# Shedding New Light on Old Data

**Huafang Li\***

## Abstract

Using old data is not a new business in nonprofit research. However, the importance and potential of using old data for new nonprofit research seemed to be underestimated. This article reemphasizes the necessity of using old data, categorizes various old data, and showcases how to use them to advance nonprofit research. This article also discusses the process and issues of making old data available, accessible, and searchable and urges nonprofit scholars to publicize their data to enhance academic transparency and accountability and benefit new research.

## Keywords

secondary data, data availability, data accessibility, data search, nonprofit research

## 1. Introduction

Using existing data for new research is an old business in scientific research in general and nonprofit research as well. Why should nonprofit scholars care about using old data that exist prior to a new study? First, data associated with published nonprofit studies are not often publicly available. When the nonprofit research field demands academic accountability and transparency, it is troubling that making such data publicly available is not a common practice for leading nonprofit research journals, such as *Nonprofit and Voluntary Sector Quarterly (NVSQ)*, *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations (VOLUNTAS)*, and *Nonprofit Management and Leadership (NML)*. Lacking data availability makes replication difficult and thus diminishes the research transparency and accountability and worsens the academic inequality by impeding scholars with less resources from catching up with research

---
\* Huafang Li, assistant professor, graduate school of public and international affairs, University of Pittsburgh.

developments.

Second, the importance of using old data is often underestimated. To study nonprofit organizations, we need data. The idea that data scarcity is one of the biggest challenges for making research progress is not unseen. However, collecting primary data could be burdensome to many scholars because it is time-consuming and often costly. Such burdens again widen the gap between resourceful scholars and those less affluent ones. On the other hand, the underutilization of collected data results in a waste of resources. Old data can inspire new research. If old data is available, scholars' time, money, and other resources that would be spent on primary data collection to conduct analyses can be better used for research.

Last but not least, nonprofit research lags behind other disciplines in utilizing old data. For example, the National Center for Charitable Statistics (NCCS) data (nccs-data.urban.org) is one of the most commonly used secondary datasets in nonprofit research in the United States and beyond. According to Google Scholar, scholars had used the NCCS data 2,470 times to research various topics during the 2010 to 2020 period. However, during the same time period, the commonly used National Administrative Studies Project (NASP) data in the public administration field and the General Social Survey (GSS) data in general social science have been used 3,830 and 1,950,000 times, respectively. There are plenty of spaces for nonprofit scholars to improve old data usage in research. The remained question is, "how can we find and utilize existing old data for new nonprofit research more effectively?"

To collect primary data, such as using surveys or conducting experiments or building on other data sources, is an effort that should be highly respected and appreciated in any fields. It is not an easy task for scholars with limited budgets and resources to develop new surveys and gather new data. It is true that existing secondary data are old and are not designed for your specific study purposes and research questions. However, access to secondary data sources that are already available is cheaper and sometimes even free. In addition to the fact that old data can be used for replications that enhance academic transparency and accountability, this article reemphasizes that old data can also be used to confirm and strengthen existing studies and develop new ones.

The rest of this article proceeds as follows: after briefly introducing the necessity of using old data, the article then describes the categories of old data and showcases how to use various old datasets. Finally, this article discusses the process and issues of making old data available, accessible, and searchable and urges nonprofit scholars to publicize their data that can advance theories and improve practices.

**2. Shedding new light on old data**

Using old data is not a new business in nonprofit research. Scholars in the field use various forms of old data, such as secondary datasets developed by others, or sets of administrative data (e.g., the United States Internal Revenue Service's 990 Forms) or big data (e.g., nonprofits' social media data) compiled and built by others. However, the importance of using old data seemed to be underestimated in nonprofit research. Using old data for research can be effective and prolific. If there is old available data that suits scholars' research purposes, then they might not need to collect costly primary data and thus can focus on analyzing the old data. Scholars then can better use their time, energy, and resources to achieve their academic goals. For example, it is inefficient and unnecessary for those scholars who are interested in researching U.S. nonprofit finance to compile the 990 forms data independently. Instead, they can use the NCCS data to answer a variety of new research questions.

