RTI Press

ISSN 2378-7929

May 2021



Social Network Analysis Methods for International Development

Eric M. Johnson and Robert Chew



At their core, international development programs attempt to catalyze new relationships and new ways of working among program stakeholders. For example, youth development programs may seek to connect educators to private sector employers to make skills training more demand-driven, increasing youth employability. Others may seek to steer at-risk youth toward positive role models and community services, and away from gangs, to improve their safety and resiliency. Reflecting this, an official at the United States Agency for International Development (USAID), the world's largest bilateral international development donor, recently stated, "the design, implementation and evaluation of development programs is intrinsically about people, institutions and the relationships between them."1 Many development programs, however, lack a sophisticated understanding of these relationships and a way to measure change in them. Social

Key Research Needs

- Incorporate social network analysis (SNA) into monitoring, evaluation, and learning processes. SNA can be conducted at various points of a project to inform program design, adaptive management, learning, and evaluation by considering network structure and network changes over time.
- Demystify the use of SNA. Increased use of SNA tools and clear presentation in widely read publications are needed to bring the analytic approach into the mainstream of international development.
- Build capacity to conduct SNA. The capacity to conduct and interpret SNA is lacking across actors in international development. Efforts by some organizations to build capacity in the community are well noted and should be built upon.
- Build understanding of relationships between social networks and development outcomes. SNA will be useful only to the extent it helps users understand the relationship between networks and development outcomes that matter.
- Establish norms for data collection and identity protection. Data about individuals and their interactions with others are inherently sensitive data. As a part of standard research ethics protocols, SNA practitioners must make carefully considered decisions about how or if to anonymize data when reporting it.

network analysis (SNA) is a quantitative research method that programs can use to help address this critical knowledge gap. This policy brief advances the case for SNA in international development, outlines general approaches, and reviews two recently conducted case studies that illustrate its potential. Included is an agenda for future work and applications in international development.

Social Network Analysis

Social networks are "a set of players and patterns of exchange of information and/or goods among these players."² The intellectual home of network analysis is in sociology, where Durkheim emphasized the study of social relationship patterns, and Granovetter advanced the importance of "weak ties" (distant acquaintances) in relational phenomena, such as successfully acquiring leads for job opportunities.³ Later, in political science, Putnam explained waning social capital in United States by the breakdown of community networks, such as bowling leagues and economic structures like labor unions.⁴ Only recently, though, have researchers drawn on the methods and empirical basis of SNA to elevate it to a mainstream analytic tool. The rapid growth in published papers and grant funding over the past 15 years demonstrates this uptake.⁵

SNA is a quantitative analysis tool used to identify and understand relationships between people or, in other words, social networks. It visually displays data so researchers can see behavioral relationships at the micro level (individual, institutional) and patterns at the macro or network level. SNA has the flexibility to treat networks as both independent and dependent variables. For example, it can help answer how differences in individuals' networks (independent variable) explain their risk for contracting COVID-19⁶ or how racially segregated schools affect a young person's friendship networks (dependent variable).⁷

The data used in SNA can include secondary sources, such as social media data (connections, likes, shares, etc.); evidence of collaboration, such as co-authoring a paper; or administrative records such as school attendance, employment history, or club membership. Surveys can also collect primary data for SNA, with respondents asked to answer questions about their relationships, exchanges, and affiliations. Surveys often ask about the level of respondents' connection to the others (i.e., frequency of communication); the nature of those exchanges (information, goods and services, collaboration); and the value the respondent assigns to them.

SNA data are typically analyzed and interpreted in two ways. First, a set of network metrics can characterize the network and quantify its dimensions. Typical network measures include density, reciprocity, transitivity, centralization, and modularity. See Figure 1 for definitions and explanations of these metrics, among others. Second, researchers can also explore and interpret social networks visually. Various software tools map the connections among network actors and produce social network graphs or "sociograms." In these graphs, colors demark different kinds of actors, or nodes on the graph. The sizes of the nodes indicate the levels of connectedness. The position and partitioning of nodes in the network maps visualizes the network structure, including central actors, isolated actors, bridging actors, and any sub-groupings or cliques. See the Annex for more on the computing software needed and approaches to network visualization.

