2-2019

Recording Chinese Grassroots NGOs: A Web Archiving Project

Zhaohui Xue

Jidong Yang

Yan Long

Follow this and additional works at: https://scholarsarchive.byu.edu/jeal

Recording Chinese Grassroots NGOs: A Web Archiving Project

Zhaohui Xue
Stanford University

Jidong Yang
Stanford University

Yan Long
University of California, Berkeley

Introduction

The Chinese NGO web archiving project at Stanford is an effort to track, collect, store and provide access to the born-digital data of selected Chinese grassroots NGOs. More than 400 NGO web sites, Weibo (the Chinese equivalent of Twitter) posts, blogs, and selected news sites related to NGOs are archived on a monthly or quarterly basis. The project received financial support from Stanford University libraries, Stanford Center on Philanthropy and Civil Society, and the Mellon Foundation-Council on East Asian Libraries Innovation Grants for East Asian Librarians.

Why collect data on grassroots NGOs in China?

The past three decades witnessed a rapid growth in the number and diversity of Chinese non-governmental organizations (NGOs) or social organizations, as they are labeled in China. By 2017, there were more than 800,000 social organizations registered in China (Huang 2018). "NGO," by the commonly accepted definition, is an organization that is independent of government control. However, there is no consensus on what constitutes an NGO in the Chinese context. Many so called non-governmental organizations are actually government-organized non-governmental organizations (GONGO), such as the All-China Federation of Trade Unions (ACFTU), the All-China Women's Federation, and the Communist Youth League. Many other NGOs, such as professional associations, chambers of commerce, and learned societies, have a close relationship with the party-state in China. In a detailed study of these organizations and their successors, Unger (2008, 9) concluded that "China's major associations were in fact founded by the state and today remain firmly under the control of a state or Party agency." The labels for such organizations vary and are often used interchangeably, such as non-governmental organizations (非政府组织), mass organizations (群众团体), people's organizations (人民团体), social organizations (社会组织), third sector organizations (第三部门组织), or voluntary organizations (志愿组织), etc.

New forms of civic organizations and real grassroots NGOs emerged in the 1990s after the Fourth World Conference on Women in Beijing in 1995. At that time, the concept of NGOs as a form of independent civic organization was officially introduced to the Chinese audience. Many grassroots organizations, not funded by the state, mushroomed, particularly in the

areas of environmental protection, rural education, HIV/AIDS patient rights, etc. The wave of new social organizations can be attributed to the following factors. First, drastic social changes, widening income disparities and environmental and public health problems, all increased public awareness and social needs for organized civil actions in China (Hsu 2014). Second, the marketization and privatization of the Chinese economy since the 1980s created "systemic pressures" that pushed the government to support the development of NGOs in order to alleviate its own overload of responsibilities in the "micromanagement of economy and social welfare" (Howell 2004). Thirdly, China's opening-up to the outside world brought not only new concepts of civic organizations and civil society, but also foreign foundations and international NGOs that facilitated and provided the know-how in the development of grassroots organizations (Spires 2012). Since then, the Chinese NGOs have taken part in reframing the nation-state both from below, by taking over social functions that the Chinese state used to perform, and from above, by their capacities to mobilize political and financial resources beyond the bounds of national politics.

Primary characteristics of the Chinese grassroots NGOs can be described as low operating costs and strong dependence on Internet technology (Zhou and Pan 2016). As most of the Chinese NGOs emerged in the digital age of the Internet, it was natural that these organizations used this low cost, effective means for communication and formation of ties across traditional socioeconomic boundaries. This is particularly salient in China, as many grassroots NGOs were not able to obtain legal non-profit status, due to government policy restrictions. Setting up a website or a Weibo site is the most effective way for the NGOs to show their legitimate presence to the public. Many Chinese NGOs only exits and operate in a virtual space through their websites, blogs, Weibo and Douban.

