

**UNDERSTANDING SOCIAL CONNECTIONS IN COMMUNITIES: HOW  
TO USE SOCIAL NETWORK ANALYSIS  
GUIDELINES DOCUMENT**

**How to cite this reference:**

**Morris, S., A. Verville and L. Vasseur. (2014). Understanding social connections in communities: how to use social network analysis. Guidelines document. Community-University Research Alliance – Coastal Communities Challenges in face of climate change, 26 p.**

## Table of Contents

INTRO TO CURA PROJECT .....	3
INTRO TO SOCIAL NETWORK ANALYSIS.....	3
WHAT IS SOCIAL NETWORK ANALYSIS? .....	3
<i>What disciplines has it been used in?</i> .....	5
<i>What to look for in your social network: Measurements and key terms</i> .....	5
<i>Why do social networks matter to community groups?</i> .....	7
HOW CAN COMMUNITY GROUPS USE IT? .....	8
<i>What can social network analysis measure or uncover?</i> .....	8
<i>Whole network or ego network</i> .....	8
METHODOLOGICAL CONSIDERATIONS: A STEP-BY-STEP PROCESS .....	9
A STEP BY STEP PROCESS.....	9
STEP ONE: DEFINE YOUR OBJECTIVE .....	9
<i>Why are you measuring?</i> .....	9
STEP TWO: GATHER THE INFORMATION .....	10
<i>How do you collect data?</i> .....	10
STEP THREE: INPUTTING DATA .....	13
<i>How do you input data?</i> .....	13
STEP FOUR: VISUALIZING YOUR SOCIAL NETWORK AND SOFTWARE .....	13
<i>Visualizing social networks</i> .....	13
<i>Software: What is Gephi?</i> .....	13
STEP FIVE: INTERPRETATION .....	19
<i>What do these sociograms mean for me, and for my organisation?</i> .....	19
REFERENCES .....	21

## **INTRO TO CURA PROJECT**

Coastal communities have been known to be highly vulnerable to climate change, especially due to sea level rise and extreme events such as storm surges which lead to flooding and accelerated coastal erosion. Most of these communities however are ill equipped to understand the issues and find solutions or strategies to adapt to these changes. In the context of resilience, there is a need for these communities to integrate concepts of social-ecological systems and governance to move forward for a more sustainable territorial development. The main challenge for most of them is to acquire the tools to be able to move forward.

The CCC-CURA (Coastal Communities Challenges Community-University Research Alliance project, funded by the Social Sciences and Humanities Research Council of Canada (SSHRC)) was initiated in 2010 to:

- 1) understand the complexity and uncertainties surrounding social-ecological systems and environmental phenomena by considering both scientific expert knowledge and local existing knowledge,
- 2) strengthen the governance of communities through participative action research and the use of interactive tools, and
- 3) improve the resilience (ability to recover after an extreme event) of individuals and communities by integrating new developed knowledge and tools.

The ultimate goal is to help improve resilience for long-term sustainability of these coastal communities. From the start, the CCC-CURA project was influenced by the December 2010 storm, which affected parts of Atlantic Canada, including the provinces of Quebec, New Brunswick and Prince Edward Island.

In this project, it was proposed to use a longitudinal multisite approach (Yin, 2003) in which half of the communities (i.e. five) involved in this project would come from those affected by the storm while the other half would come from communities that had not been affected. One of the main approaches used in this project was Participatory Action Research (PAR) in which, over the past three years, in each of these communities, researchers have acted as moderators and facilitators to fill the "adaptation deficit" of local actors (public and economic sector, civil society). At the same time, as one of the main goal of the CCC-CURA, tools were developed to help communities and non-governmental groups to acquire better skills and understanding on how to work with the actors in communities to build adaptation strategies and improve resilience. The current guideline was produced with this goal in mind. In this case, organizations wanted to have a tool to help them map the connections among organizations, institutions, governmental agencies, etc. and how through these networks it is possible to improve the participation of actors in building resilience. This guideline on social network analysis (SNA) is meant to be used to inform organizations and support their activities with the additional tools already available or developed by CCC-CURA.

## **INTRO TO SOCIAL NETWORK ANALYSIS**

### **WHAT IS SOCIAL NETWORK ANALYSIS?**

A social network represents the linkages or connections found among people. An example of a social network can be your family or the people you interact with at work. Facebook is a good example of how networks can become very complex.

When you hear the term "social network analysis", the first thing that may stick out in your mind is "social network", and you may be visualising the structure of a group of people. An important part of SNA is uncovering not only who is part of the social network, but also why they are part of the network. Are they in your network because they are a family member or a friend; perhaps a co-worker; or maybe someone that you only interacted with only once? The relationship that you have with each of these

people will be different, and the importance of your relationship with them will depend on a number of things. SNA looks at the social structure of a group of people based on their relationships, behaviour and/or actions. See Box 1 for important terms to know before moving ahead.

Often social networks are built around a specific goal or common interest. This is the case for people working in the same community, such as healthcare, natural resources, or business. The Canadian Water Network is a great example of a social network where the members are connected in order to promote similar values, in this case, water management research in Canada.

[Box 1: **Important Terms**

**Social Network** – “a set of actors and the relations (such as friendship, communication, advice) that connect them” (Kilduff and Tsai, 2003, p. 135)

**Actors** – An actor is an individual, or organisation. Actors are also referred to as **nodes, alters** or **egos**

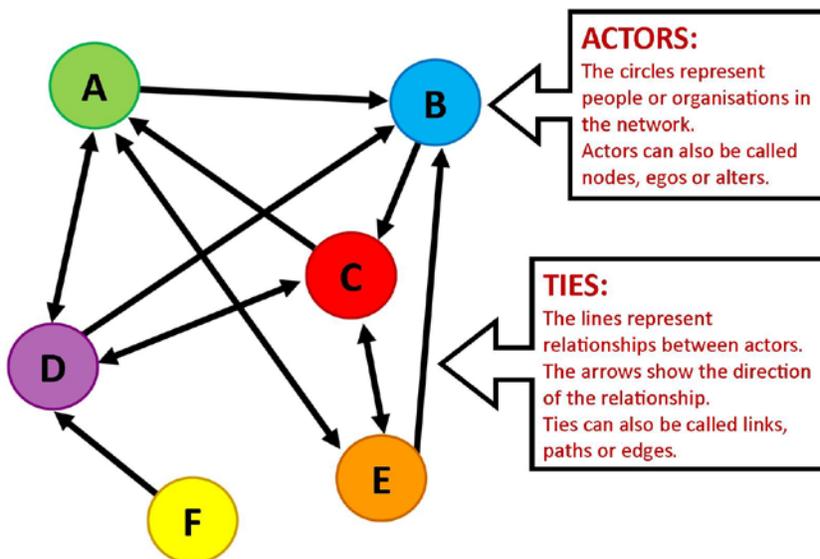
**Ties** – A tie is also known as a relationship. You can think of it as something that ties two individuals (actors) together. Ties are also referred to as **links, paths, or edges**