SNA metrics are often grouped into two categories to characterize networks by their level of (1) **network closure or cohesion**, measured by levels of density, reciprocity, transitivity, degree centrality, and shortest path, and (2) **network heterogeneity**, measured by modularity. Networks with higher levels of closure are associated with higher familiarity, trust, and social capital, and moreefficient exchange of information, goods, or services. Moreheterogeneous networks are considered more effective at mobilizing resources, given that network exchanges often require coordination of many skills and various inputs across different types of actors.

For SNA to be useful or effective, it should be methodologically well-aligned to answer clear research questions. In international development, it presents a potentially useful tool for understanding a range of network relationships, including levels of collaboration and exchange; the existence of central actors; excluded populations; and absent connections among individuals, organizations, or groups. Two recent applications of SNA to international development used SNA to understand collaboration among different types of actors that programs had been unable to assess adequately in the past.

El Salvador: Research and Development Clusters for Innovation-Led Economic Growth

In 2018, a team from RTI International and Duke University conducted SNA in El Salvador to assess network connections among university, private sector, and government collaborators in distinct economic sector clusters (energy, light manufacturing, information and communications technology [ICT], and agroindustry and food processing). These clusters had been formed as a part of a 5-year USAID program in El Salvador to advance economic growth through improved higher education performance. At the time in El Salvador, university faculty generally lacked knowledge of industry trends, had few connections with employers, and, thus, were largely unaware of in-demand competencies needed for student employment. Further, faculty engaged in little research to develop applied solutions to the challenges of private industry. On the private sector side, Salvadoran employers reported difficulty in finding new hires with appropriate technical and soft skills and had difficulty sourcing innovation.8

Figure 1. Common social network analysis measures

Network metric	Explanation		
Network Size	Network size refers to the number of nodes in the network. In this example there are two nodes. They are different colors to represent people from different groups of interest, like gender. The size of the node is typically related to how many connections it has.	AB	
Edges	Edges are connections among nodes. Edges can be directed, meaning we care if the connection goes both ways (i.e., is reciprocal), or undirected, where we only seek to understand if there is any connection between A and B.		A B Undirected
Density	Density indicates how close a graph is to being fully connected. $0 = no$ connections at all; $1 = all$ possible connections among nodes are made. In this example the solid lines represent connections and the dotted lines potential connections that are not realized. In this directed network there are 6 possible connections, with 4 connections made, or a density of 66%.	A B	
Reciprocity	Reciprocity is the fraction of all possible connections in which nodes are mutually connected. Using the same example as above, we see one reciprocal connection out of the three connected nodes, or 33%.	B	C
Transitivity	Transitivity is the proportion of all possible triads connected within a graph. Among the nodes A, B, C, D, there are four possible triads that can be formed. In this example, we see one connected triad with solid lines, or 25% transitivity.	A B C D	
Shortest Path	Shortest path is the path that represents the shortest distance between a given pair of nodes. To summarize how isolated a node is, we can take the average of a node's shortest paths to all other nodes in a network. For example, the average shortest path for A is 1.33 as it takes 1 edge to get to nodes B and D and two edges (in red) to get to node C, an average distance of 1.33 edges.	A	`
Degree Centralization	Degree centralization is a measure of how much "influential" nodes impact network structure. 1 = star shaped network with all connections flowing through one node. 0 = fully connected graph or a graph with no connections, respectively seen in the second examples here.	C D High Degree Centrality	B C C Low Degree Centrality
Modularity	Modularity is a measure of partitioning between groups in a network. Scores above 0 indicate more within-group ties than between group ties. Lower scores indicate greater mixing across groups. Scores of 0 indicate a mixing expected in a random graph.	High Modularity	Low Modularity

To catalyze mutually beneficial relationships and overcome these challenges, the USAID program promoted a cluster approach that brought academic, private, and government actors together to develop new curricula, collaborate in research and development, and provide new career pathways for students. Each cluster included anchor and affiliate universities, industry associations and affiliated businesses, and relevant government counterparts. These cluster members met regularly across a 5-year period and accessed grant funding and technical support. Ultimately, collaborations led to 30 new industry-aligned academic degree programs. Cluster collaborators also conducted twenty-six applied research and development programs promoting technology transfer from universities to industry.

