Module 4

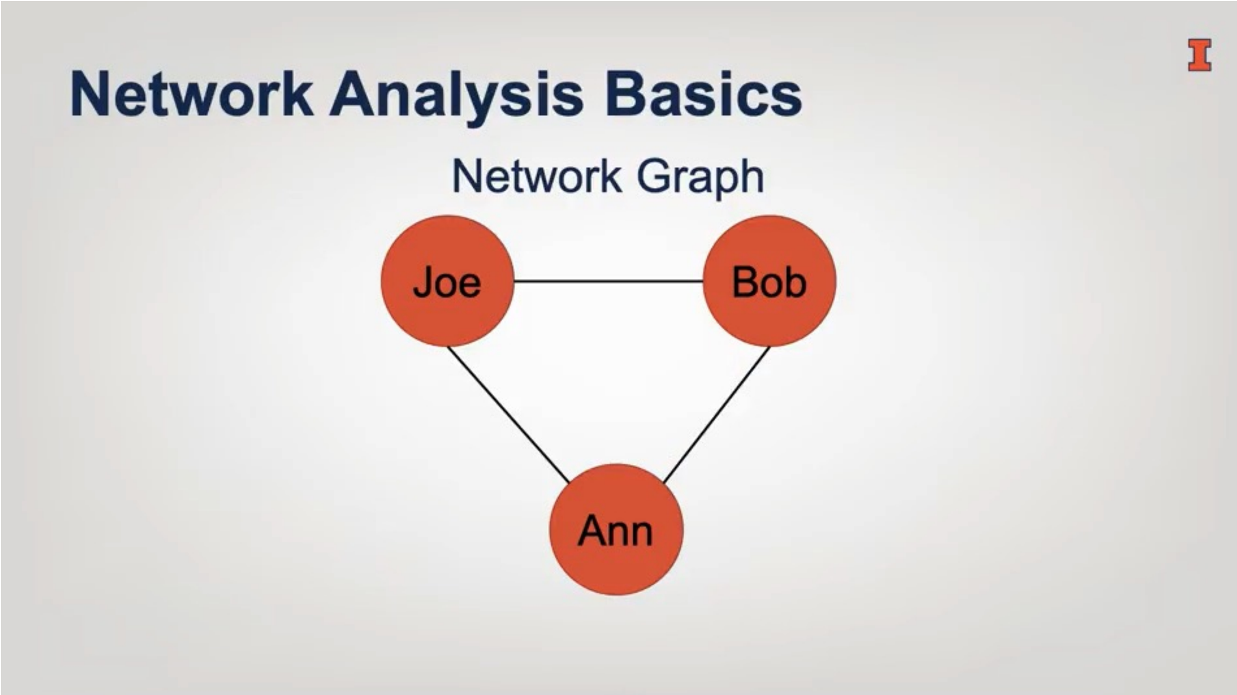
Table of Contents

Introduction to Valuation Analysis 2
- Network Analysis Basics 2
- Analyzing Influence with Network Analysis 7
- Network Analysis with the SMM 14
- Influencer Brand Personality Analysis Part 1 20
- Influencer Brand Personality Analysis Part 2 24

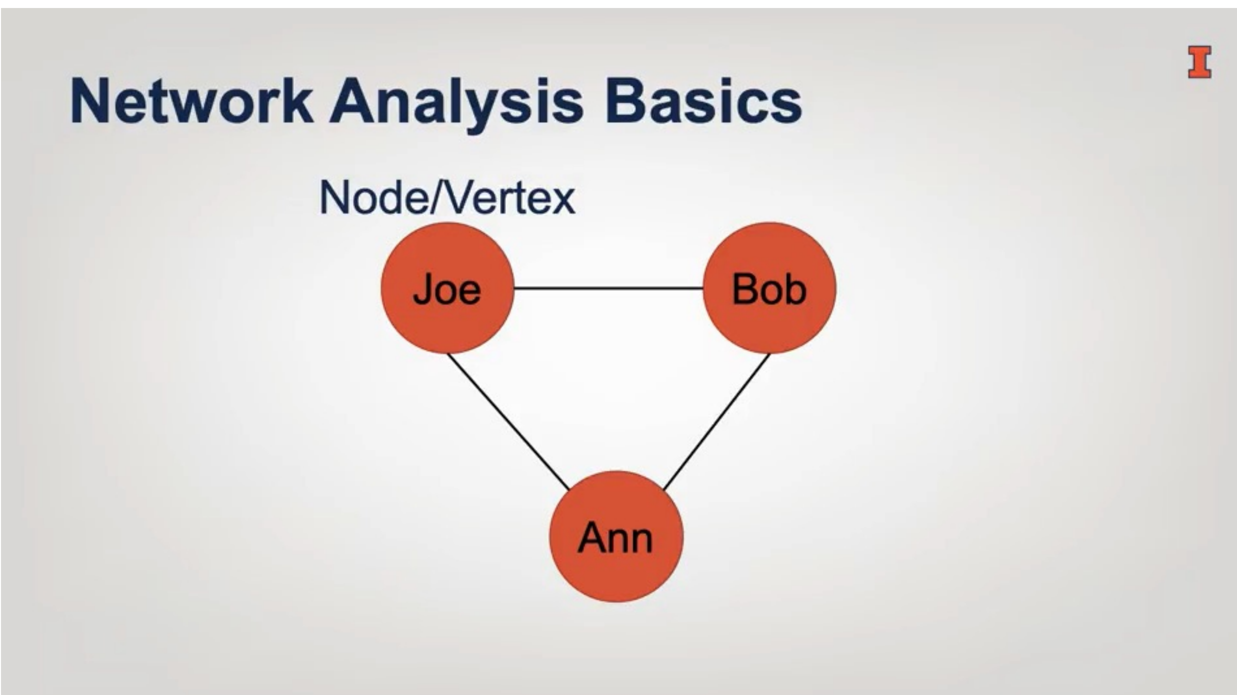
Introduction to Valuation Analysis

Network Analysis Basics

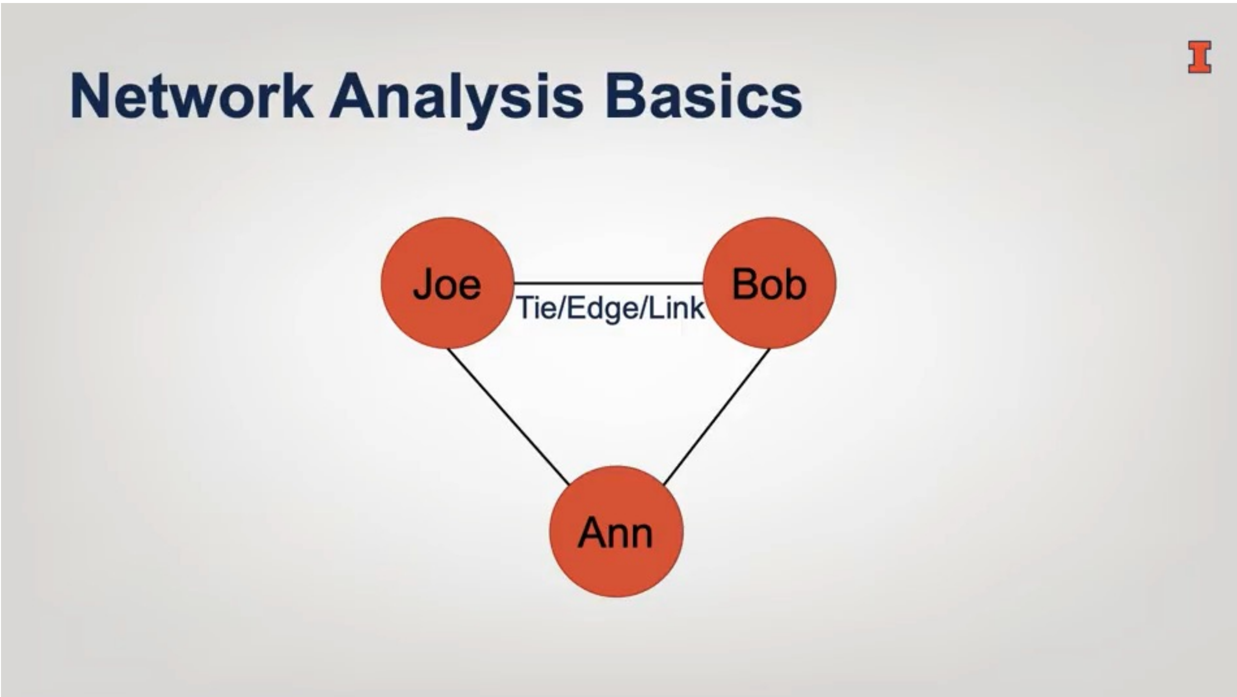
In this module, we will now look at customer satisfaction influence analysis by using this tool of network analysis. This is a bit of an intro to network analysis, some basics. First of all, when we're engaging in network analysis, what types of data should we analyze? Well, the list is not that long, at least here. I suggest social media data, because if we're trying to understand how people are interacting with other people and forming networks, then really social networks are perfect dataset to analyze this phenomenon. But there's, of course, more datasets, I'm just suggesting that those datasets are probably not easily accessible publicly. Your company might have datasets of how people interact with one another regarding customer satisfaction. But for this course, we will look primarily or solely at social media data. Now, a definition or a conceptual understanding of network analysis can be found here. I'm not going to read it all, but just the first few sentences. The basic idea of a social network is very simple. A social network is a set of actors that may have relationships with one another.



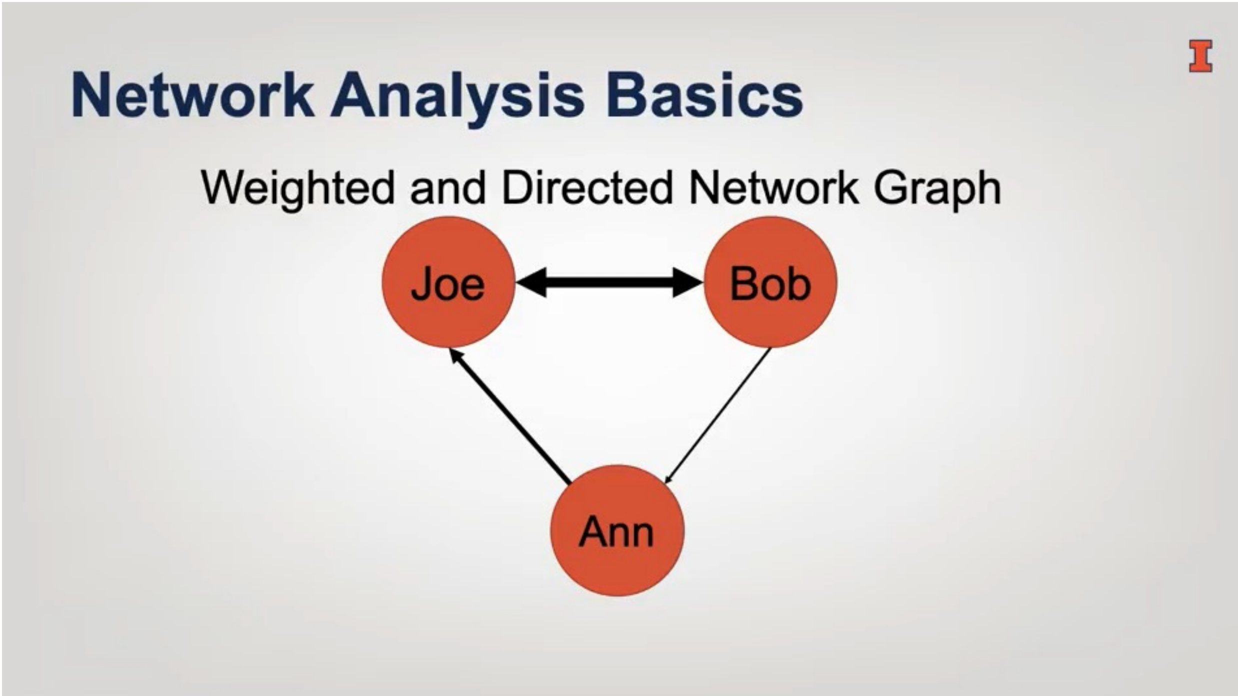
Now, in network analysis, usually these networks are visualized. This is a visual that signifies a very simple network. Now there's various terms used in network analysis that signify the same concepts, and so I'll try to use as many of them as possible just so that you know all of them or at least many of them that are used commonly. This is typically called either a network, or a graph, or they combine the words and they say network graph.