**Social Network Analysis (SNA)** - “the mapping and measuring of relationships and flows between people, groups, organizations, computers, web sites and other information/knowledge processing entities. The nodes in the network are the people and groups while the links show relationships or flows between the nodes. Social network analysis provides both a visual and a mathematical analysis of human relationships” (Divjak and Peharda, 2010, p.69)

**Formal vs. Informal Networks** - Social networks can be formal, meaning that you know all of the actors in the network (ex. all of the workers in a company), or informal, meaning that you don’t necessarily know who all of the actors are in the network (ex. all of the people in a town who like to eat fish)]

Sociograms are graphs that visualise a social network (also known as social network diagrams). Figure 1 is a visual example of a sociogram. The circles represent different actors in the network, while the lines with arrows represent the ties between them.

**Figure 1:** An example of a sociogram



SNA can be a very useful tool for many individuals, organisations, and groups if used properly. The aim of this guideline document is to help you understand how SNA can help you as a water or environmental management professional, and understand how you can apply it to your situation. First, it's important to explore SNA a little more.

***What disciplines has it been used in?***

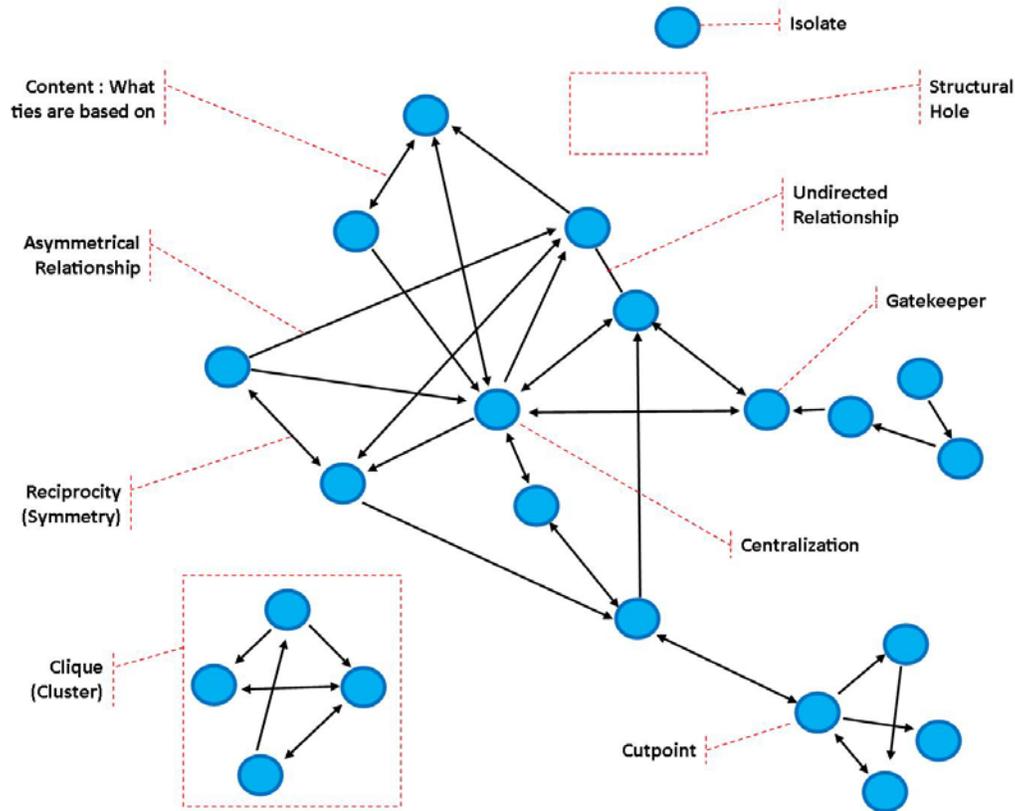
SNA has been used to study all sorts of topics in a wide variety of disciplines. Here are some examples:

<i>Disciplines:</i>	<i>Topics:</i>
Sociology	Spread of Illness
Education	Terrorist Networks
Social Psychology	Diffusion of Ideas
Anthropology	Organisation Design
Politics	Knowledge Management
Communications	Industry Ecosystem Analysis
Organisation Studies	Discovering Key Opinion Leaders
Natural Resource Management	Network Vulnerability Assessment
Biology	Discovering Communities
Security	Mapping and Measuring Information
Physics	
Linguistics	
Psychotherapy	
Public Health	
Epidemiology	

***What to look for in your social network: Measurements and key terms***

The first step for any organisation interested in analysing its social network is to answer the question: why do we want to do a SNA? A SNA can uncover many different things, so before you start your SNA it is important to know what you can measure or look for. Figure 2 visualizes some of the possible relationships, measurements, or actors that you may want to analyse. Each of these is explained further on the following page.

