# Linker QuickStart Guide

The Linker is a tool for taking structured data and converting it to a meta-network representation (Carley, 2003). It differs from imports available from other tools, including ORA (Carley, Pfeffer, Reminga, Storrick, & Columbus, 2013) or UCINET (Borgatti, Everett, & Freeman, 2002), because it allows for unique nodes to be found by using multiple columns of the source-data natively within the tool. It is designed for analysts and scientists comfortable with their source data to generate useful meta-networks from their data.

Meta-networks, a collection of node-sets and graphs, are useful for understanding connections between actors, groups, resources, and knowledge. Converting survey data to meta-network representations allows us to generate new features which can be used to supplement other analyses, such as statistical regression analyses or machine learning classification.

Introduction

By default, the Linker assumes that the data-files given are Tab-Separated Values, with a header row providing titles for each column. You can think of each header as the title of a flash-card, with the data provided as the answer on each flash-card. A new set of flash-cards is generated for each row of data.
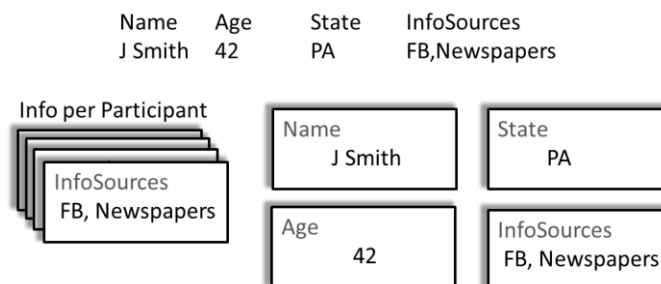

Manipulating and using this set of flash-cards is central to using the Linker. Often, entities are listed in surveys, and we can use the Linker to draw those entities out into a network representation. The figure, below, shows how different survey elements can be used by Linker.

Header Name	Meaning	Value	
ParticipantName	Name of the respondent	J Smith	Respondent: J Smith
Age	Age of respondent	42	
State	State of Respondent	PA	Info: FB
InfoSources	Where do they get info from?	FB, Newspapers	Info: Newspapers
PartnerA_Name	Name of PartnerA	Name12	
PartnerA_Role	Role of PartnerA	Grower	Partner: Name12+Grower
PartnerB_Name	Name of PartnerB	Name20	
PartnerB_Role	Role of PartnerB	Beekeeper	Partner: Name20+Beekeeper

By default, everyone is connected to the respondent, J Smith, resulting in a network from this small set of data that looks like this:
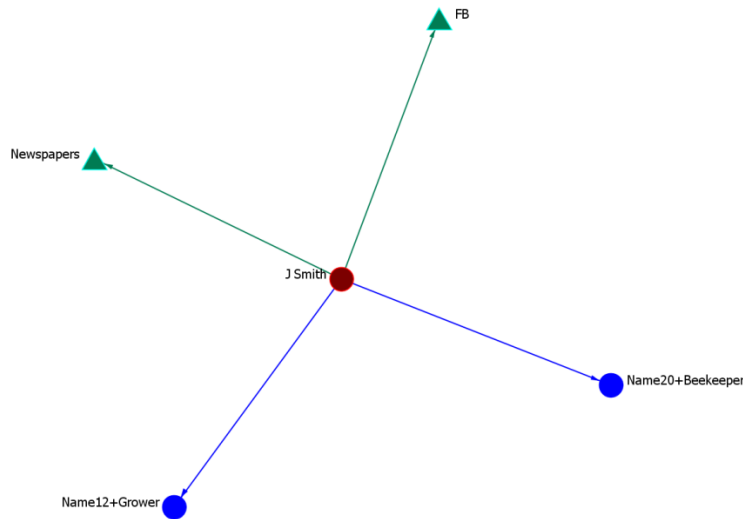
How do we do this? That's the goal of this guide.

MS Excel and Tab-Separated Value (TSV) Data

A Microsoft Excel spreadsheet can be converted to a tab-separated file which can be processed by the Linker. The resulting file, which will also be a tab-separated file, can also be read by Microsoft Excel. This section will give the steps necessary to both output a tab-separated value, and to read a tab-separated value into Excel.