**Figure 1: Framework**
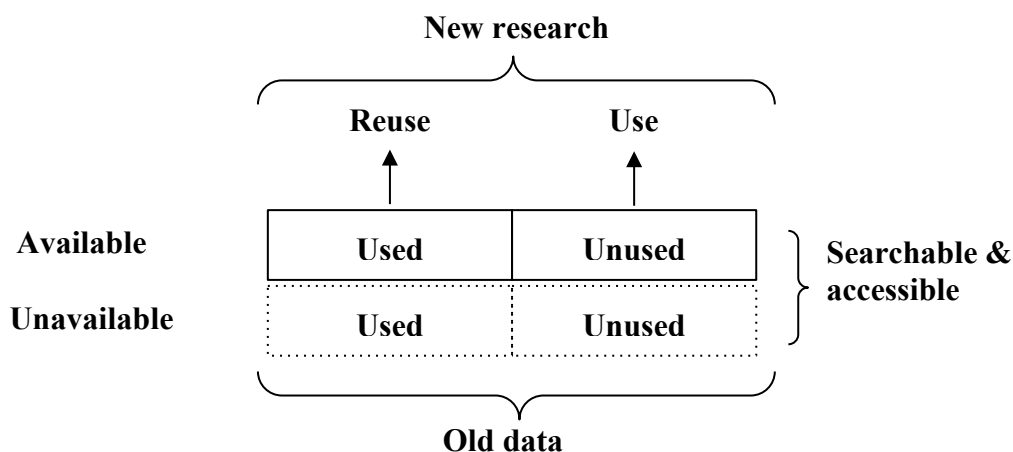
Based on two dimensions - old data usage (used or unused) and availability (available or unavailable), we can categorize old data into four different types: available used old data, available unused old data, unavailable used old data, and unavailable unused old data. Scholars need different strategies to access old datasets because the degree of their usage and availability varies and tailor old data to suit their research purposes (Figure 1). This article will showcase the use of old data and discuss data access and privacy issues. The following section first shows how to use old data to answer new research questions.

### 2.1 Reusing the used old data

Scholars reuse old data to answer new research questions because different scholars view the same dataset in varying ways. For example, the 2007 San Diego County (SDC) survey data has been used by several studies.[1] The SDC data was collected by surveying residents (n=1,002) in San Diego County, California, to research a variety of issues relating to public attitudes toward and involvement in local nonprofit organizations. The Social Science Research Laboratory (SSRL) at San Diego State University (SDSU) fielded the survey from December 2007 to January 2008. The survey consisted of 29 questions, including involvement in volunteering and donating, information sources used prior to donating and volunteering, confidence in nonprofit organizations, and socio-demographic items.

We illustrate how new research can benefit from using the same old dataset by highlighting McDougle and Handy (2014)[2] and Li and McDougle (2017)[3]. Li and McDougle (2017) [4] were curious about where individuals get information on nonprofit organizations and how their reliance on different information sources can influence their charitable giving decisions. They used the

---

[1] McDougle, L. M., & Lam, M. (2014), "Individual- and Community-Level Determinants of Public Attitudes Toward Nonprofit Organizations", Nonprofit and Voluntary Sector Quarterly, 43(4), 672–692, https://doi.org/10.1177/0899764013479830; McDougle, L. M., & Handy, F. (2014), "The Influence of Information Costs on Donor Decision Making", Nonprofit Management and Leadership, 24(4), 465–485; Li, H., & McDougle, L. (2017), "Information Source Reliance and Charitable Giving Decisions", Nonprofit Management and Leadership, 27(4), 549–560.

[2] McDougle, L. M., & Handy, F., 465–485.

[3] Li, H., & McDougle, L., 549–560.

[4] Li, H., & McDougle, L., 549–560.

same SDC dataset that is associated with an earlier study. [1]The SDC survey was designed for other research, but the data suit well for examining the effects of relying on different information sources on individuals' giving decisions.

McDougle and Handy (2014) [2] used information costs to predict information-gathering strategies that individuals used for making donation decisions. In their study, outcomes were measured by gathering information from sources such as (1) word of mouth from friends or colleagues, (2) organizational websites, (3) past experiences or connections to the organization, (4) news coverage, and (5) accrediting organizations. They used information costs and confidence in nonprofit organizations to predict individuals' information-gathering strategies for donation decision-making. Information costs were measured by individuals' marital status, race/ethnicity, educational background, income, and age, which influence how individuals gather nonprofit information to facilitate their charitable decisions. They used two different questions to measure the level of individual confidence in nonprofit organizations: (1) How much confidence would you say that you have in the ability of local nonprofits to effectively provide quality services? And (2) How much confidence would you say that you have in the ability of local nonprofits to spend money wisely? However, they did not use individual donating and volunteering decisions in their study. They found that information source reliance varies by different individual characteristics that represent information costs.

**Table 1: Commonly used variables between two studies**

| McDougle & Handy, 2014 [3] | Variables | | | Li & McDougle, 2017 [4] |
|---|---|---|---|---|
| | | Donating and volunteering decisions | Dependent variables | |
| | | | | |
| | Dependent variables | Information gathering strategies | Independent variables | |

---

[1] McDougle, L. M., & Handy, F., 465–485.

[2] McDougle, L. M., & Handy, F., 465–485.

[3] McDougle, L. M., & Handy, F., 465–485.

