

This is a transcript of Video 3 in William I. Wolff's webtext "Twitter Archives: A Discussion of Systems, Methods, Visualizations, and Ethics," published in *Kairos: Rhetoric, Technology, Pedagogy*, 22.2, available at http://praxis.technorhetoric.net/tiki-index.php?page=PraxisWiki%3A_%3ATwitter+Archives

Hello, this tutorial is going to show you how to create a very easy mentions map from exporting data from DMI-TCAT and opening it in Gephi and it will walk you through the process. To complete this, you will need to have Gephi open and you will need to have open in a browser the capture page of your DMI-TCAT if you do not recall how to get to this page you can go to your document where you store the information when you set up DMI-TCAT and you will want to go to the analysis URL and there's the username and the password and that should bring you right there. So when you come back to DMI-TCAT choose the archive that you're going to be mapping, here these are the Terrence Crutcher hashtag tweets, choose a date range, you're not going to want to have the entire archive that's much too long but choose a date range that you think makes sense.

If you have a TV show that you're looking at more than one episode several weeks would probably be enough now notice that there are over seven hundred and ninety eight thousand tweets from 400,000 distinct users, that's a nice range we're only going to be taking out very small sample of that so I could have even broken it down here but for some reason I want a little longer. Okay, so once you've determined the date range and the hashtag that you're going to be grabbing scroll down to the network's section and you're going to be getting a social graph by mentions and these are all times when there is a mention within a tweet, including, including retweets so this does include the retweets that are not direct and replies to, these are when my user is mentioned in the tweet and this is good for finding hubs and communities and, and looking for some patterns within there.

So you're going to click launch, and this little pop-up window will appear and you want to choose a thousand. That's a good number. This is the total top users you want to get if you enter zero at this point we would have 400,000 users and that is much too big, so we want to get a thousand as a nice round number and we'll be able to tune out, when you do that it will take a second to save while that is creating the account, I'm sorry- the file I'm just going to show you the Gephi area over here very quickly there are three different sections of Gephi there's the overview, which is where we do all the creating of the network map, there is a data laboratory where you can go in and look at the actual data the numbers we're not going to be taking a look at that today and then there was the preview area where you create your final image.

This is the the big graph areas where we're going to be seeing the graph tape manifest itself and take place the appearance is where we adjust the node size and we determine what we want to highlight in the map, down here is the width is where we choose which layout we want to use in their variety of different ones available to us we're going to be using one for this tutorial and here there are filters that we can apply to the Data to reduce the number of nodes that we're looking at and statistics that we can apply and we're going to be just running a couple of those I should at this point thank Jen Colbeck, who has wonderful YouTube videos on how to use Gephi and much of what I'm going to be showing you today is learned from her. Let's go back to see if

our file is ready it is so we can just right click on it and save link as, I'll save it somewhere where you know is going to be, I'm going to save it as 3, this is the third time I'm using this one, and there it is it's safe down there... Ok now I'm going to go to Gephi, and I'm going to go to file, open, and I'll choose the one that I'm going to be using which is right here, at the one that I just opened, I created and then I click open.

You are going to get this import report, "There are no issues found during this import" which is nice, because this is a mentions graph we have a, it is a directed graph because they're going from one person to another we can see the number of nodes a thousand because we took the top thousand users and the number of edges which means it's 3826 connections among those 1000 users, so that's a nice nice amount. We click OK, and we get a nice square blob that looks like the Borg from Star Trek. So we really can't do anything with this it is just a big square mass, we need to apply some patterns to it and the first thing that we're going to do is to see how this works is we're going to go to the layout area and we're going to scroll down to Yifan Hu, and we'll click run, and we'll see what happens oh, and we'll see we see it exploding out like that, and what this is doing is it's starting to create relationships between different users here we have the different nodes and you can see if you move in you can see all the different nodes are in there everything is sort of black and white grey on the outside and the outer rings we see nodes that are not really connected to anybody else, these are just us you know solo nodes they don't really have many connections to the larger mass, and they're really not very interesting for what we need... Were much more interested in what's going on down here particularly, in this little area over here, and in some other places. So we want to filter out some of this data so we can begin to work with what we really want to work with so what we do is we go to the filter section over on the right and in the topology area we - we click the little carrot and we'll go and we'll find the giant component and we take this giant component and we drag it down, to where it says drag filter here.