**Figure 2.** What to look for in your SNA



*Relationships:*

- Cliques (clusters) – this occurs when a group of actors only has direct ties within the group and no connection outside of this group (Kilduff and Tsai, 2003)
- Content – what a relationship is based on, for example exchanging information, resources, etc. (Haythornthwaite, 1996)
- Asymmetrical relationships – the relationship between two actors flows only one way (Haythornthwaite, 1996)
- Reciprocity/symmetry – this measure looks at the nature of a relationship between actors, mainly if the relationship flows both ways or is reciprocated (Kilduff and Tsai, 2003)
- Undirected relationships – the relationship between two actors either is not measured or is irrelevant (ex. Joint membership on a board) (Haythornthwaite, 1996)
- Homophily – the tendency of actors to network with actors that have similar backgrounds, behaviours and opinions (Kilduff and Tsai, 2003)
- Strength – measures “the intensity of a relationship” (Haythornthwaite, 1996, p. 327)
- Tie strength – Similar to “strength”, it measures the intensity of a relationship between two actors. It can help to determine how important an actor is in a network (Haythornthwaite, 1996)
- Range – the resources that an actor can access through the network (Haythornthwaite, 1996)
- Reachability – the path between two actors determines how “reachable” they are directly to one another or through another actor (Kilduff and Tsai, 2003)

*Actors:*

- Gatekeeper – an actor in a network that communicates information into the network from an external source (Kilduff and Tsai, 2003)

- Prominence – this measure identifies actors that hold power or influence in the network
- Bridging organisation – an organisation that strategically links two or more organisations together (Crona and Parker, 2012)
- Isolate – an actor in a network that does not have any connections to any other actor(s) in the network, or a degree of 0 (Haythornthwaite, 1996)
- Cutpoint – an actor in a network “whose removal from the network results in subsets of actors between whom there is no connection” (Kilduff and Tsai, 2003, p.133)

*Measurements:*

- Cohesion – this measure applies to the whole network and can identify strong relationships between actors in the network, and the level of access network members have to information/resources (Haythornthwaite, 1996)
- Density – this measure shows the degree that actors in the network are connected to all other members in the network (Haythornthwaite, 1996)
- Centralization – this measure looks at how actors are “organized around one central point” in the network (Haythornthwaite, 1996, p. 333)
  - Global centrality/closeness – this is a similar measure to centralization, however this measures the shortest path length between an actor and every other individual in the whole network. The actor that has the lowest total path length in the network has the highest global centrality (Haythornthwaite, 1996)
- Embeddedness – this measure is defined as “a situation of very intense interrelationship that influences social actions, creates opportunities and constraints on actors’ behaviour” (Maggetti et al., 2013, p.7)
- Structural equivalence – this measure “identifies actors with similar roles” (Haythornthwaite, 1996, p.334)
- Betweenness – this measure identifies how likely it is that an actor is the most direct route between two other actors within the network (Butts, 2008)
- Degree – this measure identifies the actor with the most connections. The higher the number of connections, the higher the actor’s degree. The actor with the highest degree is known to be most central in the network (Haythornthwaite, 1996)
- Brokerage – this measure identifies areas in the network where there are gaps between actors and opportunity for an actor to become a “broker” (an actor that acts as an intermediary between actors or groups of actors and therefore gains control of the information being passed through them). It can be measured by using betweenness which will show the extent that an actor sits between other actors in that network (Haythornthwaite, 1996)
  - Structural holes – this is what you call the gaps left between actors or groups of actors that haven’t been filled by a broker (Haythornthwaite, 1996)

*Why do social networks matter to community groups?*

SNA can benefit community groups in many ways. It can:

- help identify key people in the network
- improve organisational structure, productivity, and planning
- identify informal networks
- uncover which groups are more connected or less connected (even isolated)
- improve collaboration and in many cases, communication
  - generate, gain and diffuse different types of knowledge
  - mobilise and allocate key resources
  - resolve conflicts
  - agree on common rules

“...by working together, community organizations can draw on the broad range of resources and expertise provided by the other organizations in the network, and, as a result, the health and well-being of

community members will be improved (Chisholm 1998; Provan and Milward 2001).” (Provan et al., 2005, p. 603)

In planning for a SNA, when defining your objectives and the types of measurements you want to use, keep in mind what benefits you may gain from doing so. It is worth writing them down and discussing them as a group.

## **HOW CAN COMMUNITY GROUPS USE IT?**

### *What can social network analysis measure or uncover?*

Now that we’ve discussed why SNA is important for community groups; it’s time to explain how your group can use it. SNA can help you to:

- Locate and develop social networks
- Develop a collaborative network structure
- Share authority
- Increase social memory
- Enhance information flows
- Increase collaboration across scales
- Communication
- Visualise and understand the governance structures, restructuring them as necessary
- Understand how resilient the network is
- See how the network changes over time – this is also known as a longitudinal study
- Specific SNA measures such as cohesion, density, centralization, embeddedness, structural equivalence, degree, brokerage, structural holes, and reciprocity/symmetry

In addition to what it can measure or uncover, an outcome of a social network can be increased collaboration.

### *Whole network or ego network*

There are three approaches that you can take to look at a SNA:

**1. Whole Network Approach** – This means that before you start the SNA, you know who all the actors are. Usually it also means that you can easily collect data from all of them. An example of a whole network is an office where it is easy to get a list of all of the people who work there. In this case, the network relationships would be work-focused. The network data would come from everyone in the network.

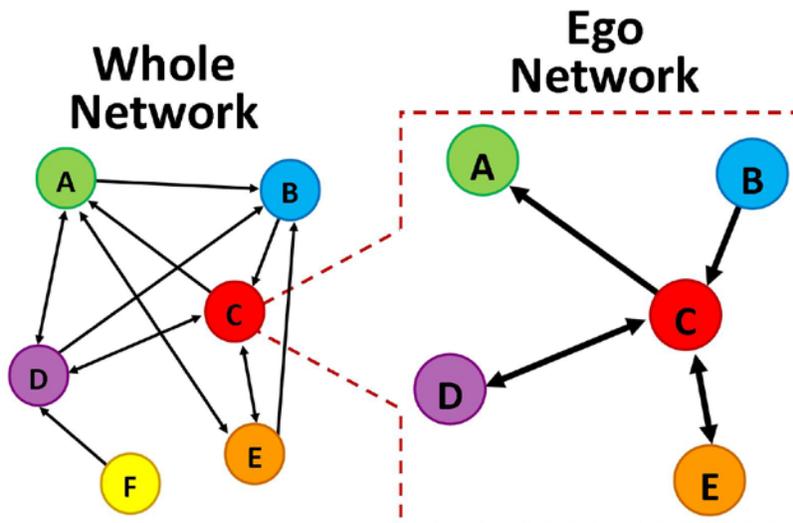
**2. Ego Network Approach** – This approach looks at the network of an individual. Individual ego-networks can be combined to gain a better understanding of the bigger network picture. Ego networks are especially helpful when you don’t know who all of the actors in the network are before you begin. An example of an ego network is an individual’s group of friends. In this case, the network relationships would be based on some sort of measurement of friendship (i.e. how often you talk to each other, how many years you have been friends, etc.). All of the network data would only come from one person.