[4] Li, H., & McDougle, L., 549–560.

|  |  |  |  |  |
|---|---|---|---|---|
|  |  | (1) word of mouth: opinions of friends/colleagues |  |  |
|  |  | (2) organizational website |  |  |
|  |  | (3) past experiences or connections/association with the organization |  |  |
|  |  | (4) news coverage |  |  |
|  |  | (5) information from third-party outside accrediting organizations |  |  |
|  |  |  |  |  |
|  | Independent variables | Information Costs | Controls |  |
|  |  | (1)Marital status |  |  |
|  |  | (2)Race/ethnicity |  |  |
|  |  | (3)Education |  |  |
|  |  | (4)Income |  |  |
|  |  | (5)Age |  |  |
|  |  | Confidence in Nonprofits |  |  |
|  |  | (1)Provide quality services |  |  |
|  |  | (2)Spend money wisely |  |  |
|  |  |  |  |  |
|  | Controls | Volunteer |  |  |
|  |  | Employment status |  |  |
|  |  | Sex |  |  |

Li and McDougle (2017) [1]were interested in a different question: will individuals who rely on different information sources making varying charitable giving decisions? McDougle and

---

[1] Li, H., & McDougle, L., 549–560.

Handy (2014)[①] did not design the survey for answering this particular research question. However, the dataset presented in McDougle and Handy (2014)[②] suited well for Li and McDougle's (2017) [③]research purposes. Table 1 shows commonly used variables by both studies.

To study the effects of different information-gathering strategies on individuals' giving and volunteering decisions, Li and McDougle used self-reported retrospective (2006) and current (2007) measures of monetary donation and volunteer proclivity, as well as self-reported predictive (2008) measures of monetary donation and volunteer intention as outcome measures. Retrospective proclivity was measured using the question: "Did you donate money to (volunteer with) any San Diego County NPOs in 2006? Current proclivity was measured using the question: Have you volunteered with (donated money to) any San Diego County NPOs in 2007? Predictive intention was measured using the question: In 2008, do you intend to volunteer with (donate money to) any San Diego County NPOs?" [④] Li and McDougle used different information-gathering strategies as their main independent variables to understand how relying on varying information sources influences individuals' giving decisions. They controlled information costs and confidence in nonprofits that McDougle and Handy used as independent variables in addition to other commonly controlled variables (Table 1).  They found that individuals' willingness to give time and money to a particular organization increases significantly when they learn about nonprofits through their personal experience. Whereas relying on other information sources, such as "word of mouth," media, and online sources, had no such effects. The example shows how different studies can use the same old dataset to answer different research questions.

### 2.2 Using the unused old data

Many scholars who collect primary data do not end up using all of the data they gather. When

---

[①] McDougle, L. M., & Handy, F., 465–485.

[②] McDougle, L. M., & Handy, F., 465–485.

[③] Li, H., & McDougle, L., 549–560.

[④] Li, H., & McDougle, L., 551.

scholars design a survey, they often include some routine questions along with their primary research questions. They then focus on those main research questions and largely ignore the data associated with routine questions. It is true that data generated by routine questions may seem less important to a particular scholar's work, but it is also true that unused old data are often just sitting there waiting for their chance to shine. One researcher's trash is another's treasure. Other scholars can use this unused or wasted data for their research.

Recently, from philanthropy scholars and professionals, our field has seen an increased interest in experiential philanthropy (also known as "learning by giving") [1]. Experiential philanthropy is a service-learning pedagogy that allows students to study social problems and then donate money funded by foundations to nonprofit organizations to solve these problems. [2] Experiential philanthropy courses teach students about theories of philanthropy and practices of philanthropic responses to social issues. Experiential philanthropy aims to advance student learning outcomes, educate future philanthropic leaders, fulfill educational institutes' goals, and benefit the communities by providing funds to nonprofit organizations that solve social problems. [3]

Scholars researching experiential philanthropy want to understand whether educating future philanthropic leaders is possible. Li and McDougle (2017)[4] showed the importance of personal experience in predicting charitable giving decisions. Therefore, can we educate future generations through an experiential philanthropy approach that exposes them to "learning by giving"

---

[1] Li, H., Xu, C., & McDougle, L. M. (2019), "Philanthropy Can Be Learned: A Qualitative Study of Student Experiences in Experiential Philanthropy Courses", Philanthropy & Education, 2(2), 29–52; Li, H., McDougle, L., & Gupta, A. (2020), "Experiential philanthropy in China", Journal of Public Affairs Education, 26(2), 205–227, https://doi.org/10.1080/15236803.2019.1667185;
McDougle, L., McDonald, D., Li, H., McIntyre Miller, W., & Xu, C. (2017), "Can Philanthropy Be Taught?", Nonprofit and Voluntary Sector Quarterly, 46(2), 330–351, https://doi.org/10.1177/0899764016662355; Xu, C., Li, H., & McDougle, L. M. (2018), "Experiential Philanthropy. In A. Farazmand (Ed.)", Global Encyclopedia of Public Administration, Public Policy, and Governance (1–7). Springer International Publishing. https://doi.org/10.1007/978-3-319-31816-5_3048-1.

