Voluntas (2023) 34:52–63
https://doi.org/10.1007/s11266-021-00414-x

ISTR International Society for Third-Sector Research

RESEARCH PAPERS

Computational Social Science for Nonprofit Studies: Developing a Toolbox and Knowledge Base for the Field

Ji Ma[1] · Islam Akef Ebeid[1] · Arjen de Wit[2] · Meiying Xu[1] · Yongzheng Yang[3] · René Bekkers[2] · Pamala Wiepking[2,3]

Accepted: 4 September 2021 / Published online: 12 October 2021
© The Author(s) 2021

Abstract How can computational social science (CSS) methods be applied in nonprofit and philanthropic studies? This paper summarizes and explains a range of relevant CSS methods from a research design perspective and highlights key applications in our field. We define CSS as a set of computationally intensive empirical methods for *data management*, *concept representation*, *data analysis*, and *visualization*. What makes the computational methods "social" is that the purpose of using these methods is to serve quantitative, qualitative, and mixed-methods social science research, such that theorization can have a solid ground. We illustrate the promise of CSS in our field by using it to construct the largest and most comprehensive database of scholarly references in our field, the Knowledge Infrastructure of Nonprofit and Philanthropic Studies (KINPS). Furthermore, we show that through the application of CSS in constructing and analyzing KINPS, we can better understand and facilitate the intellectual growth of our field. We conclude the article with cautions for using CSS and suggestions for future studies implementing CSS and KINPS.

Keywords Computational social science · Nonprofit · Philanthropy · Knowledge Infrastructure of Nonprofit and Philanthropic Studies · KINPS

Since computational social science (CSS) was coined in 2009 (D. Lazer et al., 2009), it has been growing exponentially in many social science disciplines and is projected to have the potential to revolutionize social science studies (D. M. J. Lazer et al., 2020). Over the past decade, the field of nonprofit and philanthropic studies has also begun to apply computational methods, such as machine learning and automated text analysis. We start this article by explaining CSS from a research design perspective and framing its applications in studying the nonprofit sector and voluntary action. Next, we illustrate the promise of CSS for our field by applying these methods to consolidate the scholarship of nonprofit and philanthropic studies—creating a bibliographic database to cover the entire literature of the research field. The article concludes with critical reflections and suggestions. This article speaks to three audiences: (1) readers without technical background can have a structural understanding of what CSS is, and how they can integrate them into their research by either learning or collaboration; (2) technical readers can review these methods from a research design perspective, and the references cited are useful for constructing a CSS course; and (3) readers motivated to study the intellectual growth of our field can discover novel methods and a useful data source. The primary purpose of this short piece is not to exhaust all CSS methods and technical details, which are introduced in most textbooks and references cited.

✉ Ji Ma
maji@austin.utexas.edu

Pamala Wiepking
pwiepki@iu.edu

[1] The University of Texas at Austin, Austin, TX, USA

[2] Center for Philanthropic Studies, Department of Sociology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[3] Indiana University–Purdue University Indianapolis, Indianapolis, Indiana, USA

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

Computational Social Science for Nonprofit Studies: A Toolbox of Methods

Though all empirical analysis methods are computational to some extent, why are some framed as "computational social science methods" (CSS) while others are not? Is it just a fancy but short-lived buzzword, or a new methodological paradigm that is fast evolving?

Empirical studies of social sciences typically include two essential parts: theorization and empirical research (Fig. 1; Shoemaker, Tankard, and Lasorsa 2003; Ragin & Amoroso, 2011, 17; Cioffi-Revilla, 2017). Theorization focuses on developing concepts and the relationship among these concepts, while empirical research emphasizes representing these concepts using empirical evidence and analyzing the relationship between concepts (Shoemaker, Tankard, and Lasorsa 2003, 51). The relationship between theorization and empirical research is bidirectional or circular—research can be either theory-driven (i.e., deductive), data-driven (i.e., inductive), or a combination of both. Quantitative and qualitative studies may vary in research paradigm and discourse, but they typically follow a similar rationale as Fig. 1 illustrates.

CSS has been widely discussed but poorly framed—an important reason causing many scholars' perception that the CSS is only a buzzword but not a methodological paradigm. We define CSS as *a set of computationally intensive empirical methods employed in quantitative, qualitative, and mixed-methods social science research for data management, concept representation, data analysis, and visualization*. What makes computational methods "social" is the objective to serve empirical social science research, such that theorization can have a solid ground, either by completing the deductive or the inductive cycle. What makes social science methods "computational" is the use of innovative and computationally intensive methods. The advantage of CSS for our highly interdisciplinary field is that it facilitates collaboration across traditional disciplinary borders, a promise that is being materialized in other fields of research (D. M. J. Lazer et al., 2020).

CSS methods primarily serve the four aspects of empirical research as included in Fig. 1: data management, concept representation, data analysis, and visualization. *Data management* methods help represent, store, and manage data efficiently. This is especially relevant when dealing with "big data"—heterogeneous, messy, and large datasets. *Concept representation* methods help operationalize concepts. For example, using sentiment analysis in natural language processing to scale political attitudes. These computational methods are complementary with traditional operationalizations such as attitude items in surveys or questions in interviews. *Data analysis* in CSS

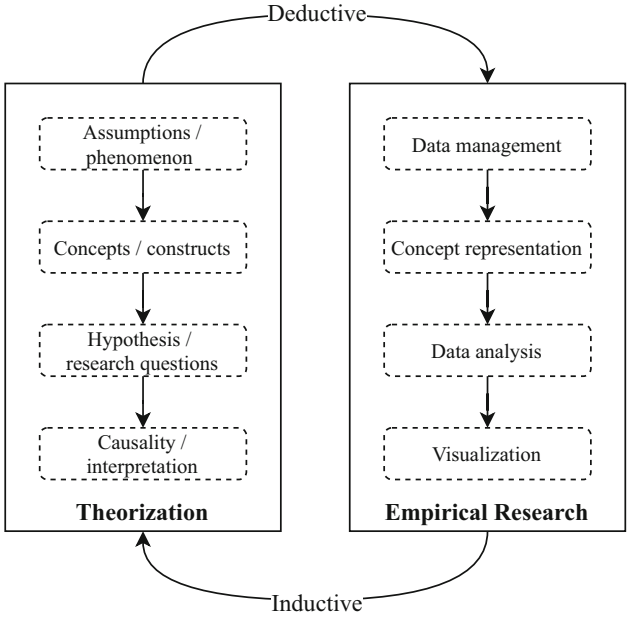


Fig. 1 Structure of empirical social science studies. A diagram summary of Shoemaker, Tankard, and Lasorsa (2003), adapted by the authors of this paper

shares many statistical fundamentals with statistics (e.g., probability theory and hypothesis testing) but typically consumes more computational resources. The *visualization* of CSS illustrates data from multiple dimensions and using graphs that enable human–data interaction, so that consumers can closely examine the data points of interests within a massive dataset.