**3. Archival Approach** – This approach is used when you have historical documents that clearly indicate relationships amongst actors. An example of an archival network could be someone mapping their family tree (or genealogy) from historical records or from a place such as ancestry.com. All of the network data would come from documents and/or websites.

Figure 3 visualizes an ego network within a whole network.

By collecting data from as many people as you can in the network/community, you can gain a more accurate representation of what relationships are actually occurring or are perceived to be occurring.

**Figure 3:** Ego network within a whole network. The ego network shown in this figure is the ego network of actor “C”.



### **METHODOLOGICAL CONSIDERATIONS: A STEP-BY-STEP PROCESS**

The aim of this guide is not only to provide a brief introduction to SNA, but to also provide you and your organisation with a tool to help you complete your own SNA.

We recommend that you follow these five steps to complete a successful SNA:

1. Define your objective and the type of network analysis you want to do
2. Gather the information and survey design
3. Input data
4. Integrate into the software – Visualizing your network
5. Interpret your results

Each of these steps is outlined further in the following section.

### **A STEP BY STEP PROCESS**

#### **STEP ONE: DEFINE YOUR OBJECTIVE**

##### *Why are you measuring?*

The first step in completing a SNA is to ask: why do I (we) want to do a SNA? As previously mentioned, if you are part of an organization, you may want to take a few minutes to write down why you want to do this, and what benefits you expect to gain from this exercise.

Typically social network analyses are used to:

- Visualise who is involved
- Uncover individuals/organisations that are involved that you might not have been previously aware of
- Learn who is talking to whom and about what

- Learn who is providing or receiving resources
- Learn who is collaborating within the network
- See if there are individuals/organisations that are not connected to the rest of the network, and determine why that is the case
  - Additionally, you may want to create an action plan to get them involved
- See if there are any individuals/organisations that act as a “bridging organisation”
- And more!

### *What are you measuring?*

Once you have determined why you are doing a SNA, it is important to decide what you want to learn from it. Some common things that are looked for using SNA include:

- Information flows/sharing
- Funding
- Decision-making
- Collaboration
- Gatekeepers
- Isolates
- Bridging organisations
- Clusters
- Evolution of the network

There are many types of relationships that can be measured and visualised using SNA. It is important that you have a clear idea of what you want to measure before you begin creating a tool to collect your data.

You can also collect data on attributes (such as gender, and age), interactions, etc. if it will be useful for you or your organisation. Attribute data simply refer to anything that describes an actor/organisation. It's important to include any attribute questions when designing your survey. The attributes section will be separate from your SNA section, and may help you when analysing and understanding your social network.

- Attribute questions can often be answered by providing multiple choice responses. Using multiple choice responses will make it easier for you when inputting the data, and will limit the variety of responses you can receive.
- **Tip:** There are a large number of questions that you may want to include. Asking too many questions might be overwhelming not only to the person you are interviewing, but also to you when you are analysing the data. Make sure to only include what you need.

## **STEP TWO: GATHER THE INFORMATION**

### *How do you collect data?*

There are several ways in which you can collect data for your SNA including: surveys/questionnaires, focus groups, workshops, planning sessions, and archival searches.

The most common methods used for data collection in SNA are surveys or questionnaires (Butts, 2008). This is because you can ask individual(s) representing an organisation very specific questions that describe their social network and relationships within their social network. Answers can be confirmed (relations) by surveying the other individuals in their network. Another approach is to develop your network analysis through exchange during meetings or during a workshop. This approach is more

interactive and requires a certain control from the organizers to ensure that all the information is captured either on flipcharts or sticky notes.

If you are taking a whole network approach, it is important that you collect data from every member in the network so that you are showing the true relationships occurring in the network. If you are taking an ego network approach, it is important that you identify a few important individuals in the network you wish to study. An approach you can take to retrieving more information about the whole network is to use a “snowball sample” method. If you are unsure of what the “snowball sample” method is, read Box 2.

**[Box 2: Snowball Sampling Approach**

Think of this approach the same way that you think of a snowball: the longer you roll it, the bigger it gets. This approach requires you to ask someone to identify people in their network based on a specific relation (ex. Exchange information with \_\_[name of person]\_\_ about fisheries). Once they have answered your questions, you can contact the individuals they identified to participate in your survey as well. NOTE: Don't forget to ask for contact information for the people identified!].

If you would like to see how your network has changed over time, it is important to collect your data at multiple points in time. You would use the same questions and ask the same people to respond to them. This is referred to as a longitudinal study.

By including a social network matrix (table) in your survey, you will make analysing the social network data much easier. By already having the data in a matrix (table), transferring it into your software will be a breeze (SNA software love matrices).

For example, you could ask:

1. Who has your organisation collaborated with on water quality projects?
  - a. Did your organisation initiate the collaboration? (y/n)
2. Who has your organisation given financial resources to for research?
3. Who has your organisation received financial resources from for research?
4. Who has your organisation given information to about cats?
5. Who has given your organisation information about cats?
6. Who has your organisation communicated with about saving forests?
  - a. How often?
7. Who has communicated with your organisation about saving forests?
  - a. How often?