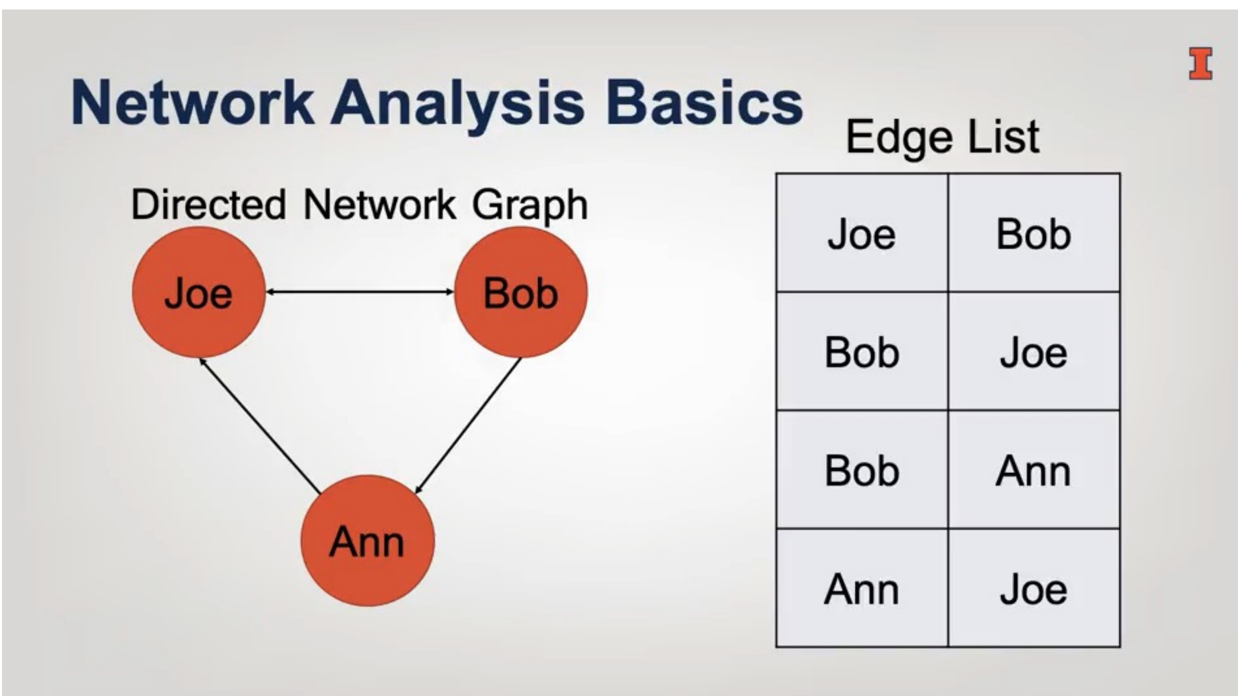
Network graphs have components, these components are, first, nodes or vertexes, and again, the same concept, different terms. Joe is either a node or Joe is a vertex. In this case, we're using nodes as people, but nodes can be anything, nodes can be organizations, nodes can be animals, anything that can interact with something else. Nodes can be bacteria, so on and so forth.



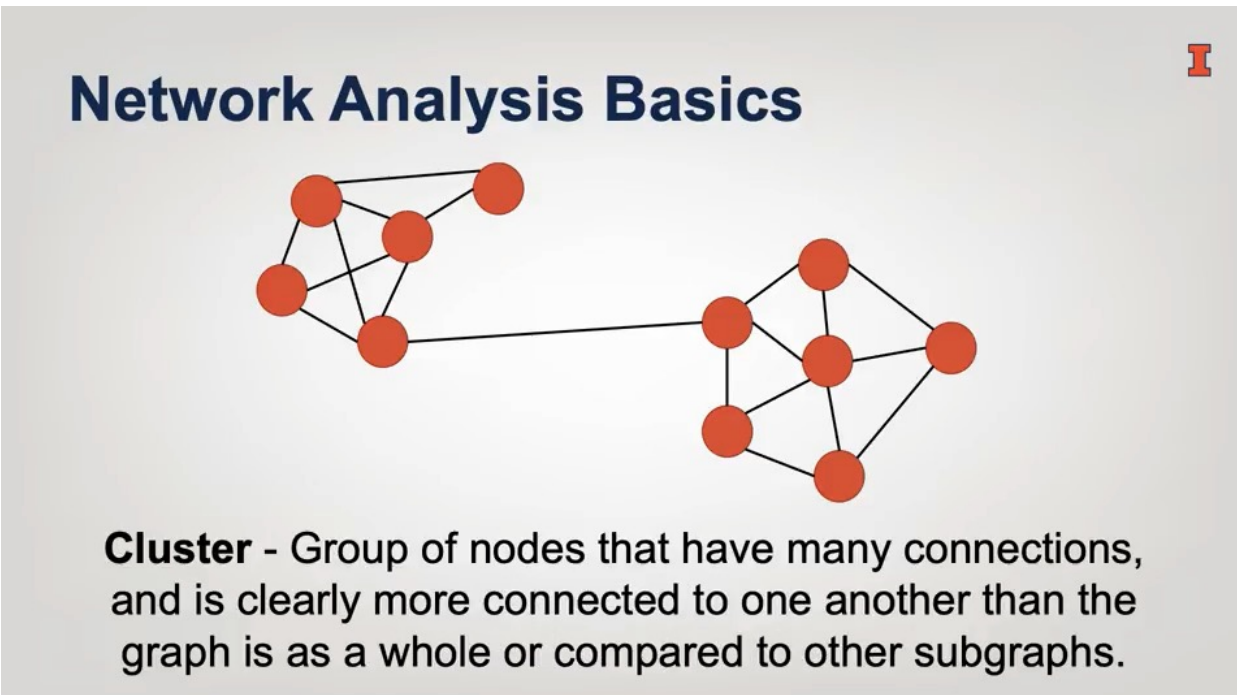
Nodes are connected via ties or edges or links. Again, same concept, different terms. Now, you can define what these ties represent, what these edges represent. Let's define these edges as friendships. This is a friendship network. Joe is friends with Bob, Bob is friends with Ann, Ann is friends with Joe. Now, edges can either be weighted or unweighted, and if they're weighted, this is called a weighted network graph or a weighted network or a weighted graph, again, the combinations are bound. But what does weighting signify? If the edge signifies a friendship, then the weighting signifies the intensity of that friendship. In this case, Joe is best friends with Bob. Bob is maybe an acquaintance with Ann, and Ann is friends with Joe but not best friends. This is what this weighted network graph signifies. Graphs can also be directed. What does that mean? Well, if this is a friendship graph, this is sad. Because what this means is Joe believes that Bob is a friend and Bob believes that Joe is a friend, it's a mutual friendship. But then in the case of Bob and Ann, Bob thinks Ann is a friend, but Ann does not think Bob is a friend. In the same way, Ann thinks Joe is a friend, but Joe does not think Ann is a friend. Now this is a sad graph, but it's a good example of a directed graph. Graphs can have direction.



Then we can combine all of this and do a weighted and directed network graph where Joe and Bob consider mutually each other to be best friends. Bob considers Ann to be an acquaintance, Ann does not consider Bob to be an acquaintance. Whereas Ann considers Joe to be a friend, but Joe does not consider Ann to be a friend. You can see you can play around with these components to represent different relationships between entities, between nodes or vertexes using edges and these edges can be weighted and they can be directed.

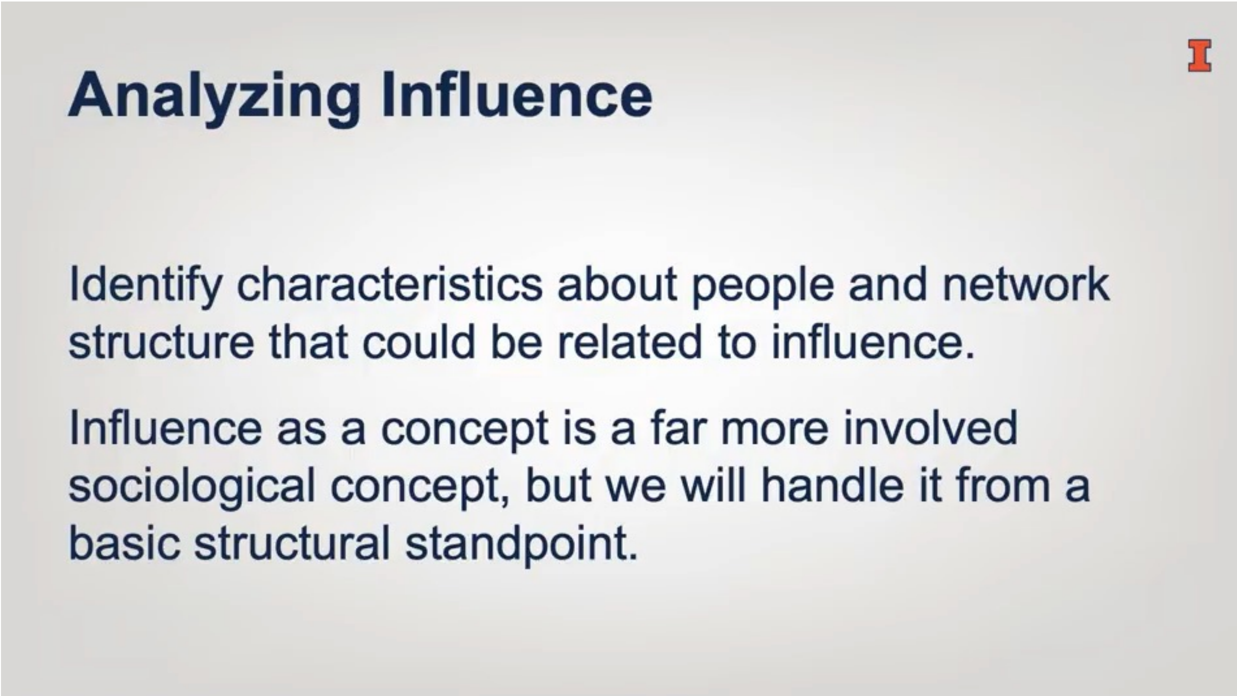


You can represent a network graph as well from a visual as well as writing down the network graph as an edge list. What you see here is a directed network graph on the left and an edge list that represents that directed network graph. You see Joe, Bob, and in the next row you see Bob, Joe, why? Joe and Bob listed twice. Well, it's signifying in a directed graph that there's a direction going from Joe to Bob, but then there's also a direction of an edge going from Bob to Joe. Then we see Bob, Ann, but we don't see Ann, Bob, because again remember Ann doesn't consider Bob a friend. Then we see Ann, Joe but not Joe, Ann because remember Joe does not consider Ann to be a friend. You can represent this visual on the left via this edge list on the right. There's also names for special networks. Now I don't know if this is special, but in the case of certain phenomenon, it could be special where you have a singleton, which you might have a point or a vertexes a node that is just by itself, it's not connected to anybody else. That might be an interesting phenomenon depending on what you're studying. That's usually called a singleton amidst all other nodes and vertexes. You might also have what is called the click. We've heard that term probably before in society. But in network analysis it's not necessarily an exclusive group of people, it's just an interesting group of nodes, and all these nodes must be connected. Let's say Joe, Bob, and Ann for whatever reason in enlarge or picture that Joe, Bob, and Ann form some special function, so they're interesting, and so they would be called a click.

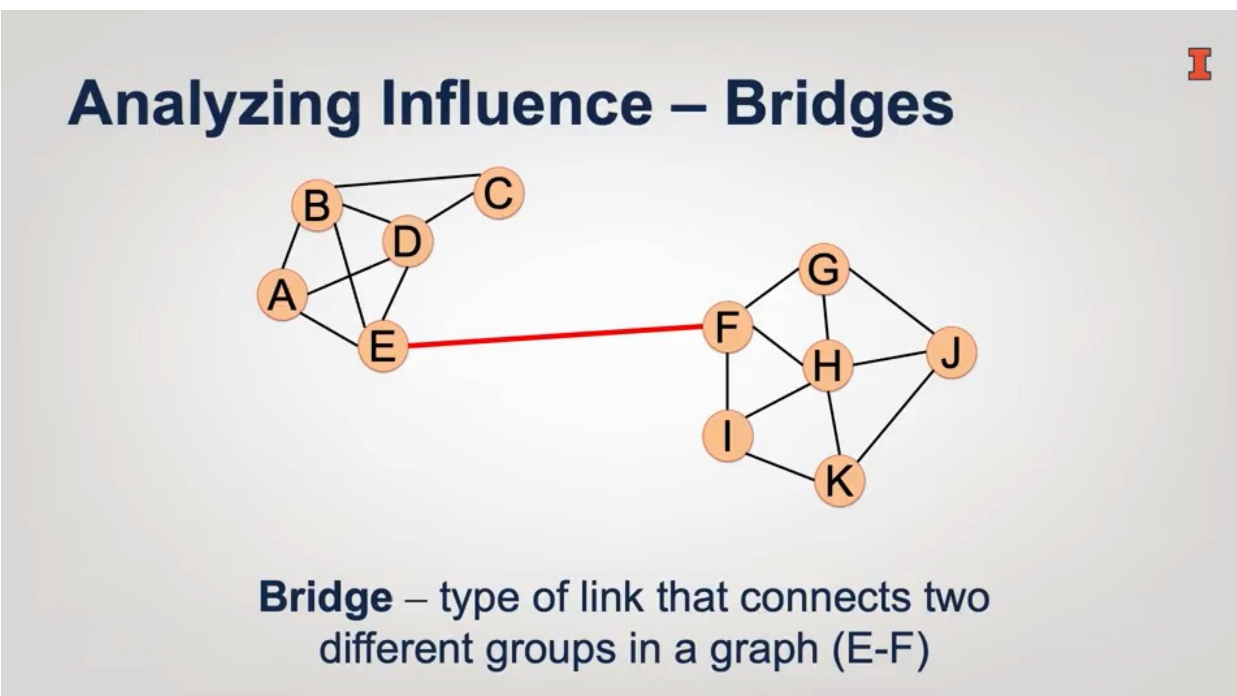


Then we'll see clusters. Let's pretend Joe, Bob, and Ann are just part of this larger graph, this larger network. You see that this larger network even visually there seems to be two parts to this larger network, and those are called clusters. They're group of nodes that have many connections and it's clearly more connected to one another than the graph is as a whole are compared to other sub-graphs. Clusters are very important phenomenon within network analysis because you'll see separation between groups of people for whatever reason, and that might be interesting to what you're studying. These are just a few things that form a basic foundation for network analysis.