[2] Xu, C., Li, H., & McDougle, L. M., 1–7.

[3] Ahmed, S., & Olberding, J. (2007), "Can Student Philanthropy Help to Address the Current Nonprofit Identity Crisis? A Case Study of a Multiyear, Multidisciplinary Project at Northern Kentucky University", *Journal of Public Affairs Education*, 13(3/4), 593–615; Millisor, J., & Olberding, J. C. (2009), "Student Philanthropy in Colleges and Universities", Academic Exchange Quarterly, 13(4), 11–16; Olberding, J. C. (2012), "Does Student Philanthropy Work? A Study of Long-term Effects of the 'Learning by Giving' Approach", Innovative Higher Education, 37(2), 71–87.

[4] Li, H., & McDougle, L., 549–560.

experience? In a study of experiential philanthropy courses "Can Philanthropy Be Taught?", [1] authors used the survey data collected from the Mayerson Student Philanthropy Project (MSPP) at the Northern Kentucky University (NKU) over a five-year period (2009-2013) to investigate such courses' effects on a series of student learning outcomes. They examined change scores for items on the pre- and post-course surveys asking students' interest in courses and schools, their awareness of social issues and local nonprofits, and their intention in future philanthropic activities. The results from their quantitative analysis showed promising positive effects of experiential philanthropy courses on students' academic, social, and philanthropic learning outcomes.

However, whether the statistically positive associations between taking experiential philanthropy courses and student learning outcomes are causal remains unclear. The increased scores from pre- to post-course survey items could result from students' self-selection. More prosocial and philanthropic students are more likely to enroll in experiential philanthropy classes and thus they increased the scores due to their natural tendency rather than the classes. The self-selection then biases the results. The positive effects of experiential philanthropy classes on student learning outcomes could result from other factors that changed during the same period. For example, an increase in family income or new friendships could also improve learning outcomes. Additional data, if available, can help rule out alternative explanations and identify whether there are any true causal effects of experiential philanthropy classes on student learning outcomes.

Li and colleagues (2019) [2] noticed that the "Can Philanthropy Be Taught" study did not utilize students' responses to several open-ended questions routinely asked on the NKU post-course surveys and decided to analyze the unused rich qualitative data to examine whether the positive effects of experiential courses on student learning outcomes are causal. The NKU post-course surveys asked the following open-ended questions: (1) Please discuss the most significant things you learned this semester about social problems or needs, nonprofit organizations, or the philanthropy process; (2) As a result of this course, do you plan to make any changes related to your level of involvement in campus life or community life? (3) What did you most like or

---

[1] McDougle, L., McDonald, D., Li, H., McIntyre Miller, W., & Xu, C..

[2] Li, H., Xu, C., & McDougle, L. M., 29–52.

appreciate about the project experience? From 2009 to 2013, the NKU post-course surveys documented 973 students' responses to these questions.

If the significant improvements of student learning outcomes were indeed a result of the experiential philanthropy courses, students would reflect the causality in their responses to open-ended questions. Li and colleagues (2019) then used computer-assisted text-mining methods in their analysis because of the large amount of qualitative data. They first created word clusters and topic models to find important themes and then identify typical student responses in the raw data. Doing so allowed them to extract essential information within the raw qualitative data and better identify the typical answers to the questions. They found that, in students' own words, "philanthropy can be learned".[1] Their research complements the "Can Philanthropy be Taught" study from a different angle and confirms experiential philanthropy courses' positive effects on student learning outcomes.

### 3. Making old data available, accessible, and searchable

The above-mentioned use of old data is only possible because authors of the original studies shared their data with authors of the new studies, not because journals require them to make datasets that are associated with publications available to the public. Making data associated with publications available to the public can increase academic transparency, accountability (replication), citation, visibility, and reputation of authors and journals. In addition, it can inspire new research. Making data associated with publications available to the public is also a regular business in some scholarly fields. For example, the American Economic Association (AEA) requires authors to make their data fully available when submitting their manuscripts.[2] The AEA specifies its data requirements in the following statement:

> *It is the policy of the American Economic Association to publish papers only if*
> *the data and code used in the analysis are clearly and precisely documented and*

---

[1] Li, H., Xu, C., & McDougle, L. M., 29–52.

[2] American Economic Association. (n.d.), Data and Code Availability Policy, American Economic Association. Retrieved October 30, 2020, from https://www.aeaweb.org/journals/data/data-code-policy.