Table 1 presents a list of the most commonly used computational methods. The following sections briefly introduce them and provide applications in nonprofit studies. Our purpose is not to be comprehensive and exhaustive, but to introduce the principles behind these methods from a research design perspective in non-jargon language and within the context of nonprofit studies.

Data Management

Science is facing a reproducibility crisis (Baker, 2016; Hardwicke et al., 2020). Since researchers using CSS methods usually deal with large volumes of data, and their analysis methods contain many parameters that need to be specified, they need to be extra cautious to reproducibility issues. Fortunately, researchers from various scientific disciplines have identified an inventory of best practices that contribute to reproducibility (Gentzkow & Shapiro, 2014; Wilson et al., 2017).

As a starting point for data management, an appropriate data structure helps represent and store real-world entities and relationships, which is fundamental to all empirical

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

Table 1 Common computational social science methods and their roles in empirical studies

Computational methods	Empirical component of social science studies			
	Data management	Concept representation	Data analysis	Visualization
Relational database and tidy data	X			
Documentation and automation	X			
Network analysis		X	X	X
Machine learning		X	X	
Natural language processing		X		X

studies. Such demands can be met by using a relational database that has multiple interrelated data tables (Bachman, 1969; Codd, 1970). There are two important steps for constructing such a database. First, store homogeneous data record in the same table and uniquely identify these records. Wickham (2014) coined the practices of "Tidy Data," which offer guidelines to standardize data preprocessing steps and describe how to identify untidy or messy data. Tidy datasets are particularly important for analyzing and visualizing longitudinal data (Wickham, 2014, 14). Second, relate different tables using shared variables or columns and represent the relationships between different tables using graphs, also known as a database schema or entity-relationship model (Chen, 1976).

Because CSS methods heavily rely on data curation and programming languages such as Python and R, documentation and automation can improve the replicability and transparency of research (Corti et al., 2019; Gentzkow & Shapiro, 2014). The best practices of documentation include adhering to a consistent naming convention and using a version control system such as GitHub to track changes. The primary purpose of automation is to standardize the research workflow and improve reproducibility and efficiency.

Knowledge about data management is not new, but it becomes particularly essential to nonprofit scholars in the digital age because they often deal with heterogeneous, massive, and messy data. For example, Ma et al. (2017) and Ma (2020) constructed a relational database normalizing data on over 3,000 Chinese foundations from six different sources across 12 years. Data from different sources can be matched using codes for nonprofit organizations (De Wit, Bekkers, and Broese van Groenou 2017) or unique countries (Wiepking et al. 2021). Without the principles of data management, it is impossible to use many open-government projects about the nonprofit sector, such as U.S. nonprofits' tax forms[1] and the registration information of charities in the UK. Furthermore, a growing number of academic journals, publishers, and grant agencies in social sciences have started to require the public access to source codes and data. Therefore, it is important to improve students' training in data management, as this is currently often not part of philanthropic and nonprofit studies programs.

Network Analysis

While the notion of social relations and human networks has been fundamental to sociology, modern network analysis methods only gained momentum since the mid twentieth century, along with the rapid increase in computational power (Scott, 2017, 12–13). A network is a graph that comprises nodes (or "vertices," i.e., the dots in a network visualization) and links (or "edges"), and network analysis uses graph theory to analyze a special type of data—the relation between entities.

Researchers typically analyze networks at different levels of analysis, for example, nodal, ego, and complete networks (Wasserman & Faust, 1994, 25). At the nodal level, research questions typically focus on the attributes of nodes and how the nodal attributes are influenced by relations. At the ego network level, researchers are primarily interested in studying how the node of interest interacts with its neighbors. At the complete network level, attributes of the entire network are calculated, such as measuring the connectedness of a network. Research questions at this level usually intend to understand the relation between network structure and outcome variables. The three levels generally reflect the analyses at micro-, meso-, and macro-levels. Researchers can employ either a single-level or multi-level design, and the multi-level analysis allows scholars to answer complex sociological questions and construct holistic theories (e.g., Lazega et al., 2013; Müller et al., 2018).

Nonprofits scholars have been using metrics of network analysis to operationalize various concepts. For example, the connectedness of a node or the entire network can be

[1] See a list of sources on Nonprofit Open Data Collective: https://web.archive.org/web/20210508182350/https://nonprofit-open-data-collective.github.io/.

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

regarded as measuring social capital of individuals or communities (Herzog & Yang, 2018; Xu and Saxton 2019; Yang, Zhou, and Zhang 2019; Nakazato & Lim, 2016). Network analysis has been also applied to studying inter-organizational collaboration (Bassoli, 2017; Shwom, 2015), resource distribution (Lai, Tao, and Cheng 2017), interlocking board networks (Ma & DeDeo, 2018; Paarlberg, Hannibal, and McGinnis Johnson 2020; Ma, 2020), and the structure of civil societies (Seippel, 2008; Diani, Ernstson, and Jasny 2018). Networks can even be analyzed without real-world data. For example, Shi et al. (2017) created artificial network data simulating different scenarios to test how different organizational strategies affect membership rates.

Using social media data to analyze nonprofits' online activities is a recent development with growing importance (Guo & Saxton, 2018; Xu and Saxton 2019; Bhati & McDonnell, 2020). However, social media platforms may often restrict data access because of privacy concerns, which encouraged "a new model for industry—academic partnerships" (King & Persily, 2020). Researchers also have started to develop data donation projects, in which social media users provide access to their user data. For instance, Bail et al. (2017) offered advocacy organizations an app with insights in their relative Facebook outreach, asking nonpublic data about their Facebook page in return.

Machine Learning

Machine learning (ML) can "discover new concepts, measure the prevalence of those concepts, assess causal effects, and make predictions" (Grimmer et al., 2021, 395; Molina & Garip, 2019). For social scientists, the core of applying ML is to use computational power to learn or identify features from massive observations and link those features to outcomes of interest. For example, researchers only need to manually code a small subset of data records and train a ML algorithm with the coded dataset, a practice known as "supervised machine learning." Then, the trained ML algorithm can help researchers efficiently and automatically classify the rest of the records which may be in millions. ML algorithms can also extract common features from massive numbers of observations according to preset strategies, a practice known as "unsupervised machine learning." Researchers can then assess how the identified features are relevant to outcome variables. In both scenarios, social scientists can analyze data records that go beyond human capacity, so that they can focus on exploring the relationship between the features of input observations and outcomes of interest.

Despite these advantages, ML methods also suffer from numerous challenges. A recurrent issue is the black-box effect concerning the interpretation of results. The trained algorithms often rely on complex functions but provide little explanation on why those results are reasonable. Along with the advancement of programming languages, ML methods are becoming more accessible to researchers. However, scientists should be cautious to the parameters and caveats that are pre-specified by ML programming packages. Human validation is still the gold standard for applying ML-devised instruments in social science studies.

Although nonprofit scholars have not yet widely employed ML in their analysis, the methods have already shown a wide range of applications. For example, ML algorithms were experimented in analyzing nonprofits' mission statements (Litofcenko, Karner, and Maier 2020; Ma, 2021) and media's framing of the Muslim nonprofit sector (Wasif 2021).