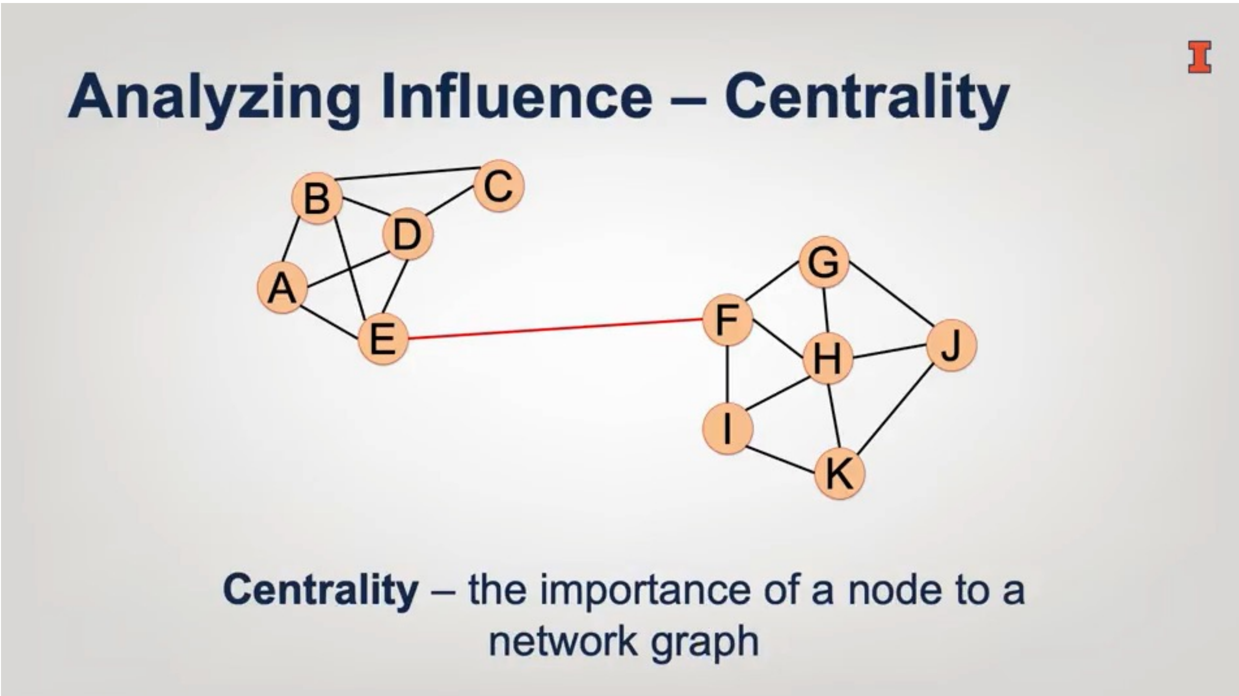
Analyzing Influence with Network Analysis



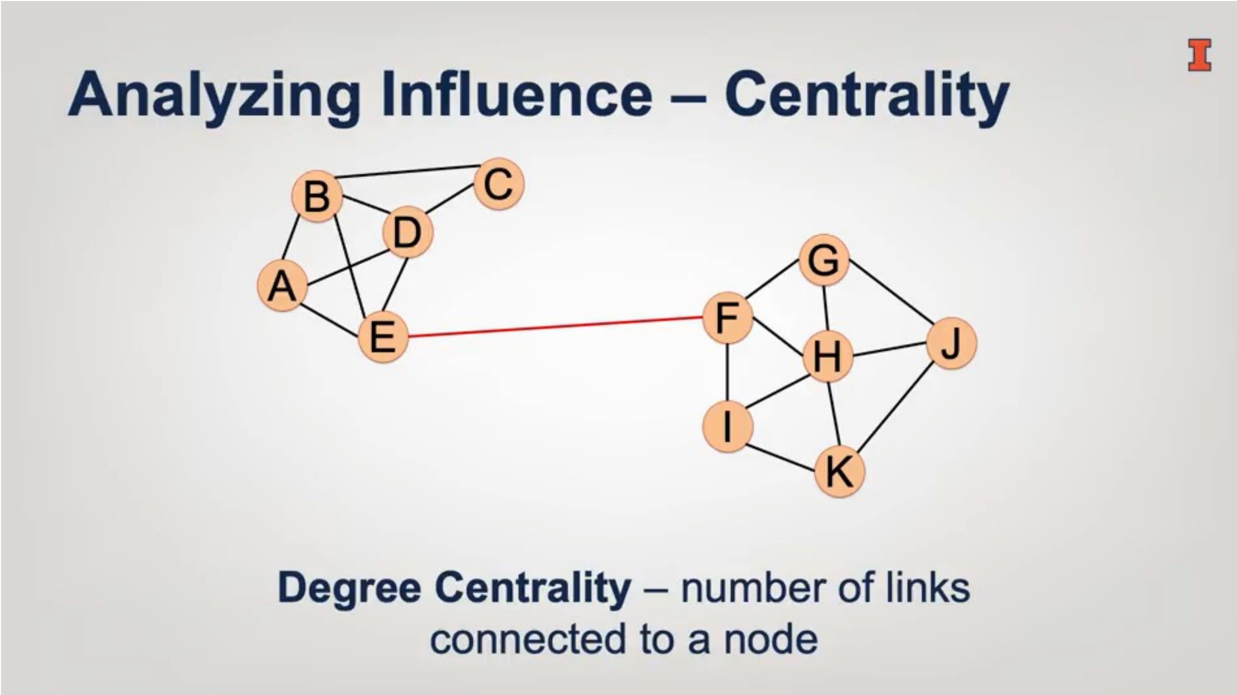
Everyone, now that we've formed the foundation for the basics with network analysis, we will look at analyzing influence with regards to customer satisfaction. Here are some things that we have to think about with regards to analyzing influence in regards to customer satisfaction. We want to identify characteristics about people and network structure that could be related to influence. Now, influence is of course, a concept that is far more involved sociologically. But we are just going to handle it in this lesson from a basic structural standpoint.



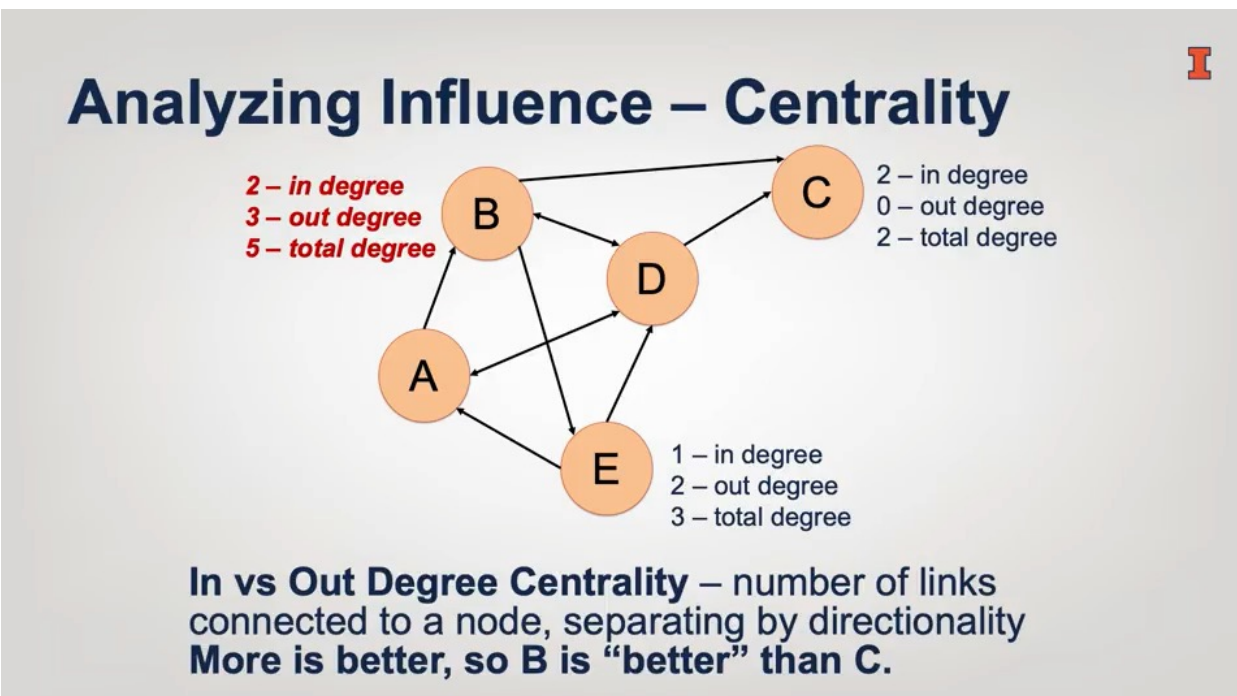
Let's think about it with regards to network analysis or network graphs. Remember this diagram or this graph from the previous lesson that we were talking about clusters. But I don't know if you noticed that when you saw this graph, there's this interesting phenomenon that's happening in this labeled graph between E and F. That between E and F, if that edge was not there, this graph would be two graphs. There'd be two subgraphs or two separate graphs. This phenomenon is usually called a bridge. A type of link that connects two different groups in a graph. We can suggest that these individuals, E and F, they hold some power in this graph because if it wasn't for their connection, these two subgraphs would be on their own. We can see from a network graph structure, potential influence or potential power positions at play here.



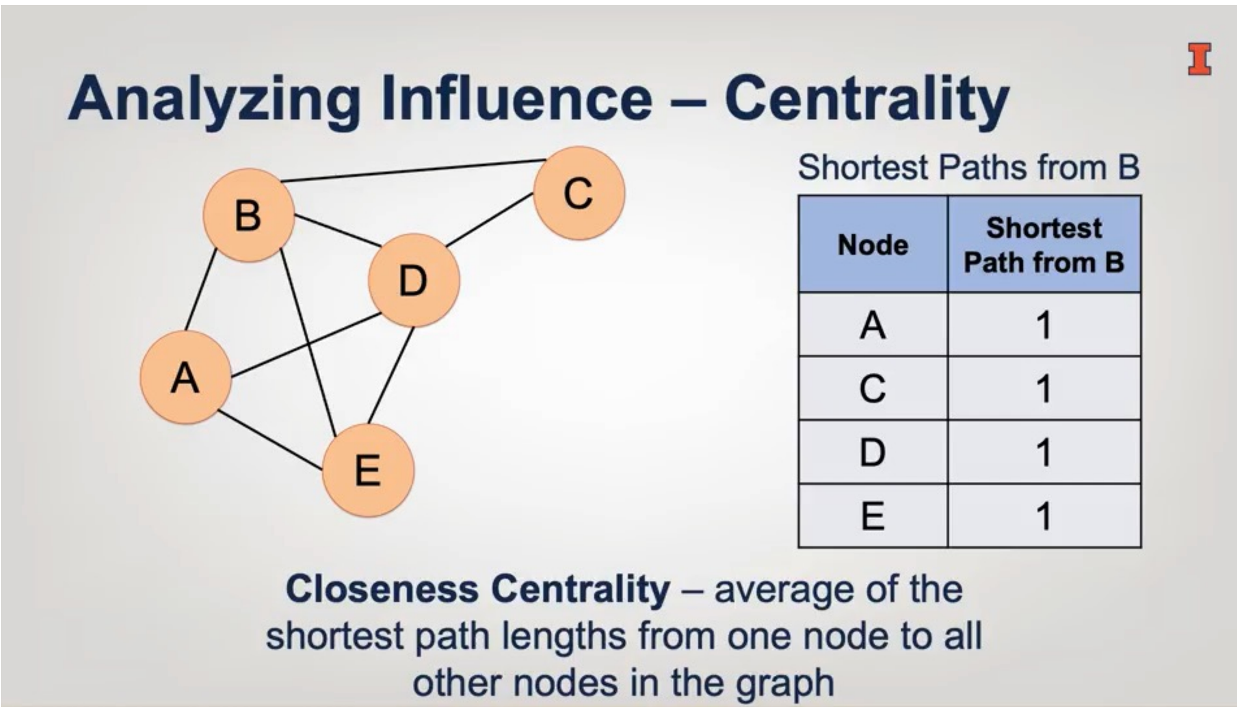
Another concept with regards to analyzing. H has five edges coming into it, and so the degree centrality is five and with J degree of centrality is three. influence when it comes to network analysis is what is called centrality. Centrality is the importance of a node to a network graph.



