

# Guide to Formatting and Uploading Data

## Contents

- Notes
- Quick Terminology Explanation
- Formatting Data
  - Person Nodes Schema
    - \* Name
    - \* Programme
    - \* Email
    - \* StaffSiteLink
    - \* SiteName
    - \* SiteLink
    - \* MediaLink
    - \* Bio
    - \* Role
    - \* Collaborator
    - \* Special Nodes
    - \* Example Table
  - Special Nodes Schema
    - \* Name
    - \* MediaLink
    - \* SiteName
    - \* SiteLink
    - \* Bio
    - \* Example Table
  - Modifiers Schema
    - \* Example Table
  - Project Folder Setup
- Uploading Your data
  - Using Your Git Service
    - \* Git GUI
    - \* Git Bash

## Notes

This guide will cover the fundamentals for structuring your datasets for the network visualisation tool. In addition, it covers how to push your data onto the git service.

## Quick Terminology Explanation

- **Node** is a circular object that represents your data in the network visualisation tool.
- **Person** nodes are nodes which represents the people in your network.
- **Special** nodes are all other nodes in your network that do not represent people.
- **Modifiers** are groups for each type of person and is helpful for showing these groups through having unique colour properties per group.

## Formatting Data

The following sections explain in detail the different schemas and the project folder setup. An important note is that all files must have their set fields, which will be defined later in this guide. **Compulsory** fields need to have a value. Whereas **Optional** fields may be left blank.

### Person Nodes Schema

Ensure that your data contains the following Schema.

#### Name

Name of the person node which labels the node on the GUI. It can be as many words or characters.

**Compulsory**

#### Programme

Name of the programme that the person belongs to. It can be as many words or characters.

**Compulsory**

#### Email

Email of the person which is shown on the details pane. It must have proper formatting of an email address.

**Optional**

#### StaffSiteLink

Website link that the person has which is shown on the details pane. It must have proper formatting of a URL

**Optional**

**SiteName**

Name of the website link provided from the previous field. Does not have to be present regardless of presence of staff site link.

**Optional**

**SiteLink**

Additional website link the person may have which is shown on the details pane.

**Optional**

**MediaLink**

Embedded video link that relates to the person which is shown on the details pane.

**Optional**

**Bio**

A description of the person node which is shown on the details pane.

**Optional**

**Role**

Modifiers information for grouping your people in the visualisation tool.

**Compulsory**

**Collaborator**

A list of names of other people nodes who have worked to together to draw connections between them.

**Optional**

**Special Nodes**

The name of an entry inside a special node. For example, if John Smith is in the Infrastructure department, then we would need a field called department where the value is Infrastructure. Having these fields draws a connection between the person node and special node.

**Compulsory if there are special nodes**

### Example as Table

An important note when defining **connected person nodes** is to follow proper *CSV formatting* to ensure your data is processed correctly.

name	program	email	staff	SiteLink	SiteName	mediaLink	bio	role	collaborators	department
John Smith	Engineering	js@sigmat.com		https://www.jst.com/	John Smith	https://www.johnblog.com/	John Smith is a hard working engineer."	Junior	"Kirt Denning"	Infrastructure
Kirt Denning	Engineering					https://youtube.com/watch?v=4S9W3NcQ	"I work to be the best"	Senior	W3NcQ, Anderson, John Smith"	Developer
Nicholas Anderson	BA	na@yahoo.com								Risk Analyst

### Special Nodes Schema

Ensure that your data contains the following Schema.

#### Special Node Name

Name of the special node which labels the node on the GUI. It can be as many words or characters. **Compulsory**

#### Special Node MediaLink

Embedded video link that relates to the person which is shown on the details pane. **Optional**

#### Special Node SiteName

Name of the website link provided from the previous field. Does not have to be present regardless of presence of staff site link. **Optional**

#### Special Node SiteLink

Additional website link the person may have which is shown on the details pane. **Optional**

#### Special Node Bio

A description of the special node which is shown on the details pane. **Optional**

### Special Node Example as Table

An important note when defining **connected special nodes** is to follow proper *CSV formatting* to ensure your data is processed correctly.

	name	mediaLink	siteName	siteLink	bio
Infrastructure	https://www.example.com			https://www.js.com	"Details.."
Developer	https://www.example.com			https://www.js.com	"Details.."
Risk Analyst	https://www.example.com			https://www.js.com	"Details.."

### Modifiers Schema

The format for modifier nodes must follow the simple structure. It only requires one field. It is currently required to be named **role**. As mentioned earlier in this guide, this purpose of this file is to define groups to help with the clarify of the network visualisation tool. An example table is provided below. It must follow *CSV formatting* to ensure your data is processed correctly.

### Role Example as Table

role
Junior
Senior
Manager

### Project Folder Setup

To ensure your data is processed correctly, we have established a folder setup that is simple to follow.

- The root folder must be called *data*
- The next level must have folders that sort data by *year*. Example:
  - 2017
  - 2018
  - 2019
- Inside each folder will contain the *data* for that year

```
-- data
|-- 2016
|-- 2017
    |-- nodes
        |-- person.csv
        |-- roles.csv
```

```
\--specialNodes
  |-- project.csv
  \-- school.csv
```

The purpose of having folders sorted by year is for the visualisation tool to show a range of data. This is to let the end user isolate data they care about.

PLEASE NOTE: the year folder is required even if there is only one year to display.

## Uploading Your data

There are two major steps for this process. The first step is to ensure that the your data is formatting accordingly to this guide. Secondly, all you need to do is to push your data onto your git service. We have chosen to use a git service to help with maintenance and data validity. Do not fret if you have not used git before, this guide explains how to use git.

## Using Your Git Service

This next section of the guide explains how to use the Git GUI and the Git Bash to push your data onto the git service.

### Git GUI

1. Navigate to the data folders
2. Select the upload files button
3. Drag and drop files
4. Submit files.

### Git Bash

1. Fork the original repository
2. git add .
3. git commit -m "pushed my data"
4. git push