Natural Language Processing

Natural language processing (NLP) aims at getting computers to analyze human language (Gentzkow et al., 2019; Grimmer & Stewart, 2013). The purposes of NLP tasks can be primarily grouped into two categories for social scientists: identification and scaling. *Identification* methods aim at finding the themes (e.g., topic modeling) or entities (e.g., named-entity recognition) of a given text, which is very similar to the grounded theory approach in qualitative research (Baumer et al., 2017). *Scaling* methods put given texts on a binary, categorical, or continuous scale with social meanings (e.g., liberal-conservative attitudes). Identification and scaling can be implemented through either a dictionary approach (i.e., matching target texts with a list of attribute keywords or another list of texts) or a machine learning approach. Although NLP methods are primarily developed in computational linguistics, they can also serve as robust instruments in social sciences (Rodriguez & Spirling, 2021).

Table 2 lists empirical studies that are relevant to nonprofit and philanthropic studies. Scholars in other disciplines offer additional examples of the potential of NLP methods. For example, researchers in public administration and political science have applied sentiment analysis and topic modeling to find clusters of words and analyze meanings of political speeches, assembly transcripts, and legal documents (Mueller & Rauh, 2018; Parthasarathy et al., 2019; Anastasopoulos & Whitford, 2019; Gilardi, Shipan, and Wüest 2020). In sociology, text mining has proven useful to extract semantic aspects of social class and interactions (Kozlowski et al., 2019; Schröder et al., 2016). As Evans and Aceves (2016, 43) summarize, although NLP methods cannot replace creative researchers, they can identify subtle associations from massive texts that humans cannot easily detect.

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

Table 2 Example articles studying nonprofits with natural language processing methods

		Purpose of natural language processing	
		Identification	Scaling
Dictionary approach		Fyall, Moore, and Gugerty (2018); Litofcenko, Karner, and Maier (2020)	Ma, Jing, and Han (2018); Paxton, Velasco, and Ressler (2020); Brandtner (2021)
Machine learning approach	Unsupervised	Kang, Baek, and Kim (2021); Wasif (2021)	Chen and Zhang (2021)
	Supervised	Ma (2021)	Wasif (2020;2021)

Applying the Methods: The Knowledge Infrastructure of Nonprofit and Philanthropic Studies

Most of the social science disciplines have dedicated bibliographic databases, for example, Sociological Abstracts for sociology and Research Papers in Economics for economics. These databases serve as important data sources and knowledge bases for tracking, studying, and facilitating the disciplines' intellectual growth (e.g., Moody, 2004; Goyal, van der Leij, and Moraga-González 2006).

In the past few decades, the number of publications on nonprofit and philanthropy has been growing exponentially (Shier & Handy, 2014, 817; Ma & Konrath, 2018, 1145), and nonprofit scholars have also started to collect bibliographic records from different sources to track the intellectual growth of our field. For example, Brass et al. (2018) established the NGO Knowledge Collective[2] to synthesize the academic scholarship on NGOs. Studying our field's intellectual growth has been attracting more scholarly attention (Walk & Andersson, 2020; Minkowitz et al., 2020; Kang, Baek, and Kim 2021).

To consolidate the produced knowledge, it is important to establish a dedicated bibliographic database which can serve as an infrastructure for this research field. CSS not only provides excellent tools for constructing such a database, but also becomes central to studying and facilitating knowledge production (Edelmann et al., 2020, 68). By applying the newest CSS advancements introduced earlier, we created a unique database: the Knowledge Infrastructure of Nonprofit and Philanthropic Studies (KINPS; https://doi.org/10.17605/OSF.IO/NYT5X). KINPS aims to be the most comprehensive and timely knowledge base for tracking and facilitating the intellectual growth of our field. In the second section of this article, we use the KINPS to provide concrete examples and annotated code scripts for a state-of-the-art application of CSS methods in our field.

Data Sources of the KINPS

The KINPS currently builds on three primary data sources: (1) Over 67 thousand bibliographical records of nonprofit studies between 1920s and 2018 from Scopus (Ma & Konrath, 2018); (2) Over 19 thousand English records from the Philanthropic Studies Index maintained by the Philanthropic Studies Library of Indiana University–Purdue University Indianapolis; and (3) Google Scholar, the largest bibliographic database to date (Gusenbauer, 2019; Martín-Martín et al., 2018).

Database Construction Methods

Constructing the database primarily involves three tasks: (1) normalizing and merging heterogeneous data records; (2) establishing a classification of literature; and (3) building a knowledge graph of the literature. As Table 3 presents, each of the three tasks requires the application of various computational methods introduced earlier. We automate the entire workflow so that an update only takes a few weeks at most.[3]

Normalizing Data Structure from Different Sources

The bibliographic records from different sources are in different formats, so the first task is to normalize these heterogeneous entries using the same database schema and following the principles of relational databases. This task is especially challenging when different data sources record the same article as Fig. 2 illustrates.

To normalize and retain all the information of an article from different sources, the schema of the KINPS should achieve a fair level of "completeness" that can be evaluated from three perspectives: schema, column, and population (Ma et al., 2017). *Schema completeness* of the KINPS measures the degree to which the database schema can capture as many aspects of an article as possible. As Fig. 2 illustrates, the schema of the KINPS includes both "Reference Table" and "Classification Table." *Column*

[2] Archived version of its official website: https://web.archive.org/web/20210506024336/https://ngoknowledgecollective.org/.

[3] It takes so "long" because most data sources have quota limits.

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

Table 3 Computational social science methods used in constructing the Knowledge Infrastructure of Nonprofit and Philanthropic Studies

Database construction tasks for KINPS	Computational social science methods				
	Relational database and tidy data	Documentation and automation	Network analysis	Machine learning	Natural language processing
Data normalization	X	X			X
Literature classification		X		X	X
Knowledge graph		X	X		

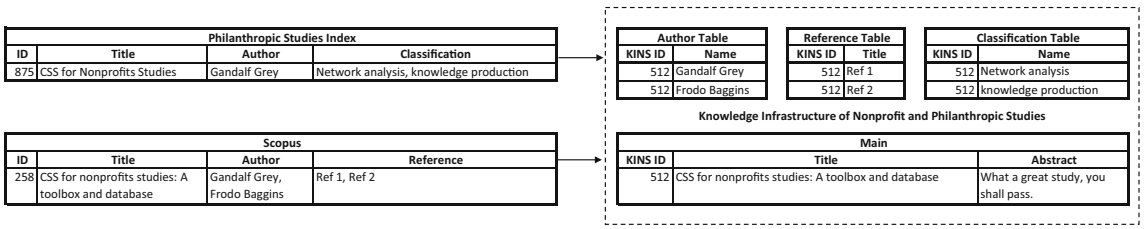


Fig. 2 An example of data normalization

completeness measures the comprehensiveness of attributes for a specific perspective. For example, only the KINPS has the "Abstract" attribute in the "Main" table. *Population completeness* refers the extent to which we can capture the entire nonprofit literature. It can be evaluated by the process for generating the corpus, which was detailed in Ma and Konrath (2018, 1142). Figure 3 shows the latest design of KINPS's database schema.

