of religion, legal system, and political system as dominant systems in all language areas in the first half of the 19th century. In the Russian, French, German, and Italian language areas, the legal system was the most dominant function system in a period between 1870/80 and 1910/20.

Science appears to be particularly important in the English, Russian, and German language area, in each of which it ranks second since World War I or II.

The economy never had a dominant position in any of the language areas at any point of time. After featuring a moderate uptrend stagnating in the second half of the 20th century in the larger number of language areas, the economy is the second most important system in the French area, number three in the German, four in the English and the Italian, and fifth in the Spanish and Russian in 2000.

Although fairly important in the early 19th century, health is most unpopular in all language areas by the end of the 20th century, with the only exception being the English, where health is more popular than art since about 1970.

The two World Wars are visible as a sometimes tremendous increase of the political graph in all language areas but the Spanish;
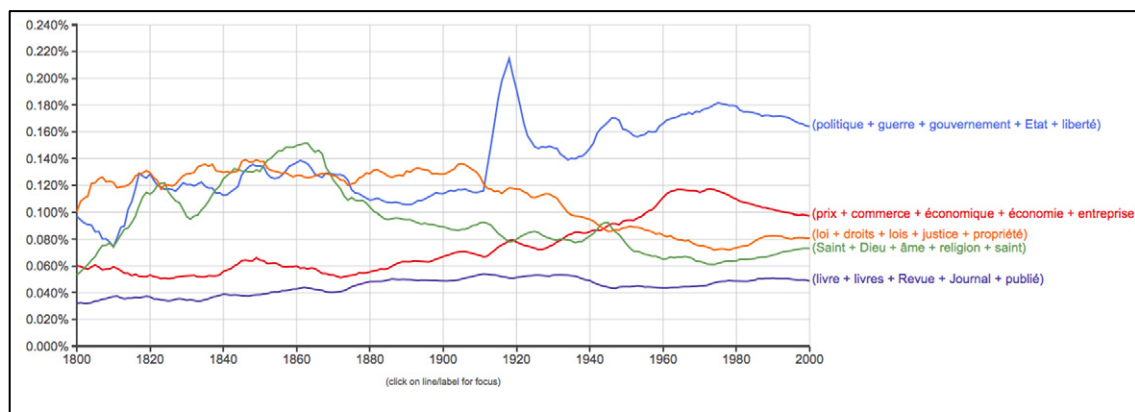

Fig. 4. Combined occurrence frequencies of the five most frequent keywords for political system (blue), economy (red), religion (green), mass media (violet), and legal system (orange) in the French language Google Books corpus (1800–2000).
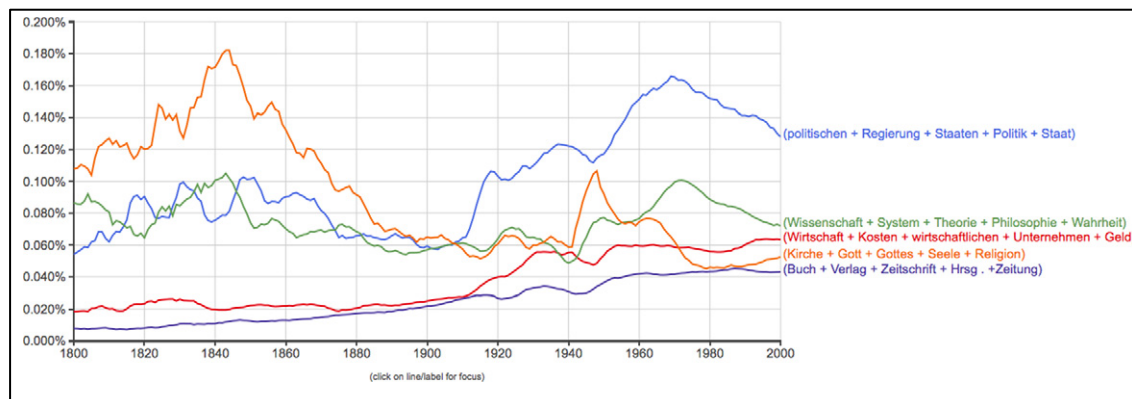
Fig. 5. Combined occurrence frequencies of the five most frequent keywords for political system (blue), economy (red), religion (orange), mass media (violet), and science (green) in the German language Google Books corpus (1800–2000).

there is also a visible influence on art and science in the German and Italian as well as on religion in the Spanish, French, German, and partly the Italian case.

It is interesting to note that the Russian language area displays the lowest overall level of functional differentiation among all areas, particularly during the 19th century. The political graph of the Spanish chart scores the highest value of all language areas around 1870 and almost reaches the same level around 1990 again.

In total, we find that each language area has its own distinct profile of function system preferences. At the same time, there are common trends featured in all or most function systems, the most striking of which are the dominant position of the political system in the 20th century; the ultimately slightly inverted decline of religion; and the intensive interaction of religion, legal system, and political system during the second half of the 19th century.

5. Discussion

Our global brain wave measurement shows that the occurrence frequencies of the function system indicators exhibit an unequal distribution both in time and across all language areas. According to the data, the global brain recalls that it did *not* treat the function systems as equally important throughout the last centuries. The null hypothesis is therefore rejected.

As we further linked the alternative hypothesis to the verification of popular social macro trend statements, it was our ambition to check

whether the global brain recalls significant trends in all of its linguistic regions. The results clearly indicate that there are trends in all language areas, which is why our second alternative hypothesis (H2) is also confirmed in its general form. We therefore proceeded to discuss the data against our sub-hypotheses H2.1–4, according to which we checked whether the global brain recalls specific social macro trends such as the secularization (H2.1), economization (H2.2), politicization (H2.3), and mediatization of society (H2.4).

5.1. Secularization of society, confirmed

Starting in the second half of the 19th century, religion presents a strong downtrend in the English, Russian, and German and a moderate downtrend in the Spanish, French, and Italian language areas. Initially the dominant system in most of the cases, religion is of little importance in the English and French and of very little importance in the Russia language area. Yet, the system remains the second most important in the Spanish and the third in the Italian language area. The German language area shows an ambivalent image: on the one hand, a dramatic downtrend between 1850 and 1940, on the other hand a short countertrend temporarily pushing religion back into the second position in the 1940s; another downtrend is inverted in the 1980s, leaving religion on rank four. In fact, as this post-1980 revival of religion is common to all language areas and, more importantly, as the downtrend is only moderate in half of the areas, with religion remaining second in one of them, the data seems to confirm the secularization hypothesis with
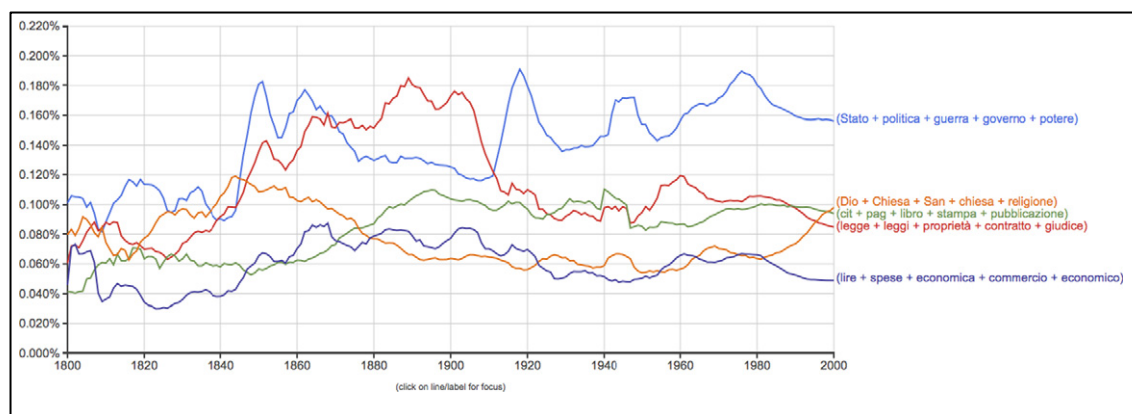

