
Software Application Profile

Network canvas: an open-source tool for capturing social and contact network data

Patrick Janulis,^{1,2*} Gregory Phillips II,^{1,2} Joshua Melville,^{1,2}
Bernie Hogan,³ Kate Banner,^{1,2} Brian Mustanski,^{1,2} Carrie B Oser,^{4,5,6}
Martha Tillson,^{4,5} John Schneider^{7,8} and Michelle Birkett^{1,2}

¹Department of Medical Social Sciences, Northwestern University, Chicago, IL, USA, ²Institute for Sexual and Gender Minority Health and Wellbeing, Northwestern University, Chicago, IL, USA, ³Oxford Internet Institute, University of Oxford, Oxford, UK, ⁴Department of Sociology, University of Kentucky, Lexington, KY, USA, ⁵Center on Drug and Alcohol Research, University of Kentucky, Lexington, KY, USA, ⁶Center for Health Equity Transformation, University of Kentucky, Lexington, KY, USA, ⁷Department of Medicine, University of Chicago, Chicago, IL, USA and ⁸Chicago Center for HIV Elimination, University of Chicago, Chicago, IL, USA

*Corresponding author. Northwestern University, 625 N Michigan Ave, Chicago, IL 60611, USA.

E-mail: patrick.janulis@northwestern.edu

Received 7 September 2022; Editorial decision 15 February 2023; Accepted 8 March 2023

Abstract

Motivation: Social influence and contact networks are extremely important for understanding health behaviour and the spread of disease. Yet, most traditional software tools are not optimized to capture these data, making measurement of personal networks challenging. Our team developed Network Canvas to provide an end-to-end workflow with intuitive interfaces to enable researchers to design and conduct network interviews.

Implementation: Network Canvas consists of three applications (*Architect*, *Interviewer* and *Server*). All applications are written in JavaScript and run on Windows, macOS and Linux; *Interviewer* also runs on Android and iOS.

General features: Network Canvas substantially reduces the complexity and technical knowledge required to collect network data via three point-and-click applications. The tool has wide applicability for measuring contact and social influence networks in epidemiology.

Availability: Network Canvas is open source and freely available [networkcanvas.com] under the GNU General Public License 3.

Key words: social networks, contact networks, measurement

Introduction

Social and contact networks have an enormous impact on the health of individuals.¹ For example, networks serve as

conduits for pathogens, information and behaviour, all of which greatly affect the distribution of disease in a population.² Accordingly, accurate measurement of these networks

is paramount to understanding population health, modelling the spread of communicable diseases and implementing effective prevention interventions.

Despite the importance of personal networks (i.e. the social networks surrounding individuals), measuring these networks is incredibly complex³ and traditional survey tools are ill equipped to capture these data.⁴ Whereas many survey software tools are flexible enough to ask traditional network survey questions, they are not optimized for this purpose, which may lead to measurement error and underreporting of network contacts. Error is particularly likely when capturing perceived ties between network members (e.g. are the participant's contacts in contact with one another?). The number of possible non-directed ties in a network equals $n * (n - 1)/2$, where n is the number of individuals in one's network. These potential ties are traditionally captured during an interview using a matrix to ask all possible connections between network members, or in an online survey with n numbers of pages, each page having $n - 1$ questions to indicate the presence or absence of a tie between each network member.⁵ Given the potential cognitive burden, these methods may lead to underreporting of network members and network ties,^{6,7} particularly if multiple network types are captured sequentially.⁸ Visual methods that let the participant draw their network offer an alternative approach,⁹ but cannot be easily implemented using traditional digital tools. For this reason, many researchers have historically relied on bespoke data collection tools that require substantial coding expertise to create and deploy.⁵

Our team developed Network Canvas to leverage the unique benefits of a visual network survey with the ease of building, deploying and collecting data which is commonly associated with traditional digital survey tools. This tool is widely applicable to epidemiological studies where connections between individuals (e.g. social, behavioural or physical proximity) play an essential role in understanding the spread of disease¹¹ or context of behaviour.² Whereas other specialized software exists for network data collection, such as Egoweb 2.0¹⁰ and Enso,⁵ Network Canvas is unique in its emphasis on visual-oriented data collection optimized for in-person interviewing and interviewer-assisted self-administration. In this paper we introduce Network Canvas and present its key benefits to epidemiological research, namely reducing technical barriers to creating and conducting network interviews, thereby increasing the ease with which non-network data specialists can collect these data. We highlight two use cases of this software: (i) a cohort study of young men who have sex with men; and (ii) a study of multilevel factors affecting opioid use disorder treatment.