Using these questions, you could set up your matrix (table) like this:

Name of Individual/ Organisation	1	1a	2	3	4	5	6	6a	7	7a
<i>Here you would list all of the actors the interviewee names</i>	Here you would check which questions apply to the interviewee for each of the actors they list						<input type="checkbox"/> Daily <input type="checkbox"/> Weekly <input type="checkbox"/> Monthly <input type="checkbox"/> Every 2 months <input type="checkbox"/> Yearly	<input type="checkbox"/> Daily <input type="checkbox"/> Weekly <input type="checkbox"/> Monthly <input type="checkbox"/> Every 2 months <input type="checkbox"/> Yearly		

Jane			x	x	x	x	x	<input type="checkbox"/> Daily <input type="checkbox"/> Weekly <input checked="" type="checkbox"/> Monthly <input type="checkbox"/> Every 2 months <input type="checkbox"/> Yearly	x	<input type="checkbox"/> Daily <input checked="" type="checkbox"/> Weekly <input type="checkbox"/> Monthly <input type="checkbox"/> Every 2 months <input type="checkbox"/> Yearly
WaterH2O	x	n	x					<input type="checkbox"/> Daily <input type="checkbox"/> Weekly <input type="checkbox"/> Monthly <input type="checkbox"/> Every 2 months <input type="checkbox"/> Yearly		<input type="checkbox"/> Daily <input type="checkbox"/> Weekly <input type="checkbox"/> Monthly <input type="checkbox"/> Every 2 months <input type="checkbox"/> Yearly
Bob	x	y		x	x	x		<input type="checkbox"/> Daily <input type="checkbox"/> Weekly <input type="checkbox"/> Monthly <input type="checkbox"/> Every 2 months <input type="checkbox"/> Yearly	x	<input type="checkbox"/> Daily <input type="checkbox"/> Weekly <input type="checkbox"/> Monthly <input type="checkbox"/> Every 2 months <input checked="" type="checkbox"/> Yearly

Or you could set up your matrix (table) like this:

Name of Individual/ Organisation	Collaborate with (initiate)	Give financial resources	Receive financial resources	Give information to	Get information from	Communicate to (Daily =1, Weekly =2, Monthly = 3, Every two months =4, yearly = 5)	Communicate from (Daily =1, Weekly =2, Monthly = 3, Every two months =4, yearly = 5)	Relationship quality
	Check if applicable							Circle one
Jane		x	x	x	x	X (3)	X (2)	12345
WaterH2O	x (n)		x					12345
Bob	x (y)	x		x	x		X (5)	12345

**Tip:** If you know everyone in the network, you can fill in the first column and add a section for “other organisations/individuals” in case you might have missed an actor or they’ve expanded the network. This will help to speed up the survey process.

#### Attribute Questions

Here are some examples of attribute questions you can ask the interviewee:

- What is the name of the organisation you are representing?
- What type of organisation is it?
  - What sector does the organisation fall into?
- What is your current position at your organisation?
- How many years have you been working at your organisation?
- How many employees does your organisation have?
- What is the main goal of your organisation?
- What are the main responsibilities of your organisation?

- What level of education do you hold?
- Do you support a specific political party?
- What is your favourite colour?

Here are some examples of attribute questions you can ask the interviewee to answer about each of the actors they identify:

- What type of organisation is it?
- How long has your organisation been in contact with this organisation?
- Do you report to this organisation?
- Who should I contact at this organisation? What is their contact information?

#### *Templates/ Examples*

Here are some examples of SNA surveys available online:

- University of Canberra: [http://www.durantlaw.info/sites/durantlaw.info/files/SNA\\_Survey.pdf](http://www.durantlaw.info/sites/durantlaw.info/files/SNA_Survey.pdf)
- Many different researchers: <http://socialnetworks.soci.ubc.ca/SocNets/Surveys/Surveys.html>

### **STEP THREE: INPUTTING DATA**

#### *How do you input data?*

Typically, data collected from the survey will need to be placed into a spreadsheet so you can easily transfer your results into SNA software. The way you input your data and layout your spreadsheet will depend on what software you use to analyse and visualise your social network. Although some SNA software allows you to input your data directly into their interface, it is easier to organise your data in a spreadsheet and import it.

There is several different spreadsheet software available such as, Microsoft Excel, Numbers, OpenOffice Calc.

### **STEP FOUR: VISUALIZING YOUR SOCIAL NETWORK AND SOFTWARE**

#### *Visualizing social networks*

As mentioned earlier, one way to display your results from your SNA is to create a sociogram. Previously, someone who wanted to do a SNA would use many complex mathematical equations to calculate different measurements. Instead of drawing and calculating relationships by hand, there are now several software packages available to analyse social network data. This makes SNA much easier!

After a search for the most appropriate software for this project, we came to the conclusion that your organisation would benefit most from using Gephi. The rest of this guideline document will focus on introducing you to Gephi and help you start the analysis part of your SNA.

#### *Software: What is Gephi?*

Gephi is open source software for network visualization and analysis created in 2008 ([www.gephi.org](http://www.gephi.org)). It is available to download on PC, Mac or Linux. It is an interactive program that helps individuals and organisations analyse and visualise networks. Visit <http://vimeo.com/9726202> to watch a short video that introduces you to Gephi and some of its features.

#### *Main functions*

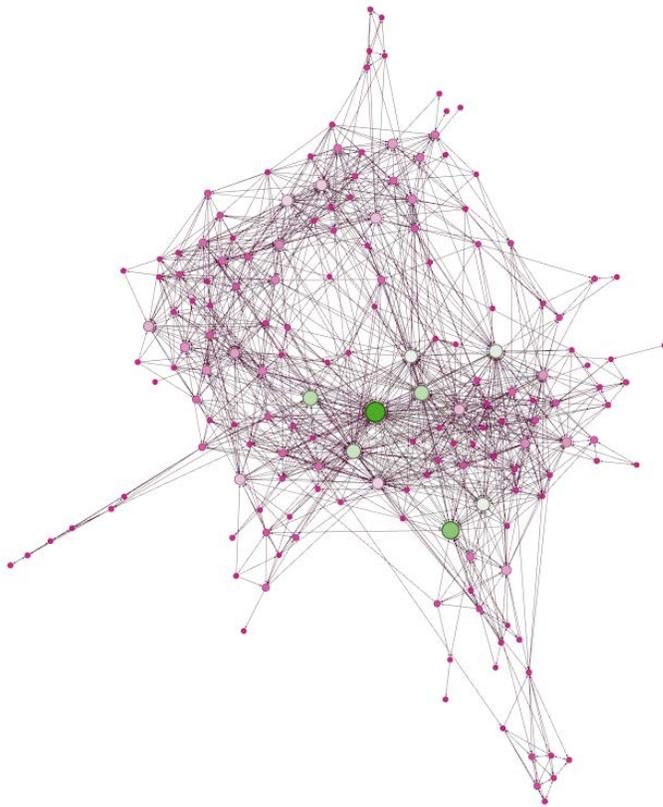
The main functions of Gephi include:

- It is capable of handling up to 50,000 nodes (actors), and 1,000,000 edges (ties).
- It is available in several languages including: English, French, Japanese, Portuguese, Brazilian, Russian, Chinese, and Czech.