Although the program outputs were impressive, program stakeholders wanted to know the structure and strength of the cluster networks and whether the collaborations might be sustained. One month after program closure, we sent an online survey to 120 participants in the program's four economic clusters. The survey presented respondents with a list of names from their cluster. For each name, respondents indicated their level of collaboration with that individual, if any. If collaboration was of a certain level, the respondent answered further questions about that individual. Receiving 80 responses (66% response rate), our research team used these data to assess (1) the importance of reported connections, (2) the level of prior collaborations, (3) the state of current collaborations, and (4) anticipated future collaborations.

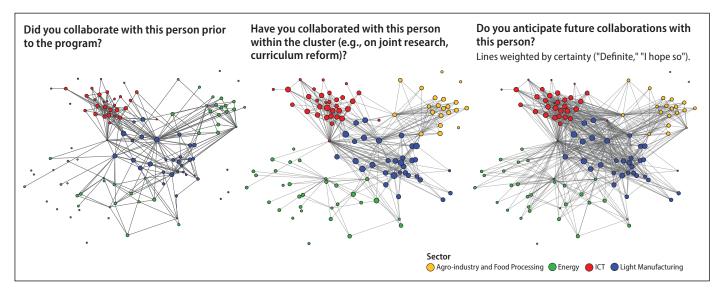
For each network, we produced several network measures (see Table 1 for the light manufacturing cluster, as an example). Network closure increased as respondents recalled past collaboration, reported on current collaborations, and projected future collaborations. For example, the network density in light manufacturing on the collaboration measure went from 7% (past collaboration), to 14% (current collaboration), to 22% (future collaboration). Reciprocity and average shortest path displayed similar patterns. Network heterogeneity decreased slightly when comparing collaboration before and after the program. The increase in modularity, combined with an increase in density, suggests that the program especially helped grow within sector collaboration, though overall connections across sectors also grew over the course of the program.

With these data, we created network maps to visualize changes in these networks over time, as seen in Figure 2. The colors represent different clusters, with the light manufacturing sector used in Table 1 in blue. Visually, we observe in the light manufacturing graph the increasing density of connections, and the presence of several star-shaped cliques where important nodes act to connect others in the network, as seen in the average shortest path score.

Table 1. Network scores for cluster connections in the lightmanufacturing sector

	Prior Collaborations	Current Collaborations	Future Collaborations
Density	7%	14%	22%
Reciprocity	10%	16%	12%
Average shortest path	2.2	1.8	1.6
Modularity	0.37	0.43	0.28

Figure 2. Network maps for prior, current, and future cluster collaborations in El Salvador



These increasing levels of network closure imply stronger communication channels, more-efficient collaboration, and improved ability to prioritize collective action.⁹ These networks also show moderate levels of heterogeneity, measured by modularity. Overall, these data indicate networks in El Salvador that have moderate organizational capacity and appear to be improving over time. The expectation is that these social networks will sustainably support the USAID program goals of improving the flow of information from the academic to the private sector, promoting technology transfer, and creating pathways for student internships and employment.

Indonesia: Knowledge System Connections for Better Evidence-Based Policy

Since 2012, RTI has directed a program in Indonesia called the Knowledge Sector Initiative (KSI), funded by the Australian Department of Foreign Affairs and Trade. KSI seeks to improve Indonesian public policy by strengthening systems that encourage the use of research and evidence in policy making.¹⁰ To do this, KSI works to strengthen and connect four parts of the evidence-to-policy system:

- Knowledge enablers: regulatory authorities, public and private funding bodies
- Knowledge producers: universities, research centers, think tanks
- Knowledge intermediaries: civil society, media
- Knowledge users: policy makers, parliamentarians, government ministries

KSI operates through policy working groups that combine knowledge enablers, producers, intermediaries, and users in a particular policy area to build a body of knowledge, promote its use, and develop policy based upon it. Indonesian stakeholders and the Australian funders of KSI have an interest in understanding the strength of these networks and their changes over time.