Now there's various forms of centrality. We'll look at degree centrality first. Degree centrality is the number of links connected to a node. What do I mean by that? Here are some examples. Let's look at C on the top left-hand corner of this slide. How many links are coming into or edges are coming into C? Two. The degree centrality of C is two. How many edges are coming into E? Four. The degree centrality of E is fourThis could be a measure of influence with regards to any one of these nodes that maybe we say H, because it's got a high degree centrality, there's some power influence that H has on this overall network as compared to other nodes. Now degree centrality can be broken out into in-degree and out-degree.



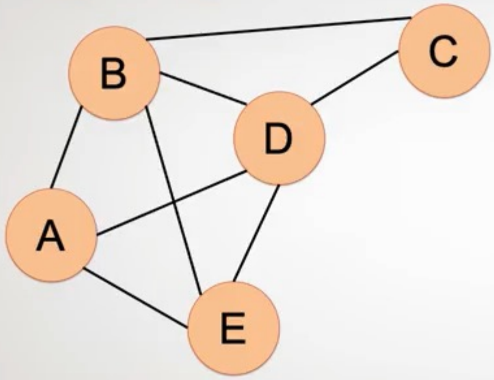
In-degree is the number of links in a directional graph that are coming into a node and out-degree is a number of links or edges coming out of a node, again in a directed graph. Let's look at B. Node B has two edges coming into it, and so it's in-degree is two. But it has three edges coming out of it. We have B to E, B to C, and then it's hard to see B to D. There's three coming out of B. The total degree of centrality of B is five, but in-degree is two, out-degree is three. Now let's look at node C. You have two edges coming in, so its two in-degree, zero edges coming out so it's zero out-degree. But the total degree is two. Now if you're looking at total degree, B seems to have more influence or power than C. But let's say for whatever reason that you're looking at a phenomenon where in-degree is the most important thing. Then B is actually equal to C because B has two in-degree and C has two in-degree. You see how when you're looking at in-degree versus out-degree with a greater degree centrality, you could look at different ways to slice and dice influence or positions of power within a network. Now, there is another centrality measure called closeness centrality.



Node	Shortest Path from B
A	1
C	1
D	1
E	1

Now closeness centrality is the average of the shortest path links from one node to all other nodes in the graph. Basically how quickly can one node get to all the other nodes within a graph? Let's look at the closeness centrality for B. What's the shortest path from B to A? You see this table on the right-hand side signifying shortest paths from B to the node on the first column, which B to A is one. There's an edge that goes directly to it. B to C is one, B to D is one, B to E is one.

Analyzing Influence – Centrality



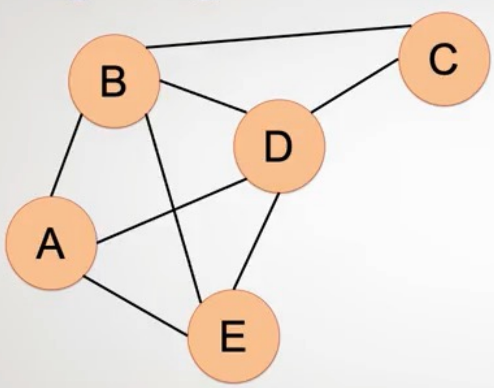
Shortest Paths from B

Node	Shortest Path from B
A	1
C	1
D	1
E	1

Closeness Centrality of B – 1+1+1+1/4 = 1

How you calculate closeness centrality is you say 1 plus 1 plus 1 plus 1, which is all the shortest paths from B to the other nodes in the graph and you divide it by the number of nodes minus one because we're not counting B to B. 1 plus 1 plus 1 plus 1 divided by 4 equals 1. The closeness centrality of B equals 1. Now what about the closeness centrality of E? It's slightly higher than the closest reality of B because from E to A is one, E to B is one. But what about E to C? We see in this table that it's actually two.

Analyzing Influence – Centrality



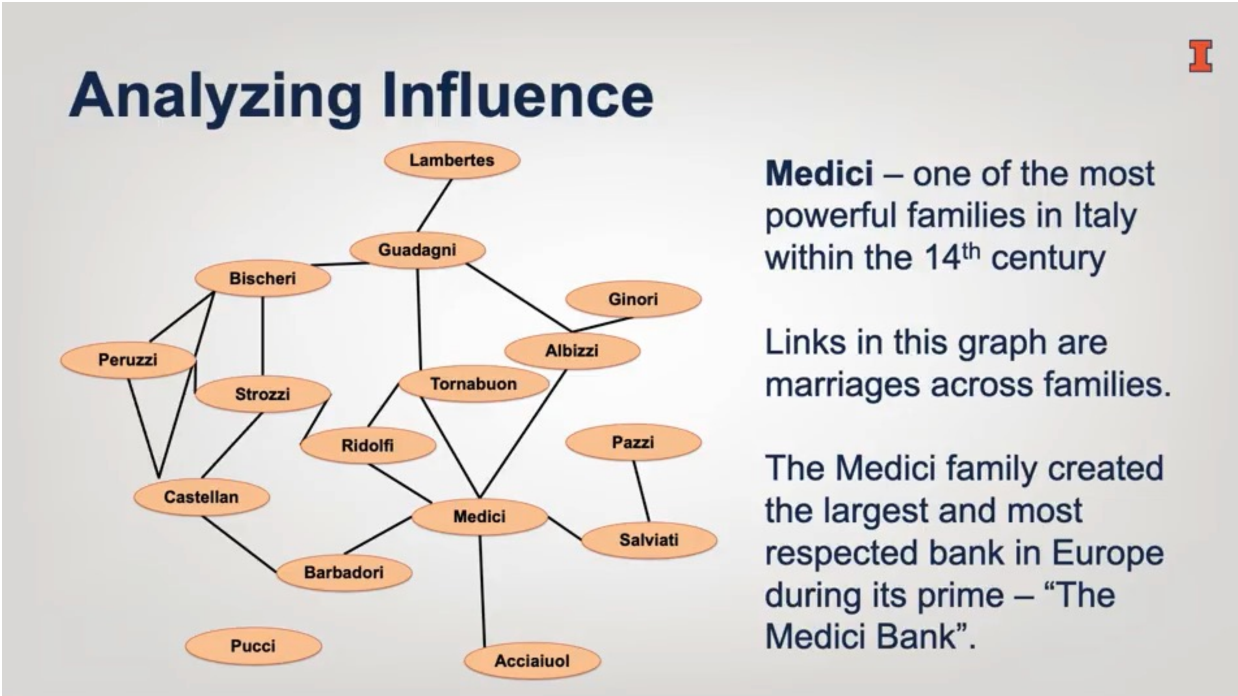
Shortest Paths from E

Node	Shortest Path from E
A	1
B	1
C	2
D	1

Closeness Centrality of E – 1+1+2+1/4 = 1.25
Less is better, so B is "better" than E in closeness centrality.

E cannot get to C without going through either D or B. You have 1 plus 1 plus 2 plus 1 divided by 4 equals 1.25. Less is better. B is better in a sense than E in closeness centrality. This all sounds theoretical, but let's say that you want to get a message out to everybody in the graph as quickly as possible. Let's say its emergency message. You ask yourself, who's the best person to place that message with first? I can tell you right now it's probably B over E because E's got to take way too many steps with regards to getting to C than B does. You might place a message with B and say, we need this emergency message to get to everybody. B is closest to everybody compared to E. That's a measure of closeness centrality. Now there's a bunch of other measures of centrality, betweenness, eigenvector, etc.

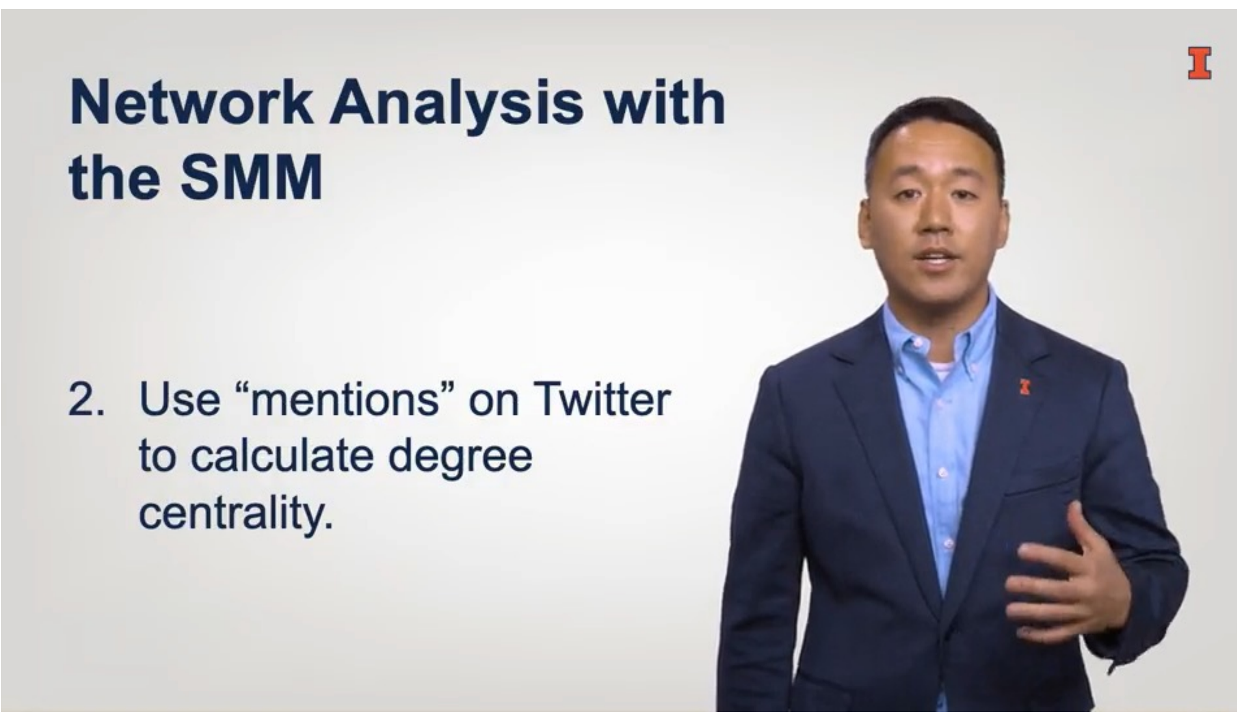
We're not going to cover those, but you can read up on those. If this is interesting knowledge to you, eigenvector is what actually calculates the relationship of pages on the Internet and Google originally used eigenvector centrality to understand how influential pages were to each other and then that changed how Google search ranked various suggestions when you search something. Eigenvector centrality is very much used online.



With all of this said, I've talked about some measures, bridges, we've looked at centrality as various ways to use network structure to understand influence. I'd like to close with this example, this very real-world example with regards to how network structure was shown to explain why a certain family, The Medici Family, which was one of the most powerful families in Italy, held so much power within the 14th century. Now, in this network graph, the nodes are families, and the links of the edges in this graph are marriages across families. Apparently at that time, they would strategically marry off their family members to other families to form a trust relationship with those families. What you see here in this network graph is by that lower right-hand quadrant where The Medici Family lives. That there are certain families that cannot get to the rest of the other families in this Italian family network without going through the trust network of The Medici Family. When you look at this graph, many scholars say that the way that they made their marriage alliances, when you look at it from a network structure perspective, they became influential. They became powerful with regards to their relationships with everybody else, all the other telling families. The Medici family actually created the largest and most respected bank in Europe during their prime, and it was called The Medici Bank. We can see that if we use network structure to understand influence, that we could potentially use that understanding of influence to understand how people could affect customer satisfaction towards your company, your product, your services, or your brand.