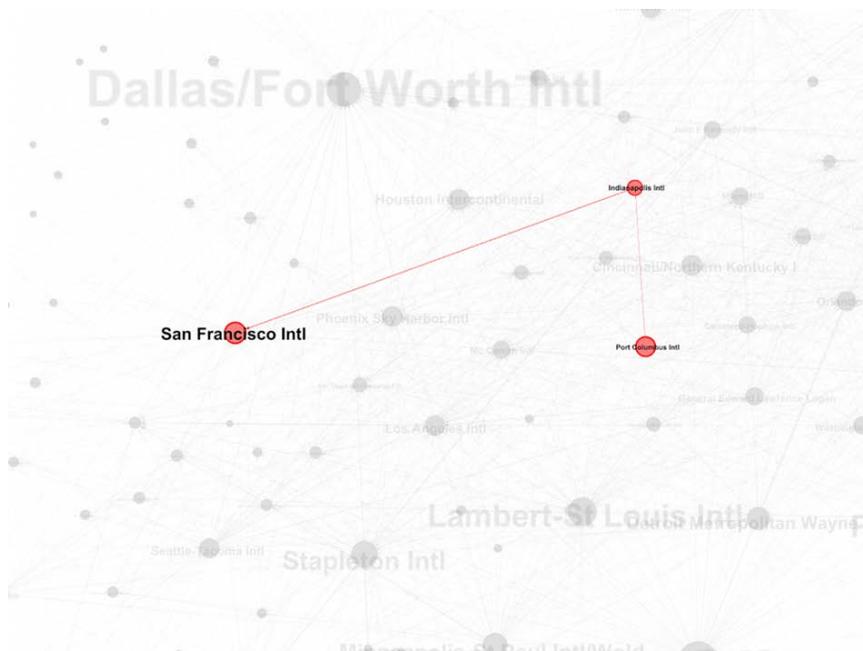
- It makes visualizing social networks **very** easy. It has many features so social networks can be visually represented in a variety of ways. For example, background colours can be changed, labels can easily be placed on nodes, social networks can be overlaid onto maps, and much more.
- It has plugins available to expand its functionality (<https://marketplace.gephi.org/>). Plugin categories include: Clusters, Data Laboratory, Examples, Exports, Filter, Generator, Imports, Layouts, Metrics, Preview, Ranking, and Tools. These have the potential to be very useful; you can even use a plugin to add a map to your social network!
- The SNA features include:
  - Metrics: betweenness, closeness, diameter, clustering coefficient, average shortest path, PageRank, and HITS
  - Community detection (modularity)
  - Random generators
  - Dynamic network analysis where you are able to import temporal graphs, and stream graphs
  - Create cartography
  - Explore multi-level graphs (expand and contract sub graphs, link and attribute clustering)
  - Dynamic filters without scripting. Build new networks from the filtering results and save your favourite queries

*Examples of how Gephi is used and what it's been used for*

This is one of many examples of a sociogram designed in Gephi (image from <https://gephi.org/features/>).



In the image below you can see the shortest path between two actors in the social network (image from <https://gephi.org/features/>).



As discussed earlier, Gephi has many more features available for you to use in your SNA. These are just two examples of sociograms that have been created using Gephi.

### *Gephi in research*

Here are some examples of research that Gephi has been used in:

- E.J.L. Chappin, A. Ligetvoet: [The difference between transition and transformation: a bibliometric analysis of two scientific networks](#)
- M. Zachariadis et al.: [Leadership of healthcare commissioning networks in England: a mixed-methods study on clinical commissioning groups](#)
- J.L. Ortega et al.: [Institutional and country collaboration in an online service of scientific profiles: Google Scholar Citations](#), Journal of Informetrics, Volume 7, Issue 2, April 2013, Pages 394–403
- A. Maireder, J. Ausserhofer, A. Kittenberger: [Mapping the Austrian Political Twittersphere](#), Proceedings of CeDem12 Conference for E-Democracy and Open Government. Krems: Danube University, pg. 151-164
- B. Rieder: [The refraction chamber: Twitter as sphere and network](#), First Monday, Volume 17, Number 11 - 5 November 2012
- N. Martínez Arqué, D. Nettleton: [Analysis of On-line Social Networks Represented as Graphs – Extraction of an Approximation of Community Structure Using Sampling](#), MDAI 2012, LNAI 7647, pp. 149–160, 2012.
- J. Kimmerle et al.: [Knowledge construction in an outsider community: Extending the communities of practice concept](#), Computers in Human Behavior, Available online 6 November 2012
- R. Nagarajana, A. T. Kalinkab, W. R. Hogan: [Evidence of community structure in Biomedical Research Grant Collaborations](#), Journal of Biomedical Informatics, Available online 6 September 2012
- E. Mustafaraj, P. Metaxas, S. Finn, A. Monroy-Hernandez: [Hiding in Plain Sight: A Tale of Trust and Mistrust inside a Community of Citizen Reporters](#)
- D. Hristova, A. Mashhadi, G. Quattrone, L. Capra: [Mapping Community Engagement with Urban Crowd-Sourcing](#), WCMWC12

### How to download and resources

You can download Gephi here: <https://gephi.org/users/download/>

- Use the quick start guide to get yourself more familiar with the software: <https://gephi.org/users/quick-start/>

They have many resources available to help you, including:

- A Wiki page to help you not only get started with Gephi, but to also give instructions, support and advice on the software ([http://wiki.gephi.org/index.php/Main\\_Page](http://wiki.gephi.org/index.php/Main_Page))
- A page on the Gephi site dedicated to help get you started (<https://gephi.org/users/>)
- A forum with information on software support, and questions sent in by users (<https://forum.gephi.org/>)
- The Gephi Consortium (<https://consortium.gephi.org/>)
- The book titled “Creating Networks Graphs with Gephi” which gives an overview of Gephi from the beginning installation to deploying visualisations in Gephi (<http://www.packtpub.com/article/creating-network-graphs-with-gephi>)
- And more (<https://gephi.org/users/support/>)

We recommend downloading Gephi and practising with their sample projects (ex. Les Miserables.gexf).