Merging Heterogeneous Data Records using NLP Methods

Another challenge is disambiguation, a very common task in merging heterogeneous records. As Fig. 2 shows, records of the same article from different sources may vary slightly. The disambiguation process uses NLP methods to measure the similarity between different text strings.

A given piece of text needs to be *preprocessed* and *represented* as numbers using different methods so that they can be calculated by mathematical models (Jurafsky & Martin, 2020, 96). The preprocessing stage usually consists of tokenization (i.e., splitting the text strings into small word tokens) and stop word removal (e.g., taking out "the" and "a"). The current state-of-the-art representation methods render words as vectors in a high dimensional semantic space pre-trained from large corpus (Devlin et al., 2019; Mikolov et al., 2013).

For the disambiguation task, after preprocessing the text strings of publications from different data sources, we converted the text strings to word vectors using the conventional count vector method (Ma, 2021, 670) and then measured the similarity between two text strings by calculating the cosine of the angle between the two strings' word vectors (Jurafsky & Martin, 2020, 105). This process helped us link over 3,100 records from different sources with high confidence (code script available at https://osf.io/pt89w/).

Establishing a Classification of Literature

Classification reflects how social facts are constructed and legitimized from a Durkheimian perspective. A classification of literature presents the anatomy of scholarly activities and also forms the basis for building knowledge paradigms in a discipline or research area (Kuhn, 1970). What is the structure of knowledge production by nonprofit scholars, how does the territory evolve time, and what are the knowledge paradigms in the field? To answer such fundamental questions, the literature of nonprofit and philanthropy needs to be classified in the first place.

We classified references in the KINPS using state-of-the-art advancements in supervised machine learning and NLP (Devlin et al., 2019). After merging data records from different sources, 14,858 records were labeled with themes and included abstract texts. We used the title and abstract texts as input and themes as output to train a ML algorithm. After the classification algorithm (i.e., classifier) was trained and validated, it was used to predict the topics of all 60 thousand unlabeled references in the KINPS (code script available at https://xxx).

The classification in KINPS should be developed and used with extreme prudence because it may influence future research themes in our field. We made a great effort to assure that the classification is relevant, consistent and representative. First, the original classification was created

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

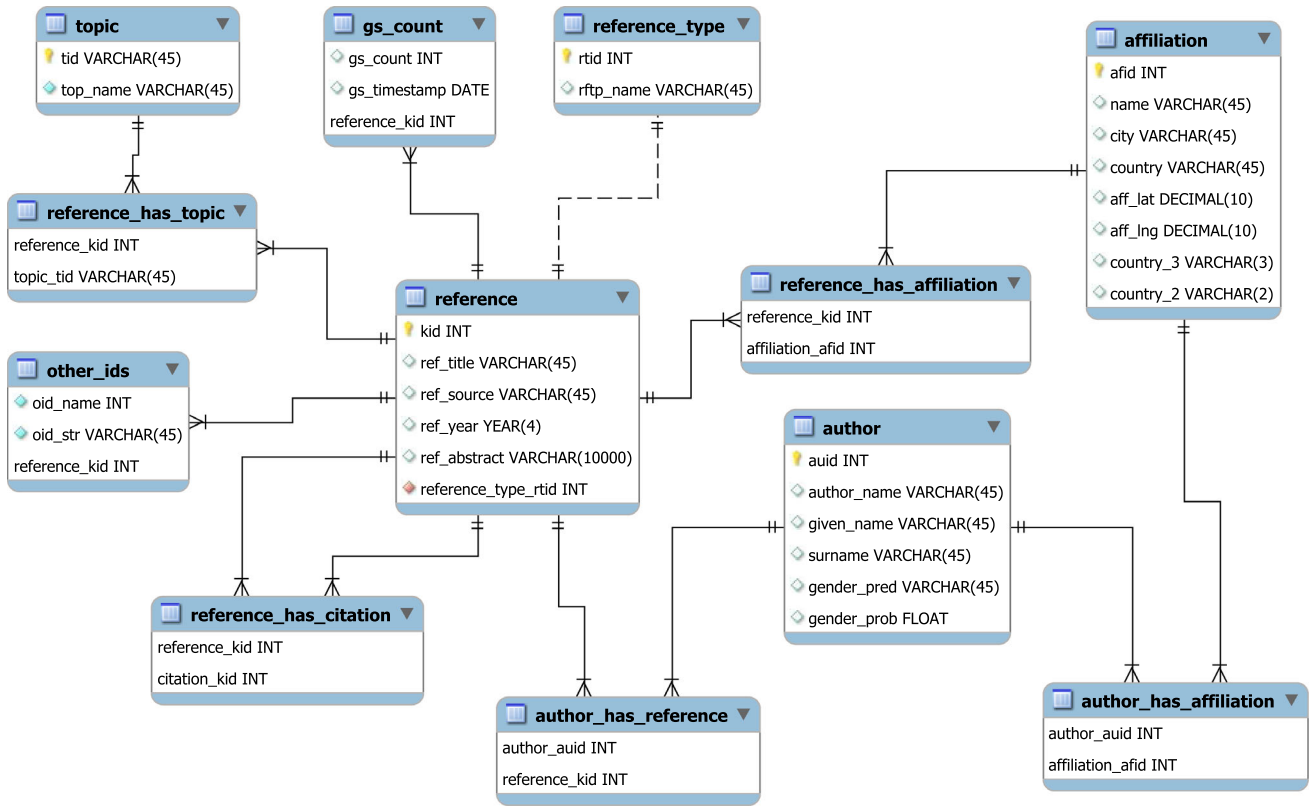


Fig. 3 Design of database schema of the Knowledge Infrastructure of Nonprofit and Philanthropic Studies (2020–12-14 update)

by a professional librarian of nonprofit and philanthropic studies[4] between the late 1990s and 2015. Second, we normalized the original classification labels following a set of rules generated by three professors of philanthropic studies and two doctoral research assistants with different cultural and educational backgrounds. Third, we invited a group of nonprofit scholars to revise the predicted results, and their feedback can be used to fine-tune the algorithm. In future use of the database, continuously repeating this step will be necessary to reflect changes in research themes in the field. Lastly, bearing in mind that all analysis methods should be applied appropriately within a theoretical context, if scholars find our classification unsatisfactory, they can follow our code scripts to generate a new one that may better fit their own theoretical framing.

Building a Knowledge Graph of the Literature

From the perspective of disciplinary development, three levels of knowledge paradigm are crucial to understand the maturity of a research field. Concepts and instruments are *construct paradigms* (e.g., social capital), which are the basis of *thematic paradigms*[5] (e.g., using social capital to study civic engagement). By organizing different thematic paradigms together, we are able to analyze the *meta-paradigms* of our knowledge (Bryant, 1975, 356).