Fig. 6. Combined occurrence frequencies of the five most frequent keywords for political system (blue), economy (green), religion (orange), mass media (green), and legal system (red) in the Italian language Google Books corpus (1800–2000).

some reservation only. The global brain seems to recall periods of clear secularization in some contexts and less clear situations in others. In any case, the global brain seems to be quite sure that there has been secularization after the mid-19th century in all language areas. H2.1 may be cautiously confirmed in the end.

5.2. Economization of society, rejected

Moderate uptrends of the economy may be observed in all language areas predominantly in the 20th century. At the end of these processes, which stopped and reversed toward the end of the century in all cases, the economy makes it to a second rank in the French and a third rank in the German in 2000 while remaining (even) less important in the other cases. There is not a single period in a single language area in which the economy has been close to being the dominant function system. Apparently, the global brain does not recall any situation in which it has been distracted or even ruled by economic principles. The economization of society hypothesis is therefore rejected.

5.3. Politicization of society, confirmed

The dominance of the political system is striking in all language areas starting with World War I at the latest. The distance between the first ranked political system and the second place systems is enormous in all cases except for the English one. Technically speaking, the English language area has not been politicized because it already was politicized. Yet, as all other language areas display a political uptrend, and as all language areas are effectively dominated by the political system in the 20th century, it is safe to assume that the global brain was increasingly politicized between 1800 and about 1920 and heavily politicized since then. The politicization of society hypothesis is confirmed.

5.4. Mediatization of society, rejected

The results concerning the mediatization hypothesis are ambivalent. There are visible uptrends of the mass media indicators in the English (since 1960) and Italian language area (1850–1890), where the mass media system ranks third and fourth in 2000. We also see a longer moderate uptrend in the Spanish case since the second half of the 19th century, a similar trend starting even earlier in the German case, and shorter moderate uptrend in the Russian case since 1960, too. The French area does not feature a significant trend at all. As the majority of the curves are comparably flat, and as the final results of the mass media are not particularly good, we decided to reject the mediatization of society hypothesis with the reservation that the global brain apparently recalls trends to the mass media system in most linguistic regions, but still does not hold the system to be particularly important.

5.5. Limitations and future research

As some of our results may appear counterfactual particularly to those colleagues who believe in a stronger importance of the economy and therefore assume that our method fails to reflect it, we wish to point at some weak points of our approach.

As much as a physiological encephalography measures electric impulses rather than thoughts or ideas (thus still giving usable results), so too does our method measure footprints of communication rather than communication, a circumstance which is further complicated by the fact that we observed word frequencies without word contexts. Although the wild card search option of the Google Ngram Viewer allowed for simple context checks, the ideal case might be a research design for the analysis of 2- or more-grams. Yet, a noncontingent selection of key phrases rather than keywords per function system required access to superior IT infrastructure as already the extraction of the top 10,000 word list required several hours per list and the effort for even just 2gram lists would be dramatically higher.

Another methodological bias in our approach was our focus on only five keywords per function system, which was necessary because of the above limits to simultaneous search entries into the Google Ngram Viewer. This approach systematically disfavored the stronger function systems that feature not only the more frequent but also simply more keywords in the word frequency lists. Particularly the dominance of the political system might therefore be even more pronounced if we had the means to trace the combined performances of all political keywords, whereas we do not have any evidence that the relative performance of the economy would be increased if we combined all economic keywords.

Despite these considerable limitations, we are confident that our research is solid enough to present a reliable approximation to the relative importance of the function systems in the observed language areas. In fact, our global brain wave measurement was able to capture many significant historical events and trends such as the secularization, the Russian Revolution, the World Wars, or the moderate information and wellness trend(s) in the concerned language area(s). In fact, the only counter-intuitive result in our research is the mediocre importance of the economy, and criticism of our method would have to address the question why the method was able to capture secularization and politicization while simultaneously failing to display the true importance of the economy and, in doing so, make proposals of how the importance of the economy may be better identified on such a large scale.

The major challenges for future global brain wave measurements will be

1) The cross-validation of the method systematically exploring interactions between the charts and established historical knowledge in the respective language areas,
2) The inclusion of the missing language areas, with a particular challenge being the Chinese where both OCR issues and specifics of the Chinese language need to be addressed,
3) A cross-media integration allowing for the combined analysis of book and Internet data particularly in view of the post-millennium period,
4) The development of research designs that allows for the trending of combined system-specific more-grams, although it is not clear yet whether and how keyword combinations and sentences might be better indicators than keywords or how word context may be better captured by any other means, respectively,
5) The integration and development of research designs for the anticipation of future social mega trends in functional differentiation.

We believe that this effort is justified and worthwhile simply because it may be used as an explorative tool that helps with generating research questions. Moreover, there is clear evidence that the data are much more than random. We can all see the tremendous impact of the two World Wars on the importance of function systems in general and the political system in particular. Another striking result is the intensive fluctuations and interactions of religion, legal system, and political system in the 19th century as well as the subsequent interregnum of the legal system, which corresponds to works by Thornhill (2008, 2010); and even more specific results such as the relationship between the political and legal systems in the Russian language area between 1900 and 2000 which interact well with pertinent research on the evolution of law under (post-) communism in Eastern Europe (Schönfelder, 2016).

The method is definitely useful to challenge our overconfidence in traditional trend statements or definitions such as the truism that modern societies are economized or capitalist.

6. Conclusion

In this article, our ambition was to verify popular social macro trend statements and to review how reasonable it is to use these statements

to characterize modern societies. To this end, we performed a global brain wave measurement in the form of word frequency analyses in the largest online body of human knowledge to screen the global brain's memory for traces of a secularization, an economization, a politicization, and a mediatization of society in six language areas from 1800 to 2000.