*access to the data and code is non-exclusive to the authors. Authors of accepted papers that contain empirical work, simulations, or experimental work must provide, prior to acceptance, information about the data, programs, and other details of the computations sufficient to permit replication, as well as information about access to data and programs.*

However, leading journals in nonprofit research, such as *NVSQ*, *VOLUNTAS*, and *NML*, do not have such requirements. Unlike AEA journals that do not consider submissions without available related datasets, *NVSQ*, *VOLUNTAS*, and *NML* still consider such manuscripts and encourage authors to deposit data associated with their submissions voluntarily. For example, *VOLUNTAS*'s Research Data Policy states that "the journal encourages authors, where possible and applicable, to deposit data that support the findings of their research in a public repository." The journal and the publisher do not have a preference for data repositories. If datasets deposited to a repository have digital object identifiers (DOIs), authors can cite them in their references. *NML* has a similar data policy. In reality, very few scholars follow the "voluntary" data disclosure guidelines.

Compared with AEA journals, the main reason that leading nonprofit journals lagged behind in promoting data availability and accessibility might be due to the interdisciplinary nature of nonprofit research. While AEA meetings mainly attract economists, the Association for Research on Nonprofit Organizations and Voluntary Action (ARNOVA) and the International Society for Third-Sector Research (ISTR) annual meetings have participants from various disciplines, such as public administration, management, political science, sociology, law, and others. It is difficult for scholars from different disciplines to have a consensus on data policies. However, it is never too late to start the conversation on data transparency, availability, and accessibility for the purpose of advancing scientific research.

Nonprofit scholars should share their data publicly for several reasons. First, as mentioned at the beginning of this article, nonprofit scholars have to increase academic transparency and accountability of the field. A more transparent and accountable research field benefits all in the

long term. Second, scholars who share their primary data publicly will be recognized and credited within and beyond nonprofit studies.[1] Last but not least, nonprofit scholars need to respond to the open data movement in many other disciplines and react to the call of "data philanthropy".[2] Data philanthropy means sharing private data voluntarily to contribute to the public good. Nonprofit scholars can be a leading force in data philanthropy to "fill knowledge gaps and turn data into insights across a broad range of pressing and timely issues."[3] McKeever and colleagues[4] documented how the Urban Institute team used its partnership with the Mastercard Center for Inclusive Growth to utilize private Mastercard transaction data to understand critical policy issues related to charitable giving and equitable development in US cities. In a field where we cherish voluntarism, data philanthropy should be a default choice.

Also, it is time to start making systematic changes in the nonprofit research field to motivate data sharing. In particular, we need to improve the current data policy in the field. *NVSQ* (2020)[5] recently changed its article submission guidelines for experimental studies to promote transparency and replicability. According to the "*NVSQ* Data Transparency Policy for Results Based on Experiments," it requires "authors of manuscripts reporting on data from experiments to provide, upon submission, access to the data and the code that produced the results reported. This will be a condition for the manuscript to proceed through the blind peer review process."[6] However, this policy is currently piloting on experimental studies only. *NVSQ* does not require other types of data to be made available to the public at this point. It may eventually require all

---

[1] Ma, J., Wang, Q., Dong, C., & Li, H. (2017a), "The Research Infrastructure of Chinese Foundations, a database for Chinese civil society studies", Scientific Data, 4, 170094, https://doi.org/10.1038/sdata.2017.94; Ma, J., Wang, Q., Dong, C., & Li, H. (2017b), "The Research Infrastructure of Chinese Foundations, a database for Chinese civil society studies [Data set]", In Harvard Dataverse, Harvard Dataverse. https://doi.org/10.7910/DVN/OTNI1L.

[2] McKeever, B., Greene, S., MacDonald, G., Tatian, P., & Jones, D. (2018), Data Philanthropy: Unlocking the Power of Private Data for Public Good. Urban Institute, https://www.urban.org/research/publication/data-philanthropy-unlocking-power-private-data-public-good.

[3] McKeever, B., Greene, S., MacDonald, G., Tatian, P., & Jones, D..

[4] McKeever, B., Greene, S., MacDonald, G., Tatian, P., & Jones, D..

[5] Nonprofit and Voluntary Sector Quarterly. (2020), NVSQ Data Transparency Policy for Results Based on Experiments: Changes in Article Submission Guidelines, https://journals.sagepub.com/pb-assets/cmscontent/NVS/NVSQ_data_policy_2020.pdf.

[6] Nonprofit and Voluntary Sector Quarterly..

types of data associated with submissions to be made publicly available.

