

A background network graph visualization consisting of numerous small, light-colored nodes connected by thin, light-colored lines, forming a complex, interconnected web structure. The nodes are distributed across the slide, with a higher density in the center and lower density towards the edges.

Introduction to Social Network Analysis For the Study of Religion in East Asia

# Visualizing Network Graphs on the Web




Simon Wiles

NUS, August 7<sup>th</sup> 2018

# Why Visualize Network Graphs?

- Visualizing network graphs for research.
  - exploring, investigating, and analysing data
  - diagnosing, finding errors, and “sanity checking”
- Visualizing network graphs for presenting and communicating data and findings.
  - static visualizations for print/web etc.
  - interactive visualizations for web and other digital platforms

# Web-based Options for Network Visualization

- “Solutions”
  -  Palladio
  -  Google Fusion Tables
  -  Gephi-based solutions
    - the Gephi `sigma.js` plugin
    - GEXF on the web

# Starting Points for Network Data

- Data formats and data wrangling
  - Tabular data and CSV files
  - XML files, GEXF, GraphML etc.
  - Other Gephi-supported formats
    - GDF
    - Netdraw VNA
    - etc...

# Introduction to Palladio

- What is Palladio and what does it do?
  - <http://hdlab.stanford.edu/palladio/>
  - Humanities + Design Research Lab at Stanford
    - Networks in History and Mapping the Republic of Letters
  - An online platform and a suite of tools for visualizing datasets
    - consumes data in tabular (spreadsheet) form, and visualizes spatial, temporal, and other connections within the dataset
    - nothing is saved online
    - no account is required or used
  - Palladio is under constant development, so there are always some features in beta, and there is always the potential for bugs



# Limitations of Palladio

- Palladio is a (set of) tool(s) for exploring datasets
  - You can save and download particular results and visualizations, but Palladio does not allow you to export interactive visualizations, and you cannot use Palladio directly to present your data to others (at least, not yet).
  - You can create interesting network diagrams and visualizations, but Palladio is not an advanced network analysis suite, like Gephi etc.

# Interlude – DILA Biographies Project Data

- Connections from the Liang Gaoseng Zhuan 梁高僧傳
  - <http://buddhistinformatics.dila.edu.tw/biographies/gis/>
  - XML/TEI from the Gaosengzhuan project

	Liang	Tang	Song	Ming	Biqiuni	CSZJJ	MingSeng	BuXu	Total
Total Number of Files (== Zhuan):	257	485	543	112	65	32	36	473	2003
Most Recent Edit:	06 Aug 2018	15 Feb 2017	01 Mar 2017	28 Nov 2012	01 Mar 2017	16 Oct 2012	22 Oct 2012	28 Nov 2012	06 Aug 2018
Total distinct persNames:	1411	2278	2774	521	261	261	117	1990	8255
Total distinct placeNames:	869	2001	1946	603	196	212	142	2134	6176
Total Nexus Points:	1790	5478	5000	591	375	474	188	3976	17872
Total Nexus Points with multiple people:	954	2179	2214	266	193	222	65	1809	7902
Total number of ties:	1876	4501	6888	326	271	544	100	2546	17052
Total distinct ties:	1677	3854	5418	305	250	412	89	2215	13790
Total number of actors:	926	1785	2132	274	212	206	91	1489	6472
Average ties per actor:	2.025918	2.521569	3.230769	1.189781	1.278302	2.640777	1.098901	1.709872	2.634734
Total number of possible ties:	856550	3184440	4543292	74802	44732	42230	8190	2215632	41880312
Network density (total ties over poss. ties):	0.002190	0.001413	0.001516	0.004358	0.006058	0.012882	0.012210	0.001149	0.000407
Network density (distinct ties over poss. ties):	0.001958	0.001210	0.001193	0.004077	0.005589	0.009756	0.010867	0.001000	0.000329

# Palladio Walk-Through

- Connections from the Liang Gaoseng Zhuan 梁高僧傳
  - <http://buddhistinformatics.dila.edu.tw/biographies/gis/>
  - XML/TEI from the Gaosengzhuan project
  - Palladio-ready CSV
  - loading lgsz\_palladio.csv into Palladio
  - filtering, faceting, etc.
  - saving visualizations, saving projects.