To contribute to that understanding, RTI researchers conducted SNA in 2020 in an area KSI has prioritized: aging and elderly policy. Improving aging and elderly policy, particularly the development of "Aging-Friendly Cities" in rapidly urbanizing environments, is a World Health Organization initiative active in Indonesia.¹¹ To conduct SNA in this area, we worked with an Indonesian research institute, SurveyMeter, to develop a list of 65 individuals (researchers, policy makers, journalists, and government funders of research) active in aging and elderly policy networks for each of three cities: Jakarta, Yogyakarta, and Denpasar (Bali region). To make the name rosters more manageable for survey respondents, the team randomly selected 30 names per city. Each respondent answered a set of questions about the 30-person roster: how frequently they communicated with each person, what level of collaboration they had on aging and elderly policy issues, and how much they trusted the person.

Table 2 lists the network data for the question on frequency of communication, and Figure 3 displays the network maps for these cities across the three questions.

Table 2. Network metrics for aging and elderly policynetworks in Indonesia

	Jakarta	Yogyakarta	Denpasar
Density	39%	17%	9%
Reciprocity	50%	32%	20%
Transitivity	48%	17%	14%
Degree Centralization	0.71	0.86	0.5
Modularity	-0.01	-0.01	-0.01
Average Distance	0.87	1.16	1.11

These networks, particularly that of Jakarta, show moderate levels of closure or cohesion (healthy levels of density, reciprocity, and transitivity; low distance between nodes; and high degree centrality). They also show high heterogeneity (i.e., low modularity), as seen visually in the between-group or multicolor nodal connections. This level of network closure and network heterogeneity is associated with strong policy network organizational capacity and could lead to efficient and effective policy formation.⁹ This SNA further reveals differences among locations and across the network connection questions.

Social Network Analysis and International Development: Key Recommendations

Within the international development field, the use of SNA to improve program design, program implementation, and program evaluation and learning is quite limited. USAID's Learning Laboratory and the World Bank's Independent Evaluation Group have both called for the increased use of SNA,^{1,12} and some recent projects have embraced these methods.¹³ Note that the examples in this brief were both conducted in relatively high-resource environments (urban areas of Indonesia and El Salvador) and among well-educated, survey-savvy respondents. Other SNA tools, such as <u>Net-Mapping</u> and <u>Collaboration Mapping</u>, can be used in lower resource environments, in situations where the network is not known, or where the network is not as bounded as those in the case examples presented here.