Making data available should not be a one-way effort from journals. The Research Infrastructure of Chinese Foundations (RICF) project shows that nonprofit scholars can also be part of the co-production of open data processes[1]. For years, research on Chinese nonprofit organizations is mainly qualitative and relies heavily on case studies because of lacking large empirical datasets[2]. Ma and colleagues (2017) developed the RICF database for Chinese civil society and nonprofit studies. They crawled, parsed, and compiled data manually or automatically by computer programs from the following six sources, which are ranked by their credibility: (1) Annual reports and audited financial reports from the Chinese government's civil affairs departments official websites. (2) Information disclosed by supervising government departments such as the Civil Organization Administration Bureau of the Ministry of Civil Affairs that lists websites of supervising government departments. (3) Information disclosed by the China Foundation Database (chinafoundation.org.cn). (4) Information disclosed by the China Foundation Center (foundationcenter.org.cn). (5) News from the foundation's official website. (6) News from credible magazines or websites. They published the RICF data in *Scientific Data*, a peer-reviewed open access scientific journal published by the Nature Publishing Group since 2014.[3] *Scientific Data* focuses on descriptions of data sets relevant to the natural sciences, which are provided as machine-readable data, complemented with a human-oriented narrative. In addition, they deposit the data at *Harvard Dataverse*, an open data repository, and keep updating the RICF in GitHub (github.com/ma-ji/RICF).[4] All these efforts make the RICF data available and accessible to the public.

To November 2022, according to the Google Scholar, 41 publications have cited the RICF data. A *Chinese Public Administration Review* symposium was based on studies using the RICF

---

[1] Ma, J., Wang, Q., Dong, C., & Li, H. (2017a).

[2] Li, H., & Chen, B. (2018), "Turning Challenges into Opportunities: Advancing Studies of Nonprofit Organizations in China", Chinese Public Administration Review, 9(1), 1–5. https://doi.org/10.22140/cpar.v9i1.172.

[3] Ma, J., Wang, Q., Dong, C., & Li, H. (2017a).

[4] Ma, J., Wang, Q., Dong, C., & Li, H. (2017a); Ma, J., Wang, Q., Dong, C., & Li, H. (2017b).

data.[1] For example, Wang and He (2018) [2]were able to evaluate the financial health of 2,763 Chinese foundations empirically because of the available open and free RICF data. They found that two-thirds of the Chinese foundations are financially unhealthy, with public foundations outperforming private ones and new foundations improving their financial health faster than older ones. In another application, authors used the RICF data to show that the charity expenditures of local foundations are non-linearly associated with the current conditions of socioeconomic, educational, and medical levels due to the diverse development stages of the cities in China. [3] Song and Fu (2018) [4]acknowledged that their spatial analysis of Chinese foundations and local needs was impossible without the RICF data. The RICF example clearly shows the importance of making data available and accessible to the public and how scholars can shed new light on old data.

**Figure 2: The process of making data available, accessible,and searchable for new research**



---
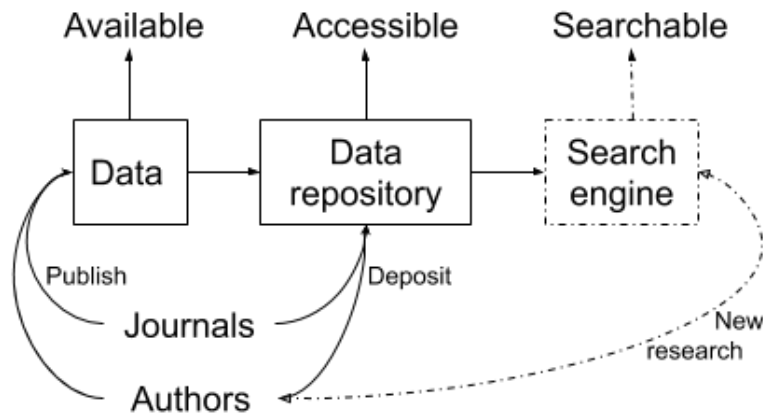
[1] Li, H., & Chen, B.; Ma, J., Jing, E., & Han, J. (2018), "Predicting Mission Alignment and Preventing Mission Drift: Do Revenue Sources Matter?", Chinese Public Administration Review, 9(1), pp. 25–33. https://doi.org/10.22140/cpar.v9i1.173; Wang, Q., & He, L. (2018), "Are the Wealthy Also Healthy? An Empirical Evaluation of the Financial Health of Chinese Foundations", Chinese Public Administration Review, 9(1), 6–24. https://doi.org/10.22140/cpar.v0i0.155.

[2] Wang, Q., & He, L..

[3] Song, Y., & Fu, L. (2018), "Do Charitable Foundations Spend Money Where People Need It Most? A Spatial Analysis of China", ISPRS International Journal of Geo-Information, 7(3), 100. https://doi.org/10.3390/ijgi7030100.