There is a sizeable literature on Chinese NGOs and the emerging civil society. Scholars have paid particular attention to the use of Internet by Chinese NGOs. An early study by Yang (2007) found that the internet has had special appeal to relatively new organizations oriented to social change and that a "web" of civic associations has emerged in China (Yang 2007). Since then many studies have examined various aspects of this phenomenon, such as how the NGOs' made use of the Internet, their Internet behaviors, networking building, as well as the effectiveness of their Internet use. Yangzi Sima (2011) showed that Internet technologies effectively empowered resource-poor activists in their self-representation, information brokering, network building, public mobilization, and their development of discourse. Zhou and Pan (2016) explored Chinese NGOs' behavior on domestic social media platforms. Sommerfeldt and Xu (2017) studied how the Chinese NGOs evaluate their digital communication. Liu (2011) used data from 19 environmental NGOs in Beijing to identify Internet-based environmental collective actions and to illustrate the specific conditions under which the NGOs employ the Internet to engage in these actions. Yang and Taylor (2010) argued that the Chinese NGOs' websites provide information to members, the public and the media but these organizations do little to organize the public to participate in environmental social movements.

There are also efforts by activists and academics to promote NGOs' adoption of Internet technology and provide training in this regard. For example, since 2009, the NGO2.0 program, launched by Professor Jing Wang of the John Hopkins University, in collaboration

with the University of Science and Technology of China, NGO Communication Net, and other institutes, have provided digital and social media literacy training to grassroots NGOs in the western and central provinces of China (Wang 2015).

Even though the research findings above point to the rich and valuable information of and about Chinese grassroots NGOs on the Internet, no library or any other organization is systemically collecting materials from and on Chinese NGOs, especially in terms of information on their websites and social media outlets. Those websites provide essential sources of information on the rise and development of an emerging civil society in China. These chronical records constitute a historical archive of valuable information that allow researchers' inquiry into the mechanisms and efforts to bring about social change in the rapidly changing Chinese society. However, many grassroots NGOs are unlikely and do not have the capacity to preserve their own records. Institutional archives and libraries also often overlook this area.

Looking even more broadly, there have been, at best, only limited efforts in web archiving and preservation of born-digital resources in the Chinese language, and in providing access to the preserved web information. This is true in China and the US as well as the rest of the world. In China, as we know, the National Library of China started their own web information collection and preservation program in 2003. Its focus is on Chinese governmental websites, news media and major events (Liu and Kong 2013). Peking University established its "Web InfoMall", which claimed to have collected more than 7 billion Chinese web pages since 2001. However, this project's website and the data it collected are confined to domestic visits only (Huang et al. 2013). Outside China, there are a number of curated collections. A three-university collaboration project by George Washington University, Georgetown University and Johns Hopkins University, is archiving social media in China's anti-corruption campaign (Ye et al. 2018). University of California, San Diego, is collecting information on selected websites of current news and Bulletin Board sites in China. In other parts of the world, University of Heidelberg and University of Leiden collaborated on the Digital Archive for Chinese Studies (DACHS) which centers on social and political discourse for Chinese studies. There are a few smaller-scale projects, related to special topics as such the 2013 Sichuan earthquake, the 2008 Olympics, etc. The above projects primarily focus on politics and major events, in contrast, the Stanford project is designed to capture and record the operation and development of the grassroots NGOs, which offers a glimpse into a rapidly changing society where the ordinary citizens are organized from bottom up and participating collectively in civic actions.

There is an urgent need to archive the information related to NGOs in China today, particularly because the Chinese government has adopted many harsh and restrictive policies toward NGOs. As a result, information on NGOs, in particular NGOs' online records, are often censored and have become more difficult to access.

The systematic collection of Chinese grassroots NGO data fits well with Stanford East Asia Library's long-established collection focus on modern and contemporary China. Stanford East Asia library's Chinese collection is particularly strong on social and political affairs for modern China. In addition to print material, databases and numeric data, the

library makes great effort in collecting unique, ephemeral and micro-level research materials. Born-digital resources from grassroots NGOs provide a great opportunity and gold mine for us to strengthen and expand the collection in an unconventional way, which will enrich the content as well as the format of the collection.