Figure 3. Networks maps for aging and elderly policy networks in Indonesia

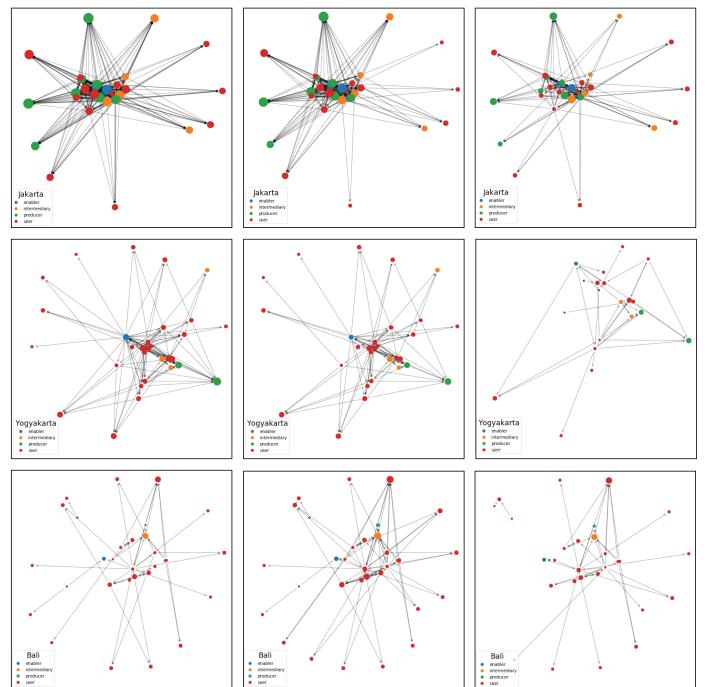
Level of Communication

In the last 24 months, how often did you communicate with the following people professionally on aging issues?

Level of Collaboration

In the last 24 months, I collaborated (worked together) with this person on aging issues.

Level of Trust I have a good level of trust in this person's work on aging issues.



The field of international development needs new tools and approaches for understanding the complex relational dynamics within which it operates and changes in the networks it seeks to affect. The field can apply SNA in many contexts, such as global health (disease spread), economic growth (market relationships), and governance (policy networks). With the right approaches and improved capacity, international development projects can conduct SNA more often to improve understanding of programs and to improve outcomes.

From experiences with SNA, including the cases presented here, we recommend the following steps for integrating SNA into international development. We present it as an agenda for SNA application in international development:

- Routinely incorporate SNA into in monitoring, evaluation, and learning processes. A project can conduct SNA at various points (i.e., baseline, middle, end) to inform program design, adaptive management, learning, and evaluation by considering network structure and network change over time. Mixing primary and secondary data and quantitative and qualitative data will augment understanding of SNA network data and visualizations. Whenever possible, SNA should be combined with other complementary analyses, such as political economy analysis. For example, SNA can add detail and texture to common political economy analyses of actors or groups with the most resources, or through whom resources are brokered. SNA can also uncover actors or groups that are relatively isolated from political-economic networks.
- Demystify the use of SNA among program participants, donors, and practitioners. Increased use of SNA tools and corresponding reports and publications should help bring the analytic approach into the mainstream. Concerted attempts to speak in plain language and display data in clear ways will help in these efforts. Toward this end, the Annex answers some frequently asked questions.
- Build capacity to conduct SNA. Many actors in international development lack the capacity to conduct and interpret SNA. At the stakeholder level in partner countries, international development researchers and practitioners should work to build stakeholder capacity to conduct and use SNA—this

could be within project teams or among other actors with the system, such as universities and research institutes. At the same time, donors, who will ultimately fund and approve the use of these methods, need SNA capacity too. We should build on efforts by some organizations to build capacity in the community, including available SNA 101 courses.^{14,15}

- Establish norms for data collection and identity protection. Data about individuals and their interactions with others are inherently sensitive. In standard research ethics protocols, such as human subject reviews, SNA practitioners must make decisions about how or whether to anonymize data when reporting it, especially in sociograms that can make individuals visible. For many maps, indication of group or organizational affiliation is sufficient, and respondents' names are not needed. In other cases, especially SNAs that seek to understand personalized networks, identifying network nodes by name may be unavoidable. In these cases, researchers should obtain explicit consent to disclose identity in advance.
- Build understanding of relationships between social networks and development outcomes. SNA will be useful only to the extent it helps users understand the relationship between networks and development outcomes that matter. Knowledge of interventions that work to improve networks is essential for meaningful use of SNA in the international development field.

Ultimately, researchers must thoughtfully consider and tailor SNA to fit the country and program environment. To be sure, there will be skeptics of these methods and concerns around the cost of the analysis, computing sophistication needed, ability of local teams to conduct SNAs, and relevance of the data to program design and evaluation. It is likely, however, that SNA can be done within reasonable costs, using open access software; combined with other ongoing analyses, such as political economy analyses or more-routine monitoring and evaluation; and simplified for use by project teams by focusing on the most important metrics or visuals.