We can use a network graph to analyze the structure and paradigms of the knowledge in our field (Boyack et al., 2005). Figure 4 illustrates the knowledge structure of nonprofit and philanthropic studies based on the KINPS. The online appendix (https://osf.io/vyn6z/) provides the raw file of this figure and more discussion from the perspectives of education, publication, and disciplinary development.

In this network graph, nodes represent the classifications labels established in the preceding section, two nodes are connected if a reference is labeled with both subjects, and the edge weight indicates the times of connection. The nodes are clustered using an improved method of community detection and visualized using a layout that can better distinguish clusters (Martin et al., 2011; Traag et al., 2019). Details and source codes are available in the OSF repository (code script available at https://osf.io/tnqkr/).

As Fig. 4 shows, there are two tightly connected meta-paradigms in our field: *humanities* and *social science*

[4] We very much appreciate Fran Huehls for her valuable and enormous work.

[5] The original study analyzes a specific discipline (i.e., sociology). We adapted the name (i.e., "sociological paradigm") to fit the study of other disciplines and research areas.

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

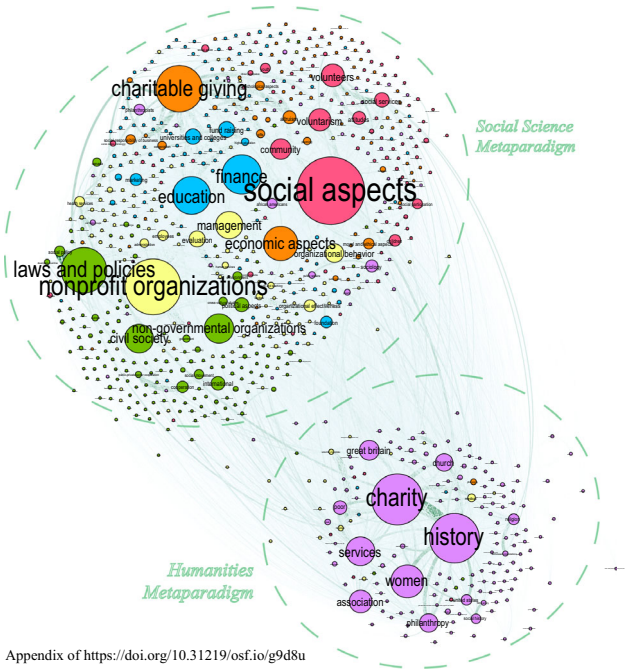


Fig. 4 A visualization of the knowledge structure of nonprofit and philanthropic studies

metaparadigms. We encourage readers to discover the key references related to the different paradigms via the KINPS's online interface. The humanities metaparadigm includes historical studies of charity, women, church, and philanthropy and many other topics. The social science metaparadigm includes five thematic paradigms represented in different colors. For each paradigm, we mention key topics: (1) the *sociological paradigm* includes the study of local communities and volunteering; (2) the *economic paradigm* includes research on giving and taxation; (3) the *finance paradigm* includes research on fundraising, marketing, and education; (4) the *management paradigm* studies evaluation, organizational behavior, and employees and prefers "nonprofit organizations" in discourse; and (5) the *political and policy paradigm* includes research on law and social policy, civil society, and social movements and prefers "non-governmental organizations" in discourse. More thematic paradigms can be found by fine-tuning the community detection algorithm (e.g., Heemskerk & Takes, 2016, 97), which will be part of future in-depth analysis of the KINPS.

Overall, the empirical examples here provide us a stimulus for studying the field's development. Nonprofit scholars have been talking about intellectual cohesion and knowledge paradigms as indicators of this field's maturity (Young, 1999, 19; Shier & Handy, 2014; Ma & Konrath, 2018). Future studies can build on existing literature, the KINPS database, and the computational methods introduced in the proceeding sections to assess the intellectual growth of our field.

Facing the Future of Nonprofit Studies: Promoting Computational Methods in Our Field

We strongly believe that computational social science methods provide a range of opportunities that could revolutionize nonprofit and philanthropic studies. First, CSS methods will contribute to our field through their novel potential in theory building and provide researchers with new methods to answer old research questions. Using computational methods, researchers can generate, explore, and test new ideas at a much larger scale than before. As an example, for the KINPS, we did not formulate a priori expectations or hypotheses on the structure of nonprofit and philanthropic studies. The knowledge graph merely visualizes the connections between knowledge spaces in terms of disciplines and methodologies. As such it is a purely descriptive tool. Now that it is clear how themes are studied in different paradigms and which vocabularies are emic to them, we can start to build mutual understanding and build bridges between disconnected knowledge spaces. Also, we can start to test theories on how knowledge spaces within nonprofit and philanthropic studies develop (Frickel & Gross, 2005; Shwed & Bearman, 2010).

Second, CSS methods combine features of what we think of as "qualitative" and "quantitative" research in studying nonprofits and voluntary actions. A prototypical qualitative study relies on a small number of observations to produce inductive knowledge based on human interpretation, such as interviews with foundation leaders. A prototypical quantitative study relies on a large number of observations to test predictions based on deductive reasoning with statistical analysis of numerical data, such as scores on items in questionnaires completed by volunteers. A prototypical CSS study can utilize a large number of observations to produce both inductive and deductive knowledge. For example, computational methods like machine learning can help researchers inductively find clusters, topics or classes in the data (Molina & Garip, 2019), similar to the way qualitative research identifies patterns in textual data from interviews. These classifications can then be used in statistical analyses that may involve hypothesis testing as in quantitative research. With automated sentiment analysis in NLP, it becomes feasible to quantify emotions, ideologies, and writing style in text data, such as nonprofits' work reports and mission statements (Farrell, 2019; J. D. Lecy et al., 2019). Computational social science methods can also be used to analyze audiovisual content, such as pictures and videos. For example, CSS methods will allow to study the use of

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

pictures and videos as fundraising materials and assess how these materials are correlated with donation.

Third, a promising strength of CSS methods is the practice of open science, including high standards for reproducibility.[6] Public sharing of data and source code not only provides a citation advantage (Colavizza et al., 2020), but also advances shared tools and datasets in our field. For instance, Lecy and Thornton (2016) developed and shared an algorithm linking federal award records to recipient financial data from Form 990 s. Across our field, there is an increasing demand for data transparency. To illustrate the typical open science CSS approach, with the current article, we not only provide access to the KINPS database, but also annotated source codes for reproducing, reusing, and educational purposes.

Implementing CSS also raises concerns and risks. Like all research and analytical methods, CSS methods are not definitive answers but means to answers. There are ample examples of unintended design flaws in CSS that can lead to serious biases in outcomes for certain populations. ML algorithms for instance can reproduce biases hidden in training dataset, and then amplify these biases while applying the trained algorithms at scale. In addition, researchers may perceive CSS to be the panacea of social science research. The ability to analyze previously inaccessible and seemingly unlimited data can lead to unrealistic expectations in research projects. Established criteria that researchers have used for decades to determine the importance of results will need to be reconsidered, because even extremely small coefficients that are substantively negligible show up as statistically significant in CSS analyses. Furthermore, "big data" often suffer from the same validity and reliability issues as other secondary data—they were never collected to answer the research questions, researchers have no or limited control over the constructs, and in particular, the platform that collects the data may not generate a representative sample of the population (D. M. J. Lazer et al., 2020, 1061). A final concern is with the mindless application of CSS methods as we have already discussed in machine learning section. Even highly accurate predictive models do not necessarily provide useful explanations (Hofman et al., 2021). Research design courses within the context of CSS methods are highly desirable. Students must learn how to integrate computational methods into their research design, what types of questions can be answered, and what are the concerns and risks that can undermine research validity.