Converting Excel to TSV Data

1. Start Excel
2. Open your data-file
3. Click the "File" tab
4. Click "Save As", a dialog "Save As" will pop up
5. Near the bottom of the dialog, there will be a field called "Save As Type"
6. Select "Text (Tab Delimited) (*.txt)"
7. Click "Save"

Reading TSV Data in Excel

By default, Excel will automatically open CSV or Excel files properly, but text files will not be read natively by Excel (instead, the file will be opened by MS Word or another program), so follow these directions to open a text file or other file in Excel.

Method One: Use "Open With" from the File Window

1. Navigate to where your file is located on your file system
2. Right-Click on the File
3. Select "Open With >"

4. Select "Microsoft Excel"
5. The file should open correctly

*If the file fails to open correctly using Method One, switch to Method Two*

Method Two: Opening the File in Excel

1. Start Excel
2. Navigate through directories to where your data-file should be
3. At the bottom of the "Open" dialog, there is a combo-box that indicates what files will be shown, change the filter from "All Excel Files" to "All Files", now your data-file should be visible!
4. Select your data-file
5. A new dialog, called "Text Import Wizard – Step 1 of 3" will come up. The choices are "Delimited" or "Fixed Width". Choose "Delimited", and click "Next"
6. Excel will scan your file and try to guess what the delimiters are – make sure only "Tab" is selected and click "Finish"

Starting the Linker

To start the Linker, double-click on its JAR file. You must have Java installed on your machine[1].

You should immediately see the following dialog box:
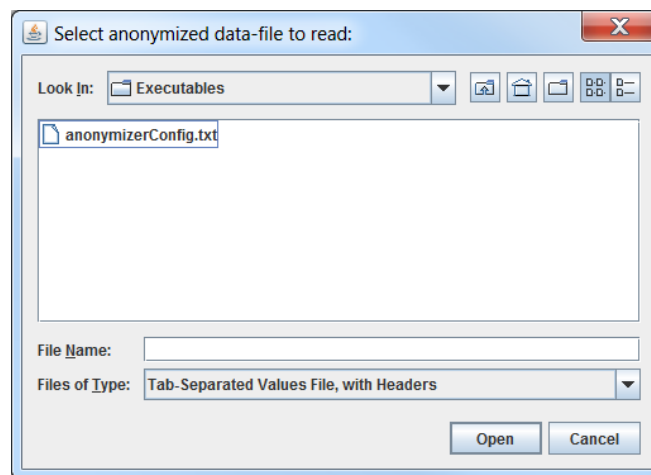
Navigate to where your tab-separated data is stored, and then click on it. Then click "Open". The following screen should then appear.

---

[1] Go to this official link for details on how to install Java:
https://www.java.com/en/download/help/download_options.xml
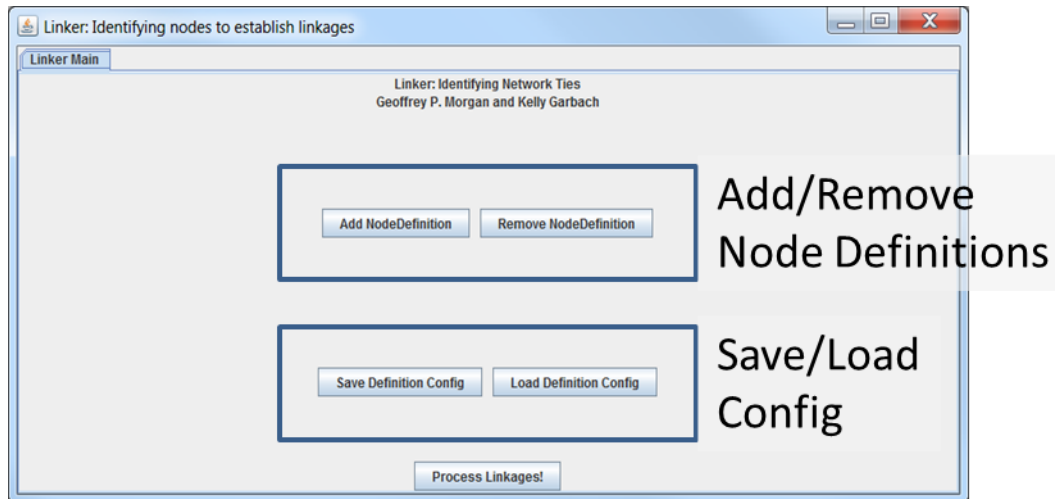
The screen has one tab, labeled "Linker Main". This tab has five buttons:

- Add NodeDefinition – This creates a new tab to support a new node definition
- Remove NodeDefinition – This lets you select a node definition and remove it from your configuration.
- Save Definition Config – By clicking this button, you can save your current configuration for future use.
- Load Definition Config – By click this button, you can load a saved configuration.
- Process Linkages! – By clicking this button, a new meta-network is created based on your data and node definitions.