The results suggest that modern societies – as far as they belong to the English, Spanish, Russian, French, German or Italian language areas – can be appropriately defined as politicized societies as of World War I. It is furthermore appropriate to highlight secularization as mega trend, although the data also suggest that this trend was most significant in the second half of the 19th century and partly reverted in the late 20th century. We might even still continue to speculate that there is an emerging mega trend towards an information society as at least some language areas feature a considerable uptrend of the mass media system towards the end of the observation period. What is not supported by our data, however, is the idea that modern societies are dominated by the economic system. Definitions of modern societies as capitalist societies therefore appear untenable as long as the corresponding definitions of capitalism imply an over-average importance of the economy in one way or another. In fact, the global brain recalls that there is not a single region of its memory in which the economy was dominant at any point of time, and that only one of its regions corresponds to a capitalist profile in terms of a society dominated by money and power. Ironically, this case is the French language area. Traditional Marxist or other "political-economic" definitions of capitalism may hence still be applied to the French special case. With regard to all other language areas, it appears that classical critical theory is more appropriate insofar as it accounts for the high importance of what it refers to as scientific-technological rationality in modern societies, which actually is reflected by our data. Yet, even critical theories fall short of accounting for the significance of other important function systems, and this holds also true for more fashionable variants of the critical political-scientific-economic gaze such as the triple helix concept, which is obviously dominant not least in foresight and futures studies

(Roth and Kaivo-oja, 2016) although some of its most prominent promoters have already stressed the need to considerably broaden the concept (Leydesdorff, 2012, 2013). Against the background of our data, the neglect of religion is particularly striking since the global brain recalls that, in 2000, the system was second in the Spanish and Italian language areas and more important than the mass media in all language areas but the English and the Russian.

It must hence be asked whether research in the past, present, and future of the world society may afford to remain so strongly focused on the political system, science, and the probably only marginally important economy, thus ignoring influences from other potentially more relevant function systems.

As this critical question emerged in the context of research on the global brain conducted in this global brain, this question literally is a question asked and to be answered by the global brain; and there is hope that the reflection stimulated by such questions eventually increases the reflexivity up to a point where a recalibrated self-concept enables the global brain to critically review, anticipate, and influence mega trends (Heylighen and Lenartowicz, 2016). The starting point to the millionfold claimed and desired shift to a post-capitalist society (Last, 2017) may hence be in a global brain that takes its own memories seriously and therefore widely ignores local obsessions with economic issues, thus taking the liberty to concentrate on more important matters. One not undesired side effect of this refocus would be a science more aware of its actually prominent role in society using its prominence to redirect its recalibrated self-esteem and methodologies to important issues, which are not primarily in the economy, but rather in the political system and religion, and not least, in science itself.

Acknowledgement

Appendix Annex. Tables A1–6 Keywords per language area

Table A1.1
Top five keywords plus ranked combined keyword frequencies per function system in the English language Google Books sub-corpus.

System	English	Frequencies/chunk
Political	(power + government + States + political + war)	545,937,001
Science	(system + method + theory + research + analysis)	346,647,669
Mass Media	(information + pp. + book + Press + published)	315,167,212
Religion	(God + St. + Church + church + religious)	300,982,802
Economy	(business + economic + money + company + cost)	300,208,287
Legal	(law + property + Court + rights + laws)	266,353,906
Education	(school + education + students + schools + learning)	242,819,722
Health	(health + disease + patients + medical + Health)	154,036,473
Art	(art + music + style + beautiful + Art)	143,214,993
Sport[10]	(success + failure + successful + failed + game)	119,162,189

Table A1.2
Top five keywords plus ranked combined keyword frequencies per function system in the Spanish language Google Books sub-corpus.

System	Spanish	Frequencies/chunk
Political	(Estado + política + gobierno + poder + Gobierno)	134,852,140
Legal	(derecho + derechos + Ley + leyes + propiedad)	75,981,764
Religion	(San + Dios + Santa + Iglesia + fe)	74,476,545
Science	(sistema + verdad + análisis + ciencia + teoría)	56,072,336
Economy	(económica + económico + comercio + empresas + empresa)	49,330,531
Mass Media	(libro + información + libros + Revista + edición)	38,558,611
Education	(educación + enseñanza + escuela + Escuela + Educación)	31,905,924
Art	(poeta + poesía + música + poema + belleza)	20,910,042
Health	(salud + enfermedad + médico + enfermedades + médicos)	19,128,267
Sport[10]	(éxito + juego + fracaso + juega + deporte)	11,839,655

Table A1.3
Top five keywords plus ranked combined keyword frequencies per function system in the Russian language Google Books sub-corpus.

System	Russian	Frequencies/chunk
Political	(СССР + войны + власти + США + государства)	76,946,803
Science	(системы + исследования + наук + теории + науки)	59,039,008
Economy	(предприятий + предприятия + экономической + экономических + экономического)	29,734,309
Mass Media	(книги + информации + печати + книге + Библиогр)	26,339,208
Legal	(право + закона + собственности + суда + закон)	21,365,792
Art	(искусства + поэта + искусство + поэзии + стихи)	18,387,362
Education	(школы + школе + обучения + учащихся + студентов)	18,038,985
Religion	(церкви + церковь + Бога + бог + религии)	11,682,835
Health	(больных + болезни + здоровья + лечения + заболевания)	9,884,313
Sport[10]	*(играет + играют + играл + игра + игру)*	*5,245,216*

Table A1.4
Top five keywords plus ranked combined keyword frequencies per function system in the French language Google Books sub-corpus.

System	French	Frequencies/chunk
Political	(politique + guerre + gouvernement + Etat + liberté)	128,660,182
Legal	(loi + droits + lois + justice + propriété)	95,442,257
Religion	(Saint + Dieu + âme + religion + saint)	81,371,288
Economy	(prix + commerce + économique + économie + entreprise)	71,997,017
Science	(système + vérité + science + analyse + méthode)	68,356,691
Art	(art + Art + beau + belle + musique)	54,578,746
Mass Media	(livre + livres + Revue + Journal + publié)	41,145,107
Education	(enseignement + école + examen + education + écoles)	40,349,608
Health	(maladie + malade + santé + maladies + maladies)	29,691,319
Sport[10]	*(succès + jeu + jouer + échec + échoué)*	*22,549,308*

Table A1.5
Top five keywords plus ranked combined keyword frequencies per function system in the German language Google Books sub-corpus.