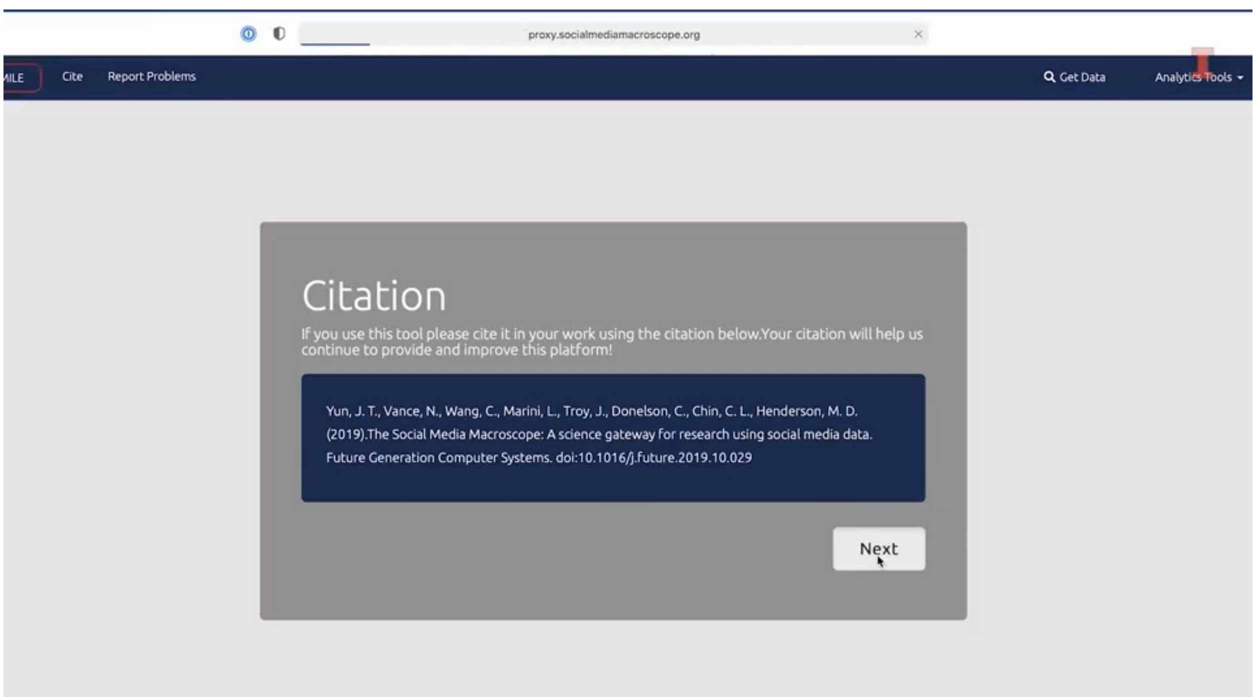
Network Analysis with the SMM

Okay everyone, now that we've gone over various techniques to understand influence with regards to network analysis. We will apply that knowledge within the social media microscope. So what we're going to do here, is that we are going to download twitter data via search term of your choice.

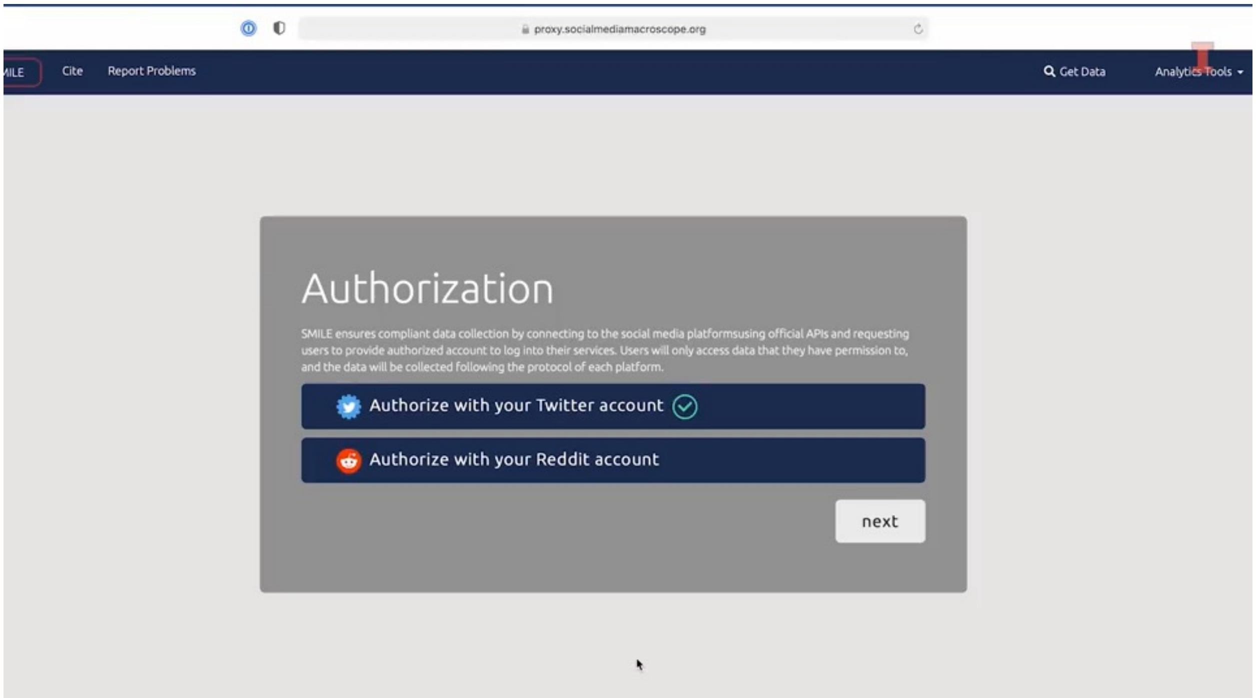


We're going to use mentions as the way to build edges with regards to the twitter data to calculate degree centrality. And then we'll find a few accounts that could be considered as influential within the conversation of your company, your brand, your product or your service.

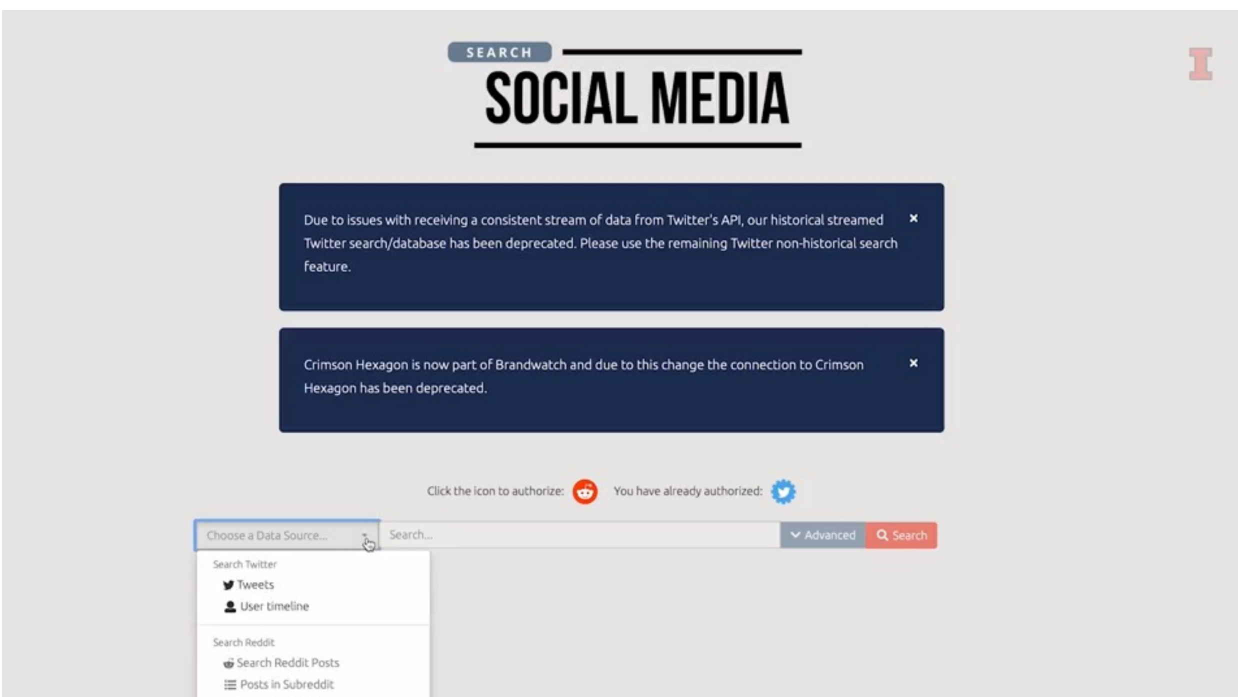
So let's get into the social media microscope and start looking at influence. >> Everyone, we are now here and smile again and we're going to do some network analysis.



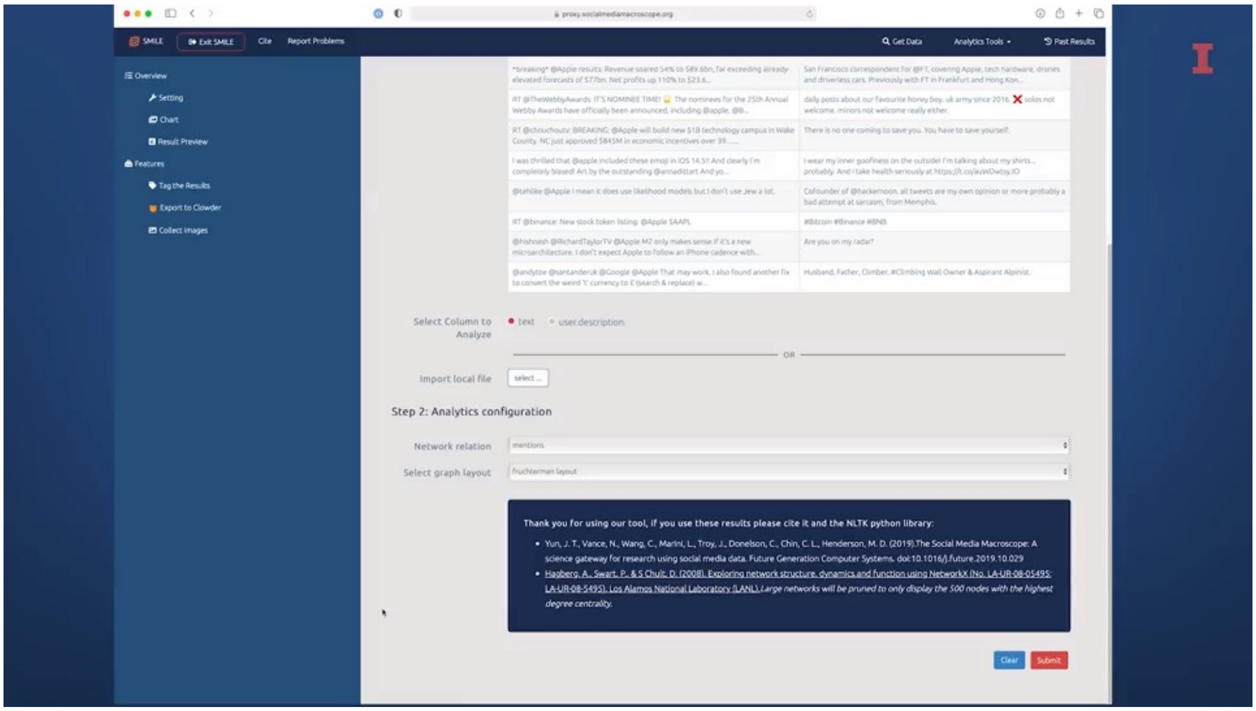
So let's go through getting started again and let's pull data again, just to make sure that we know how to walk through this process. We'll authorize with twitter again, you don't have to do this each time but every time you load up smile again, you'll have to walk through this again.