For future research implementing CSS in nonprofit and philanthropic studies and with a larger community of international scholars, we will be working to expand the KINPS to include academic publications in additional languages, starting with Chinese. We encourage interested scholars to contact us to explore options for collaboration. Furthermore, the KINPS is an ideal starting point for meta-science in our field. For example, with linked citation data, it is possible to conduct network analyses of publications, estimating not only which publications have been highly influential, but also which publications connect different subfields of research. Furthermore, by extracting results of statistical tests, it is possible to quantify the quality of research—at least in a statistical sense—through the lack of errors in statistical tests, and the distribution of p values indicating p-hacking and publication bias. In future, algorithms may be developed to automatically extract effect sizes for statistical meta-analyses. We highly encourage scholars to use KINPS and advance nonprofit and philanthropic studies toward a mature interdisciplinary field and a place of joy.

Acknowledgements The authors thank the Revolutionizing Philanthropy Research Consortium for suggestions on keywords, Sasha Zarins for organizing the Consortium, and Gary King for commenting on interdisciplinary collaboration. We appreciate the constructive comments from Mirae Kim, Paloma Raggo, and the anonymous reviewers, and thank Taco Brandsen and Susan Appe for handling the manuscript. The development of KINPS uses the Chameleon testbed supported by the U.S. National Science Foundation.

Author contributions J.M. drafted the manuscript; designed definitions, selected the methods to be discussed in the toolbox section, constructed the database and online user interface. I.A.E. drafted the technical details. A.D.W. collected example studies and drafted the review sections. M.X. and Y.Y. collected partial data for training the machine learning algorithm. R.B. and P.W. helped frame the manuscript and definitions, and drafted the concluding sections. J.M., A.D.W, R.B., and P.W. finalized the manuscript.

Funding J.M. was supported in part by the 2019–20 PRI Award and Stephen H. Spurr Centennial Fellowship from the LBJ School of Public Affairs and the Academic Development Funds from the RGK Center. P.W.'s work at the IU Lilly Family School of Philanthropy is funded through a donation by the Stead Family; her work at the VU University Amsterdam is funded by the Dutch Charity Lotteries.Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly

[6] Some caveats regarding the concerns of privacy and intellectual property, see Lazer et al., (2020, 1061).

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Anastasopoulos, L. J., & Whitford, A. B. (2019). Machine learning for public administration research, with application to organizational reputation. *Journal of Public Administration Research and Theory, 29*(3), 491–510. https://doi.org/10.1093/jopart/muy060

Bachman, C. W. (1969). Data structure diagrams. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 1*(2), 4–10. https://doi.org/10.1145/1017466.1017467

Bail, C. A., Brown, T. W., & Mann, M. (2017). Channeling hearts and minds: advocacy organizations, cognitive-emotional currents, and public conversation. *American Sociological Review, 82*(6), 1188–1213. https://doi.org/10.1177/0003122417733673

Baker, M. (2016). 1,500 Scientists lift the lid on reproducibility. *Nature News, 533*(7604), 452. https://doi.org/10.1038/533452a

Bassoli, M. (2017). Catholic versus communist: An ongoing issue—The role of organizational affiliation in accessing the policy Arena. *VOLUNTAS International Journal of Voluntary and Nonprofit Organizations, 28*(3), 1135–56. https://doi.org/10.1007/s11266-016-9708-1

Baumer, E. P. S., Mimno, D., Guha, S., Quan, E., & Gay, G. K. (2017). Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? *Journal of the Association for Information Science and Technology, 68*(6), 1397–1410. https://doi.org/10.1002/asi.23786

Bhati, A., & McDonnell, D. (2020). Success in an online giving day: The role of social media in fundraising. *Nonprofit and Voluntary Sector Quarterly, 49*(1), 74–92. https://doi.org/10.1177/0899764019868849

Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics, 64*(3), 351–374.

Brandtner, C. (2021). Decoupling under scrutiny: consistency of managerial talk and action in the age of nonprofit accountability. *Nonprofit and Voluntary Sector Quarterly*. https://doi.org/10.1177/0899764021995240

Brass, J. N., Longhofer, W., Robinson, R. S., & Schnable, A. (2018). NGOs and international development: A review of thirty-five years of scholarship. *World Development, 112*(December), 136–149. https://doi.org/10.1016/j.worlddev.2018.07.016

Bryant, C. G. A. (1975). Kuhn, paradigms and sociology. *The British Journal of Sociology, 26*(3), 354–359. https://doi.org/10.2307/589851

Chen, P.-S. (1976). The entity-relationship model—toward a unified view of data. *ACM Transactions on Database Systems, 1*(1), 9–36. https://doi.org/10.1145/320434.320440

Chen, H., & Zhang, R. (2021). Identifying nonprofits by scaling mission and activity with word embedding. *Voluntas*. https://doi.org/10.1007/s11266-021-00399-7.

Cioffi-Revilla, C. (2017). Computation and Social Science. In *Introduction to Computational Social Science*, 35–102. Texts in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-50131-4.

Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM, 13*(6), 377–387. https://doi.org/10.1145/362384.362685

Colavizza, G., Hrynaszkiewicz, I., Staden, I., Whitaker, K., & McGillivray, B. (2020). The citation advantage of linking publications to research data. *PLoS ONE, 15*(4), e0230416. https://doi.org/10.1371/journal.pone.0230416

Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2019). *Managing and Sharing Research Data: A Guide to Good Practice* (2nd ed.). SAGE Publications Ltd.

De Wit, A., Bekkers, R., & Broese van Groenou, M. (2017). Heterogeneity in crowding-out: When are charitable donations responsive to government support? *European Sociological Review*, *33*(1), 59–71.

Devlin, J., Chang M. W., Lee K., & Toutanova K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. ArXiv:1810.04805 *[Cs]*, May. http://arxiv.org/abs/1810.04805.

Diani, M., Ernstson, H., & Jasny, L. (2018). Right to the city and the structure of civic organizational fields: evidence from cape town. *VOLUNTAS International Journal of Voluntary and Nonprofit Organizations, 29*(4), 637–52. https://doi.org/10.1007/s11266-018-9958-1.