Okay, and once we do that we click on that and we click on the bottom right hand corner filter. Now watch and see what happens in this area over here, after I click the filter, okay, and that filtered out a considerable number of those nodes and in fact if we look in the right upper right hand corner we can see that it's filtered down from a hundred percent visible, to eighty-one percent visible right? and I can unfilter and then re-filter okay, there are now ninety nine point two nine percent visible edges, and 810 to eighty-one percent visible nodes. Now we want to filter this a little bit more by a degree range so if you put this arrow we can see a sub-filter area and if we take the degree range and we drag it down on top of that, or just below, we will then, the queries will now have our degree range if we click on that you'll see in the bottom and right hand corner that there is a range of degrees that we have available to us. 0, 2, 118.

118 is the node, or nodes, that are connected to 118 other nodes so that is those users who have that range of 118 will be connected to the most nodes now that's, that's nice to have if and will so we'll be able to see who's really at the center of these conversations that are going on with this hashtag. To filter out a little bit more, we take the degree range setting slider and we drag to the right we can stop at five, that's sort of these are users who are connected to five other users, you could go and we can sort of zoom in here a little bit, if you want to you could go and say okay I want to do up to 48, but you can see that it's pretty significantly reduced right, and this is like the key people who are involved in this in this discussion and if you wanted to check on who

that is actually you can go in the bottom of the the graph area click on the “T” and you can see some of the usernames popping up, but it's a bit tight so we don't want want to see that right now we could run the Yifan Hu again, okay it's not really doing much for us, that way we need to spread things out a little bit, and to do that we can change the optimal distance and the relative strength between the nodes and sort of the forces that are being used and we can say changes to 500 anything press run and we can see the main people here in this area... and we can see DeRay is here and we can see Shaun King and some other, some more prominent people or is advised Kaepernick, yep, Kaepernick 7 the 49ers quarterback, we see it's also Tulsa police, so a lot of interesting things that we're seeing in here. Okay? But 48 is really too too too- it gives us too few people so we want to slide this back to, you know, maybe five and say and you can choose whichever you want for yours, depending on the size, and what you've got five is a nice number here I'm going to run this again, whoa, and we can see everything just sort of burst out and that's really just cool watching that happen okay now we saw we took a look at some of these samples in class so you'll recognize this grouping right here based on what we talked to that in class the other day and you'll see some of the other things appearing then, so that'll look a little familiar to you. So here we have a nice, you know, grouping so after we've done the filters and we've created the we sort of got the map in a nice sort of starting place over here, we go over to our statistics and we want to run some statistics on this, on this, so that the software can interpret it for us and help us generate, generate it. So we want to calculate the network diameter, so we just click run and okay, and again we want to make sure it's as the directed okay I may see we have network down there, 14, and that's great, average path is 4 users, and then we want to also click on the modularity.

Now modularity we have some options on the resolutions, the lower the number, the more communities we're going to get, uh if you've got about, it's good to start with number one and see what it gives you, you always go lower than one or higher than one if there are too many. A good range is between like nine and... and thirteen communities is a nice range and get some nice color distribution so, okay, and this one gives us a community of 10 so that's great, and close that then we go over to partition, nodes, and partition, we choose modularity class. And you can see it gives us a range of colors right here and this is what will then color are different communities the modular class and the different relationships between the different ones if there are a couple that don't have colors you can just click on it and it will bring up a nice little color map and you can choose which colors you want to use.

I want try to make sure they're unique, not the same as the others, I don't like them to be too dark because they overwhelm, and then once you do that you can apply and you can see that some of the, you know, the nodes and the edges have started to turn, and have been assigned a different color based on their relationships to one another now sometimes the colors are a little too light like this one, number five, it's a little too light so let's make it that and this one's a little too gray for my taste so I can change that to something like that, a little to pastel-y, right, you can play around with them all day long, okay, so that is that, now what we want to do is want to click on this sort of the rings this indicates that these are the node sizes we can change the node size based on what we're looking at and we click on ranking and we want to choose for our purposes for this tutorial I want to show you the difference with in-degree and out-degree...

So these are tweets where a user is mentioning another person or is being mentioned by another person. So you can change the sizes I like to have a nice distinct size relationship, and so minimum size the node is 10, maximum size of 100, I click apply and you can see that the users who are being mentioned quite often are larger than the users who are not mentioning, all that much. And you can see here that a lot of them are now overlapping one another are on top of each other because the sizes have changed, so you can go back to the layout area and then pull down menu click no overlap and click run and it will just move things around and out of each other's way so we can see what's going on.