System	German	Frequencies/chunk
Political	(politischen + Regierung + Staaten + Politik + Staat)	62,414,914
Legal	(Recht + Gesetz + Rechte + Gesetze + Gesetzes)	40,407,209
Science	(Wissenschaft + System + Theorie + Philosophie + Wahrheit)	40,078,241
Religion	(Kirche + Gott + Gottes + Seele + Religion)	36,619,670
Education	(Bildung + Schule + Ausbildung + Schüler + Lehrer)	25,699,368
Economy	(Wirtschaft + Kosten + wirtschaftlichen + Unternehmen + Geld)	24,474,056
Mass Media	(Buch + Verlag + Zeitschrift + Hrsg. + Zeitung)	21,346,370
Art	(Kunst + Dichter + Musik + Künstler + Schönheit)	20,451,747
Sport[10]	*(Erfolg + gewonnen + spielen + Spiel + Sport)*	*13,180,185*
Health	(Krankheit + Patienten + Kranken + Arzt + Krankheiten)	10,590,675

Table A1.6
Top five keywords plus ranked combined keyword frequencies per function system in the Italian language Google Books sub-corpus.

| System	Italian	Frequencies/chunk
Political	(Stato + politica + guerra + governo + potere)	50,776,746
Legal	(legge + leggi + proprietà + contratto + giudice)	36,448,377
Mass Media	(cit + pag + libro + stampa + pubblicazione)	30,276,212
Science	(sistema + ricerca + verità + scienza + filosofia)	29,094,653
Religion	(Dio + Chiesa + San + chiesa + religione)	23,916,703
Economy	(lire + spese + economica + commercio + economico)	19,951,296
Art	(arte + poesia + poeta + disegno + musica)	17,599,946
Education	(scuola + scuole + educazione + insegnamento + Scuola)	10,291,174
Health	(malattia + medico + salute + malattie + medici)	7,626,616
Sport[10]	*(successo + gioco + fallimento + fallito + giochi)*	*5,424,142*

A2. Additional figures
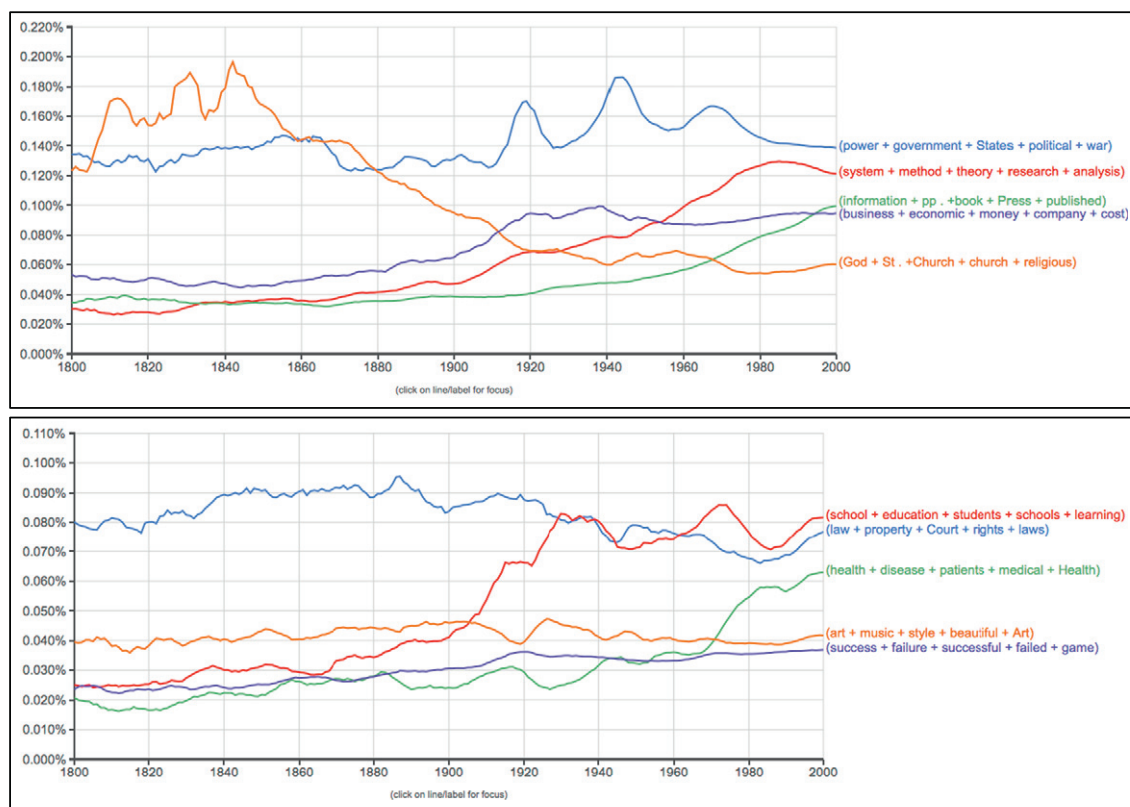
Fig. A2.1. Combined occurrence frequencies of the five most frequent keywords for all ten function systems in the English language Google Books corpus (1800–2000).
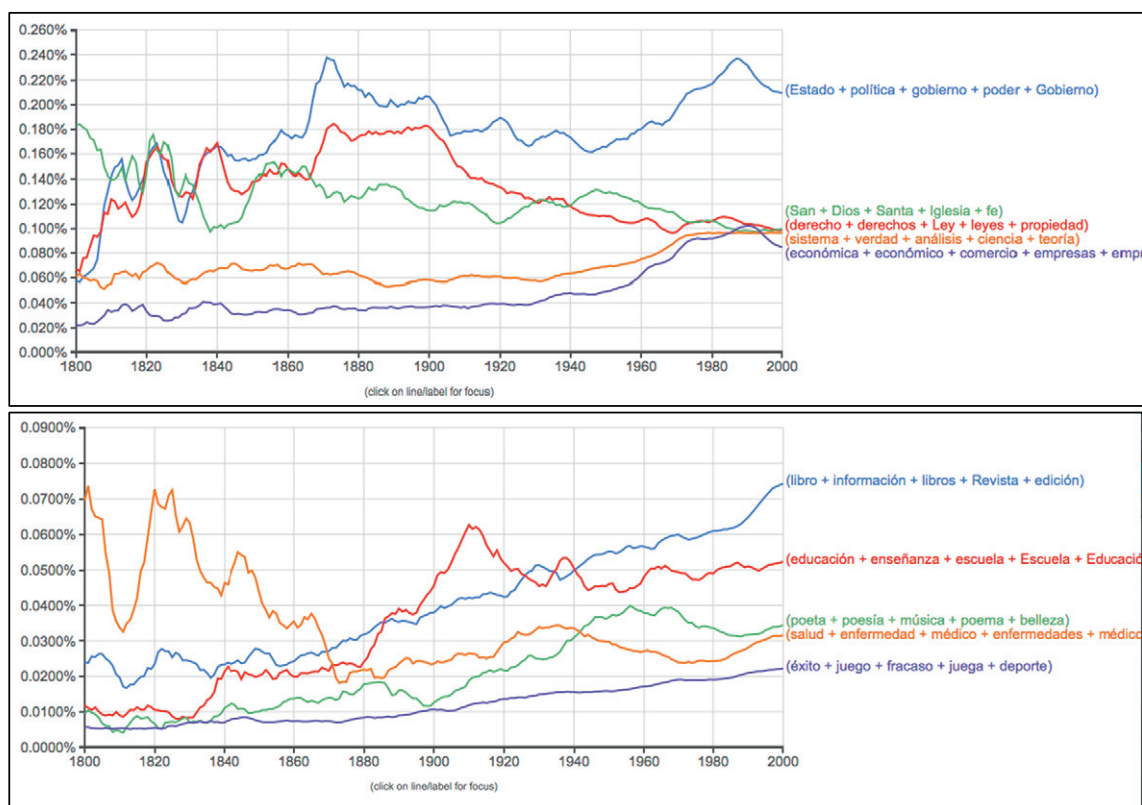