Edelmann, A., Wolff, T., Montagne, D., & Bail, C. A. (2020). Computational social science and sociology. *Annual Review of Sociology, 46*(1), 61–81. https://doi.org/10.1146/annurev-soc-121919-054621

Evans, J. A., & Aceves, P. (2016). Machine translation: Mining text for social theory. *Annual Review of Sociology, 42*(1), 21–50. https://doi.org/10.1146/annurev-soc-081715-074206

Farrell, J. (2019). The growth of climate change misinformation in us philanthropy: Evidence from natural language processing. *Environmental Research Letters, 14*(3), 034013. https://doi.org/10.1088/1748-9326/aaf939

Frickel, S., & Gross, N. (2005). A general theory of scientific/intellectual movements. *American Sociological Review, 70*(2), 204–232. https://doi.org/10.1177/000312240507000202

Fyall, R., Kathleen Moore, M., & Gugerty, M. K. (2018). Beyond NTEE codes: Opportunities to understand nonprofit activity through mission statement content coding. *Nonprofit and Voluntary Sector Quarterly, 47*(4), 677–701. https://doi.org/10.1177/0899764018768019

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature, 57*(3), 535–574. https://doi.org/10.1257/jel.20181020

Gentzkow, M., & Shapiro, J M. (2014). *Code and Data for the Social Sciences: A Practitioner's Guide*.

Gilardi, F., Shipan C. R., & Wüest, B. (2020) Policy Diffusion The Issue-Definition Stage. *American Journal of Political Science* n/a (n/a). https://doi.org/10.1111/ajps.12521.

Goyal, S., van der Leij, M. J., & Moraga-González, J. L. (2006). Economics: An emerging small world. *Journal of Political Economy, 114*(2), 403–412. https://doi.org/10.1086/500990

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). Machine learning for social science: An agnostic approach. *Annual Review of Political Science, 24*(1), 395–419. https://doi.org/10.1146/annurev-polisci-053119-015921

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*(3), 267–297. https://doi.org/10.1093/pan/mps028

Guo, C., & Saxton, G. D. (2018). Speaking and being heard: how nonprofit advocacy organizations gain attention on social media. *Nonprofit and Voluntary Sector Quarterly, 47*(1), 5–26. https://doi.org/10.1177/0899764017713724

Gusenbauer, M. (2019). Google scholar to overshadow them all? comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics, 118*(1), 177–214. https://doi.org/10.1007/s11192-018-2958-5

Hardwicke, T. E., Wallach, J. D., Kidwell, M. C., Bendixen, T., Crüwell, S., & Ioannidis, J. P. A. (2020). An empirical assessment of transparency and reproducibility-related research

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

practices in the social sciences (2014–2017). *Royal Society Open Science, 7*(2), 190806. https://doi.org/10.1098/rsos.190806

Heemskerk, E. M., & Takes, F. W. (2016). The corporate elite community structure of global capitalism. *New Political Economy, 21*(1), 90–118. https://doi.org/10.1080/13563467.2015.1041483

Herzog, P. S., & Yang, S. (2018). Social networks and charitable giving: Trusting, doing, asking, and alter primacy. *Nonprofit and Voluntary Sector Quarterly, 47*(2), 376–394. https://doi.org/10.1177/0899764017746021.

Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., et al. (2021). Integrating explanation and prediction in computational social science. *Nature, 595*(7866), 181–188. https://doi.org/10.1038/s41586-021-03659-0.

Jurafsky, D., & Martin, J. H. (2020). *Speech and Language Processing*. 3rd draft. https://web.stanford.edu/~jurafsky/slp3/ed3book_dec302020.pdf.

Kang, C. H., Baek Y. M., & Kim, E. H. J. (2021). Half a Century of NVSQ: Thematic Stability Across Years and Editors. *Nonprofit and Voluntary Sector Quarterly*, June. https://doi.org/10.1177/08997640211017676.

King, G., & Persily, N. (2020). A new model for industry-academic partnerships. *PS: Political Science & Politics, 53*(4), 703–9. https://doi.org/10.1017/S1049096519001021

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word Embeddings. *American Sociological Review, 84*(5), 905–949. https://doi.org/10.1177/0003122419877135

Kuhn, T. S. (1970) *The Structure of Scientific Revolutions*. Second Edition, Enlarged. International Encyclopedia of Unified Science. Foundations of the Unity of Science, v. 2, No. 2. Chicago: University of Chicago Press.

Lai, C. H., Tao, C. C., & Cheng, Y. C. (2017). Modeling resource network relationships between response organizations and affected neighborhoods after a technological disaster. *VOLUNTAS International Journal of Voluntary and Nonprofit Organizations, 28*(5), 2145–75. https://doi.org/10.1007/s11266-017-9887-4

Lazega, E., Jourda, M. T., & Mounier, L. (2013). Network lift from dual alters: Extended opportunity structures from a multilevel and structural perspective. *European Sociological Review, 29*(6), 1226–1238. https://doi.org/10.1093/esr/jct002

Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N., Freelon, D., et al. (2020). Computational social science: obstacles and opportunities. *Science, 369*(6507), 1060–1062. https://doi.org/10.1126/science.aaz8170

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., et al. (2009). Computational social science. *Science, 323*(5915), 721–723. https://doi.org/10.1126/science.1167742

Lecy, J. D., Ashley, S. R., & Santamarina, F. J. (2019). Do nonprofit missions vary by the political ideology of supporting communities? Some preliminary results. *Public Performance & Management Review, 42*(1), 115–141. https://doi.org/10.1080/15309576.2018.1526092

Lecy, J., & Thornton, J. (2016). What big data can tell us about government awards to the nonprofit sector: using the FAADS. *Nonprofit and Voluntary Sector Quarterly, 45*(5), 1052–1069. https://doi.org/10.1177/0899764015620902

Litofcenko, J., Karner, D., & Maier, F. (2020). Methods for classifying nonprofit organizations according to their field of activity: a report on semi-automated methods based on text". *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations, 31*(1), 227–37. https://doi.org/10.1007/s11266-019-00181-w

Ma, J. (2020). Funding nonprofits in a networked society: Toward a network framework of government support. *Nonprofit Management and Leadership, 31*(2), 233–257. https://doi.org/10.1002/nml.21426

Ma, J. (2021). Automated coding using machine learning and remapping the U.S. nonprofit sector: A guide and benchmark. *Nonprofit and Voluntary Sector Quarterly, 50*(3), 662–687. https://doi.org/10.1177/0899764020968153

Ma, J., & DeDeo, S. (2018). State power and elite autonomy in a networked civil society: The board interlocking of Chinese nonprofits. *Social Networks, 54*(July), 291–302. https://doi.org/10.1016/j.socnet.2017.10.001

Ma, J., Jing, E., & Han, J. (2018) Predicting mission alignment and preventing mission drift do revenue sources matter? *Chinese Public Administration Review* 9 (1): 25–33. https://doi.org/10.22140/cpar.v9i1.173.

Ma, J., & Konrath, S. (2018). A century of nonprofit studies: Scaling the knowledge of the field. *VOLUNTAS International Journal of Voluntary and Nonprofit Organizations, 29*(6), 1139–58. https://doi.org/10.1007/s11266-018-00057-5

Ma, J., Qun W., Chao, D., & Huafang, L. (2017). The research infrastructure of chinese foundations, a database for chinese civil society studies. *Scientific Data* 4 (July): sdata201794. https://doi.org/10.1038/sdata.2017.94.

