

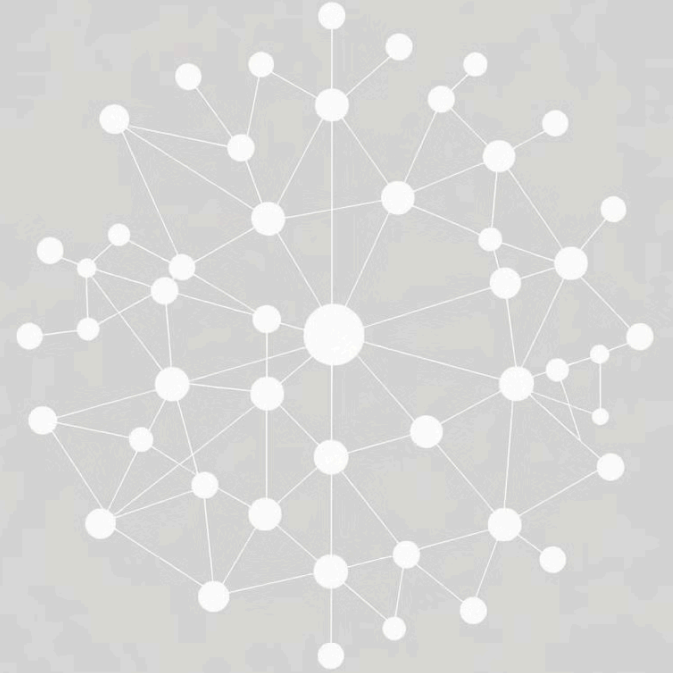
# Análise de comunidades no Instagram

Este trabalho explora a rede de seguidores do Instagram utilizando conceitos de redes complexas para entender as interações e influências entre os usuários.

O estudo extrai e modela dados da plataforma para investigar a estrutura da rede e as preferências esportivas dos estudantes de computação da Universidade Federal de São João del-Rei (UFSJ).



**by João Antônio**





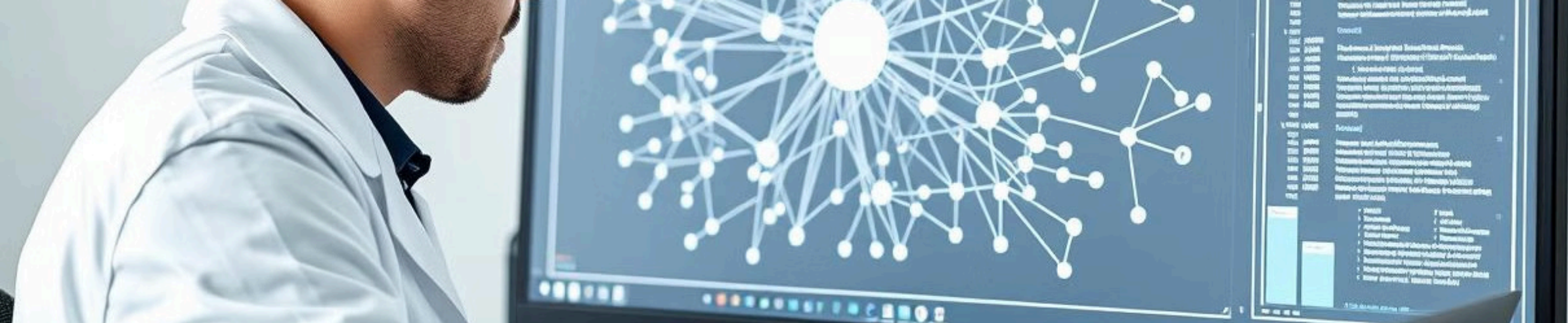
# Objetivo do Estudo

1

## Análise da Rede

Analisar a rede de seguidores do Instagram de um conjunto específico de usuários, explorando características estruturais, percebendo comunidades e calculando métricas de centralidade.

Com isso, identificar quais times de futebol são mais apoiados pelos estudantes de computação da UFSJ.



# Extração e Modelagem de Dados

## Coleta de Dados

Os dados foram obtidos por meio de raspagem de dados (web scraping) do Instagram, coletando os identificadores de perfil dos seguidores de usuários específicos. Foram utilizadas bibliotecas como Selenium para automatizar o processo de login e navegação.