This project is especially important to social science researchers. As we know, scholars in the humanities and social sciences who study China face great difficulties to collect data. Access to archival documents and activists in the field is restricted, and often times such obstacles are too difficult to overcome. In addition, the extent and scope of digital data are often beyond any scholar's capability to collect and digest on an individual basis. This is where the library plays a fundamental role. As a depository of information and knowledge, the library de-sensitizes the political aspects of the materials and could stay away from ideological debates. The library's organizational infrastructure provides scholars with a stable and lasting source of information for long-term projects, even when they move away from one university to another, as their academic careers develop.

Collaboration is the key to success

In order to carry out the web archiving project and to accomplish the goal of building a unique digital collection, we had to overcome some major challenges. First, there were significant resource limitations in terms of funding and staff time. In particular, we needed continuous funding sources that would sustain the project for a significant period of time. Second, the target of this project—Chinese grassroots NGOs—was evolving over time. Some organizations dissolved and disappeared over the years. This raises the question: How do we identify grassroots NGOs that are historically significant and their data would have potential research value? Thirdly, given a large number of grassroots NGOs on the one hand and our limited resources for archiving data on the other, we needed to set scope and boundary of the project carefully, with meaningful criteria. Finally, to carry out this non-traditional form of collection development, we needed technical knowledge and skills that were not available to us in the past.

The answer to these questions was to build a collaborative partnership among multiple parties—librarians, researchers and IT experts, as well as a variety of funding entities. The Stanford East Asia Library and the researcher on campus specializing in Chinese social organizations took the lead in forming partnerships with multiple organizations and teams inside and outside library. On campus, we received great support from library administration and the Stanford center for Philanthropy and Civil Society, as well as technical guidance and assistance from the newly established web archiving services under Stanford Digital Library Systems and Services (DLSS). We also worked with the metadata unit in the library for creating metadata. Beyond Stanford, we sought and received funding support from various foundations, as well as valuable input and advices from scholars and activists in China.

Close collaboration between the library and researcher proved essential for the project's success, as these parties offered different expertise to make this project a great collaborative experience. Librarians navigate and build connections; they coordinate

communication and discussion among the contributors, and implement and manage resources. Stanford East Asia Library provided the organizational infrastructure and platform for the project.

Researchers were the project's brain and mobilizers. They helped establish the academic legitimacy of this project. Indeed, it was the anticipation of research interest in the online activities of Chinese grassroots NGOs that initiated the development of the project. Researchers brought their expertise and knowledge of the field to the project and helped set up scope and selection criteria. Moreover, it is through their fieldwork, conferences, collaborations with other research partners and consultation with activists and practitioners in China that we received valuable input to and positive feedback on our ongoing project, which led to a series of improvements over time. Through this process of trial and error, we set the scope and criteria of the project as follows:

- Focus on grassroots NGOs that were founded and operated by individuals
- Grassroots NGOs that had operated for at least three years
- Grassroots NGOs that had gained recognition among their peer organizations
- Areas of focus: environment, rural development, labor and migrating workers, women rights, sexuality and gender, public health, education, community services, and at-risk websites.
- Targeted outlets: Selected major news sites, websites of selected NGO incubators and foundations

The criteria are clearly defined and easy to follow. As of 2018, we have archived more than 400 sites, including websites, Weibo, blog and Douban, with 359 continuing to be crawled on an ongoing basis.

The collaborative team also has been actively involved in fundraising. Funding for this project came from multiple sources: library collection development budget, Stanford Center for Philanthropy and Civil Society, Leping Foundation, and the Mellon Foundation-Council on East Asian Libraries Innovation Grants for East Asian Librarians.