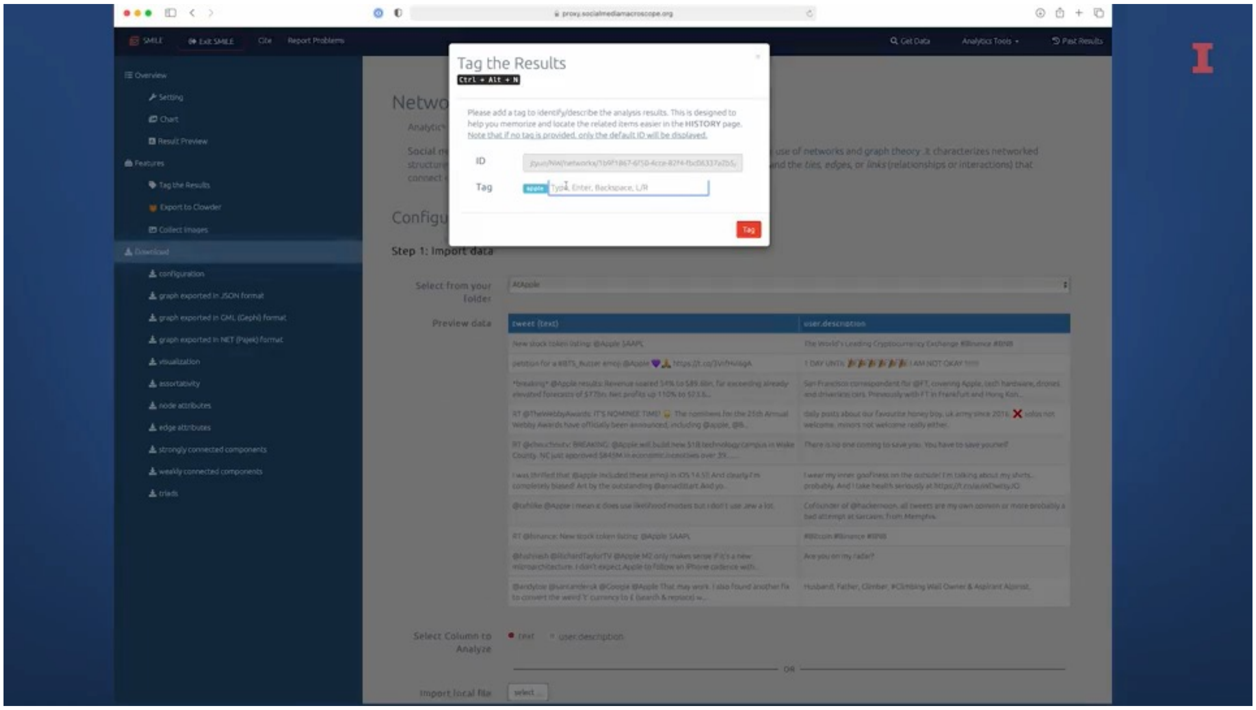
So that's what we'll do. We're going to do tweets regarding Apple again. Search, we see that's correct. Now, actually what you can do at this point since I've already pulled this data and smile will hold your past data pools for quite some time. And so if I go to past results in the top right hand corner and I go to twitter tweet. We'll actually see my at Apple here that I pulled some time ago. I can click on that and that will actually bring up a preview of some of these results.



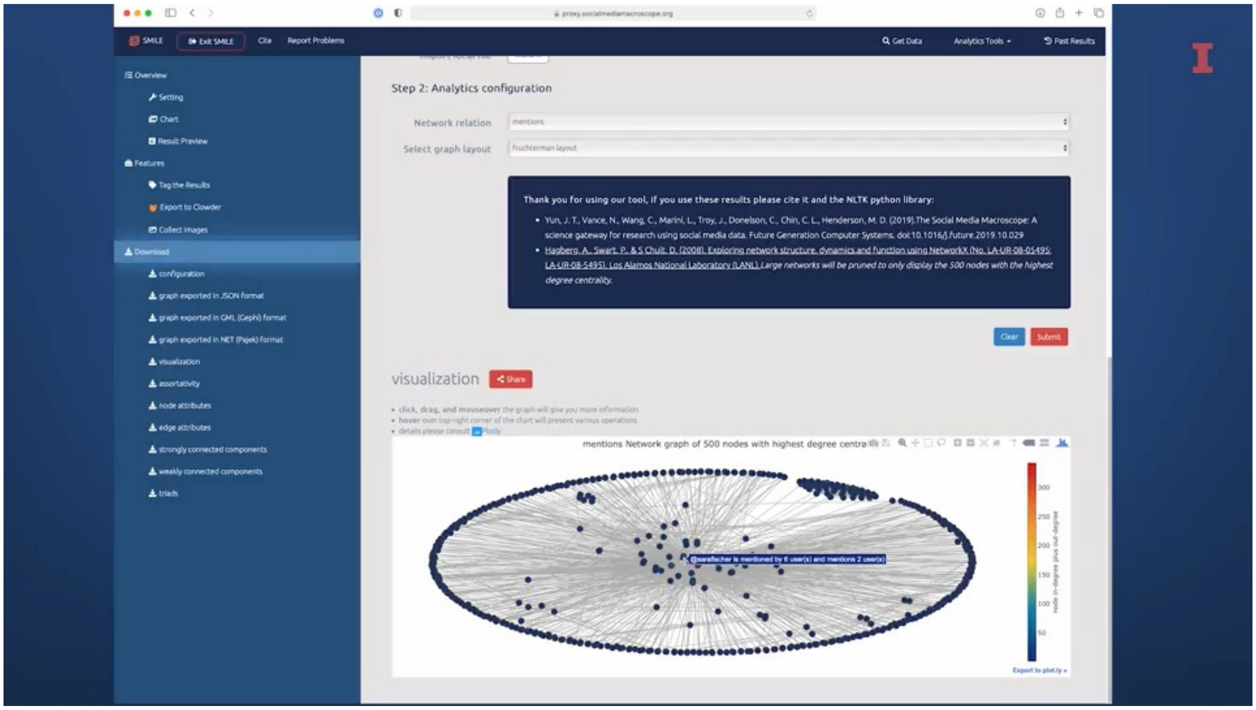
So we'll just use that directly. We can go into analytics tools here, click on network analysis, we'll select at Apple, we'll see some previews of the tweets. We're going to look at mentions and mentions if you've used twitter before. It's when you actually kind of tag somebody by saying at Apple and then you start speaking as if you're talking to the twitter account at Apple. That's how mentioning works on twitter. And so we'd like to use those mentions as edges in a network graph and of course the vertices of the nodes would be the actual twitter accounts. And so the twitter accounts will gain edges built into the network graph if they are mentioning the other edge or the other node, the other twitter account. We're going to select a graph layout.



A front German layout, not to go too much into what that means but basically what a front German layout does brings things with high centrality or nodes with high centrality into the middle of the graph and then nodes with lower centrality out to the outside of the graph. And that's kind of helpful so that we can see which notes have higher centrality. We're also using the network X tool. So if you'd like to know more about the tool that's actually building these network structures, you can follow the link that's found here in this citation box. So let's actually click on submit and then this will take a little bit of time. It's taking all of the tweets that were pulled from Apple and it's running it through the network X, actually python package. And then what it's going to do is it's going to build a very nice via a front German layout network graph for us, so that we could look at degree centrality with regards to mentions as people are interacting with at Apple. And so this computation should take just a few seconds. It may go into a minute or so, but depending on how heavy the traffic is for the site, but you should get results pretty quickly. Okay, that actually took a little bit longer. And so I cut the time or the video during the waiting period. And so what you'll get before you can see the results is another tag box.



So we'll tag it as Apple, we'll, tag it as network and we'll tag it as centrality just so that we can remember what we've done in the past. And what you'll see here is a visualization of the network structure of discussions regarding at Apple. And you see that those with the lowest centrality are pushed to the edges here to the outside, sorry, not to use the term edge. That's confusing since we have all these edges. And then you have these vortices that are close to the middle, these nodes that are very central to this conversation and so you have at Netflix at Walmart and if you hover over these, you can see these. You have at Google, you have at Microsoft, you have at Sarah fisher. Now I'm going to stop here at Sarah fisher because all those others were brand names that we recognize Sarah fisher.



I don't even know who add Sarah fisher is and so what you could do is that you could search Sarah fisher on twitter and try to figure out why is she so central to this conversation? Because clearly she has a higher degree centrality, she's potentially an influencer in this conversation. And so this is just a really simple way. You don't have to use the social media microscope to do this. But once you're able to pull social media data somehow using various a piece and you can pull it into and the network analysis package, it could be with an R or Python Network Access in Python. And then run these kind of centrality analysis to understand who are potential in this case twitter accounts that could be influential to the conversation about your brand.

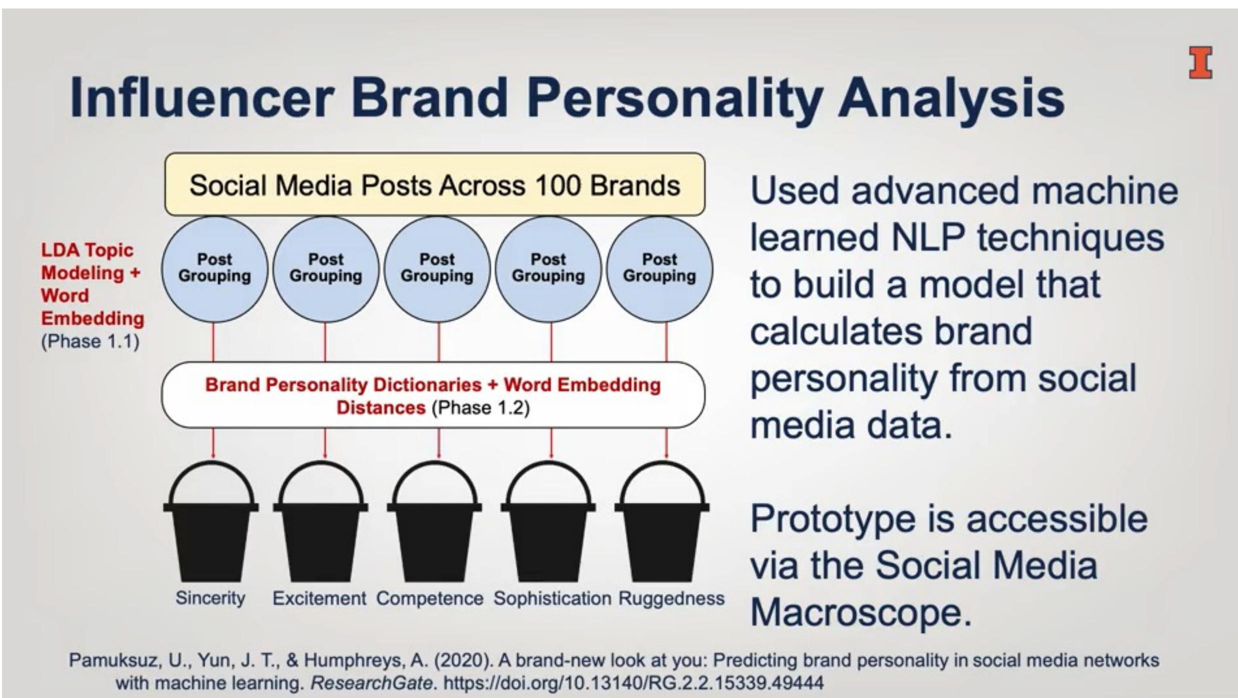
[Influencer Brand Personality Analysis Part 1]

Okay everyone, now that we've looked at centrality measures using the Social Media Macroscope to understand influence, we'll now jump into another topic with regards to influence specifically surrounding this concept of brand personality. Now, what is brand personality? Brand personality as defined by Jennifer Aaker in 97, is the set of human characteristics associated with a brand.

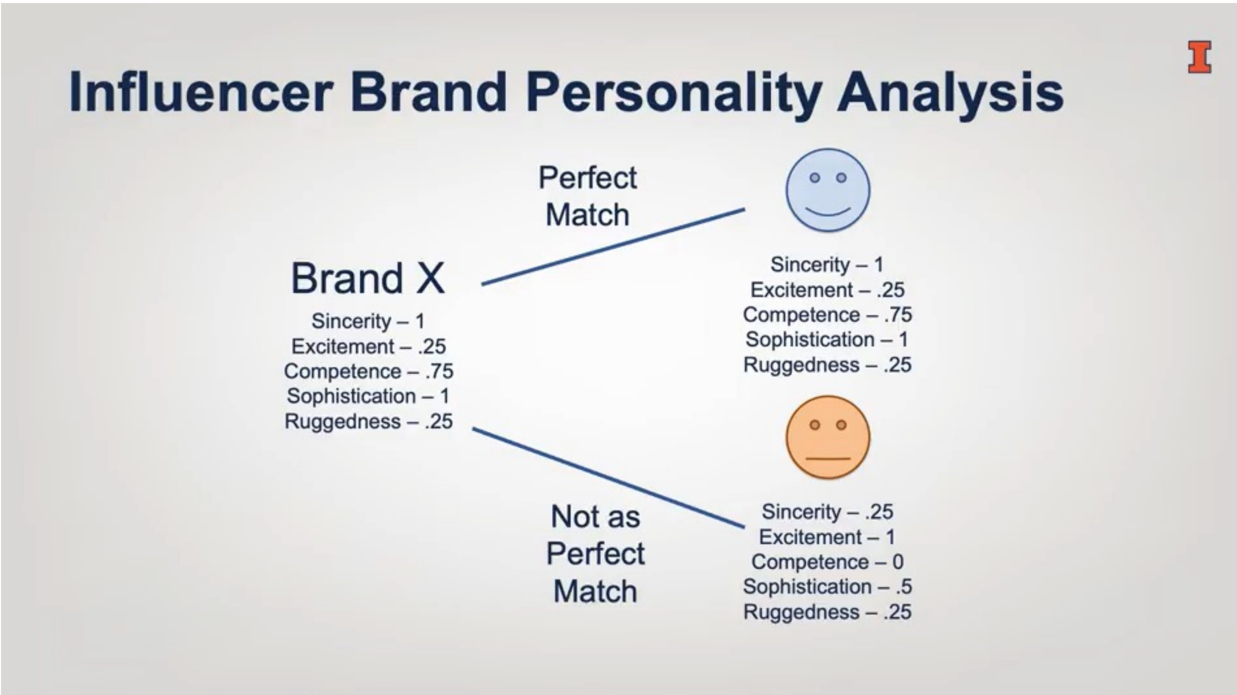
Now she suggested that much like human personality, which is measured in five factors, I'm sure you've heard of openness, conscientiousness, extroversion, agreeableness, neuroticism, she suggested that brands also have a personality, but they have different dimensions. The dimensions are sincerity, excitement, competence, sophistication, and ruggedness. For example, there's just some adjectives that we could look at under each one of these. This is not comprehensive but just exemplary. For example, with sincerity, you have adjectives like domestic, honest, genuine, cheerful. Brands have different ratings according to these different dimensions, how sincere they are, how exciting they are, so on and so forth. Now, a group of researchers and myself, we did a study where we did not use brand personality but human personality and we used machine-learned models that calculate human personality via Twitter feeds. You can read about that in the paper that we provide in this course.