# Introduction to Google Fusion Tables

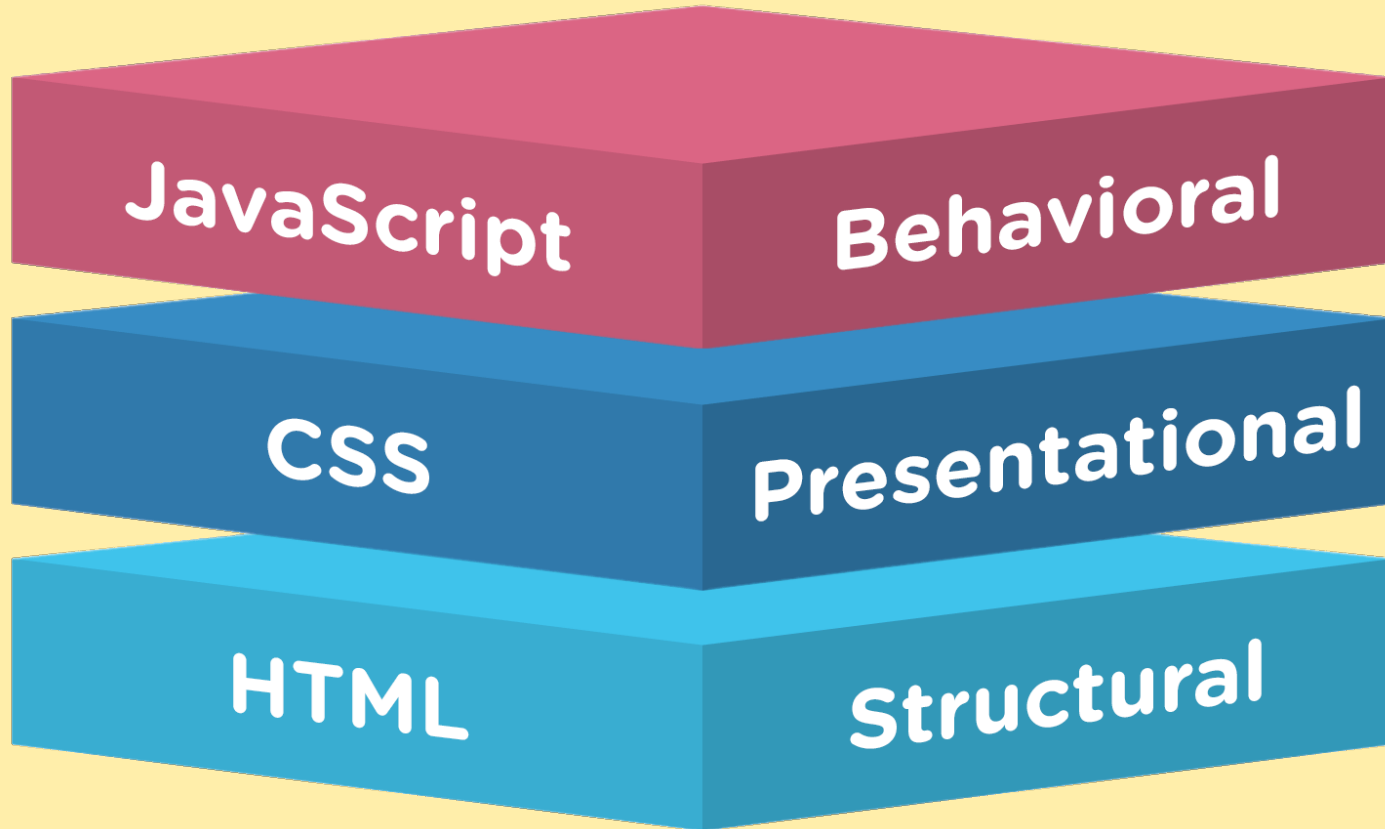
- <https://fusiontables.google.com/>
- Fusion Tables is concerned with storing, serving and visualizing tabular data, especially in a collaborative way
- Fusion Tables is pretty limited in terms of what can be done with Network Graphs
  - maximum of 100,000 rows can be visualized
  - display options are very limited
- Hosted, embeddable and publishable