Martin, S., Brown, W. M., Klavans, R., & Boyack, K. W. (2011). "OpenOrd: An Open-Source Toolbox for Large Graph Layout." In *Visualization and Data Analysis 2011*, 7868:786806. International Society for Optics and Photonics. https://doi.org/10.1117/12.871402

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & López-Cózar, E. D. (2018). Google scholar, web of science, and scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics, 12*(4), 1160–1177. https://doi.org/10.1016/j.joi.2018.09.002

Mikolov, T., Chen, K., orrado, G & Dean, J. (2013). "Efficient estimation of word representations in vector space." ArXiv: 1301.3781 *[Cs]*, January. http://arxiv.org/abs/1301.3781.

Minkowitz, H., Twumasi, A., Berrett, J. L., Chen, X., & Stewart, A. J. (2020). Checking in on the State of Nonprofit Scholarship: A review of recent research. *Journal of Public and Nonprofit Affairs, 6*(2), 182–208.

Molina, M., & Garip, F. (2019). Machine learning for sociology. *Annual Review of Sociology, 45*(1), 27–45. https://doi.org/10.1146/annurev-soc-073117-041106

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American Sociological Review, 69*(2), 213–238. https://doi.org/10.1177/000312240406900204

Mueller, H., & Rauh, C. (2018). Reading between the lines: Prediction of political violence using newspaper text. *American Political Science Review, 112*(2), 358–375. https://doi.org/10.1017/S0003055417000570

Müller, T. S., Grund, T. U., & Koskinen, J. H. (2018). Residential segregation and 'Ethnic Flight' vs. 'Ethnic avoidance' in sweden. *European Sociological Review, 34*(3), 268–285. https://doi.org/10.1093/esr/jcy010

Nakazato, H., & Lim, S. (2016). Evolutionary process of social capital formation through community currency organizations: The Japanese case. *Voluntas International Journal of Voluntary and Nonprofit Organizations, 27*(3), 1171–94.

Paarlberg, L. E., Hannibal, B., & Johnson, J. M. (2020). Examining the mediating influence of interlocking board networks on grant making in public foundations. *Nonprofit and Voluntary Sector Quarterly*. https://doi.org/10.1177/0899764019897845

Parthasarathy, R., Rao, V., & Palaniswamy, N. (2019). Deliberative democracy in an unequal world: A text-as-data study of south

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press

India's village assemblies. *American Political Science Review, 113*(3), 623–640. https://doi.org/10.1017/S0003055419000182

Paxton, P., Velasco, K., & Ressler, R. W. (2020). Does use of emotion increase donations and volunteers for nonprofits? *American Sociological Review*. https://doi.org/10.1177/0003122420960104

Ragin, C. C., & Lisa M. A. (2011). *Constructing Social Research: The Unity and Diversity of Method*. Pine Forge Press.

Rodriguez, P., & Spirling, A. (2021). Word embeddings: What works, what doesn't, and how to tell the difference for applied research. *The Journal of Politics, May*. https://doi.org/10.1086/715162

Schröder, T., Hoey, J., & Rogers, K. B. (2016). Modeling dynamic identities and uncertainty in social interactions: Bayesian Affect Control Theory. *American Sociological Review, 81*(4), 828–855. https://doi.org/10.1177/0003122416650963

Scott, J. (2017). *Social Network Analysis* (4th ed.). SAGE Publications Ltd.

Seippel, Ø. (2008). Sports in civil society: networks, social capital and influence. *European Sociological Review, 24*(1), 69–80. https://doi.org/10.1093/esr/jcm035

Shi, Y., Dokshin, F. A., Genkin, M., & Brashears, M. E. (2017). A member saved is a member earned? The recruitment-retention trade-off and organizational strategies for membership growth. *American Sociological Review, 82*(2), 407–434. https://doi.org/10.1177/0003122417693616

Shier, M. L., & Handy, F. (2014). Research trends in nonprofit graduate studies: A growing interdisciplinary field. *Nonprofit and Voluntary Sector Quarterly, 43*(5), 812–831. https://doi.org/10.1177/0899764014548279

Shoemaker, P. J., James, W. T, & Dominic L. L. (2003). *How to Build Social Science Theories*. Thousand Oaks, UNITED STATES: SAGE Publications. http://ebookcentral.proquest.com/lib/utxa/detail.action?docID=996273.

Shwed, U., & Bearman, P. S. (2010). The Temporal structure of scientific consensus formation. *American Sociological Review, 75*(6), 817–840. https://doi.org/10.1177/0003122410388488

Shwom, R. (2015). Nonprofit-business partnering dynamics in the energy efficiency field. *Nonprofit and Voluntary Sector Quarterly, 44*(3), 564–586. https://doi.org/10.1177/0899764014527174

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain To Leiden: Guaranteeing well-connected communities. *Scientific Reports, 9*(1), 1–12. https://doi.org/10.1038/s41598-019-41695-z

Walk, M., & Andersson, F. O. (2020). Where do nonprofit and civil society researchers publish? Perceptions of nonprofit journal quality. *Journal of Public and Nonprofit Affairs, 6*(1), 79–95.

Wasif, R. (2020). Does the Media's Anti-Western Bias Affect Its Portrayal of NGOs in the Muslim World? Assessing Newspapers in Pakistan. *VOLUNTAS International Journal of Voluntary and Nonprofit Organizations*. https://doi.org/10.1007/s11266-020-00242-5.

Wasif, R. (2021). Terrorists or Persecuted? The Portrayal of Islamic Nonprofits in US Newspapers Post 9/11." *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, February. https://doi.org/10.1007/s11266-021-00317-x.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Wickham, H. (2014). Tidy Data. *The Journal of Statistical Software* http://www.jstatsoft.org/v59/i10/.

Wiepking, P., Handy, F., Park, S., Neumayr, M., Bekkers, R., Breeze, B., de Wit, A., et al. (2021). Global philanthropy: Does institutional context matter for charitable giving? *Nonprofit and Voluntary Sector Quarterly*. https://doi.org/10.1177/0899764021989444

Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L., & Teal, T. K. (2017). Good enough practices in scientific computing. *PLOS Computational Biology, 13*(6), e1005510. https://doi.org/10.1371/journal.pcbi.1005510

Xu, W., & Saxton, G. D. (2019). Does stakeholder engagement pay off on social media? A social capital perspective. *Nonprofit and Voluntary Sector Quarterly*, *48*(1), 28–49.

Yang, Y., Zhou, W., & Zhang, D. (2019). Celebrity philanthropy in China: An analysis of social network effect on philanthropic engagement. *VOLUNTAS International Journal of Voluntary and Nonprofit Organizations, 30*(4), 693–708. https://doi.org/10.1007/s11266-018-9997-7

Young, D. R. (1999). Nonprofit management studies in the United States: Current developments and future prospects. *Journal of Public Affairs Education, 5*(1), 13–23.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

https://doi.org/10.1007/s11266-021-00414-x Published online by Cambridge University Press