There are three main workspaces in Gephi:

- The “Overview” space: where you can see the sociogram (the social network visualised).
- The “Data Laboratory”: where you can view and change the data imported for this social network.
- The “Preview” space: allows you to adjust how you want to visualise your network.

### How to input your social network data into Gephi

Gephi requires two spreadsheets with social network data:

1. *Node spreadsheet* – This spreadsheet tells the software who all of the actors are in your social network

ID	Label	Attribute(s)	Social Network Measure(s)	Time Interval
1	John	Male	2.341	1, 2001, 2012
2	Jane	Female	4.121	2, 2005, 2012

- ID - This needs to be unique number for each individual the label (actor). Using numbers helps Gephi know the difference between all of the people/organisations included in your social network.
- Label – Otherwise known as a name or the actor represented in your SNA (can be a company or organisation name as well). The names you place in this column will be the labels that appear on your sociogram.
- *Optional Columns:*
  - Attribute(s) – You can put multiple columns of attribute data. Attributes can include things like gender, age, race, type of organisation, how many employees, and more. Attributes aren’t required for analysing social network data, but you may have some use for specific attributes.
  - Social Network Measure(s) – This may be for more advanced analysts who know already how to calculate network measurements. If you have already calculated some social network measures such as centrality, you can enter it into one or more of these columns.

- Time Interval – You can use this column if you would like to analyse your social network over time. This is especially useful when you know that interactions between certain groups occurred only during a project’s lifetime. The format for this column will need to look like this:
  - [ID, starting year, ending year]

2. *Edge spreadsheet* – This spreadsheet tells the software about the relationships between the actors in your social network.

Source	Target	Type	ID	Label	Weight	Attribute(s)	Time Interval
2	1			Information	1	Co-worker	
3	7			Funding	10	Watershed organisation	

- Source – The ID of the actor that the relationship starts at
- Target – The ID of the actor that the relationship ends at
- *Optional Columns:*
  - Type – whether the relationship is directed (i.e. that source 2 contacts target 1) or undirected (i.e. that source 2 knows target 1). This column will be auto-filled by Gephi when you upload your data.
  - ID – Gephi will assign a unique ID for the relationship.
  - Label – You can use this column to categorise what the relationship type is.
  - Attributes – Similar to the node table, however you would assign attributes for the relationship (ex. family, friend, boss, co-board member, etc.).
  - Weight – You can assign a higher weight to relationships that you believe are more important. This should be first decided upon by your group to make sure that everyone understands the assigned weights.
    - Examples of how you can use weights: how often two actors interact, how strong an actor thinks their relationship is with another actor, etc.
  - Time Interval – You can also assign a time interval to the relationship. You would use the same format used in the Node table:
    - [ID, starting year, ending year]

**Example:**

John interviewed someone from the WaterH2O organisation. He asked them who they provide research funding to (if applicable), and who they receive research funding from (if applicable). WaterH2O told him that they provide funding to the following organisations: Water99 (\$40,000), I Love Water (\$20,000), and River Water (\$15,000). They also told him that they receive funding from the following organisations: Provincial government (\$100,000), Research Council (\$50,000). This is how John would set up his spreadsheets in Gephi.

*Node spreadsheet:*

ID	Label	Number of Employees (attribute)	Organisation type (attribute)
1	WaterH2O	100	Watershed Organisation
2	Water99	80	Watershed Organisation
3	I Love Water	45	Watershed Organisation
4	River Water	20	Watershed Organisation
5	Provincial Government	1000	Government

6	Research Council	40	Research Organisation
---	------------------	----	-----------------------

*Edge spreadsheet:*

Source	Target	Label	How much funding (\$)
1	2	Research Funding	40000
1	3	Research Funding	20000
1	4	Research Funding	15000
5	1	Research Funding	100000
6	1	Research Funding	50000

Your survey can have several different relationships that happen in the social network.

**Continuing on our example:** John also asked WaterH2O whom they collaborate with on research projects. WaterH2O said: Water99, River Water, CleanWater, and BlueWater. John would then add to his spreadsheets he had already started:

*Node spreadsheet:*

ID	Label	Number of Employees (attribute)	Organisation type (attribute)
1	WaterH2O	100	Watershed Organisation
2	Water99	80	Watershed Organisation
3	I Love Water	45	Watershed Organisation
4	River Water	20	Watershed Organisation
5	Provincial Government	1000	Government
6	Research Council	40	Research Organisation
7	CleanWater		NGO
8	BlueWater		NGO

*Edge spreadsheet:*

- This one is a little different because WaterH2O collaborates with the other organisations, meaning that the relationship flows both ways. This means that you will have to input WaterH2O as the source and the target for each of the collaborations.

Source	Target	Label	How much finding (\$)
1	2	Research Funding	40000
1	3	Research Funding	20000
1	4	Research Funding	15000
5	1	Research Funding	100000
6	1	Research Funding	50000
1	2	Collaboration on Project	
2	1	Collaboration on Project	
1	4	Collaboration on Project	
4	1	Collaboration on Project	
1	7	Collaboration on Project	
7	1	Collaboration on Project	
1	8	Collaboration on Project	

8	1	Collaboration on Project	
---	---	--------------------------	--

We recommend that you input your data into the spreadsheets after each interview you complete or each time you collect data. This will make your life a lot easier when it comes time to start analysing.

Once you have input all of your data into your spreadsheet software and are ready to put it into Gephi, you will have to take the following steps:

- Save/Export your spreadsheet as a “.csv” file. This is the only kind of file Gephi can read.
- Open Gephi and start a new project.
- Then when you are in the “Data Laboratory” you will click “Import CSV”. You will need to import your node spreadsheet **and** your edge spreadsheet.
- This Wiki page gives more information on importing your spreadsheets:  
[http://wiki.gephi.org/index.php/Import\\_CSV\\_Data](http://wiki.gephi.org/index.php/Import_CSV_Data)

You also have the option of changing or updating the data once you’ve imported it into Gephi by typing directly into the node and edge spreadsheets. Once your data are imported you are able to start your analysis using Gephi’s SNA functions.