But what we found was when we looked at human personality, machine-learned measures via people's Twitter feeds, and then we calculated human personality of the brands that they followed, we found that people tend to follow brands that are closer in human personality congruent to themselves versus brands that are not.

We thought that this could be a very interesting method for brands to find influencers that are similar in personality to themselves and they could potentially project marketing that would affect customer satisfaction because their brand personality is much more similar to the brand personality of those brands.



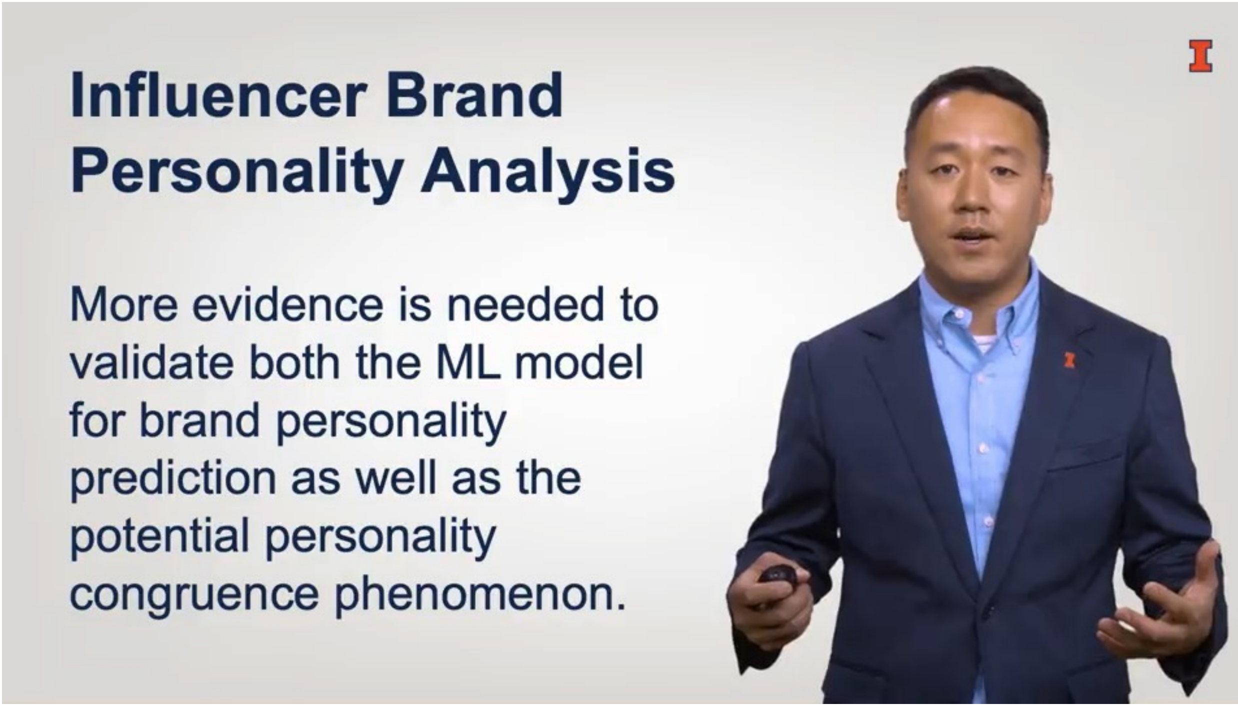
Now in this slide, this is just one of the diagrams from our paper because what we were trying to do is we were trying to create not a human personality model, a machine-learned model, but we wanted to create a brand personality, machine-learned model. This diagram is basically showing that we used advanced machine learning Natural Language Processing techniques to build a model that calculates brand personality from social media data. A prototype of this model is accessible via the Social Media Macroscope of which we will use for this course.



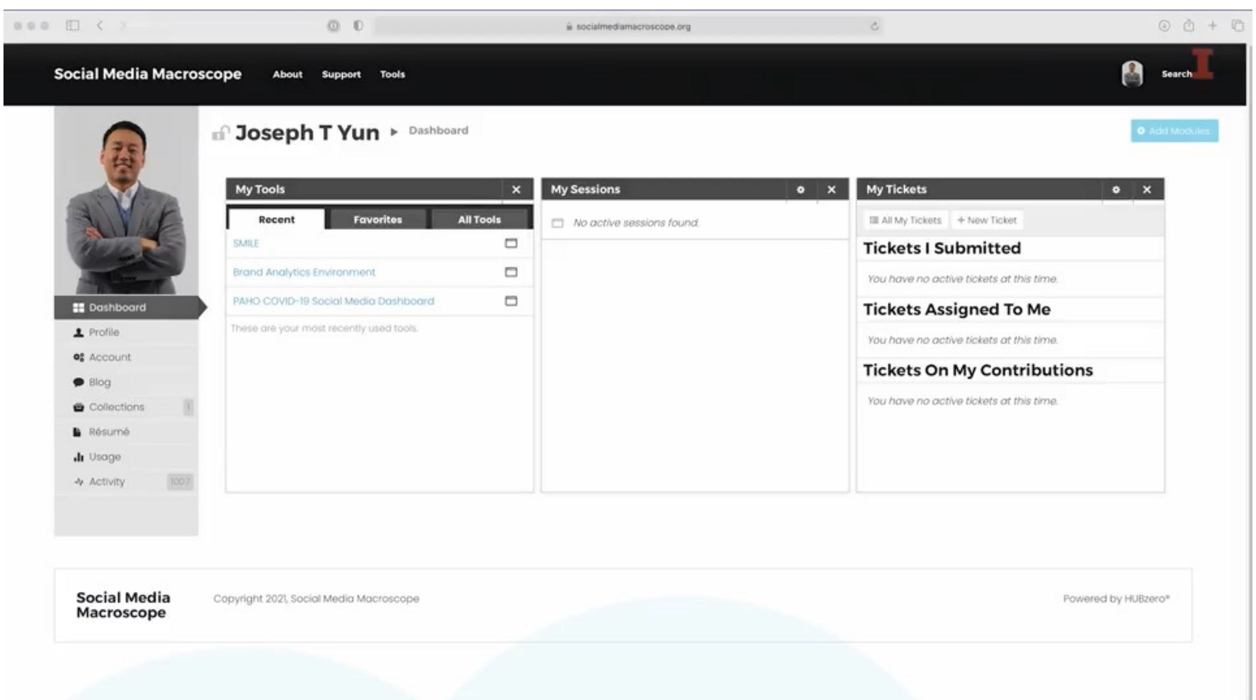
Basically how it would work is something like this, that we have brand X, and let's say that we have measured their brand personality. Sincerity is one, and this is from a scale from 0-1, their excitement is 0.25, their competence is 0.75, their sophistication is one, and their ruggedness is 0.25. Then we have a blue individual. That person's individual has sincerity one, excitement 0.25, competence 0.75, so on and so forth. Then we have a red individual whose sincerity is 0.25, excitement is one, competence is zero, so on and so forth. If we do a mathematical comparison of these essentially vectors, these five value vectors of brand personality, and you can use a mathematical function, typically called cosine similarity. You can use other similarity measures, but you try to figure out how similar is blue to brand X compared to how similar is red to brand X. What we find here, this is just a contrived example so that we have a perfect match, is that blue is a perfect match in brand personality to brand X, whereas red is not a perfect match to brand X in brand personality. The thought could be that maybe we should engage with this blue individual to be an influencer because they so represent our brand with regards to brand personality. Now just as many of you know that many times influencers on social media, they're not always celebrities that everybody knows. Sometimes they are everyday people just like you and me. But for whatever reason, they built a great following on social media. Maybe they could represent our brand. But how do we know if they're fitting for our brand? How do we know if they could be a good fit to represent what we represent? This brand personality machine-learned modeling might be one way to look at that. In the next lesson, we'll actually dive into looking at calculating this in the Social Media Macroscope.

[Influencer Brand Personality Analysis Part 2]

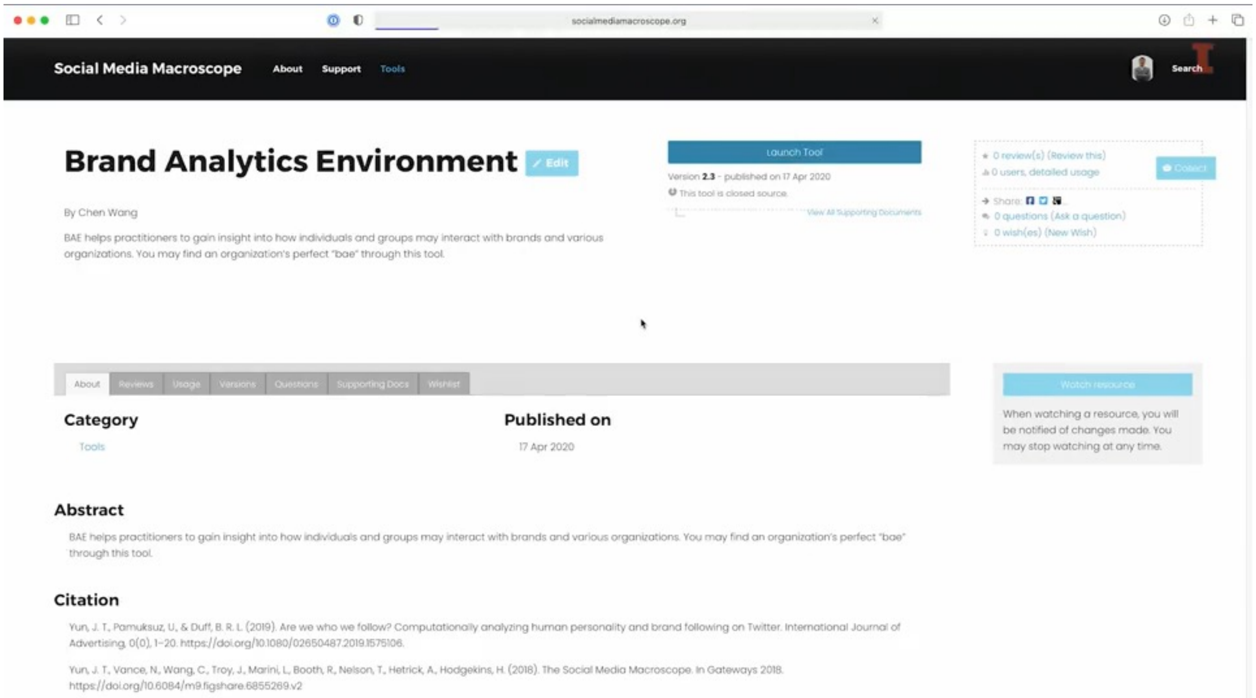
Now that we talked about the theoretical foundation for brand personality and also talked about a machine learning model that can estimate brand personality from Twitter data, let's jump into the Social Media Macroscope and look at some brand personalities. But before we do so, please know that we will be using a different tool, not the SMILE tool in the Social Media Macroscope, but the BAE tool, the brand analytics environment tool. I also want to say before we jump in, that more evidence is needed to validate both the machine learning model for brand personality prediction, as well as this potential personality congruence phenomenon where people seem to follow brands that they're close in personality to than brands that they're not. This is all just stemming from one study that we did.