Unlike traditional collection development practice, web archiving demands involvement of and interaction among a large number of different library units and individuals, not only subject librarians and researchers, but also IT specialists and experts. Collaboration with DLSS and the metadata unit in the Stanford library brought experts and talents in developing and implementing programs for web archiving, digital depository and preservation, data representation and metadata creation. We followed the general guideline and policy for web archiving established by the Stanford web archiving service. We sent permission notices to the web content owners of our intent to archive and publicly represent their web content after a six-month embargo before proceeding with the crawling. The metadata were created using Metadata Object Description Schema (MODS) developed by the metadata specialist for the Stanford Digital Repository.

With the tremendous efforts and large-scale collaboration, this data is now taking

shape and provides an opportunity for innovative research. With a wide variety of information types, the data—which ranges from historical gazetteers, maps, to news items, images, and social medial posts—is a rich volume of information that can be considered a sealed treasure trove that, once opened, has the potential to illuminate the past and present in ways hitherto thought closed for lack of practical methods to give the findings shape and meaning.

Innovation in collection development, preservation and representation of data

As digital scholarship becomes more important and dynamic, and information technology increasingly complex, libraries have been exploring and experimenting with new and creative ways to better serve the research community in a rapidly changing environment. This project on Chinese grassroots NGOs web archiving aimed to innovatively curate a subject-specific digital collection that would open doors to do innovative research.

Traditionally, East Asian librarians collect print and electronic resources produced by commercial publishers, and then catalog and add them to the library's holdings. Few of them have tried to collect Internet-based information that plays an increasingly central role in transforming contemporary societies in East Asia. We used Archive-It service to automatically and periodically capture rich data from the web and social media sites of grassroots NGOs. Although the crawl quality for these sites varies, due to the web crawler's limitations in crawling some more dynamitic Weibo sites, and more importantly, the often unstable, limited Internet connection in China, the data being collected and saved has tremendous value for understanding the Chinese society, as the data recorded the evolution of NGOs over time. Many websites have disappeared in recent years because of Chinese government's crackdown on grassroots NGOs, particularly for those rights advocacy NGOs. Our archived data are perhaps the only source of information to track their web existence and information. Although it is beyond the capacity of this project to archive all web documentations related to China's growing civil society, we hope that this project provides a valuable experiment in creating new primary sources for East Asian studies.

Discoverability is a key issue we have to address whenever we acquire any special collections. The data are accessible from the SearchWorks, Stanford library's main discover tool. We also adopted Spotlight exhibit to represent and provide access to the data we captured. Developed by Stanford University Libraries, Spotlight (http://spotlight.projectblacklight.org/) is an innovative open-source software solution that enables libraries and other cultural heritage institutions to build online exhibits from content in their repositories to better highlight their digital collections. Spotlight functions as a plugin for Blacklight (http://projectblacklight.org), a popular open-source platform for building library discovery environments, but it enhances Blacklight by providing a self-service form-based user interface that allows librarians or faculty to customize the search interface and homepage and to build media-rich feature pages, providing historical context for their collections. In this way, Spotlight enables an exhibit builder to create pages giving viewers a deeper understanding of the collection and its items. In addition, it allows both for keyword search and for browsing of custom-built facets, such as resource type, date,

language, author, topic, region, era, and organization. Spotlight has similarities to existing exhibit solutions but seeks to expand on current models to more tightly integrate with repository infrastructures and bring equally strong focus on search results, objects, and supporting intellectual scaffolding. By using this easy-to-use tool, the longitudinal data of the Chinese grassroots NGOs that was captured by Archive-It, then transferred and preserved in the Stanford Digital Repository, is re-opened to the scholars and practitioners with an interest in China's civil society movement.

Conclusion

Recent years witnessed a historical moment of change as the Chinese government has passed a series of laws to exert greater regulatory control over the emerging civil society. In spite of such stricter conditions, however, the growth of NGOs has not stopped and there is a wide range of choices available for them. There are important questions regarding the NGOs in China: What will be the role of NGOs in China's transformation, especially the emergence of a civil society? Are the NGOs going to become harbingers of democracy, alternative social service providers, the state's collaborators, the engine for social experiments, or a combination of these possibilities? We hope that our effort in developing the web archiving project can provide innovative data and information that can facilitate researchers to address these and other important questions.