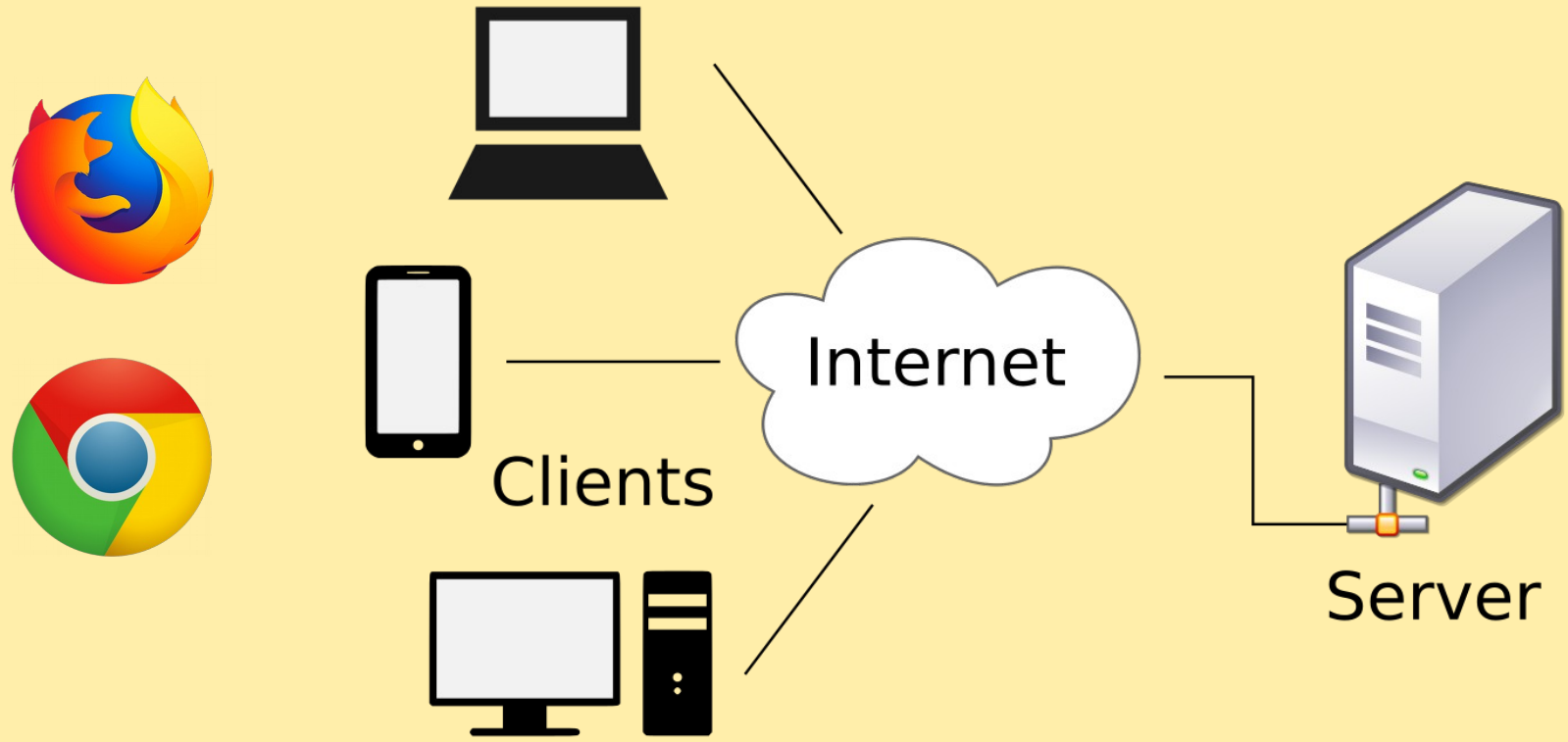
# Google Fusion Tables Walk-Through

- Connect Fusion Tables to Drive
- Connections from the Liang Gaoseng Zhuan 梁高僧傳
  - loading lgsz\_fusion\_tables.csv
  - creating a network chart
  - publishing/embedding the chart

# A Brief Interlude – How do web pages work?



# A Brief Interlude – How do web pages work?



# Publishing your Gephi-created Networks

- The SigmaExporter Plug-in for Gephi
  - <https://github.com/oxfordinternetinstitute/gephi-plugins/wiki>
  - Oxford Internet Institute / JISC
  - Produces a self-contained “website” which can be used locally (Firefox) or uploaded to a website.
    - poking around inside the package and changing a few things

# Online visualizations based on GEXF Files

- The GEXF file format
  - <https://gephi.org/gexf/format/>
  - the VIZ module
    - <https://gephi.org/gexf/format/viz.html>
- Javascript Libraries

# Online visualizations based on GEXF Files

- **sigmajs**

- <http://sigmajs.org/> – Alexis Jacomy / Guillaume Plique

-  Data Driven Documents

- <https://d3js.org/> – Mike Bostock

- <https://github.com/emeeeks/gexfd3> – Elijah Meeks

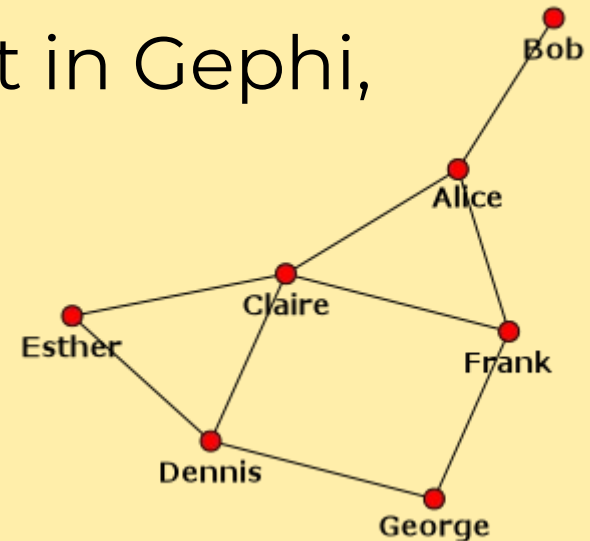
- **gexf-js**

- <https://github.com/raphv/gexf-js> – Raphaël Velt

- <http://sna2018.mbingenheimer.net/>

# Creating Graphs from Your Own Data

- Edge lists and node lists
  - the simplest possible cases
  - create the edge list, and try it in Gephi, Palladio, Fusion Tables etc.





# Creating Graphs from Your Own Data