Implementation

Network Canvas is written in JavaScript. It uses the React¹² framework for its core components, with wrapper technologies to enable publication as point and click desktop and tablet apps. Network Canvas has three distinct applications: *Architect*, *Interviewer* and *Server*. *Architect* facilitates the creation of interview protocol files (stored as files with the 'netcanvas' extension), *Server* coordinates the deployment of these files to interview devices and retrieves response data from them, and *Interviewer* is used to conduct interviews. All three applications can be run on Windows, macOS and Linux, and *Interviewer* also runs on Android and iOS devices. While basic knowledge of behavioural data collection is required, no technological expertise is necessary to use this software, as all applications can be downloaded and run with a click of a button and include extensive graphical user interfaces. For example, readers can conduct a network interview by launching *Interviewer* and loading either of the protocol files included as *Supplementary Material*, or load them in *Architect* as the foundation of a new custom protocol. Although comprehensive documentation of the features within Network Canvas can be found elsewhere,¹³ below we highlight key features of this software.

Architect

Architect streamlines the creation of a network interview protocol via a point-and-click graphical user interface. Design of a protocol is focused on the creation of interview 'stages'. Most stages focus on a single variable type (e.g. ordinal or categorical) to reduce cognitive burden of each task. Many stages also allow multiple 'prompts' that collect participant responses to different questions within the same variable type but with different response options across prompts. Whereas most features in *Architect* are typical in traditional web-based survey creation, there are key features that are unique to the design of a network interview. Users may create stages that enable the collection of multiple types of network entities (e.g. people and places), node characteristics in all classic variable types (e.g. categorical, ordinal, number etc.), types of connections between entities (e.g. sexual contact, drug use, social ties), and characteristics of those connections (e.g. edge weights). *Architect* also has many traditional survey design features such as real-time editing of question wording and variable names. Similarly, skip logic and filtering can be created both at the participant (i.e. only asking questions to a subset of participants) and network member (e.g. only asking questions about network members with certain characteristics) levels. Completed protocols are saved to

local storage as a ‘.netcanvas’ file format, which is interpretable by all three applications for later editing and deployment.

Server

Server provides a central hub on a local machine for protocol deployment to and data extraction from interview devices. To transfer a research protocol to *Interviewer*, researchers must first pair their instance of *Server* with device(s) running *Interviewer* on a local network (or virtual private network). This pairing process mirrors the convenience of Bluetooth pairing in consumer electronic devices such as automatic device detection and secure pairing through use of uniquely generated PINs. A single instance of *Server* can be paired to multiple interview devices for larger projects and, once a secure pairing is established, protocols loaded into *Server* can be downloaded and deployed on any paired device. Furthermore, completed interview data can be transferred to *Server* where they can be monitored and exported. Data can be exported in two file formats: GraphML¹⁴ and comma-separated values (CSV) files. GraphML is a specialized markup language using XML syntax to promote standardized interoperability of graph data and is easily importable to most common software tools used in network analysis (e.g. iGraph¹⁵ R package). For both export formats, interviews can be exported as separate files (a single personal network) or as a combined network with multiple disconnected networks representing each personal network.

Interviewer

Interviewer serves as the participant interfacing application used on interview devices to collect data. As noted, interviews are conducted in discrete stages, some with multiple prompts, to simplify the network interview process and focus participants on a single screen. *Interviewer* is optimized for an interviewer-assisted experience where an interviewer first guides a participant through the basic mechanics of the application. However, the tool is flexible to allow for a more self-guided interview experience (i.e. where the participants take control of input on the device) for participants who prefer more autonomy or privacy when entering their responses.