[4] Song, Y., & Fu, L..

Figure 2 illustrates the process of making data available, accessible, and searchable for new research. First, during the submission and publication process, when authors submit manuscripts to journals, they should make the data, if any, associated with the manuscripts available for peer review and replication. Authors can choose a public data depository to deposit their data if the journals' publishers do not provide one to make sure that others can freely access the data. Even though some scholars make data available and accessible to the public, and journals (e.g., *Scientific Data*) introduce and promote data for research, many scholars might be unaware of such open and free datasets that they can use for their own studies. In addition, due to the volume of data repositories available on the Web, it can be extremely difficult to determine not only where is the dataset that has the information that you are looking for, but also the veracity or provenance of that information. [1]

We should also make searching for data much easier. For example, using Google Dataset Search can find datasets wherever they're hosted, whether it's a publisher's site or a data repository. However, authors who provide their searchable datasets should describe the data in a way that Google (and other search engines) can better understand. The Google Dataset Search's guidelines include salient information about datasets: who created the dataset, when it was published, how the data was collected, what the terms are for using the data, etc. The Google Dataset Search then analyzes where data versions of the same dataset might be and finds publications that may be describing or discussing the dataset. The Google Data Search uses an open standard (e.g., schema.org) for data description. If scholars want their datasets to be searchable in Google Data Search, they should adopt this common standard so that all datasets are part of the robust data ecosystem .[2] To illustrate, you can use Google Dataset Search ([datasetsearch.research.google.com](datasetsearch.research.google.com)) to search "research infrastructure of Chinese foundations (RICF)", you will find the description of RICF and where you can download the data (Figure 3). To make available and accessible old data searchable is critical to facilitating the reproducibility of research results and enabling scholars to

[1] Noy, N., & Brickley, D. (2017), "Facilitating the discovery of public datasets", Google AI Blog. http://ai.googleblog.com/2017/01/facilitating-discovery-of-public.html.

[2] Noy, N. (2018, September 5), Making it easier to discover datasets, Google. https://blog.google/products/search/making-it-easier-discover-datasets/.

do new research.

**Figure 3: A screenshot of searching "RICF" in Google Data Search**



## 4. Concluding remarks

This article shows how scholars can shed new light on old data by reusing used old data[1] and using unused old data[2]. This article also uses the RICF example to show the importance and benefits of making old data available, accessible, and searchable [3]. The article concludes that in addition to increasing academic transparency and accountability, using old data can improve existing studies and inspire new research.

This article is not without limitations. First, we must acknowledge the limitation of old data itself. Old data cannot reflect the changing nature of the real world. For example, the RICF data

---

[1] Li, H., & McDougle, L.; McDougle, L. M., & Handy, F., 465–485.

[2] Li, H., Xu, C., & McDougle, L. M., 29–52; McDougle, L., McDonald, D., Li, H., McIntyre Miller, W., & Xu, C..

[3] Google. (n.d.), Google Dataset Search. Retrieved March 28, 2020, from https://datasetsearch.research.google.com/; Ma, J., Wang, Q., Dong, C., & Li, H. (2017a); Ma, J., Wang, Q., Dong, C., & Li, H. (2017b).

mentioned above will not be useful for researchers who want to study the influence of the new laws and regulations on nonprofits in mainland China. Scholars need to decide whether to use old data when starting their new research. Suppose the existing data will not help the new study because the new one aims at answering questions under new situations (not included in old data). In that case, scholars should collect new data, which will be helpful for future research.

The study uses a convenient sampling strategy to select the examples of studies for illustration and therefore is very subjective. Using a convenient sample of studies could bias and underestimate the open data progress in nonprofit research because of our ignorance of other significant progress. A systematic review in the future will provide a better picture of how scholars in the field use old data and guidelines for the better utilization of old data.

To ensure privacy is a big challenge during the process of making data available, accessible, and searchable. Before submitting the data to journals and depositing them to repositories, scholars should comply with their institution's Institutional Review Board or Ethics Review Committee's requirements if they collect data from human participants. Datasets publicly available should not contain any personal information that may identify individuals' identities. This article also suggests making necessary data transformations before submission and deposit to avoid leaking the meta information of the dataset. For example, if the original data was stored in a spreadsheet format (e.g. Microsoft Excel), after removing essential identification, scholars can use a different statistical software (e.g. STATA) to transform the Excel format into a "dta" format to eliminate meta information. Doing so further improves data privacy.

Making old data available, accessible, and searchable can not only improve research transparency and accountability but also benefit new research. Many publishers and funding agencies now also require making data available publicly to improve research accountability and productivity.[1] This article urges more nonprofit scholars to be role model volunteers by voluntarily making their data available, accessible, and searchable to enhance academic transparency and

---

[1] Noy, N., & Brickley, D..

accountability and benefit new research.