References

- Baker M. Demystifying social network analysis in development: five key design considerations. 2019 Mar 26. In: USAID Learning Lab Blog [Internet]. Available from: https://usaidlearninglab.org/ lab-notes/demystifying-social-network-analysis-developmentfive-key-design-considerations
- Annen K. Social capital, inclusive networks, and economic performance. J Econ Behav Organ 2003;50(4):449–63. https://doi. org/10.1016/S0167-2681(02)00035-5
- Granovetter MS. The strength of weak ties. Am J Sociol 1973;78(6):1360–80. https://doi.org/10.1086/225469
- Putnam RD. Bowling alone: America's declining social capital. J Democracy 1995;6(1):65–78. https://doi.org/10.1353/ jod.1995.0002
- Popelier L. A scoping review on the current and potential use of social network analysis for evaluation purposes. Evaluation 2018;24(3):325–52. https://doi.org/10.1177/1356389018782219
- Firth JA, Hellewell J, Klepac P, Kissler S, Kucharski AJ, Spurgin LG; CMMID COVID-19 Working Group. Using a real-world network to model localized COVID-19 control strategies. Nat Med 2020;26(10):1616–22. https://doi.org/10.1038/s41591-020-1036-8
- Moody J. Race, school integration, and friendship segregation in America. Am J Sociol 2001;107(3):679–716. https://doi. org/10.1086/338954
- Navarro RA, Barbarasa E, Thakkar A. Addressing workforce needs by disrupting traditional industry–higher education relations: the case of El Salvador. Ind High Educ 2019;33(6):391– 402. https://doi.org/10.1177/0950422219875886
- Sandström A, Carlsson L. The performance of policy networks: the relation between network structure and network performance. Policy Stud J 2008;36(4):497–524. https://doi.org/10.1111/j.1541-0072.2008.00281.x
- Hertz JC, Brinkerhoff DW, Bush R, Karetji P. Knowledge systems: evidence to policy concepts in practice. Research Triangle Park (NC): RTI Press; 2020. https://doi.org/10.3768/rtipress.2020. pb.0024.2006

- 11. Suriastini W, Buffardi AL, Fauzan J. What prompts policy change? Comparative analyses of efforts to create age-friendly cities in 14 cities in Indonesia. J Aging Soc Policy 2019;31(3):250–70. https:// doi.org/10.1080/08959420.2019.1589889
- Vaessen J, Hamaguchi K. Understanding the role of the World Bank Group in a crowded institutional landscape. 2017 Nov 14. In: World Bank Independent Evaluation Group Blog [Internet]. Available from: https://ieg.worldbankgroup.org/blog/ understanding-world-bank-groups-role
- Pilliard-Hellwig R. Ma3an's social network analysis: a wellspring of CLA. 2020 Jun 30. In: USAID Learning Lab Blog [Internet]. Available from: https://usaidlearninglab.org/lab-notes/ ma3an%E2%80%99s-social-network-analysis-wellspring-cla
- Valente TW, Palinkas LA, Czaja S, Chu KH, Brown CH. Social network analysis for program implementation. PLoS One 2015;10(6):e0131712. https://doi.org/10.1371/journal. pone.0131712
- Laesecke A, de García D. Visualizing what connects us: social network analysis in M&E. 2017 Oct 17. [Internet]. Available from: https://www.irex.org/insight/visualizing-what-connects-us-socialnetwork-analysis-me
- Kang U, Papadimitriou S, Sun J, Tong H. Centralities in large networks: algorithms and observations. 2011. In: Proceedings of SIAM International Conference on Data Mining. https://epubs. siam.org/doi/abs/10.1137/1.9781611972818.11
- Perry BL, Pescosolido BA, Borgatti SP. Egocentric network analysis: foundations, methods, and models. Volume 44. Cambridge University Press; 2018. https://doi. org/10.1017/9781316443255
- Snijders TA, Baerveldt C. A multilevel network study of the effects of delinquent behavior on friendship evolution. J Math Sociol 2003;27(2–3):123–51. https://doi.org/10.1080/00222500305892

Annex. Frequently Asked Questions Concerning Social Network Analysis

What software programs are popular for SNA?