In addition to the *Server* pairing process discussed above, *Interviewer* also enables protocol importation via two alternative methods for those who are unable (e.g. security constraints on local networks) to use the *Server* application. The first method is direct import via URL. The researcher may simply provide a URL directly to a publicly available ‘.netcanvas’ file and the software will directly

import that protocol onto the local device. Similarly, protocol files may be loaded from local device storage. In both cases, protocols are checked for validity and users are presented with an appropriate error message if the file is not a correctly formatted Network Canvas protocol file. After an interview, data can be transferred to a paired *Server* instance or downloaded directly to the local storage on the interview device.

Usage Case 1: RADAR

An early predecessor to Network Canvas (i.e. netCanvas-R) was used to measure the sexual, drug use and social networks in the RADAR study, a large longitudinal study of over 1000 young men who have sex with men and young transgender women being followed for up to 10 years.^{16,17} These data have been used in a growing number of publications examining the impact of sexual network structure on disease spread^{16,18} and the influence of social networks on health.^{19,20}

We created an interview protocol ([Supplementary Material](#)) inspired by this study. [Figure 1a](#) displays the name generation stage where participants are asked to provide names of close social, drug and sex contacts. Individuals added during the name generation stage can then be visually arranged on the sociogram ([Figure 1b](#)) where participants position their network members on the

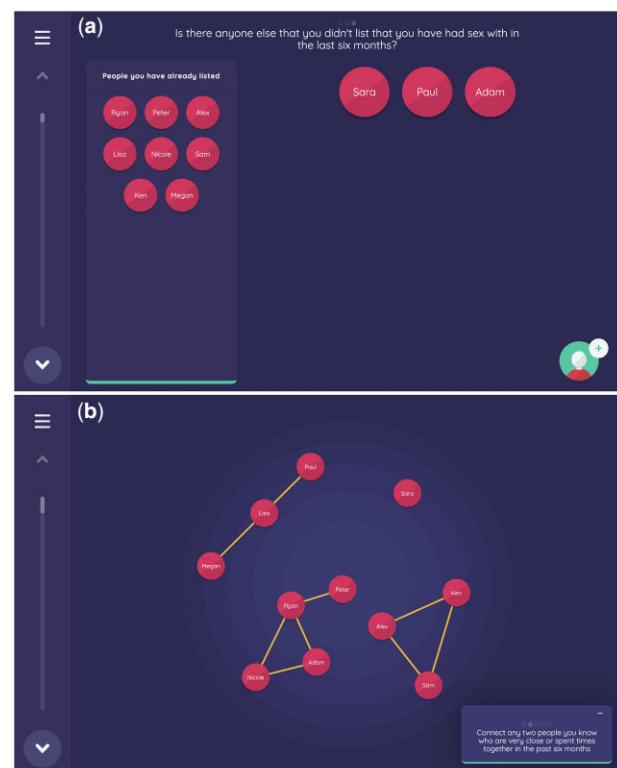


Figure 1 Name generation (a) and sociogram (b) stages



Figure 2 Categorical (a) and ordinal bin (b) stages

screen and indicate ties between members. Furthermore, the categorical bin (Figure 2a) and ordinal bin stages (Figure 2b) enable quick sorting of network members into bins to easily classify network members on categorical or ordinal attributes. This approach contrasts with a typical survey where these data are captured using multiple choice grids with a single row for each network member. Evaluations have shown that this method provided similar quality data to a traditional computer-assisted self-interview, was 50% faster than a traditional network interview and was rated highly usable by participants on a standardized System Usability Scale.¹⁷ Furthermore, analysis of longitudinal use of the Network Canvas approach revealed minimal levels of panel conditioning, such as reduction in network member nominations in subsequent visits,²¹ a key concern in longitudinal network research.²²