## References

Ahmed, S., & Olberding, J. (2007). Can Student Philanthropy Help to Address the Current Nonprofit Identity Crisis? A Case Study of a Multiyear, Multidisciplinary Project at Northern Kentucky University. Journal of Public Affairs Education, 13(3/4), 593–615.

American Economic Association. (n.d.). Data and Code Availability Policy. American Economic Association. Retrieved October 30, 2020, from https://www.aeaweb.org/journals/data/data-code-policy

Google. (n.d.). Google Dataset Search. Retrieved March 28, 2020, from https://datasetsearch.research.google.com/.

Li, H., & Chen, B. (2018). Turning Challenges into Opportunities: Advancing Studies of Nonprofit Organizations in China. Chinese Public Administration Review, 9(1), 1–5. https://doi.org/10.22140/cpar.v9i1.172

Li, H., & McDougle, L. (2017). Information Source Reliance and Charitable Giving Decisions. Nonprofit Management and Leadership, 27(4), 549–560.

Li, H., McDougle, L., & Gupta, A. (2020). Experiential philanthropy in China. Journal of Public Affairs Education, 26(2), 205–227. https://doi.org/10.1080/15236803.2019.1667185

Li, H., Xu, C., & McDougle, L. M. (2019). Philanthropy Can Be Learned: A Qualitative Study of Student Experiences in Experiential Philanthropy Courses. Philanthropy & Education, 2(2), 29–52.

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., Wang, Q., Dong, C., & Li, H. (2017a). The Research Infrastructure of Chinese Foundations, a database for Chinese civil society studies. Scientific Data, 4, 170094. https://doi.org/10.1038/sdata.2017.94

Ma, J., Wang, Q., Dong, C., & Li, H. (2017b). The Research Infrastructure of Chinese Foundations, a database

for Chinese civil society studies [Data set]. In Harvard Dataverse. Harvard Dataverse. https://doi.org/10.7910/DVN/OTNI1L

McDougle, L. M., & Handy, F. (2014). The Influence of Information Costs on Donor Decision Making. Nonprofit Management and Leadership, 24(4), 465–485.

McDougle, L. M., & Lam, M. (2014). Individual- and Community-Level Determinants of Public Attitudes Toward Nonprofit Organizations. Nonprofit and Voluntary Sector Quarterly, 43(4), 672–692. https://doi.org/10.1177/0899764013479830

McDougle, L., McDonald, D., Li, H., McIntyre Miller, W., & Xu, C. (2017). Can Philanthropy Be Taught? Nonprofit and Voluntary Sector Quarterly, 46(2), 330–351. https://doi.org/10.1177/0899764016662355

McKeever, B., Greene, S., MacDonald, G., Tatian, P., & Jones, D. (2018). Data Philanthropy: Unlocking the Power of Private Data for Public Good. Urban Institute. https://www.urban.org/research/publication/data-philanthropy-unlocking-power-private-data-public-good

Millisor, J., & Olberding, J. C. (2009). Student Philanthropy in Colleges and Universities. Academic Exchange Quarterly, 13(4), 11–16.

Nonprofit and Voluntary Sector Quarterly. (2020). NVSQ Data Transparency Policy for Results Based on Experiments: Changes in Article Submission Guidelines. https://journals.sagepub.com/pb-assets/cmscontent/NVS/NVSQ_data_policy_2020.pdf

Noy, N. (2018, September 5). Making it easier to discover datasets. Google. https://blog.google/products/search/making-it-easier-discover-datasets/

Noy, N., & Brickley, D. (2017). Facilitating the discovery of public datasets. Google AI Blog. http://ai.googleblog.com/2017/01/facilitating-discovery-of-public.html

Olberding, J. C. (2012). Does Student Philanthropy Work? A Study of Long-term Effects of the "Learning by Giving" Approach. Innovative Higher Education, 37(2), 71–87.

Song, Y., & Fu, L. (2018). Do Charitable Foundations Spend Money Where People Need It Most? A Spatial Analysis of China. ISPRS International Journal of Geo-Information, 7(3), 100. https://doi.org/10.3390/ijgi7030100

Wang, Q., & He, L. (2018). Are the Wealthy Also Healthy? An Empirical Evaluation of the Financial Health o
f Chinese Foundations. Chinese Public Administration Review, 9(1), 6–24. https://doi.org/10.22140/c
par.v0i0.155

Xu, C., Li, H., & McDougle, L. M. (2018). Experiential Philanthropy. In A. Farazmand (Ed.), Global Encyclop
edia of Public Administration, Public Policy, and Governance (pp. 1–7). Springer International Publis
hing. https://doi.org/10.1007/978-3-319-31816-5_3048-1