Fig. A2.2. Combined occurrence frequencies of the five most frequent keywords for all ten function systems in the Spanish language Google Books corpus (1800–2000).
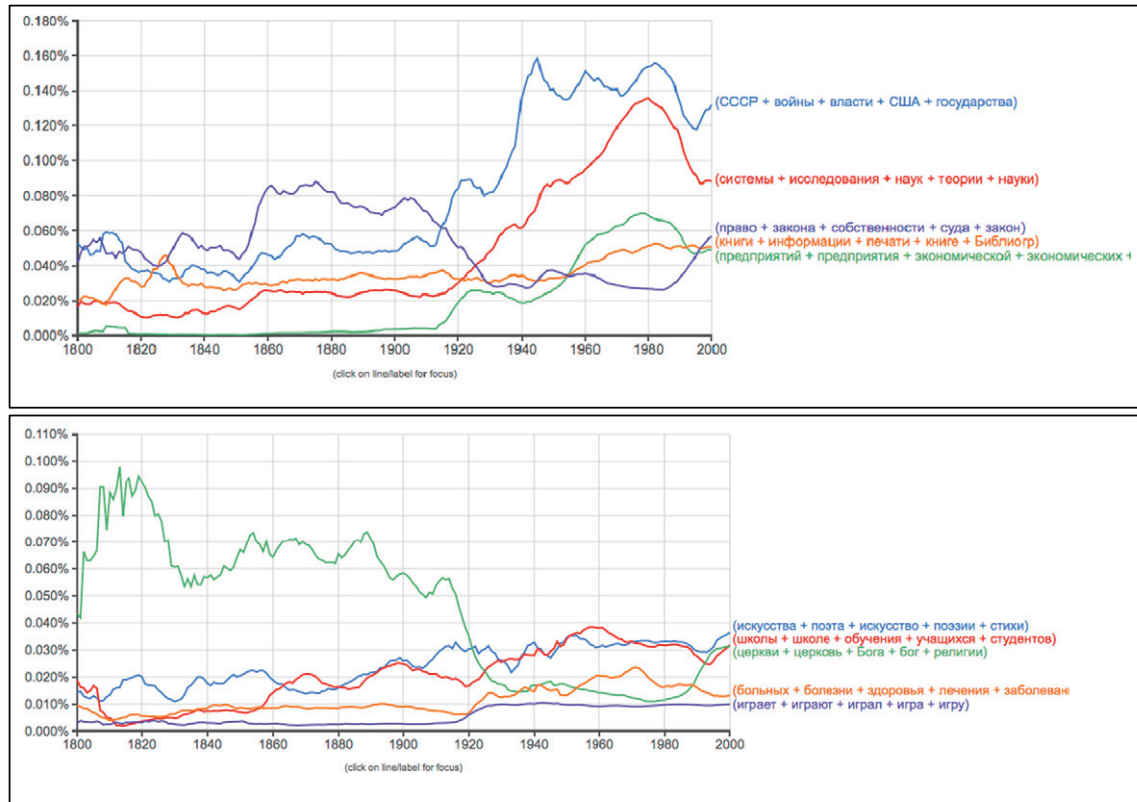
Fig. A2.3. Combined occurrence frequencies of the five most frequent keywords for all ten function systems in the Russian language Google Books corpus (1800–2000).
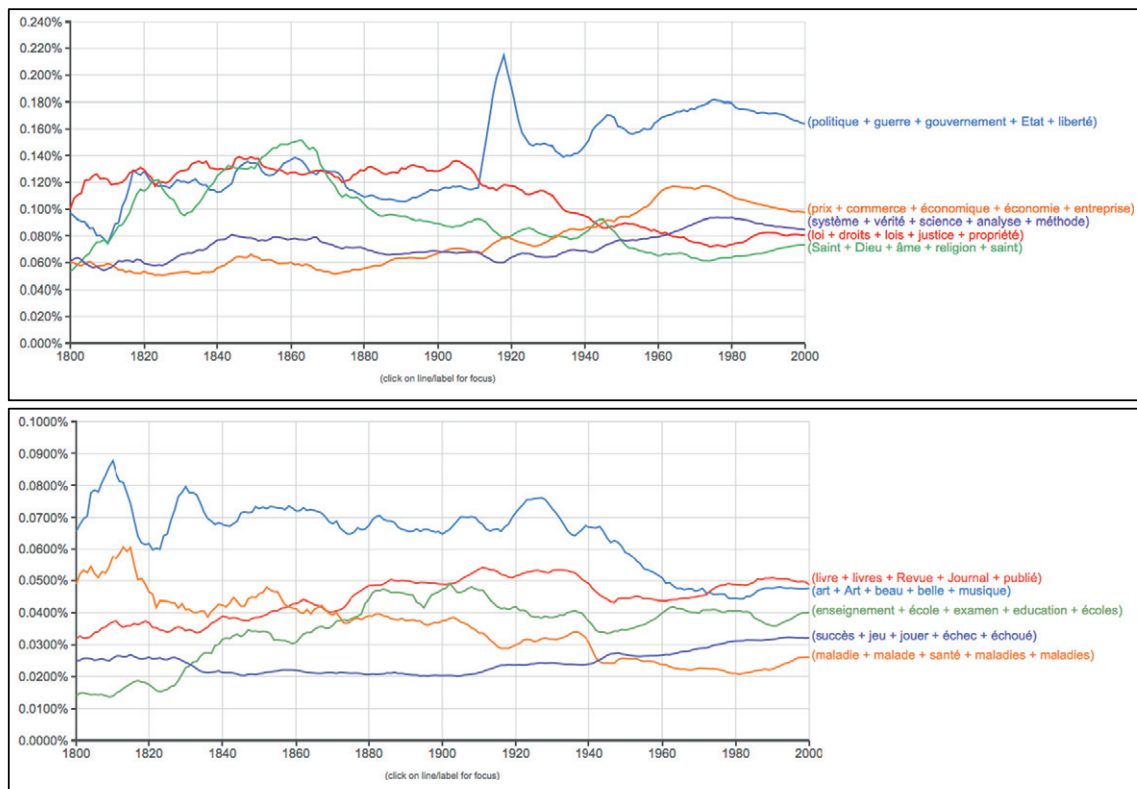