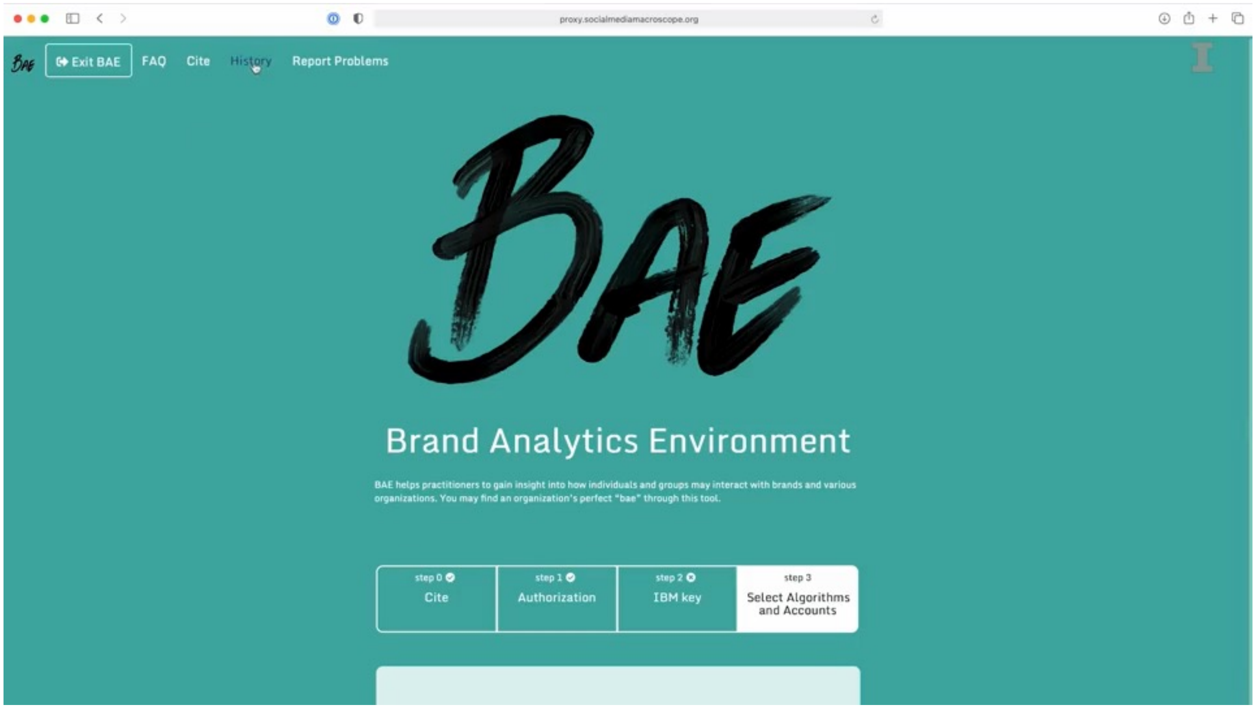
Many more studies are necessary. But it's interesting enough where, I believe, we can try to use some of these methods to see if that can affect how we find influencers that might affect customer satisfaction towards our company brand, products, and services. With that all said, let's jump into the Social Media Macroscope.



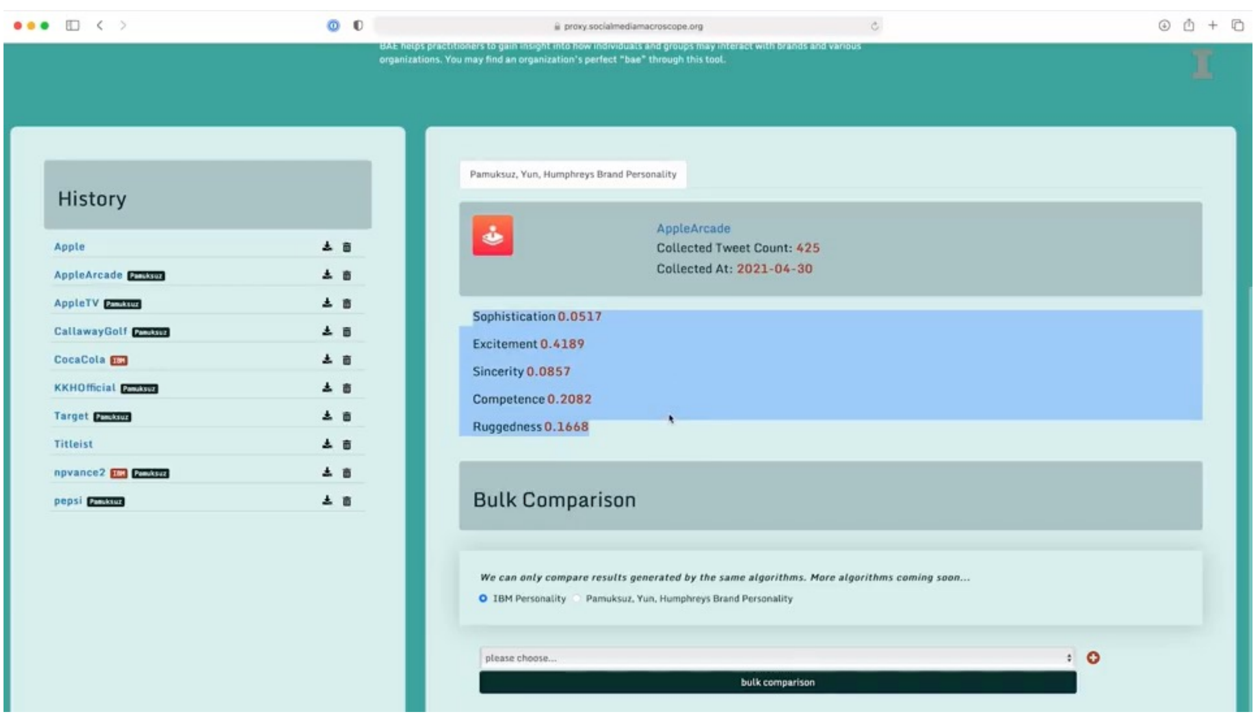
Here's the dashboard for the Social Media Macroscope. You've been here before because you've launched SMILE, and so you've either click on this box icon next to SMILE or you've clicked directly on SMILE and gone into the launch app button. But we are going to use the brand analytics environment, which was built for that study that I had mentioned in the lesson previously, that involved calculating people's human personality using IBMs tool, and then testing the congruence of personality with brand Twitter accounts that these people followed versus brand Twitter concept they did not follow. Let's start brand analytics environment. Let's click on this box either to the right or you can click on Brand Analytics Environment here and click on "Launch Tool", and just shortly the tool will launch itself.



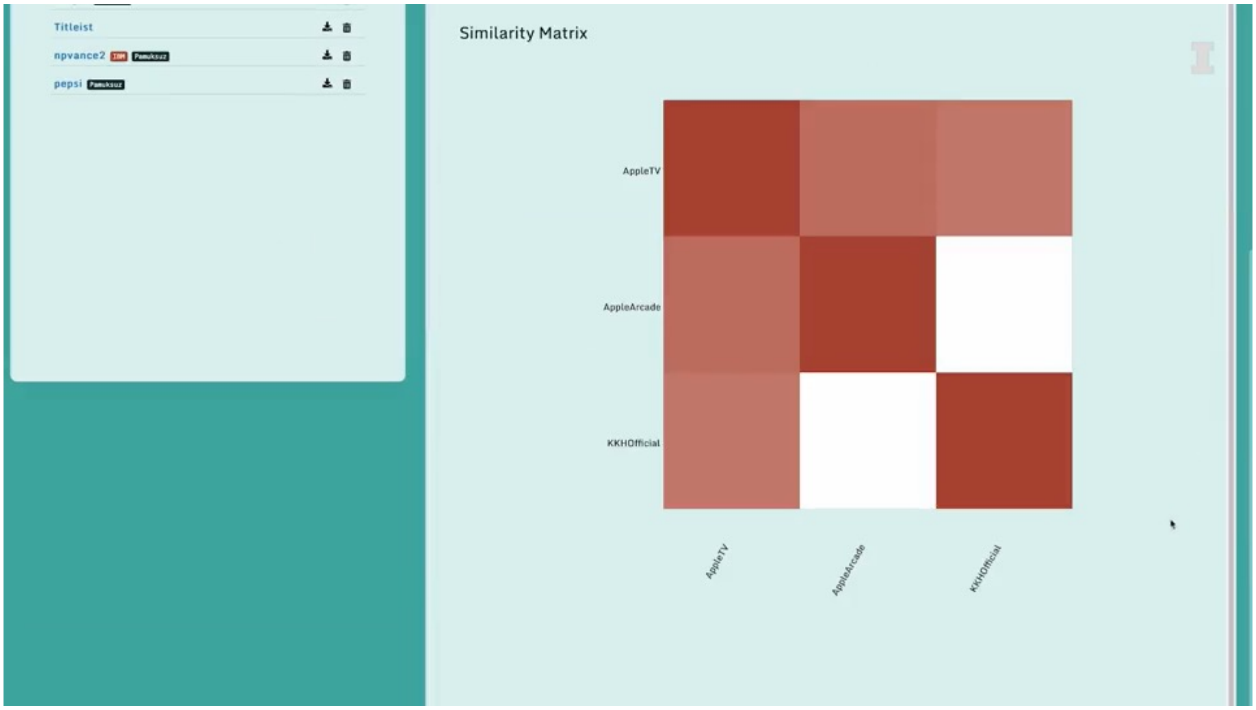
Now there's currently two personality models in BAE. One is the human personality model that IBM created. We will not be using that model, but rather the model that my team created. Right here, again, we start with citations. This is actually the paper, if you'd like to reference it with regards to our previous study. We also asked for a citation because we are trying to present these tools for free, and so we need support and proof of applicability to the real-world as well as to research. If you are doing research, will you please consider citing us. Here, we authorize our Twitter account. Let's click on "Authorize App", let's grab the pin, spring the pin and this is just like SMILE. But then for this IBM stuff, because we're not going to be using IBMs personality model as well as you need model API keys any ways that you do not have, let's click on "Skip" here. Then what we'll do is we'll choose the Pamuksuz, Yun, and Humphreys brand personality model. Let's compare a influencer on Twitter. We'll choose Kim Kardashian. As you type in letters, it'll just autofill from Twitter. Then let's choose Apple. Interesting thing about Apple's main account is they don't tweet from that account, so it has zero tweets. Maybe what we can choose is we can choose something like Apple TV or Apple Arcade. Then what you'll do here is you'll click on "Start Analysis". When you click on Start Analysis, it will ask you for your email address because this computation takes quite a bit of time. You'll put in your email address, and then you'll click on "Submit". Then once you click on Submit, you'll get an email once it's done. Now, I've already ran this analysis so I'll just go to my history. But just FYI, it will take quite some time for you to get an email to allow this to fully run.



The way I'll access my previous results is to go to history up here in this top left-hand corner. I'll click on History, and you'll see here that I have the Kim Kardashian Twitter account official Pamuksuz results. Here are the values for Kim Kardashian in sophistication, excitement, sincerity, competence, and ruggedness. Again, these values by themselves don't mean that much because really what we need to do is compare them to other accounts.



Let's say we look at Apple Arcade here, and we see values also for sophistication, excitement, sincerity, competence, and ruggedness. Remember, this is now taking tweets that Apple Arcade has presented on their Twitter account and running that through our brand personality model. How we can test congruence is we'll choose the Pamuksuz, Yun, and Humphreys brand personality model, and then we'll choose a few different accounts. We'll choose KKHOfficial. Let's choose Apple Arcade. Let's add another one even, Apple TV, and we click on "Bulk Comparison" here.



What we'll see here is that we've got a matrix that compares the personality congruence using cosine similarity of various accounts. If you compare, for example, Apple TV to Apple TV, it's a perfect, 100 percent, 1.0 comparison, and that makes sense. Apple Arcade to Apple, Arcade 1.0. Kim Kardashian to Kim Kardashian, 1.0. Now let's look at Kim Kardashian to Apple TV. It's a 94 percent or 0.94 similarity. Let's also look at Kim Kardashian to Apple Arcade, it's 81 percent, 0.81 percent similarity. Again, drawing from previous research, there might be some validity to Kim Kardashian potentially participating with Apple TV because the way that they tweet is very similar in brand personality, that there might be some potential interaction that could occur here. Of course, it could not just be Kim Kardashian, but you could choose anybody. Maybe you find somebody using network analysis for your brand on Twitter. It might be somebody that might not necessarily be a big celebrity but somehow they've got very high influence with regards to network structure on Twitter. Then you run their brand personality and see if it matches the brand personality of your brand. If it's a high number like this, 94 percent, then you may consider contacting them and partnering with them to market your brand because, again, this is all about potentially trying to move the needle with regards to customer satisfaction. We know that influencers have a pretty large effect on customer satisfaction towards a brand, especially on social media.