## Modelagem da Rede

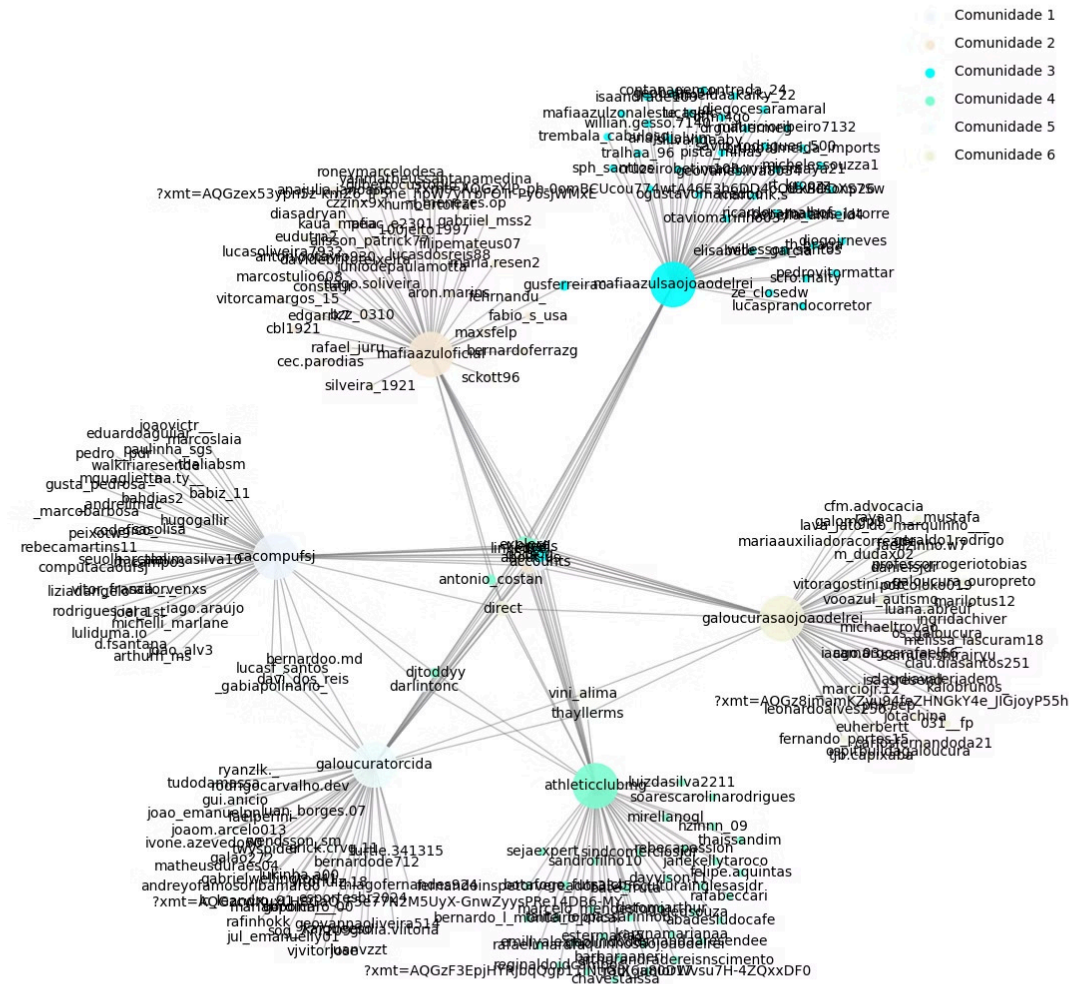
A rede foi modelada como um grafo não direcionado, onde cada nó representa um usuário e cada aresta representa uma relação de seguimento entre dois usuários. Os vértices são os identificadores dos usuários e as arestas representam as conexões de seguimento.

## Restrições e Privacidade

Devido às políticas de privacidade do Instagram, foi necessário considerar um tempo de espera adequado entre as requisições para evitar bloqueios automáticos da plataforma. Além disso, nenhuma informação pessoal dos usuários foi coletada, apenas seus identificadores.

# Visualização da rede

## Rede de Seguidores do Instagram



# Análise Descritiva da Rede

## Métricas Básicas

Foram calculados o número total de nós (usuários) e arestas (conexões de seguimento) na rede.

## Distribuição de Graus

A distribuição de graus no grafo parece ser assimétrica, com a maioria dos nós tendo um grau baixo (seguem apenas uma conta), enquanto um pequeno número de nós tem graus mais elevados (seguem múltiplas contas).

## Coeficiente de Clustering

O coeficiente de clustering mede a tendência dos nós a formar triângulos (conexões fechadas) no grafo. Neste caso, os grupos maiores parecem ter coeficientes de clustering baixos, pois a maioria das conexões são entre um usuário e uma única conta, com poucas interconexões entre diferentes comunidades.

# Análise Descritiva da Rede

## Densidade

A densidade do grafo refere-se à proporção de arestas existentes em relação ao número máximo possível de arestas. Considerando a visualização, o grafo parece ter uma densidade baixa, sugerindo que a maioria dos usuários seguem apenas uma conta, o que é evidenciado pelas grandes concentrações isoladas de nós.