Now that you have imported your spreadsheets, it is time to start visualising your social network! Use Gephi’s many features to adjust the layout, add labels, add colours, group actors together, and much more!

## STEP FIVE: INTERPRETATION

### *What do these sociograms mean for me, and for my organisation?*

As discussed before, there are a number of ways that you can use your sociogram. Although it is important to determine what you are looking for before you start collecting data, sometimes you can uncover things in your social network that you may not have been looking for when you started. Don’t forget to look back to the “What to look for in your social network: Measurements and key terms” section of this document (especially figure 2) for more ideas of what you can look for in your social network.

Here are some specific questions you can ask yourself about the social network:

- Are there any actors who are isolated from the network? How can we help them become more connected?
  - Isolates are easy to identify once you have created your sociograms. They will not have any identified connections within the network.
  - One way to get actors to be more connected within the network is by inviting all the actors together for an event, workshop, or meeting. This kind of activity may be formal or informal, depending on your preferences. By facilitating a group activity, actors will be able to connect with others in the network that they might not have even known were in the network to begin with.
- Are there actors who are really central in the network (key players)? What factors put them in this position?
  - In Gephi, you are able to make people with higher degree (more connections) larger in size. This way, you can visualise the actors who are most central in the network. To apply this, click “window” on the top menu, then select “ranking”. There will be a box on the left portion of the Overview screen that will be titled “ranking”, here you will choose the diamond symbol “size/weight”. From the dropdown menu, select “Degree”, then choose what you want your minimum and maximum sizes to be. You can play around with this until you are happy with

what your sociogram looks like. The larger the circle, the more connected the actor is.

- Once you find out which actors are most central, you can reflect on some of the survey responses to see why they are a most central. Sometimes it may be obvious, and other times it may not be so obvious. For example, it would be expected that a government funding agency would be a central actor in a research network where the interviewee was asking about source of funding. However, there may be some key players that emerge that might surprise you. Keep an eye out for this! Do the central actors have more “power” in the network (even just “perceived power”)? What factors put them in this position? Are other actors intimidated by this “power”?
  - Your definition of power will determine what actors you are looking at in this question. Perhaps power is measured by more connections, giving more resources, having many asymmetrical relationships flowing away from the actor, etc. It may be something that came up in your interviews (intentionally, or unintentionally). Either way, you may want to determine who is “powerful” in your network, what puts them in this position, and what kind of influence they have on other actors in the network.
  - You may also want to request follow-up interviews with the network members if you need more information or require clarification from an interviewee.
- Are most relationships in the network asymmetrical? Is reciprocity desired? If so, how can we increase it?
  - Whether relationships are asymmetrical or not will become very obvious once you have created your sociograms. Some relationships are naturally asymmetrical (for example, funding agencies will rarely (if ever) receive funding from someone in the network), however, there may be some desire for communication flows to become more reciprocal. You will have to determine what approach to take to increase reciprocity within the network (ex. host networking events, create an emailing listserv, create a shared repository for documents, etc.).
- Are there actors who don’t have access to resources such as money, and information? Do these actors need access to these resources? If so, how can we increase their access?
  - There may be some actors in your network that have very few connections that provide them with resources (if any at all). For example, if they are a new company, they may not know where they can get information or funding from. You will need to determine their level of need, and what you can do to increase their access to actors in the network that can provide them with such resources.
  - Isolates, clusters (cliques), and actors that are only involved in the network through gatekeepers or cutpoints are good examples of where to look first.
- Is information flowing to all of the actors who need the information? If not, how can we change that?
  - This is similar to the question above. You can look at your social network to determine which actors do not have access to information. To know the level of need, you will either have to ask a question on your survey (ex. do you wish to receive more information on watershed management from other actors in the network?), or you will have to use your discretion after you’ve completed your interviews.
- Do any of the actors in the network act as a bridging organisation or gatekeeper? Are these actors aware of this situation and its implications?
  - As discussed in the “What to look for in your social network: Measurements and key terms” section of this document, identifying actors that act as bridging organisations or gatekeepers is much easier to do once you have input all of your social network data and created your sociograms. Some actors might not even be

aware that they are the only connection between another actor and the rest of the network. In some cases, this situation may be perfectly fine; however, the disconnected actors may wish to be more connected within the network.

- How has the network changed over time? Were our efforts effective?
  - This is a great question that can be answered by completing a longitudinal study. After a certain period of time, go back to the actors you interviewed the first time and ask them the same questions. You will be able to see how the network has changed over time, and evaluate any efforts you have made to connect more people. You will have to determine how you want to measure success (ex. are the isolates from the original sociogram connected now?).

These are just a few of the many questions you can ask yourself when you have created your sociograms. Just remember that each individual SNA is unique, so you will have to determine what will work best for you, your organisation, and the network!

## REFERENCES

- Bastian M., Heymann S., and Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.
- Butts, C. (2008). Social network analysis: A methodological introduction. *Asian Journal of Social Psychology*, 11: 13-41.
- Crona, B.I., and Parker, J.N. (2012). Learning in support of governance: Theories, methods, and a framework to assess how bridging organizations contribute to adaptive resource governance. *Ecology and Society*, 17(1): 32. <http://dx.doi.org/10.5751/ES-04534-170132>
- Divjak, B., and Peharda, P. (2010). Social network analysis of study environment. *JIOS*, 34(1): 67-80.
- Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *LISR*, 18: 323-342.
- Kilduff, M, and Tsai, W. (2003). *Social Networks and Organizations*. Sage Publications, Thousand Oaks, California.
- Maggetti, M., Ingold, K., and Varonse, F. (2013). Having your cake and eating it, too: Can regulatory agencies be both independent and accountable? *Swiss Political Science Review*, 19(1): 1-25.
- Provan, K.G., Veazie, M.A., and Staten, L.K. (2005). The use of network analysis to strengthen community partnerships. *Public Administration Review*, 65(5): 603-613.
- Yin, R. K. (2003). *Case study research, design and methods*. Newbury Park, Sage Publication.

### How to reference this guide:

Morris, S., A. Verville and L. Vasseur. 2014. Understanding social connections in communities: how to use social network analysis - Guidelines document. Coastal community Challenges – Community University Research Alliance. 21 pages.