We go through each of these operations in turn. Each node definition is opened on a new tab. Return to the "Linker Main" whenever needed.

## Adding and Editing Node Definitions

To create a new node definition, click "Add Node Definition". When you do, a dialog box appears.
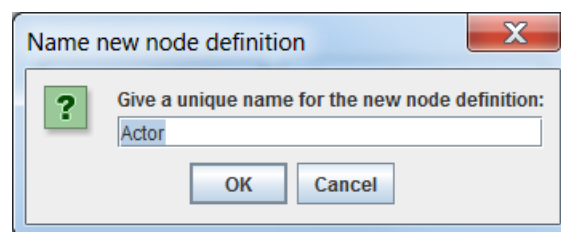
The default node unique name for a new node definition is "Actor", but this should be edited to suit your data. The tool will reject a duplicate node definition name with the following warning message.
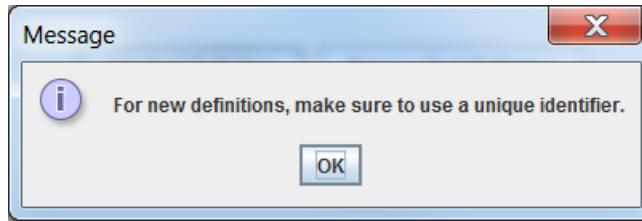
This creates a new tab matching the name of your new node definition, when you click on this tab, it looks like so:
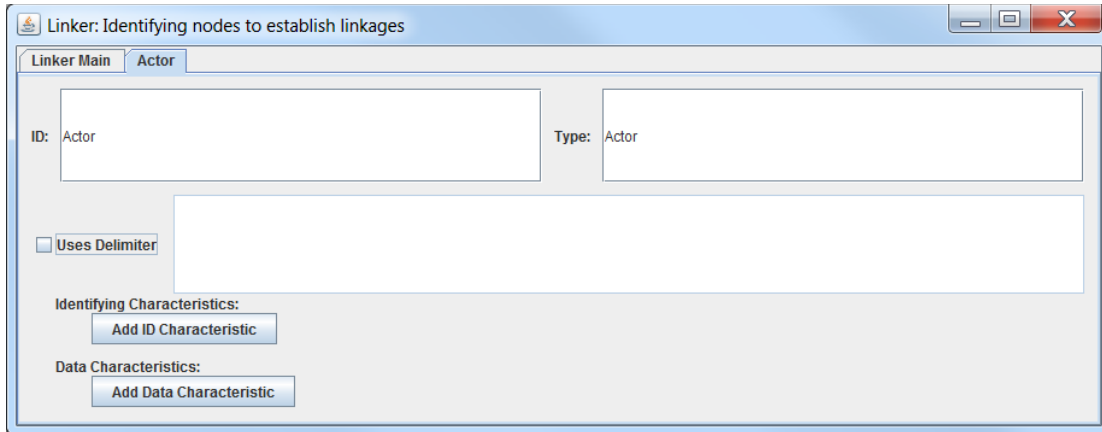
There are several fields on the tab:

- ID: The unique ID of the node definition
- Type: The type of node that should be created, note that types may be the same across node definitions
- Uses Delimiter: If the data field has multiple entries in the column, then check this box
- Delimiter Field: Next to the "Uses Delimiter" checkbox, you specify how those values are separated; a common choice is the comma ','.
- Identifying Characteristics: Columns of the data that are used to identify nodes
- Data Characteristics: Columns of the data that add information to the nodes, but are not used to identify the node

Earlier in the Introduction, we identified some sample data and how we could use that to inform node definitions. That figure is reproduced here, and then the settings you would use for each given node-type follow.