Usage Case 2: JCOIN

Network Canvas has also been implemented within the Justice Community Opioid Network (JCOIN), a \$150-M multisite research cooperative supported by the National Institute on Drug Abuse, to address the opioid epidemic in justice-involved populations. JCOIN comprises 11 research hubs, three of which have added a personal network protocol while conducting clinical trials examining an array of linkage facilitation strategies to connect people with

opioid use disorder to community-based treatment after leaving jail or prison.^{23,24} Data collection is underway within the University of Kentucky hub and is planned for implementation in University of Chicago and Yale University hubs in 2022. Network Canvas is used to capture social network data at baseline (i.e. in jail/prison), and at 6 and 12 months following release. The study has currently gathered data from 259 baseline observations, with recruitment and data collection ongoing. The network protocol focuses on capturing the social context of substance use and opioid use disorder treatment, with an emphasis on social support and drug use networks. Participants were asked usability questions about Network Canvas on a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree), with the majority reporting feeling confident using the application (mean = 4.98, SD = 0.18) and agreeing that most people would learn to use the application very quickly (mean = 4.92, SD = 0.44). Finally, participants were also asked which two words described the interview experience, with common responses including ‘easy’, ‘fast’, ‘simple’ and ‘good’. The protocol (available in the *Supplementary Material*) shares many themes with the RADAR study protocol but also uses unique features such as the ‘per-alter form’ to collect follow-up data on the drug use behaviours of each network member who uses drugs.

Conclusion

One reason network research remains rare in epidemiology is due to the difficulty in collecting network data. The goal of Network Canvas is to provide an end-to-end workflow for the collection of network data via in-person interviewing, to substantially reduce the technical expertise required to collect these data and to expand the ability of researchers to conduct this work beyond those with specific expertise in these methods. Limitations of the software include its focus on in-person interviewing. Other network survey software^{4,10} may be more appropriate for studies that require remote administration or self-interviews. The suite is also not optimized for communicating with external databases. Despite these limitations, we believe this tool holds tremendous potential for studying the network context for a wide variety of communicable and non-communicable diseases. Furthermore, as an open-source project, we encourage and look forward to additional feedback and contributions that will maximize the usefulness of this tool for epidemiological research.

Ethics approval

The Network Canvas project was reviewed and determined to be non-human subjects research by the Northwestern University

Institutional Review Board (IRB; STU00202963). The RADAR study was reviewed and approved by the Northwestern University IRB (STU00087614). The University of Kentucky JCOIN study was approved by the University of Kentucky Medical IRB (Protocol #51656).

Data availability

Data referenced in this article cannot be shared for the privacy of participants in these studies and because of the sensitive nature of the corresponding data. Data sharing will be considered on reasonable request to the corresponding author.

Supplementary data

Supplementary data are available at *IJE* online.

Author contributions

P.J., G.P. II, J.M., B.M., K.B. and M.B. all contributed to the design and development of the software as well as reviewing and editing drafts of the manuscript first drafted by P.J. B.M., C.B.O., M.T. and J.S. all provided data from the usage cases and reviewed, edited and approved the manuscript.

Funding

This work was supported by grants from the National Institute on Drug Abuse (R01DA042711; U01DA036939; UG1DA050069) at the National Institutes of Health.

Conflict of interest

None declared.