And we see now, that we have we can tell who is at the center of this conversation this one right here, these other people, people are mentioning this one quite a bit, because it's in-degree so the arrows are going in, so these are people who have been retweeted for example quite often, we can highlight who that person is, and just that person, if you click on the little up-arrow down the bottom over here, we can click on the "T" which will give us all the user names but then under the labels tab, click "hide non-selected" so that means only the one who is selected shows up, I should say that it is much easier to use the software when you do have a mouse because you can zoom in and out very easy if you want to move around the entire, if you want to move everything around like, that you right click down and you can and you can move it it just much much easier when you have that mouse. So we can always go and take a look and see who the other large ones, so Deray has mentioned quite a bit, Kaepernick, and I don't know who a lot of these people are, but again as we talked about in class, you can then go and take a look and see what's going on. Here's Hillary Clinton being mentioned quite often in relation, and trying to figure out okay why is Hillary Clinton in the same, here's a CNN same color so we would go and try to figure all that out by, going back to the data. Okay so this is the in-degree and we can see Shaun King is the the one who's being mentioned quite often, or most likely being retweeted quite a bit.

If we do the out-degree, these are people who are tweeting mostly, most often, we can click apply and you can see how that changes. Right? I'll do it again, and in-degree go out-degree, and you see how this changes. Shaun King has now been reduced down quite a bit, and other people are doing the people who are doing the retweeting considerable amount are talking about things quite a bit are now much larger. And again to avoid all this overlap, we can click go to the overlap area and click run. And that will spread things out very nicely for us. And we can see that there are distinct communities that are sort of appearing. This again is the one that we mentioned in class the other day, talked about that grouping, you can see this grouping is here we would want to consider why this grouping... you will see and some of these that there are little arcs that are, go from itself to itself, this means that they mentioned themselves in a tweet. And either they retweeted themselves or they've applied and their username was in there. We can also see that there are arrows pointing, so we know that what the direction of the arrows, then we know what direction it is.

Okay so now we this is a nice this is a nice map it might be a little crowded for some, for some, but for our purposes it's a it's a nice starting starting place once you're you're happy with this layout I strongly suggest you going, and saving. Going to file and save, and now put it somewhere where you'll be able to find it again- I'll just throw it onto my desktop for now, hm, I had a bachelor archive in there at some point. And here I would save it as Terence Crutcher, and

it will save it as a Gephi file. Now when you do this work, you want to take notes on exactly what you have done. You want to note down, what the settings are for Yifan Hu, that you created you know you want to write down your degree range is you want to make sure you write down all this stuff, so that you know exactly, in your notes, assign it to a specific file name, so you know what you need to complete it. So i'll save that and then i could open this again on this computer or another computer one I need to work on it later on. To create to get the to get a map that we can now export, will go to the preview area, and it will bring you here it's blank you just come in, you click refresh, and here it will have the map for you. And you can adjust the size slightly- like that.

Okay, for some reason I'm getting a rendering error right now on this map it was doing that before I'm not sure if that is it's just sort of the complexity of the map it's getting at that error right here. I'm not sure if it will do that, or keep it the way when it saves it or not, but you can if you don't like these curvy lines you can uncurve them by clicking over here and click refresh and then I will show you the arrows like where the direction of the conversations that are taking place again the thickness of the line indicates the number of the size of the connections between the two, I tend to like the curvy arrows better... you can rescale the weight of the lines so that they're all the same, but I tend to like the the weight. You can even turn off the edges if you just want to see the nodes so go "constellation effect." I like to see that, you can change the thickness if you want to, making it about 10, and then you get some thing... it's ridiculous, doesn't really do anything for you.. Refresh that'll bring it back down.

You can show the labels if you want to and these are proportional, so the bigger more often a username, the bigger the node is the bigger the username is, that's a little annoying but you can go and edit the size and the fonts you know whichever one you want typeface that you want to use and you can choose a default size so they're a little bit bigger, click refresh, not too much bigger, but there they are, okay and when you're happy with what you've got and you're happy with this situation you go down here and you click on the export the bottom left hand corner, click on export, and you can go and you know "TC" I think this was out-degree, and I'll say 5 to 118 because that's the degree range that I have, and I want to save this as a .png or because they're whatever you want but for to be able to upload it to the website and stuff you'll save as .png, and click save. This is out because I change the file type, "TC out-degree", and that's it. I could save it, I can tweet it, I can do anything I want with it. If you're not happy with how this looks here, you can always go back to the overview and play around with the colors and change things. But that's a very quick way of creating an in-degree and an out-degree mentions map using using Gephi from DMI-TCAT. Good luck, and I look forward to seeing yours.