Fig. A2.4. Combined occurrence frequencies of the five most frequent keywords for all ten function systems in the French language Google Books corpus (1800–2000).
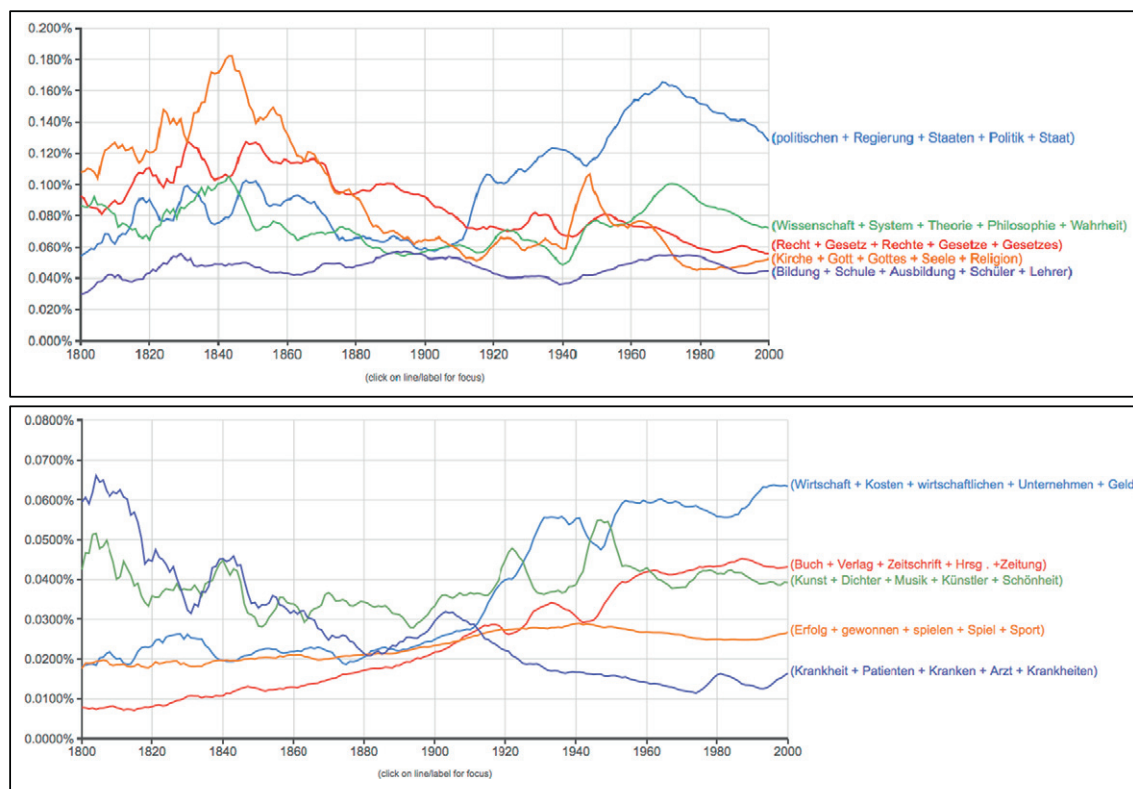
Fig. A2.5. Combined occurrence frequencies of the five most frequent keywords for all ten function systems in the German language Google Books corpus (1800–2000).
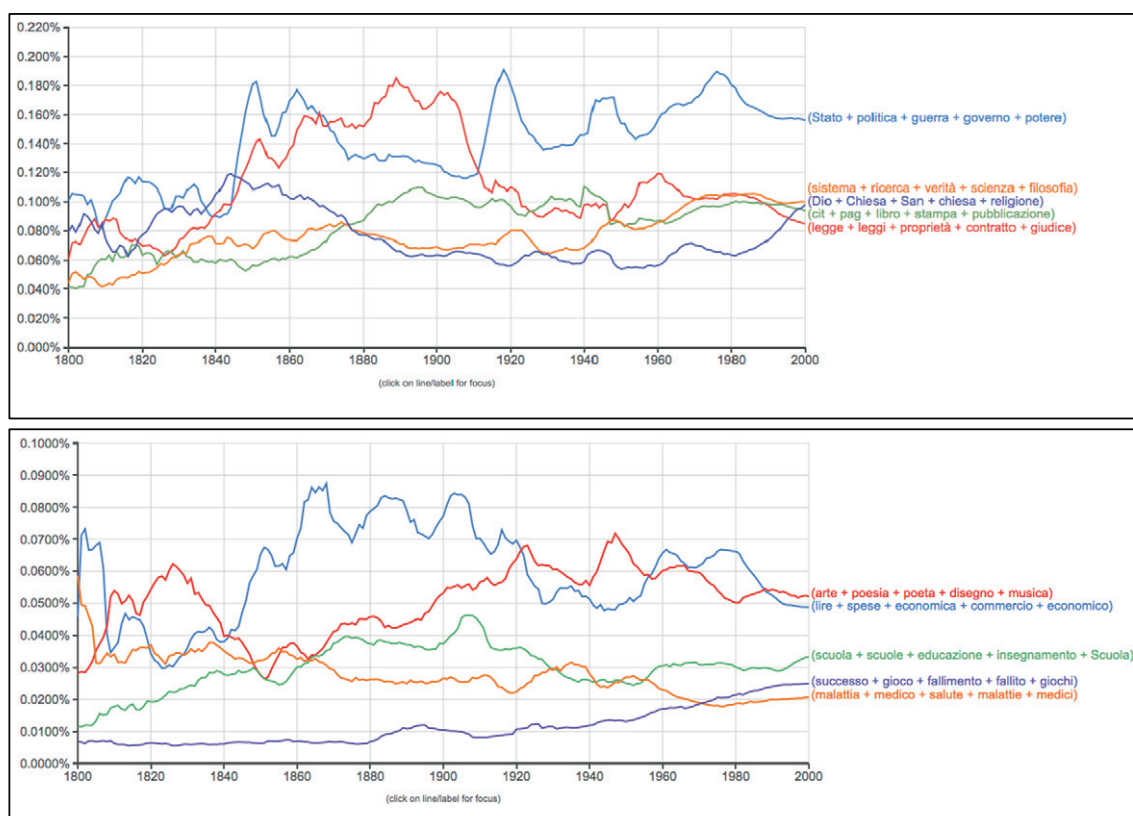

Fig. A2.6. Combined occurrence frequencies of the five most frequent keywords for all ten function systems in the Italian language Google Books corpus (1800–2000).