References

Howell, Jude. 2004. "New Directions in Civil Society: Organising Around Marginalised Interests." In *Governance in China*, edited by Jude Howell, 143-71. Lanham, MD: Rowman & Littlefield).

Hsu, Jennifer J.Y. 2014. "Chinese Non-governmental Organisations and Civil society: A Review of the Literature." *Geography Compass* 8 (2): 98-110. DOI: 10.1111/gec3.12117.

Huang, Lianen, Jinping Li and Xiaoming Li. 2013. "InfoMall: A Large-Scale Storage System for Web Archiving." In *Web-Age Information Management. WAIM 2013. Lecture Notes in Computer Science, Vol. 7901*, Yunjun Gao et al. eds., 89-98. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-39527-7_11.

Huang, Xiaoyong (eds). 2018. Zhongguo she hui zu zhi bao gao (2018). She hui ke hui chu ban she, 2018. 中国社会组织报告 (2018). 社会科学文献出版社，2018.

Liu, Jingfang. 2011. "Picturing a Green Virtual Public Space for Special Change: A Study of Internet Activism and Web-based Environmental Collective Actions in China." *Chinese Journal of Communication* 4 (2): 137-66.

Liu, Qing and Fanlian Kong. 2013. "Zhongguo wang luo xin xi cun dang ji qi yu guo wai de bi jiao: ji yu guo jia tu shu guan WICP xiang mu de yan jiu." *Tu shu xin bao gong zuo* 57(18), 80-93. 中国网络信息存档及其与国外的比较：基于国家图书馆 WICP 项目的研究. *图书情报工作* 57 (18): 80-93.

Sima, Yangzi. 2011. "Grassroots Environmental Activism and the Internet in China: Constructing a Green Public Sphere in China." *Asian Studies Review* 35 (4): 477-97.

Sommerfeldt, Erich J. and Sifan Xu. 2017. "Making a 'Difference' with Digital Media? The Evaluation Perspectives, Practices, and Challenges of Chinese NGOs." *Chinese Journal of Communication* 10 (1): 89-107.

Spires, Anthony. 2012. "Lessons from Abroad: Foreign Influences on China's Emerging Civil Society." *China Journal* 68: 125-46.

Unger, Jonathan. 2008. "Chinese Associations, Civil Society, and State Corporatism: Disputed Terrain." In *Associations and the Chinese State: Contested Spaces*, edited by Jonathan Unger, 1-13. New York: Sharpe.

Wang, Jing. 2015. "NGO2.0 and Social Media Praxis: Activist as Researcher." *Chinese Journal of Communication* 8 (1): 18-41.

Yang, Aimei and Maureen Taylor. 2010. "Relationship Building by Chinese ENGOs' Websites: Education, Not Activation." *Public Relations Review* 36: 342-51.

Yang, Guobin. 2007. "How Do Chinese Civic Associations Respond to the Internet? Findings from a Survey." *The China Quarterly* 189: 122-43.

Ye, Ding, Daniel Kerchner, Yan He, Cathy Zeljak, and Yunshan Ye. 2018. "Web Archiving for Librarians: A Three-University Collaboration to Preserve Non-Official Voices in China's Anti-Corruption Campaign." *ACRL 2017 Proceedings*, American Library Association, March 13, 2017. http://www.ala.org/acrl/conferences/acrl2017/papers (Accessed February 27, 2019). Document ID: 38d9d5a4-f7e0-27b4-c104-a7056038b90b.

Zhou, Huiquan and Pan Quanxiao. 2016. "Information, Community, and Action on Sina-Weibo: How Chinese Philanthropic NGOs Use Social Media." *Voluntas: International Journal of Voluntary & Nonprofit organizations* 27: 2433-57.