Header Name	Meaning	Value
ParticipantName	Name of the respondent	J Smith
Age	Age of respondent	42
State	State of Respondent	PA
InfoSources	Where do they get info from?	FB, Newspapers
PartnerA_Name	Name of PartnerA	Name12
PartnerA_Role	Role of PartnerA	Grower
PartnerB_Name	Name of PartnerB	Name20
PartnerB_Role	Role of PartnerB	Beekeeper

Respondent: J Smith
Info: FB
Info: Newspapers
Partner: Name12+Grower
Partner: Name20+Beekeeper

Respondent:

- ID: Respondent
- Type: Respondent
- Uses Delimiter: No
- Delimiter Field: N/A
- Identifying Characteristics:
  ParticipantName
- Data Characteristics: Age, State

InfoSource:

- ID: InfoSource
- Type: InfoSource
- Uses Delimiter: Yes
- Delimiter Field: ,
- Identifying Characteristics:
  InfoSources
- Data Characteristics: None

PartnerA:

- ID: PartnerA
- Type: Partner
- Uses Delimiter: No
- Delimiter Field: N/A
- Identifying Characteristics:
  PartnerA_Name, PartnerA_Role
- Data Characteristics: None

PartnerB:

- ID: PartnerB
- Type: Partner
- Uses Delimiter: No
- Delimiter Field: N/A
- Identifying Characteristics:
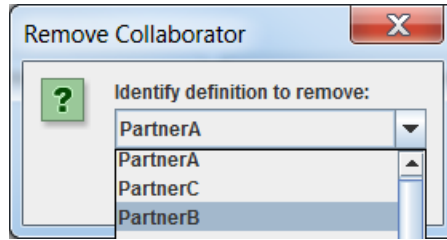 PartnerB_Name, PartnerB_Role
- Data Characteristics: None

A filled-out form for an example partner would like so:

Removing Node Definitions

To remove an existing node definition, return to the Linker Main tab and click "Remove NodeDefinition", a dialog box will open.

Click "Ok" when you've selected the node definition you want to remove. This removes the definition from the program entirely.

Saving and Loading Configurations

This section covers how to save and load configuration data. The configuration saved is always the current configuration as defined. A loaded configuration replaces all prior data.

Saving the current configuration

To save the current configuration, go to the "Linker Main" tab, and click "Save Definition Config", a dialog box will pop up.
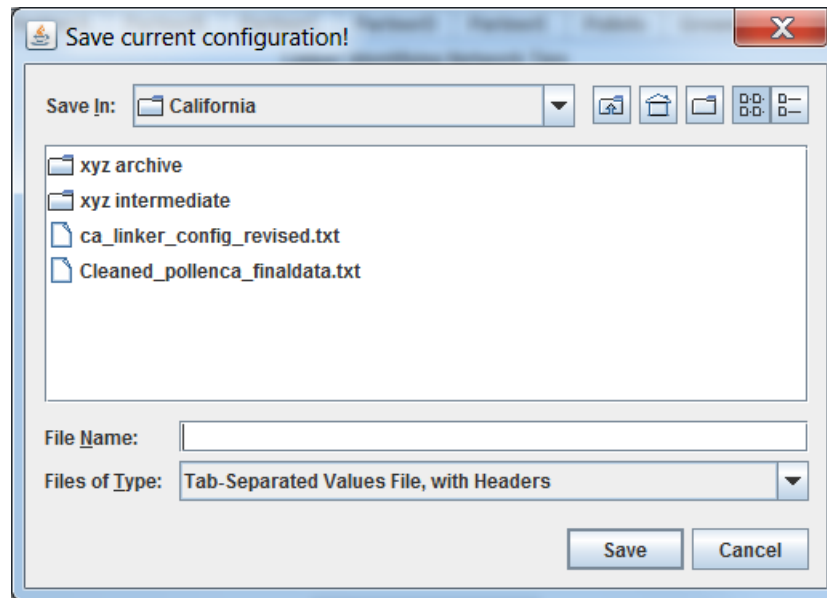

Navigate to the appropriate directory, and then type in the name of the file. Given here is an example file.

```
ID          TYPE        Delimiter ID-Chars Data-Chars
Respondent              Respondent                  opername_Anonymous County___Almond___Watermelon
PartnerA  Partner                Quest14a_Name_Anonymous___Quest14a_RoleCategory          Quest14a_Role_Cleaned___Quest14a_Role
PartnerB  Partner                Quest14b_Name_Anonymous___Quest14b_RoleCategory          Quest14b_Role_Cleaned___Quest14b_Role
PartnerC  Partner                Quest14c_Name_Anonymous___Quest14c_RoleCategory          Quest14c_Role_Cleaned___Quest14c_Role
PartnerD  Partner                Quest14d_Name_Anonymous___Quest14d_RoleCategory          Quest14d_Role_Cleaned___Quest14d_Role
PartnerE  Partner                Quest14e_Name_Anonymous___Quest14e_RoleCategory          Quest14e_Role_Cleaned___Quest14e_Role
PolInfo   PollinatorInfo     ,          Quest10
GrowerOrg GrowerOrganizations            ,          Quest15
ConservationOrg    ConservationOrganizations   ,          Quest16
```