References

- Smith KP, Christakis NA. Social networks and health. *Annu Rev Sociol* 2008;34:405–29.
- Zhang J, Centola D. Social networks and health: new developments in diffusion, online and offline. *Annu Rev Sociol* 2019;45: 91–109.
- Eames K, Bansal S, Frost S, Riley S. Six challenges in measuring contact networks for use in modelling. *Epidemics* 2015;10: 72–77.
- Eddens KS, Fagan JM, Collins T. An interactive, mobile-based tool for personal social network data collection and visualization among a geographically isolated and socioeconomically disadvantaged population: early-stage feasibility study with qualitative user feedback. *JMIR Res Protoc* 2017;6:e124.
- Eddens KS, Fagan JM. Comparing nascent approaches for gathering alter-tie data for egocentric studies. *Soc Netw* 2018;55: 130–41.
- Eagle DE, Proeschold-Bell RJ. Methodological considerations in the use of name generators and interpreters. *Soc Netw* 2015;40: 75–83.
- Paik A, Sanchagrin K. Social isolation in America: an artifact. *Am Sociol Rev* 2013;78:339–60.
- Pustejovsky JE, Spillane JP. Question-order effects in social network name generators. *Soc Netw* 2009;31:221–29.
- Jaspersen LJ, Stein C. Beyond the matrix: visual methods for qualitative network research. *Br J Manage* 2019;59:107.
- Kennedy DP, Hunter SB, Chan Osilla K, Maksabedian E, Golinelli D, Tucker JS. A computer-assisted motivational social network intervention to reduce alcohol, drug and HIV risk behaviors among Housing First residents. *Addict Sci Clin Pract* 2016;11:4.
- Stattner E, Vidot N. Social network analysis in epidemiology: current trends and perspectives. In: *2011 Fifth International Conference on Research Challenges in Information Science*. 2011. <http://dx.doi.org/10.1109/rcis.2011.6006866> (3 March 2023, date last accessed).
- Facebook Inc. *React: A JavaScript Library for Building User Interfaces*. 2019. <https://reactjs.org/> (3 March 2023, date last accessed)
- Network Canvas Documentation Site. *Network Canvas*. 2019. <https://documentation.networkcanvas.com/> (3 March 2023, date last accessed)
- Brandes U, Eiglsperger M, Lerner J, Pich C. *Graph Markup Language (GraphML)*. 2013. <https://kops.uni-konstanz.de> (3 March 2023, date last accessed).
- Csardi N. *The igraph Software Package for Complex Network Research*. 2006. https://www.researchgate.net/profile/Gabor-Csardi/publication/221995787_The_Igraph_Software_Package_for_Complex_Network_Research/links/0c96051d301a30f26500000/The-Igraph-Software-Package-for-Complex-Network-Research.pdf (3 March 2023, date last accessed).
- Mustanski B, Morgan E, D'aquila R, Birkett M, Janulis P, Newcomb ME. Individual and network factors associated with racial disparities in HIV among young men who have sex with men: results from the RADAR cohort study. *J Acquir Immune Defic Syndr* 2019;80:24–30.
- Hogan B, Melville JR, Phillips GL et al. Evaluating the paper-to-screen translation of participant-aided sociograms with high-risk participants. *Proc SIGCHI Conf Hum Factor Comput Syst* 2016;2016:5360–71.
- Janulis P, Phillips G, Birkett M, Mustanski B. Sexual networks of racially diverse young MSM differ in racial homophily but not concurrency. *J Acquir Immune Defic Syndr* 2018;77:459–66.
- Janulis P, Birkett M, Ii PG, Mustanski B. Not getting high with a little help from your friends: Social versus drug network correlates of marijuana use among YMSM. *Addict Behav* 2019;92: 180–85.
- Phillips G 2nd, Neray B, Birkett M, Felt D, Janulis P, Mustanski B. Role of social and sexual network factors in PrEP utilization among YMSM and transgender women in Chicago. *Prev Sci* 2019;20:1089–97.
- Hogan B, Janulis P, Phillips G II et al. *Assessing the Stability of Ego-centered Networks over Time Using the Digital Participant-aided Sociogram Tool Network Canvas*. 2019. <https://ora.ox.ac.uk/objects/uuid:13d53984-a847-4bc0-8bab-a6460c8c45b3> (3 March 2023, date last accessed).

22. Silber H, Schröder J, Struminskaya B, Stocké V, Bosnjak M. Does panel conditioning affect data quality in ego-centered social network questions? *Soc Netw* 2019;56:45–54.
23. Staton M, Matthew Webster J, Leukefeld C *et al.* Kentucky Women's Justice Community Opioid Innovation Network (JCOIN): a type 1 effectiveness-implementation hybrid trial to increase utilization of medications for opioid use disorder among justice-involved women. *J Subst Abuse Treat* 2021;128:108284.
24. Ducharme LJ, Wiley TRA, Mulford CF, Su Z-I, Zur JB. Engaging the justice system to address the opioid crisis: the Justice Community Opioid Innovation Network (JCOIN). *J Subst Abuse Treat* 2021;128:108307.