## Comunidades

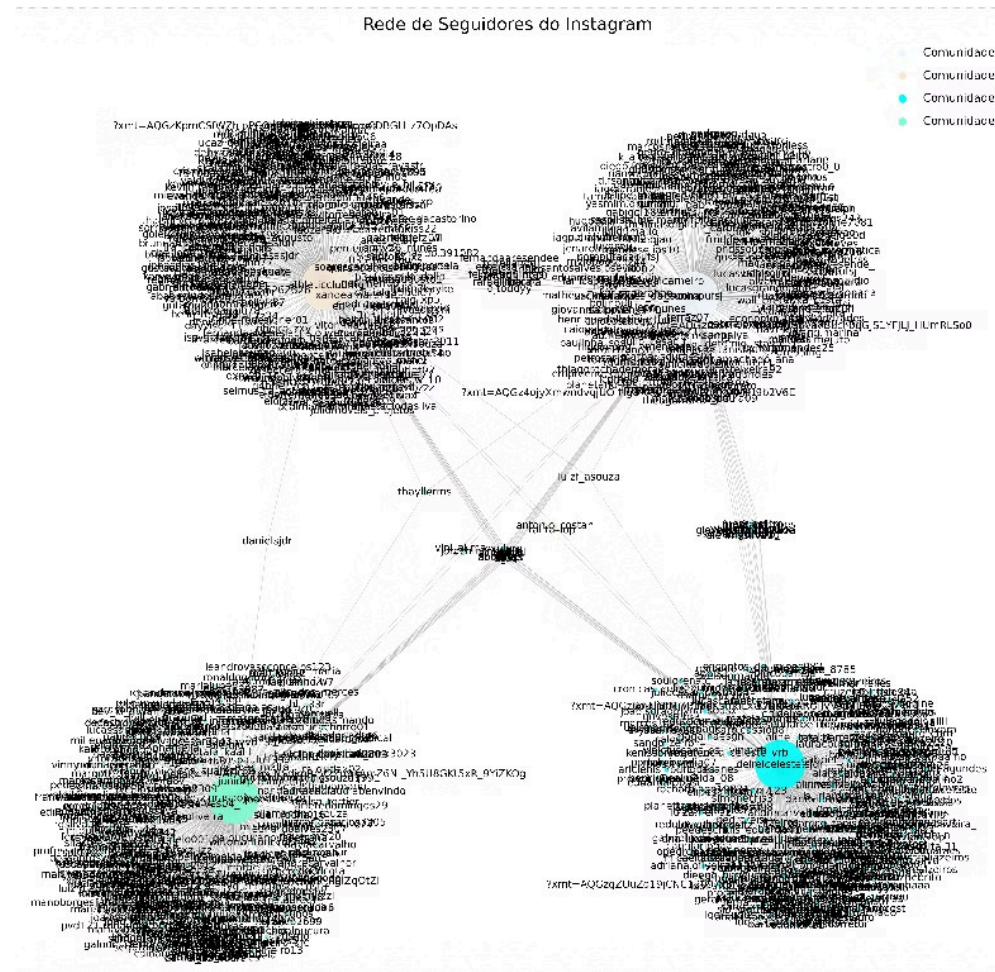
O grafo mostra quatro grandes comunidades principais (CACOMP, Athletic, Cruzeiro, Atlético Mineiro), identificadas pelas cores no grafo. As comunidades são bem definidas, com pouca sobreposição entre elas. Além disso, há pequenas subcomunidades que conectam duas ou mais das principais comunidades, sugerindo usuários com interesses mistos.

## Centralidade

A centralidade no grafo pode ser observada nos nós que conectam diferentes comunidades. Esses nós servem como "pontes" e têm alta centralidade, desempenhando um papel crucial na interconexão das diferentes comunidades. Esses nós centrais são os principais pontos de convergência de seguidores que têm interesses variados, e sua remoção poderia fragmentar ainda mais o grafo.



# Visualização da rede



# Discussão dos Resultados

1

A concentração de cima mostra uma quantidade relevante de seguidores em comum das contas do Athletic e do CACOMP, mostrando uma simpatia entre os alunos e o time de São João del Rei

2

O grupo da direita mostra uma bolha de alunos da computação que também seguem a torcida do Cruzeiro de São João del Rei

3

Há uma pequena concentração no meio, onde mistura a comunidade de Cruzeirenses que simpatizam com o Athletic e estudantes da computação que seguem a torcida do Atlético de São João del Rei

4

Além dessas concentrações, há arestas isoladas que representam relações incomuns, como seguir todas as contas estudadas, por exemplo.



# Limitações e Trabalhos Futuros

## Limitações Técnicas

Devido a limitações de tempo e recursos computacionais, além de restrições impostas pelo Instagram, não foi possível realizar a análise de redes muito grandes.

## Expansão da Análise

No futuro, seria interessante expandir essa análise para incluir dados de outras universidades ou cursos, ou até mesmo comparar as preferências esportivas em diferentes períodos do ano, observando como eventos esportivos influenciam essas redes.

## Aprofundamento da Pesquisa

Além disso, uma análise mais aprofundada das interações e da formação de comunidades na rede de seguidores poderia fornecer insights ainda mais valiosos sobre a dinâmica da plataforma.





# Conclusão

1

## Extração e Análise de Dados

Este estudo conseguiu extrair e analisar dados dos seguidores do perfil @cacompufsj, modelando a rede de interações e realizando diversas análises para entender as preferências esportivas dos estudantes de computação da UFSJ.

2

## Identificação de Preferências

O estudo também identificou a relação entre os torcedores do Athletic Club e sua possível simpatia por outros times grandes da região.