Many standalone software programs for SNA have a graphical user interface and do not require programming experience. Some popular options of this type include Gephi and Ucinet. Additionally, most statistical programming languages include packages, libraries, or modules for SNA. Currently, among opensource options, R and Python have particularly good SNA support. A popular choice for Microsoft Excel users is NodeXL, a feature-rich Excel add-in for SNA.

How computationally intensive is SNA?

SNA tends to be more computationally intensive than traditional statistical methods, as many of the core algorithms do no scale linearly with network size.¹⁶ For example, betweenness centrality requires computing the shortest paths between all pairs of actors, which exhibits cubic growth with network size. Creating efficient network algorithms and ways of calculating SNA metrics in parallel are active areas of research. However, for most network studies using survey responses, the networks are generally not large enough to create excessive computational burden. For example, we analyzed the case studies presented in this paper using a consumer-grade laptop. If researchers know they will be collecting a large amount of network data from administrative, bibliographic, or social media records, additional computational resources or careful choice of algorithms may be required to efficiently perform SNA.

Is there a minimum sample size required for SNA?

Data collection considerations are highly dependent on the type of SNA analysis planned. The most popular SNA design is the whole network or "network census" approach that assumes all members of a population of interest are represented in your data. In practice, defining the population and relevant network boundaries can be challenging, and the approach requires high response rates, unless additional assumptions can be made about the missing data. Other SNA designs have different sample considerations that reflect the types of data collected and research questions of interest. For example, egocentric network analysis has the actor as the unit of analysis and therefore has the same sample size considerations as traditional statistical analyses.¹⁷ Multiple network designs, which compare network metrics across many mutually exclusive networks (e.g., classrooms), have sample size considerations similar to multilevel regression modeling,¹⁸ where both individual- and group-level sample sizes are important. Conversations between stakeholders and social network researchers can often help clarify these requirements before data collection.

Where can I take a class on SNA?

The sociology departments of large universities typically offer SNA classes. Faculty in schools of social sciences, medicine, business, communication, or public health in dedicated centers, such as Duke University's Network Analysis Center, may also offer such classes. A good place to start is by looking at the course catalog of your local university or a directory of graduate programs from a professional society such as the International Network for Social Network Analysis. If in-class instruction is not available in your location, platforms such as Coursera or EdX offer several SNA massive online open courses (MOOCs).

About the Authors

Eric M. Johnson, PhD, is the research director of the Governance and Youth Economic Opportunities division.

Robert Chew, MS, is a research data scientist in the Center for Data Science at RTI

RTI Press Associate Editor

Barbara Biesecker

Acknowledgments

The authors acknowledge the collaboration of Jim Moody, Maria Cristina Ramos Flor, and Joe Quinn of Duke University, and Reina Duran of RTI on the El Salvador research, and Pingkan Umboh of Knowledge Sector Initiative, Wayan Suriastini of SurveyMETER, and Olga Khavjou, Alan O'Connor, and Lena Leonchuk of RTI on the Indonesia research. We thank Jana Hertz, Sarah Frazer, and Noah Flessel for their reviews

Cover Photo: Dewi Fortuna Anwar of the Indonesia Institute of Sciences listens to Allaster Cox, Charge d'Affaires of the Australian Embassy, during a KSI policy dialogue. Photo credit: Leander Gemilang Aswendro.

RTI Press Research Briefs and Policy Briefs are scholarly essays on policy, methods, or other topics relevant to RTI areas of research or technical focus.

RTI International, 3040 East Cornwallis Road, PO Box 12194 Research Triangle Park, NC 27709-2194 USA

+1.919.541.6000 rtipress@rti.org www.rti.org

© 2021 RTI International. RTI International is a trade name of Research Triangle Institute. RTI and the RTI logo are U.S. registered trademarks of Research Triangle Institute.



This work is distributed under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 license (CC BY-NC-ND), a copy of which is available at https://creativecommons.org/licenses/ by-nc-nd/4.0/legalcode

www.rti.org/rtipress