The text-file can be edited directly if preferred to working with the interface, but this is not recommended. Multiple values within a single tab-separated column are separated by "___".

Loading a configuration

If you have previously used the Linker and saved your configuration, you can load that configuration from the "Linker Main" tab. Click "Load Definition Config", and a dialog box pops up.
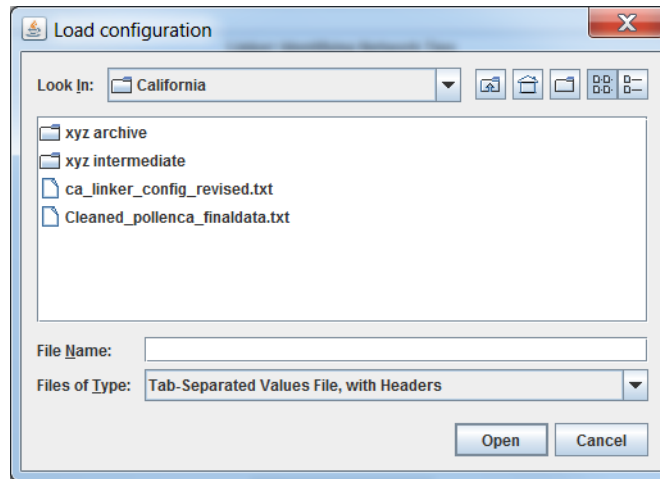

Navigate to the directory you saved your configuration file, click on it, and click "Open". New tabs will auto-populate based on your data.

Process Linkages

When you've configured the various node definitions to your liking, you can return to the "Linker Main" tab and click "Process Linkages". The tool will identify all unique node-types and inform you of their number.
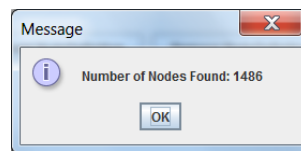

Then, the tool will identify all possible relationships and present them to you for review. Shown below is just the top of a list of these relationships. For each node-type, a network definition will be presented, and then a checkbox indicating whether the network should be outputted at all, and whether the relationship between these two node-types should be flipped to better the fit the analyst's intent.
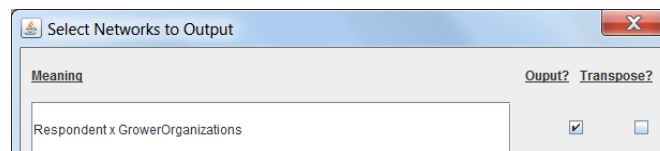
At the bottom of this dialog there will be a "Submit" button. Once you are happy with your network output configuration settings – click Submit. Another dialog box will appear, asking you to name the

resulting collection of nodes and edges – you may want this name to reflect choices made in the output and cleaning.

After you offer a name, a final dialog will appear:
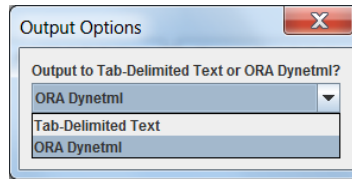
This choice controls the resulting format of the output, if you choose "ORA Dynetml" all data will be contained in a single well-formed XML file. If you choose "Tab-Delimited Text", the data will be written to a directory with the name you gave in the step previous. Each network and node-set will be written to file separately, with the networks written as edge-lists.

References

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). UCINET for Windows: Software for social network analysis. Harvard, MA: Analytic Technologies.

Carley, K. M. (2003). Dynamic Network Analysis. In R. Breiger, K. M. Carley, & P. Pattison (Eds.), *Summary of the NRC workshop on social network modeling and analysis* (pp. 133-145): Committee on Human Factors, National Research Council.

Carley, K. M., Pfeffer, J., Reminga, J., Storrick, J., & Columbus, D. (2013). *ORA User's Guide 2013*. Retrieved from Carnegie Mellon University: