

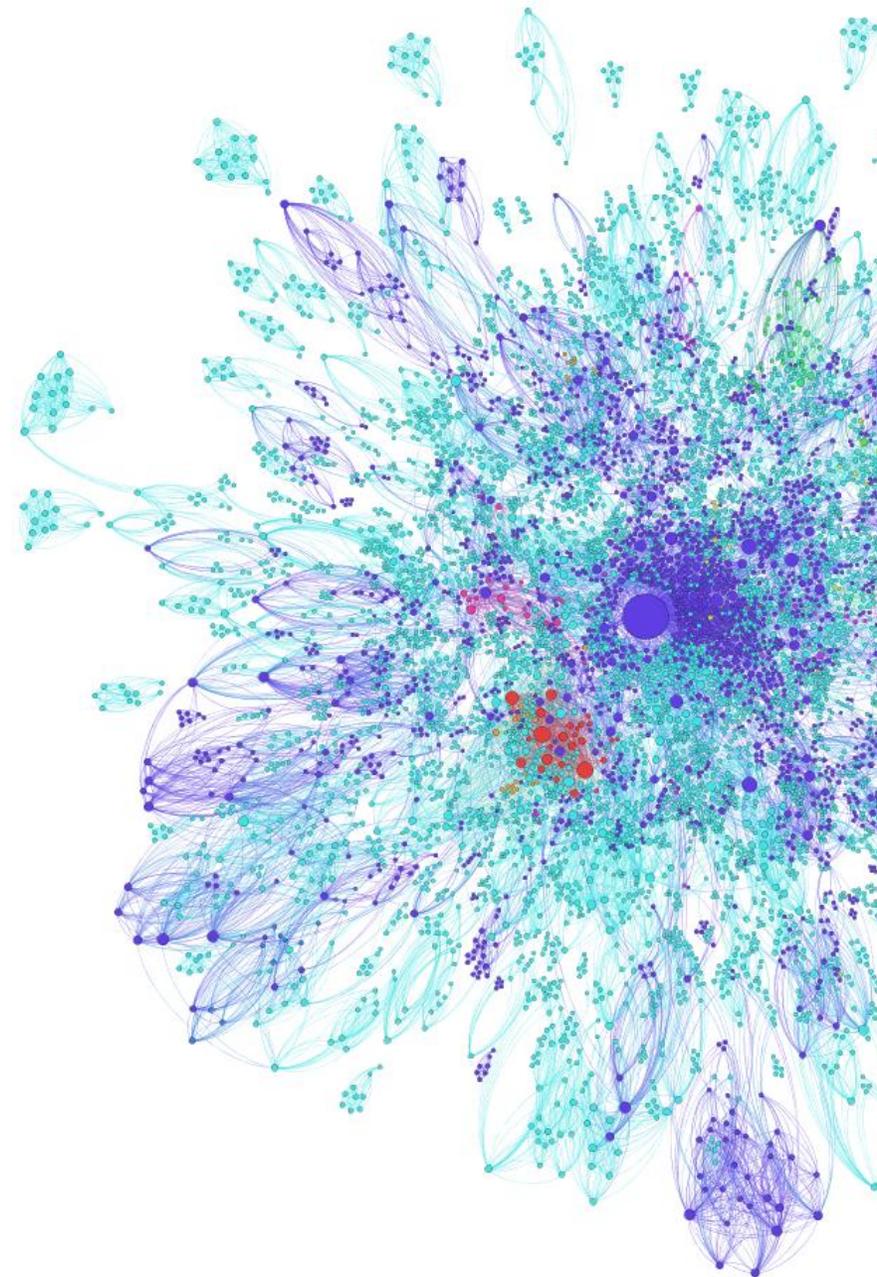
Análise de Redes Sociais

Introdução ao Gephi

Eduardo Silva

easilva@gmail.com

Laboratório de Políticas Públicas Participativas
Gestão da Informação - UFG



Análise de Redes Sociais?

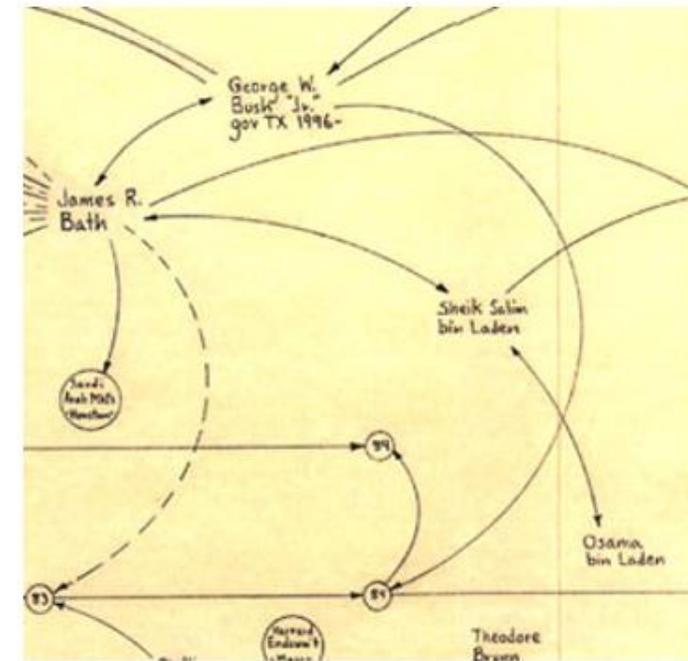
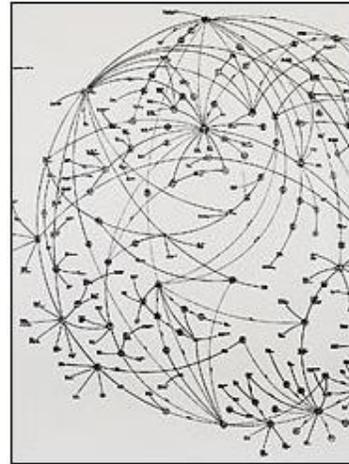
Ciências Sociais = **avaliam-se os atributos individuais e correlacionam-se entre si.**

Análise de redes Sociais = **estuda um conjunto de atores na rede e as relações entre si, dando foco ao todo.**

Possibilidade de se gerar uma visão mais ampla do contexto social.

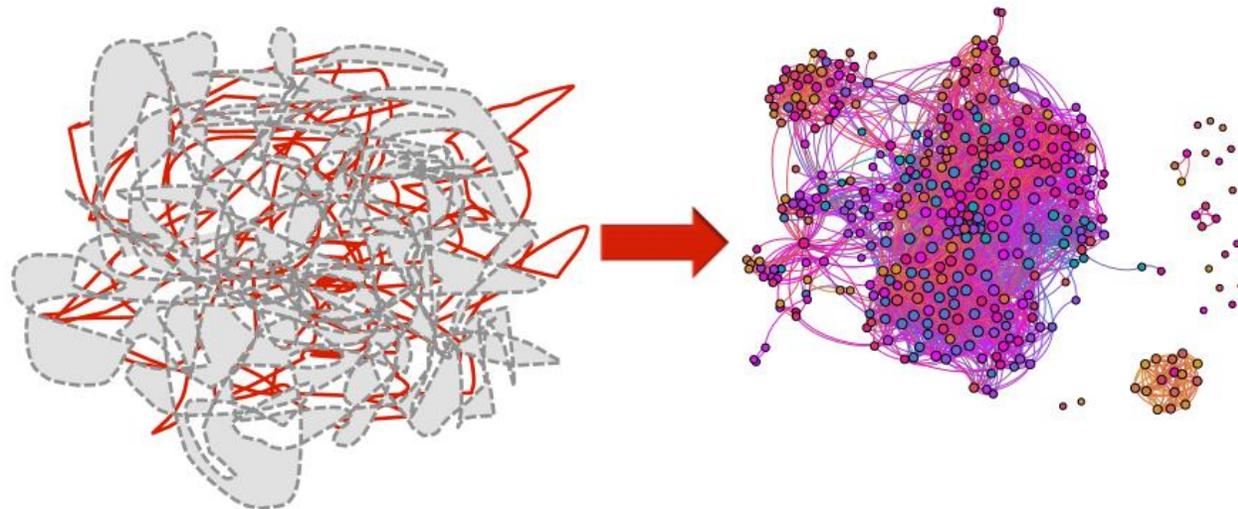
Exemplos: Redes Políticas/Financeiras

Mark Lombardi: rastreou e mapeou o fisco financeiro global nos anos 80 e 90 a partir de fontes públicas como artigos e jornais.

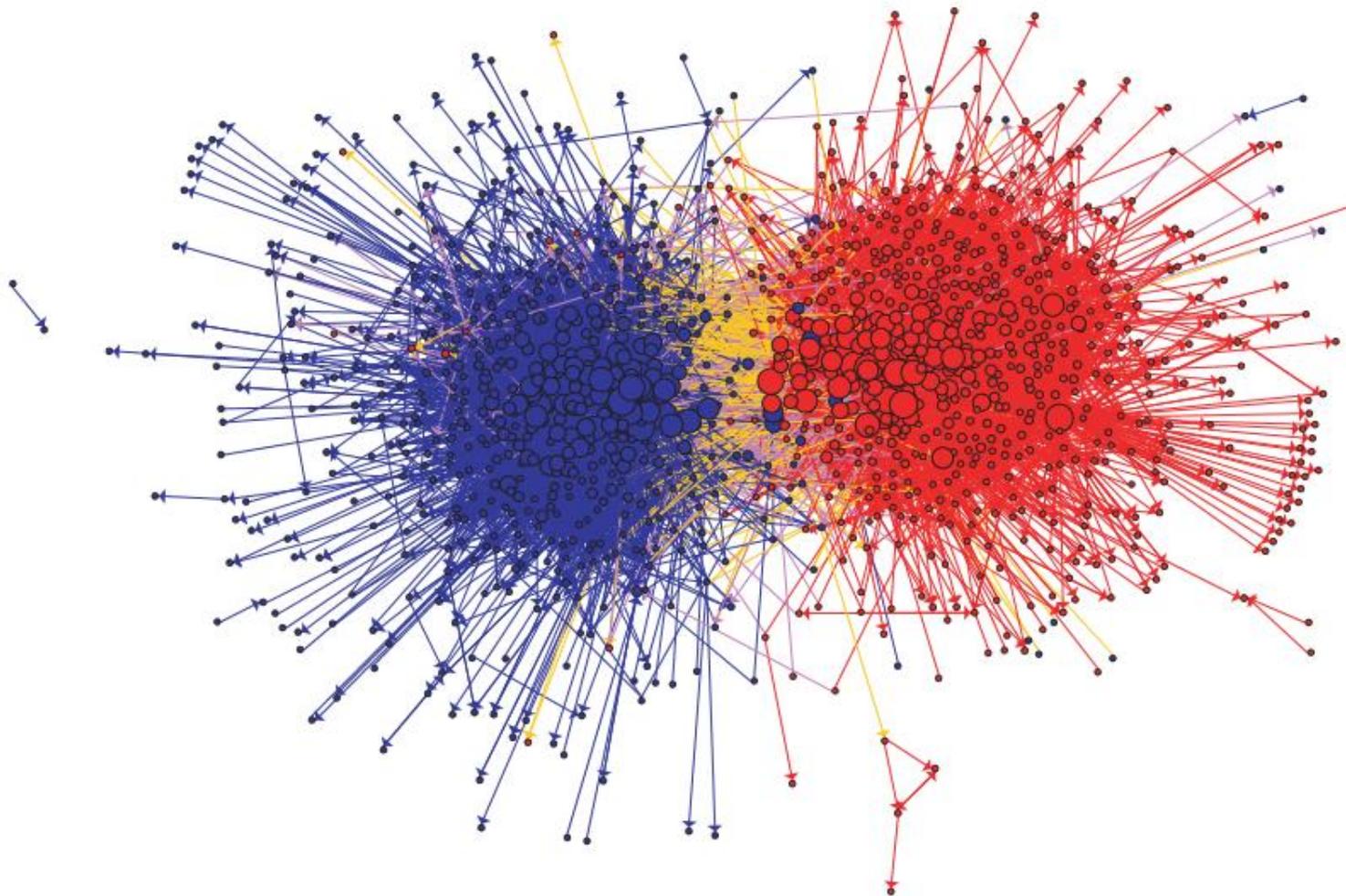


Compreender através da visualização

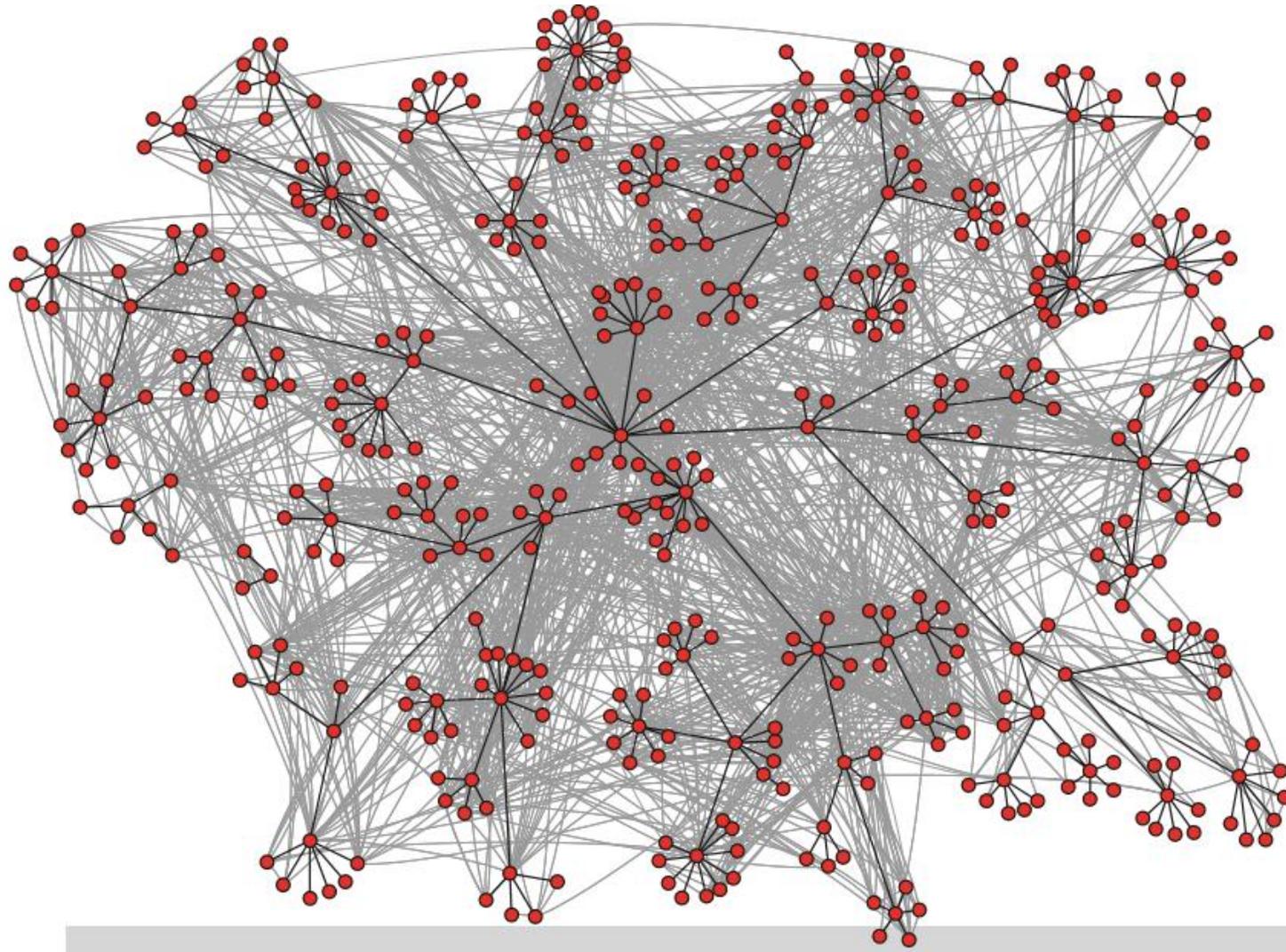
Um dos fatores mais impressionantes da análise de redes é a possibilidade de leitura e compreensão das visualizações de rede que são geradas. As métricas e todo o contexto da ARS, faz com que se torne simples gerar *insights* e compreender o que ocorre na visualização.



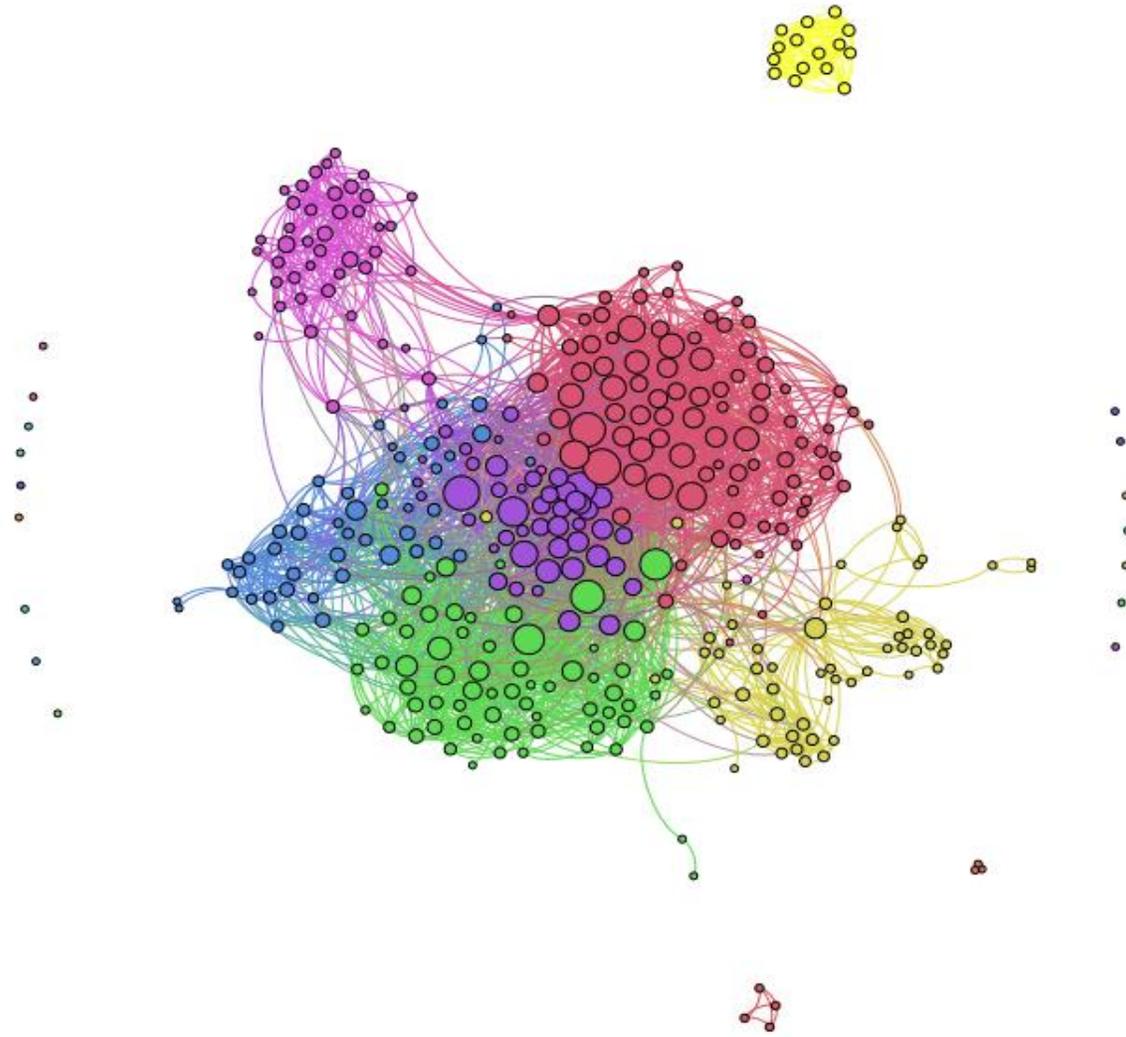
Blogs políticos



Rede de e-mails



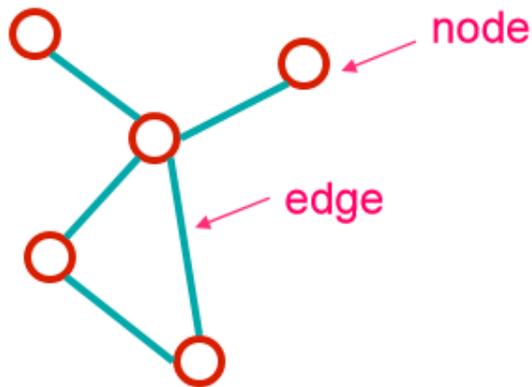
Rede do Facebook



O que é uma rede?

Redes são um conjunto de nodos conectados por arestas.

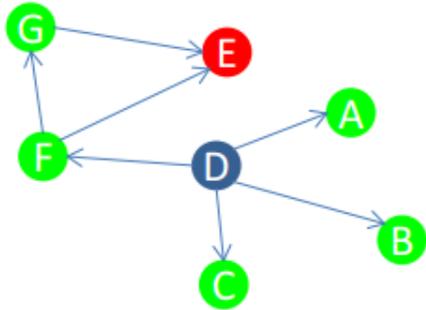
“Redes” = “Grafos” / “Sociograma”



Pontos	Linhas	
Vértices	Arestas, arcos	Matemática
Nodos	Ligações	Ciências da computação
Sites	Vínculos	Física
Atores	Ligações, relações	Sociologia

Rede direcionada ou não?

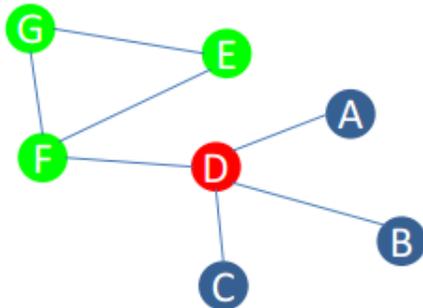
1 - Rede direcionada



Imagine uma rede “de quem segue quem” no Twitter.

Se estivermos interessados em quem tem mais seguidores, temos apenas que verificar o nodo que recebe mais conexões. **Esse nodo é o E.**

2 - Rede não direcionada



Legenda:



- centralidade +

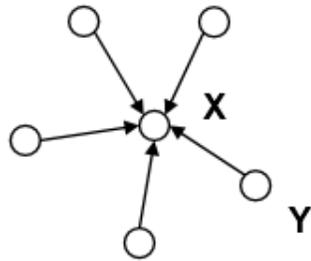
Agora imagine a mesma rede, mas como uma rede não direcionada. O nodo com maior centralidade de grau é o D (4 conexões), **não o E** (com 2 conexões).

Conclusão: no geral, as métricas retornam resultados diferentes dependendo da maneira como a rede esta direcionada ou não. A interpretação pode ser completamente errônea caso não se tome atenção a isso.

Métricas de rede

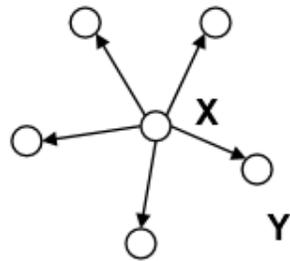
Diferentes tipos de centralidade;

Em cada uma das visualizações a seguir, o X tem mais centralidade do que o Y de acordo com as particularidades de cada métrica.



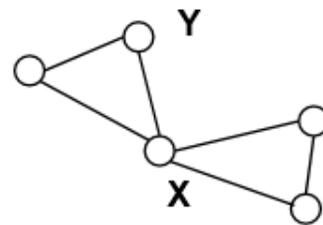
indegree

Grau de
Entrada



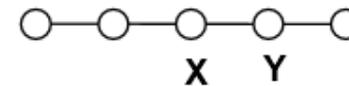
outdegree

Grau de
Saída



betweenness

Intermediação



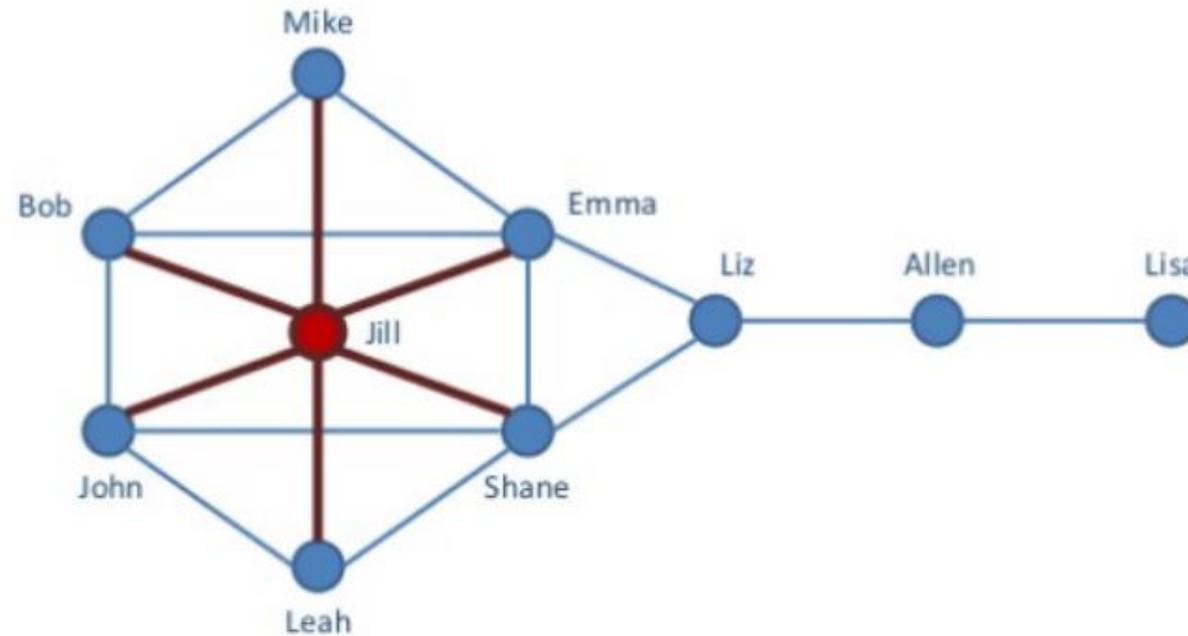
closeness

proximidade

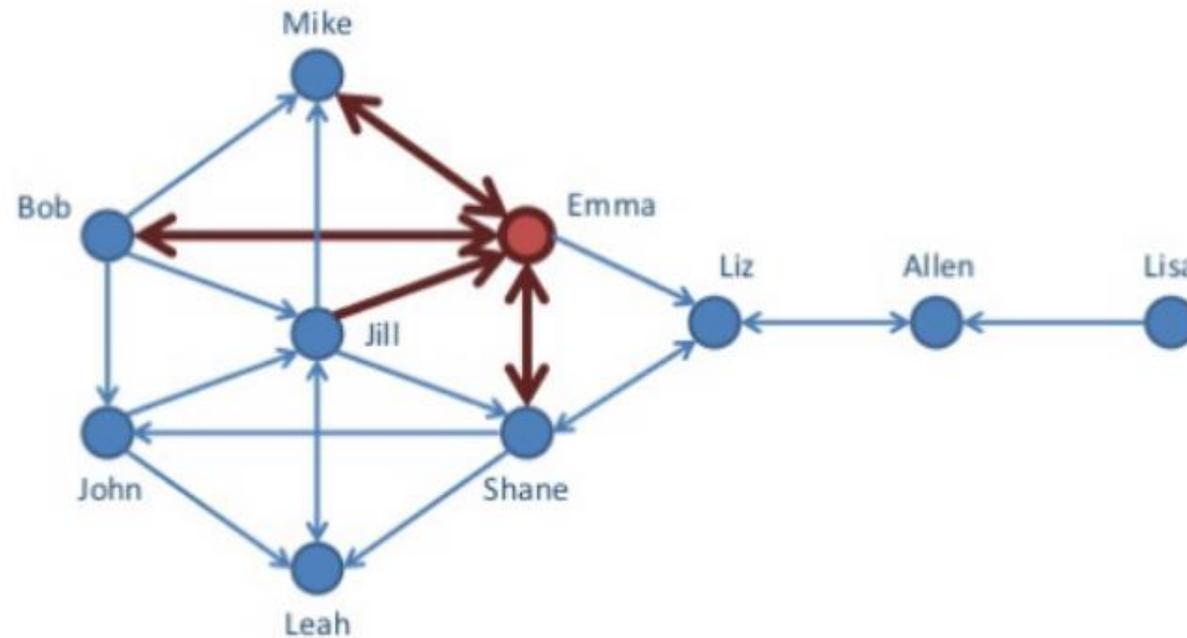
Métricas de rede (centralidade)

- **Grau:** o critério de cálculo leva em consideração a quantidade absoluta de vínculos que um nó estabelece. Não analisa como se dão essas conexões, se mais localizadas, mais distribuídas. Apenas o total.
- **Intermediação:** O critério de cálculo leva em consideração a posição de um nó entre o caminho que conecta todos os outros nós. O nó que tiver maior valor de intermediação é aquele que se encontra mais no meio do caminho entre todos os outros nós possíveis;
- **Proximidade:** O critério de cálculo leva em consideração o quão próximo é um nó de todos os outros. O algoritmo calcula a distância entre todos os nós para cada nó e aquele que tiver o menor valor é o que se encontra mais próximo de todos os outros.

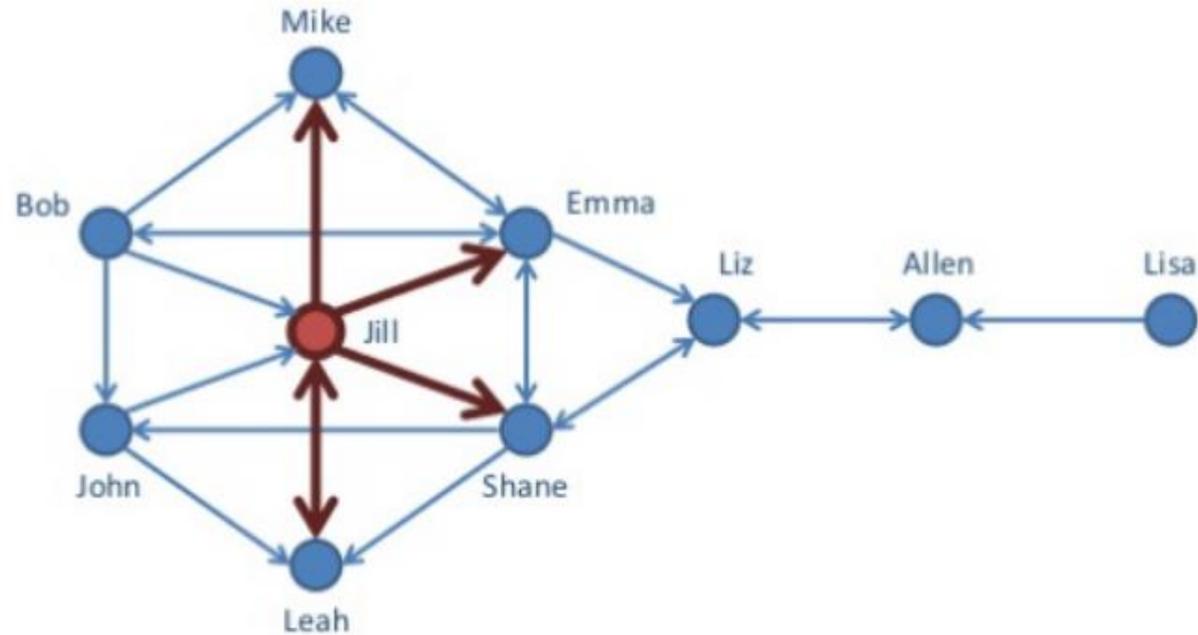
Centralidade de grau



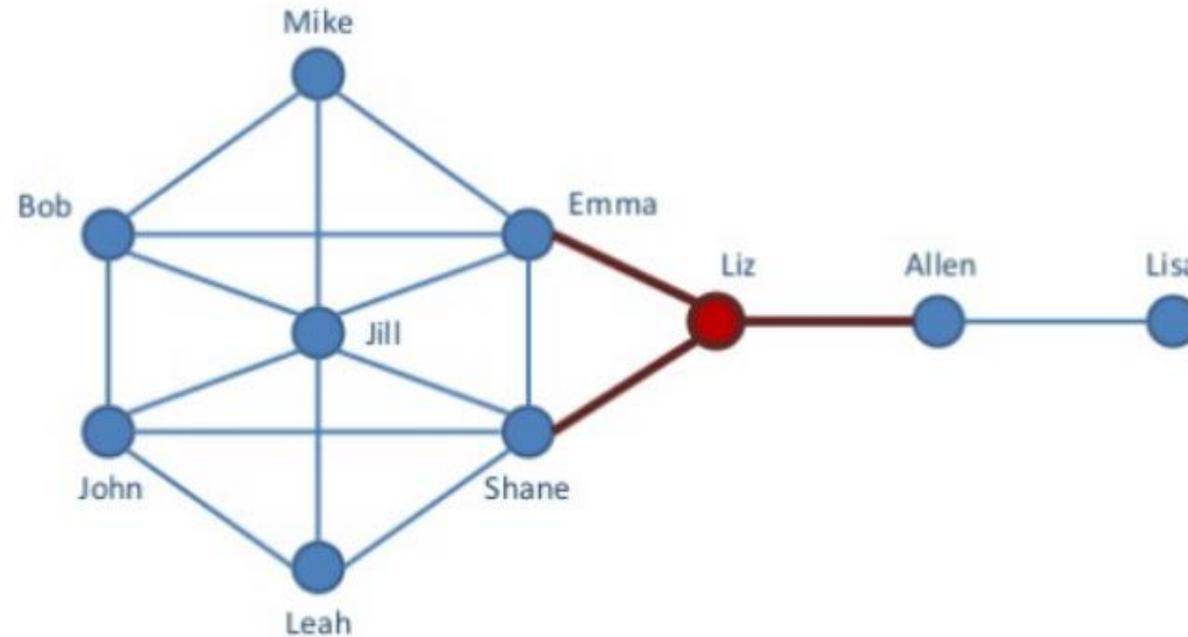
Centralidade de grau de entrada



Centralidade de grau de saída

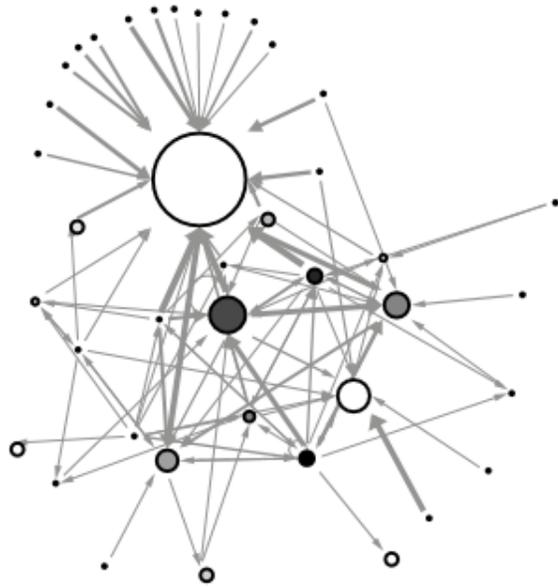


Centralidade de Intermediação

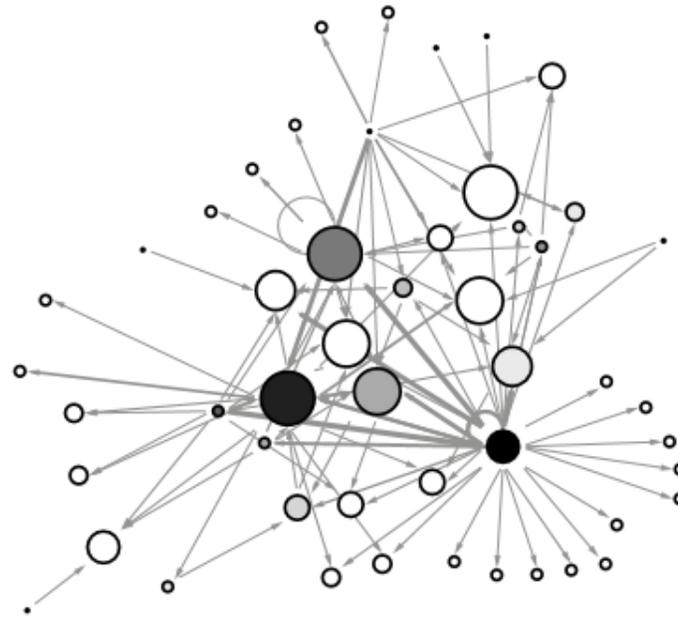


Exemplos do mundo real

Exemplo de rede de negociação financeira



Alta centralidade:
um nodo compra de
vários outros

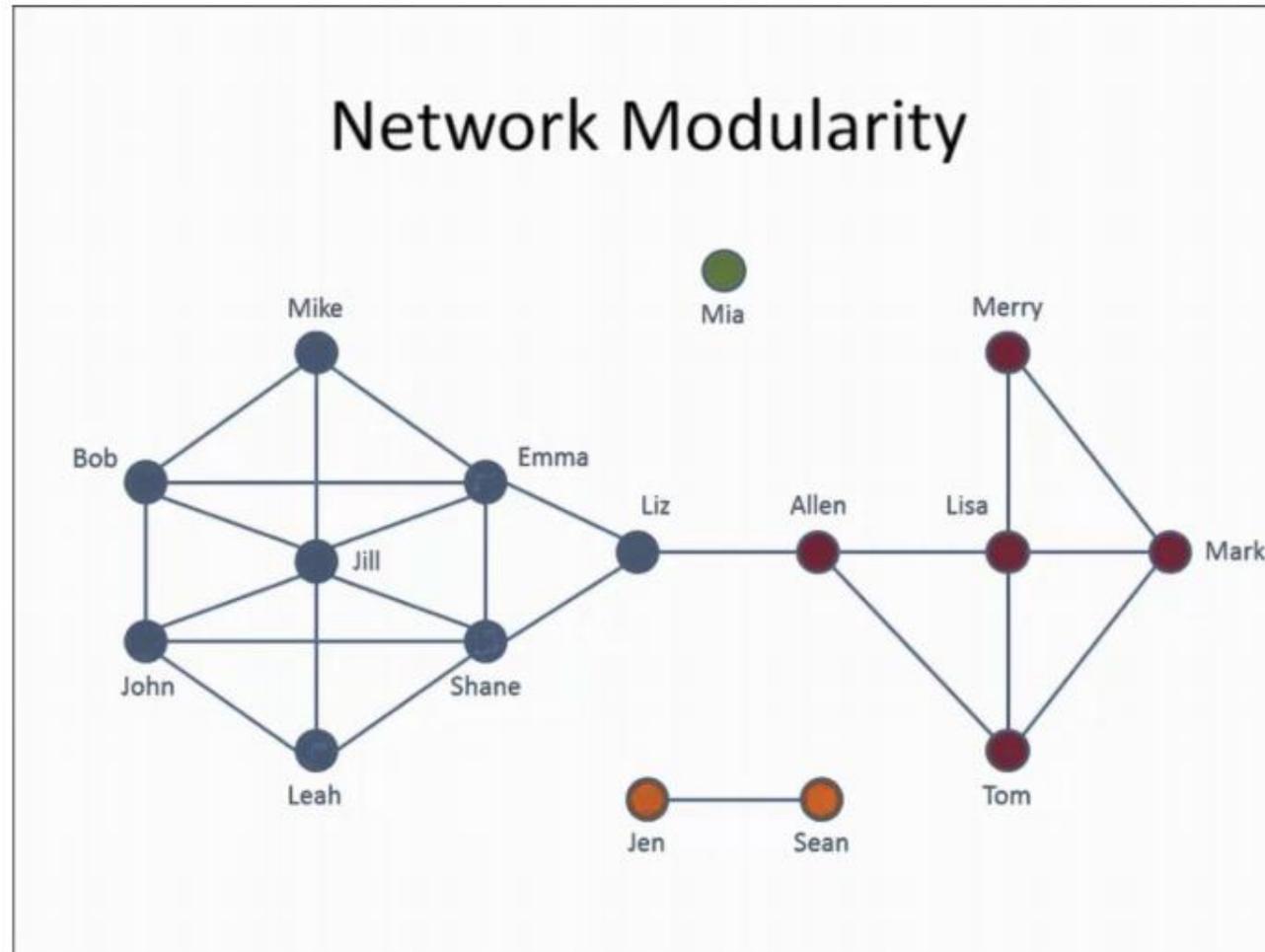


Baixa centralidade:
as compras são mais
bem distribuídas

Modularidade

- É uma medida da estrutura da rede;
- Foi desenhada para mensurar a força da divisão da rede em módulos (também chamados de grupos, clusters ou comunidades);
- Redes com alta modularidade possuem conexões densas entre os nós das comunidades e baixa densidade entre os nós de diferentes comunidades;
- É uma medida que leva em consideração a relação dos nós com seus vizinhos, ou seja, uma medida de vizinhança, levando em consideração o quanto um nó tende a aparecer em determinado grupo.

Modularidade



Gephi

Download: <https://gephi.org/users/download/>

- Aplicações
- **Análise exploratória de dados:** análise orientada por manipulação de redes em tempo real.
- **Análise de Links:** revelando as estruturas subjacentes de associações entre objetos/sites.
- **Análise de redes sociais:** criação fácil de conectores de dados sociais para mapear organizações comunitárias e redes de pequenos mundos.
- **Análise de redes biológicas:** representa padrões de dados biológicos.
- **Criação de cartazes:** promoção de trabalhos científicos com mapas imprimíveis de alta qualidade.

Caixa de Visualização

The screenshot displays the Gephi 0.9.1 interface. At the top, the menu bar includes 'Arquivo', 'Área de Trabalho', 'Ferramentas', 'Janela', and 'Ajuda'. Below it, the 'Workspace 0' tab is active, with sub-tabs for 'Visão geral', 'Laboratório de dados', and 'Visualização'. The central area shows a complex network graph with nodes of various sizes and colors (blue, purple, green, orange) connected by edges. To the left, the 'Aparência' panel allows for node and edge styling, including a color picker set to '#c0c0c0'. Below it, the 'Distribuição' panel shows the 'ForceAtlas 2' layout algorithm settings, such as 'Tolerância (velocidade)' at 1.0 and 'Número de threads' at 2. To the right, the 'Contexto' panel displays network statistics: 'Nós: 7911' and 'Arestas: 33944'. Below that, the 'Filtros' panel lists various network metrics like 'Grau médio', 'Modularidade', and 'PageRank', each with an 'Executar' button. At the bottom, a toolbar contains icons for zooming and other visualization controls.

Abas de dados
E visualização

Caixa de aparência
onde é possível
mudar as confi-
gurações dos nodos
e das arestas, além
de ser responsável
por aplicar a
modularidade

Caixa de layouts
onde é possível
mudar o layout
da rede utilizando
diferentes algoritmos e
configurações.

Dados da rede
quantidade de
nodos e arestas

Aba de filtros e
métricas de rede
aplicação de
estatísticas de
rede, como a
- Modularidade
- Grau médio
- Grau de prox.
- Grau de inter.

Aba de configurações de rótulos
da rede.

Aparência da rede

É uma parte importante do Gephi, pelo fato de dar diferentes possibilidades em relação aos nodos e arestas da rede, sendo possível, particionar a rede em cores, aumentar os nodos mais importantes.

São opções simples que costumam gerar fortes impactos na visualização da rede.

Layout da rede

Layouts baseados em força de repulsão:

- Force Atlas e Force Atlas 2 (os mais comuns)
- Fruchterman-Rheingold
 - O clássico. Nodos que estão conectados se atraem.
- Kamada – kawai
 - Mais elaborado: a força de atração entre 2 nodos é proporcional ao menor tamanho das arestas que conectam esses nodos.
- Yifan Hu
 - Agrupa os nodos e depois aplica uma lógica de força nos grupos.

Plugins do Gephi

Todos plug-ins, são encontrados na Market place do site do Gephi. Alguns tem funções interessantes como:

Plug-ins	Descrição
SigmaJS	Cria a rede a partir de uma livreria em javascript que possibilita apresentar suas redes online com interação.
Geo Layout	Plota um mapa no Gephi e a partir da latitude e longitude, distribui os nodos no mapa

Gerar rede do zero

É preciso ter atenção, são necessários dois arquivos, um para os nodos e outro para as arestas.

O erro mais comum se dá pelo nome das colunas, que devem seguir o padrão do Gephi, independente da ordem dos dados.

Gerar redes a partir de base de dados

As redes podem ser geradas a partir de diversas plataformas, sejam plataformas sociais como Facebook e Youtube, ou mesmo de um banco de dados convencional em SQL.

- Netvizz – Facebook
- Digital methods – Diversas ferramentas de coleta de dados. (youtube, linkedin, outros)