A3.1 google_word_frequency.py

```python
#!/usr/bin/env python
""" Creates top-N frequency lists based on Google's Ngram datasets published on
    http://storage.googleapis.com/books/ngrams/books/datasetsv2.html
    Only 1-grams are taken into account.
"""

from google_ngram_downloader import readline_google_store
from itertools import groupby
from string import ascii_lowercase
import json
import sys

VALID_LANGUAGES = ['fre', 'ger', 'eng', 'eng-us', 'eng-gb', 'eng-fiction',
                   'chi-sim', 'heb', 'ita', 'rus', 'spa']

DEFAULT_N = 3000
MIN_YEAR = 0
MAX_YEAR = 2012
FILTER_THRESHOLD = 3000


def all_records(lang, indices=[x for x in ascii_lowercase], verbose=True):
    """Returns a generator producing all records for language & indices"""
    if lang not in VALID_LANGUAGES:
        raise Exception('Invalid language %s' % lang)
    for _, _, records in readline_google_store(ngram_len=1, lang=lang,
                                               indices=indices,
                                               verbose=verbose):
        for record in records:
            yield record


def count_ngrams(grouped_records):
    return ((ngram, sum(r.match_count for r in records))
            for ngram, records in grouped_records)


def filter_high_frequency(ngram_counts, threshold):
    return ((word, count) for (word, count) in ngram_counts
            if count > threshold)


def record_filter(record, config):
    ret = record.year in range(config.startyear, config.endyear+1)
```

```python
    if config.debug and ret:
        print("matched: %s" % str(record))
    return ret


def ngram_input(config):
    records = all_records(config.lang, verbose=config.debug)
    return (r for r in records if record_filter(r, config))


def top_n_words(n, records, threshold=FILTER_THRESHOLD):
    """returns the n most frequent words for a given language"""
    return sorted(list(filter_high_frequency(
        count_ngrams(groupby(records, key=lambda r: r.ngram)), threshold)),
        key=lambda word_count: -word_count[1])[:n]


def write_json_file(words, config):
    timespec = "-y%d-%d" % (config.startyear, config.endyear)
    filename = "wordlist-%s-n%d%s.json" % (config.lang, config.n, timespec)
    with open(filename, 'w') as f:
        json.dump(words, f, indent=True)


def parse_config(args):
    from argparse import ArgumentParser
    parser = ArgumentParser(
        description='Extract top frequency words from Google Ngram data')
    parser.add_argument(
        '-n', dest='n', type=int, default=DEFAULT_N,
        help='number of words to extract (default %d)' % DEFAULT_N)
    parser.add_argument(
        '--startyear', dest='startyear', type=int, default=MIN_YEAR,
        help='only include words after specified year (default %d)' % MIN_YEAR)
    parser.add_argument(
        '--endyear', dest='endyear', type=int, default=MAX_YEAR,
        help='only include words up to specified year (default %d)' % MAX_YEAR)
    parser.add_argument(
        '--lang', dest='lang', required=True, choices=VALID_LANGUAGES,
        help='language code')
    parser.add_argument('--debug', action="store_true")
    config = parser.parse_args(args)

    if config.startyear < 0:
        raise Exception('invalid startyear, must be > 0')
    if config.endyear > MAX_YEAR:
        raise Exception('invalid endyear, must be < %d' % MAX_YEAR)
    if config.endyear < config.startyear:
        raise Exception('endyear < startyear')
    return config

if __name__ == '__main__':
    config = parse_config(sys.argv[1:])
    top_n = top_n_words(config.n, ngram_input(config))
    write_json_file(top_n, config)
```

A3.2. post_process.py

```python
#!/usr/bin/env python
""" Parses word lists, and performs some prepocessing and cleanup
    The processed list is written to stdout.
"""

import json
import re
from collections import OrderedDict

POS_PATTERN = re.compile("_[A-Z]+$")


def sort_list(word_list):
    return sorted(word_list, key=lambda word_count: -word_count[1])


def deduplicate(word_list):
    result = OrderedDict()
    for (word, count) in word_list:
        if word not in result:
            result[word] = count
    return list(result.items())


def strip_pos_tags(word_list):
    def strip_pos_tag(word):
        matcher = POS_PATTERN.search(word)
        if matcher:
            return word[:matcher.start()]
        else:
            return word

    return [(strip_pos_tag(word), count) for (word, count) in word_list]


def normalize(word_list):
    return deduplicate(sort_list(strip_pos_tags(word_list)))


if __name__ == '__main__':
    import sys

import codecs

with sys.stdin as f:
    writer = codecs.getwriter('UTF8')(sys.stdout)
    writer.write(json.dumps(normalize(json.load(f)), indent=True))
```

References

Blumler, J.G., Kavanagh, D., 1999. The third age of political communication: influences and features. Polit. Commun. 16:209–230. http://dx.doi.org/10.1080/105846099198596.
Bohannon, J., 2011. The Science Hall of Fame. Science 331, 143.
Castells, M., 1996. Rise of The Network Society. Blackwell Publishers, Cambridge.
Chen, Y., Yan, F., 2016. Centuries of sociology in millions of books The Sociological Review online first 31 July 2016:n/a-n/a. http://dx.doi.org/10.1111/1467-954X.12399.
Gibbs, F.W., Cohen, D.J., 2011. A conversation with data: prospecting Victorian words and ideas. Vic. Stud. 54, 69–77.
Godet, M., 1986. Introduction to 'la prospective': seven key ideas and one scenario method. Futures 18, 134–157.
Heylighen, F., 2011. In: Grinin, L.E., Carneiro, R.L., Korotayev, A.V., Spier, F. (Eds.), Conceptions of a Global Brain: An Historical Review Evolution: Cosmic, Biological, and Social, pp. 274–289.
Heylighen, F., Lenartowicz, M., 2016. The Global Brain as a Model of the Future Information society: An Introduction to the Special Issue Technological Forecasting and Social Change Online First (accessed on July 29, 2016). http://dx.doi.org/10.1016/j.techfore.2016.02.004.
Johnson, C.Y., 2010. In billions of words, digital allies find tale.
Kjaer, P.F., 2010. The metamorphosis of the functional synthesis: a continental European perspective on governance, law, and the political in the transnational space. Wis. Law Rev. 2010, 489–533.
Kloumann, I.M., Danforth, C.M., Harris, K.D., Bliss, C.A., Dodds, P.S., 2012. Positivity of the English language. PLoS One 7, e29484.
Last, C., 2017. Global Commons in the global brain. Technol. Forecast. Soc. Chang. 114, 48–64.
Leetaru, K., 2011. Culturomics 2.0: forecasting large-scale human behavior using global news media tone in time and space. First Monday 16.
Lenartowicz, M., 2016. Creatures of the Semiosphere. A Problematic Third Party in the 'Humans Plus Technology' Cognitive Architecture of the Future Global Superintelligence Technological Forecasting and Social Change Online First (accessed on July 29, 2016). http://dx.doi.org/10.1016/j.techfore.2016.07.006.
Lenartowicz, M., Weinbaum, D.R., Braathen, P., 2016. Social systems: complex adaptive loci of cognition. Emergence Complex. Organ. 18 (doi:http://10.emerg/10.17357.23db2216ba4fc080e77b2a3352a60761).
Leydesdorff, L., 2012. The triple helix, quadruple helix, ..., and an N-tuple of helices: explanatory models for analyzing the knowledge-based economy? J. Knowl. Econ. 3: 25–35. http://dx.doi.org/10.1007/s13132-011-0049-4.
Leydesdorff, L., 2013. N-tuple of helices. Encyclopedia of Creativity, Invention, Innovation and Entrepreneurship. Springer, pp. 1400–1402.
Lin, Y., Michel, J.-B., Aiden, E.L., Orwant, J., Brockman, W., Petrov, S., 2012. Syntactic annotations for the google books Ngram corpus. Proceedings of the ACL 2012 system demonstrations. Association for Computational Linguistics, pp. 169–174.
Luhmann, N., 1977. Differentiation of society. Can. J. Soc. 2:29–53. http://dx.doi.org/10.2307/3340510.
Luhmann, N., 1990. The paradox of system differentiation and the evolution of society. In: Alexander, J.C., Colomy, P.B. (Eds.), Differentiation Theory and Social Change: Comparative and Historical Perspectives. Columbia University Press, New York, pp. 409–440.
Luhmann, N., 1995. Social Systems. Standford University Press, Stanford.
Luhmann, N., 1997. The control of intransparency. Syst. Res. Behav. Sci. 14, 359–371.
Luhmann, N., 2012. Theory of Society vol. 1. Stanford University Press, Palo Alto.
Luhmann, N., 2013. Theory of Society vol. 2. Stanford University Press, Palo Alto.
Luhmann, N., Rasch, W., 2002. Theories of Distinction: Redescribing the Descriptions of Modernity. Stanford University Press.
Michel, J.-B., et al., 2011. Quantitative analysis of culture using millions of digitized books. Science 331:176–182. http://dx.doi.org/10.1126/science.1199644.
Nicholson, B., 2012. Counting culture; or, how to read Victorian newspapers from a distance. Victorian Culture]->J. Vic. Cult. 17, 238–246.
Ophir, S., 2010. A new type of historical knowledge. Inf. Soc. 26:144–150. http://dx.doi.org/10.1080/01972240903562811.
Pechenick, E.A., Danforth, C.M., Dodds, P.S., 2015. Characterizing the Google Books Corpus: strong limits to inferences of socio-cultural and linguistic evolution. PLoS ONE 10:e0137041. http://dx.doi.org/10.1371/journal.pone.0137041.
Roth, S., 2014. Fashionable functions. A Google ngram view of trends in functional differentiation (1800–2000). Int. J. Technol. Hum. Interact. 10, 88–102.
Roth, S., 2015. Free economy! On 3628800 alternatives of and to capitalism. J. Interdiscip. Econ. 27, 107–128.
Roth, S., Clark, C., Berkel, J., 2016. The fashionable functions reloaded. an updated Google Ngram view of trends in functional differentiation (1800–2000). In: Mesquita, A. (Ed.), Research Paradigms and Contemporary Perspectives on Human-Technology Interaction. IGI-Global, Hershey.
Roth, S., Kaivo-oja, J., 2016. Is the future a political economy? Functional analysis of three leading foresight and futures studies journals. Futures 81, 15–26.
Roth, S., Schütz, A., 2015. Ten systems: toward a canon of function systems. Cybern. Hum. Knowing 22, 11–31.
Russell, P., 1982. The Awakening Earth: The Global Brain. Taylor & Francis.
Schönfelder, B., 2016. The evolution of law under communism and post-communism: a system-theory analysis in the spirit of Luhmann. Financ. Theory Pract. 40.
Sparavigna, A., Marazzato, R., 2015. Using Google Ngram Viewer for Scientific Referencing and History of Science arXiv preprint arXiv:151201364.
Thornhill, C., 2008. Towards a historical sociology of constitutional legitimacy. Theory Soc. 37:161–197. http://dx.doi.org/10.1007/s11186-007-9048-7.
Thornhill, C., 2010. Niklas Luhmann and the sociology of the constitution. Journal of Classical Sociology 10:315–337. http://dx.doi.org/10.1177/1468795x10385181.
White, H.C., 1992. Identity and Control: A structural Theory of Action. Princeton UP, New York.

Steffen Roth, *La Rochelle Business School, France, and Yerevan State University, Armenia*
Prof. Dr. Dr. Steffen Roth is an Associate Professor of Strategic Management at La Rochelle Business School, France, and permanent Visiting Professor of Sociology at the Yerevan State University, Armenia. He was awarded a PhD in economics and management from Chemnitz University of Technology, Germany, and a PhD in sociology at the University of Geneva, Switzerland. Steffen Roth is mentor of the Early Career Colloquium of the European Academy of Management (EURAM), corresponding chair of the EURAM conferences track "Next Management Theory", and Associate Editor of *Kybernetes*.

Carlton Clark, *University of Wisconsin-La Crosse, US*
Carton Clark, PhD, is a lecturer at the University of Wisconsin-La Crosse, US, where he is teaching and doing research on fields such as Rhetoric and Composition, Hypertext Theory, and South American Literature. He holds a PhD from Texas Woman's University, US.

Nikolay Trofimov, *Institute for the Study of Science of the Russian Academy of Sciences, Russia*
Nikolay Trofimov is a Senior Researcher at the Institute for the Study of Science of the Russian Academy of Sciences, performing research in the field of science, technology and innovation.

Artur Mkrtichyan, *Yerevan State University, Armenia*
Prof. Dr. Habil. Artur Mkrtichyan is the Dean of the Faculty of Sociology at Yerevan State University, Yerevan, Armenia. He is a Visiting Professor at Humboldt University Berlin and has held Fellowships at the Universities of Bielefeld and Innsbruck.

Markus Heidingsfelder, *Habib University, Pakistan*
Prof. Dr. Markus Heidingsfelder is an Assistant Professor of Media Studies at Habib University of Kararchi Department of Communication Studies and Design. He holds a PhD in literature studies from LMU Munich. He has been the chief of editorial offices at the German Music TV station VIVA as well as MTV Germany.

Laura Appignanesi, *University of Macerata, Italy*
Laura Appignanesi is a PhD student in Economics, Management, and Social Sciences at the University of Macerata, Italy. She holds a MA degree in Business Administration from the University of Ancona.

Miguel Pérez-Valls, *University of Alméria, Spain*
Prof. Dr. Miguel Pérez-Valls is an Associate Professor at the University of Alméria School of Business and Economics, where he also earned his PhD. He has been a Visiting Scholar at the University of Surrey and at Cardiff University. Miguel Pérez-Valls is mentor of the Early Career Colloquium of the European Academy of Management (EURAM).

Jan Berkel, *Independent, Portugal*
Jan Berkel is an independent coder living in Costa Rica.

Jari Kaivo-oja, *Turku School of Economics, Finland*
Prof. Dr. Jari Kaivo-oja is the Research Director at the Finland Futures Research Centre of the Turku School of Economics as well as an Adjunct Professor at the University of Helsinki and at the University of Lapland. He has worked for the European Commission (FT6, FP7, H2020), the European Foundation, the Nordic Innovation Center (NIC), the Finnish Funding Agency for Technology and Innovation (TEKES), EUROSTAT, RAND Europe, and for the European Parliament. Currently Dr Jari Kaivo-oja is a researcher at RISCAPE (Horizon 2020), at EUFORIE (Horizon 2020), at EL-TRAN (Academy of Finland) and at TRYOUT! (The European Regional Development Fund, ERDF).