

Análise de Grafos no Azure SQL DB

Wagner Crivelini



Patrocínio:

GOLD



SILVER



Apoiadores:



BRONZE

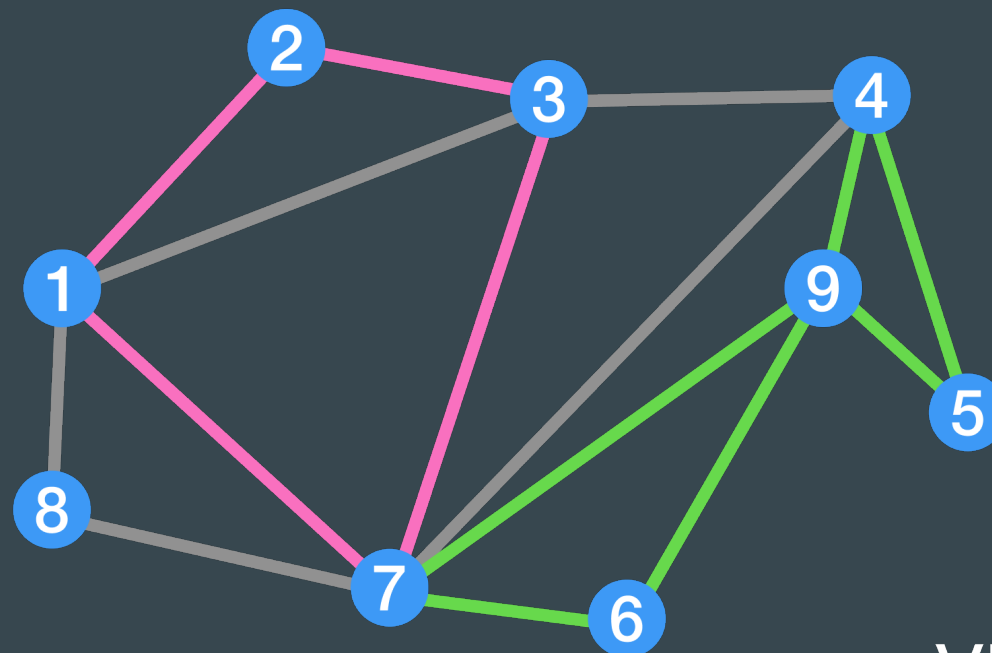


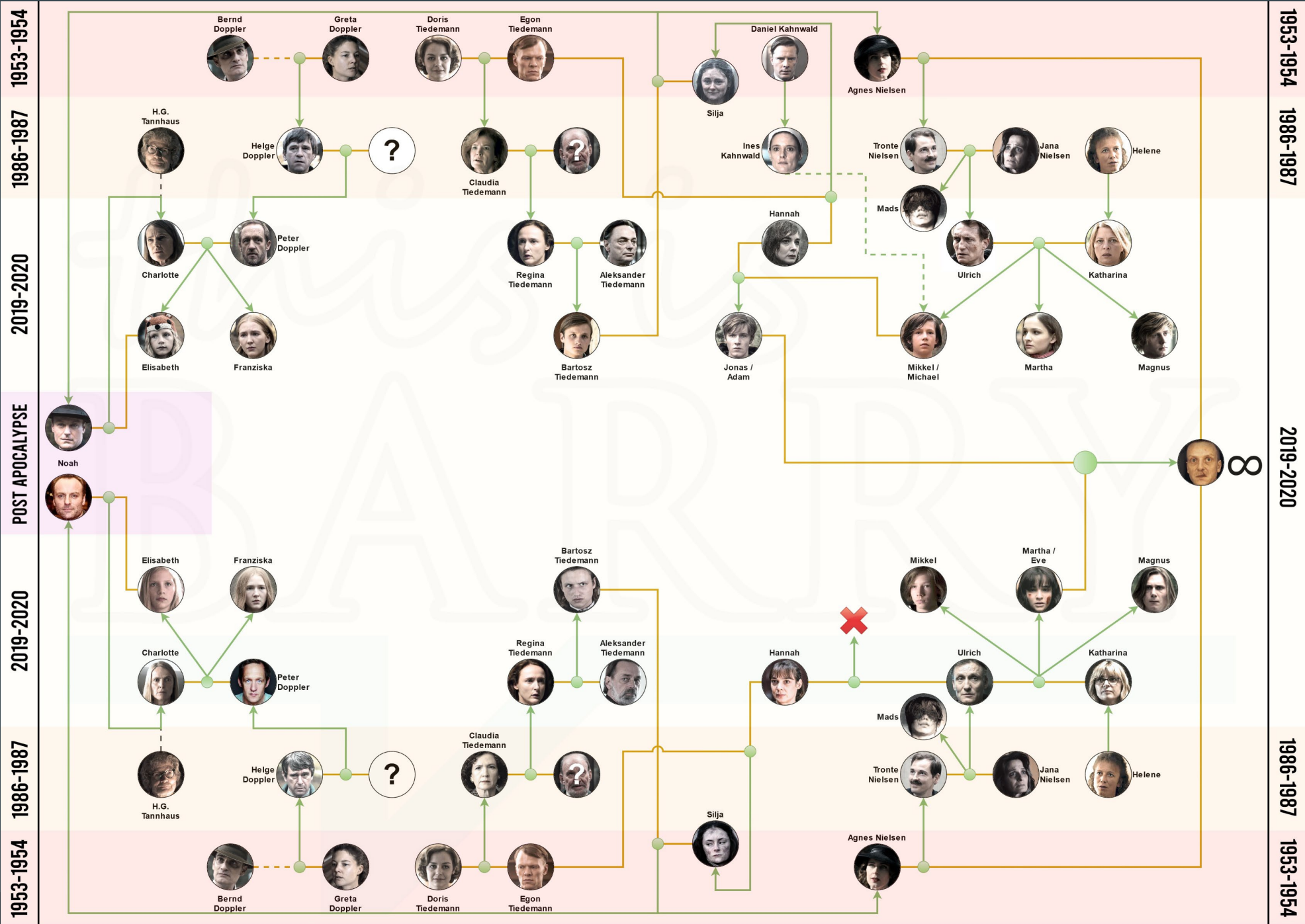
BIO: Wagner Crivelini

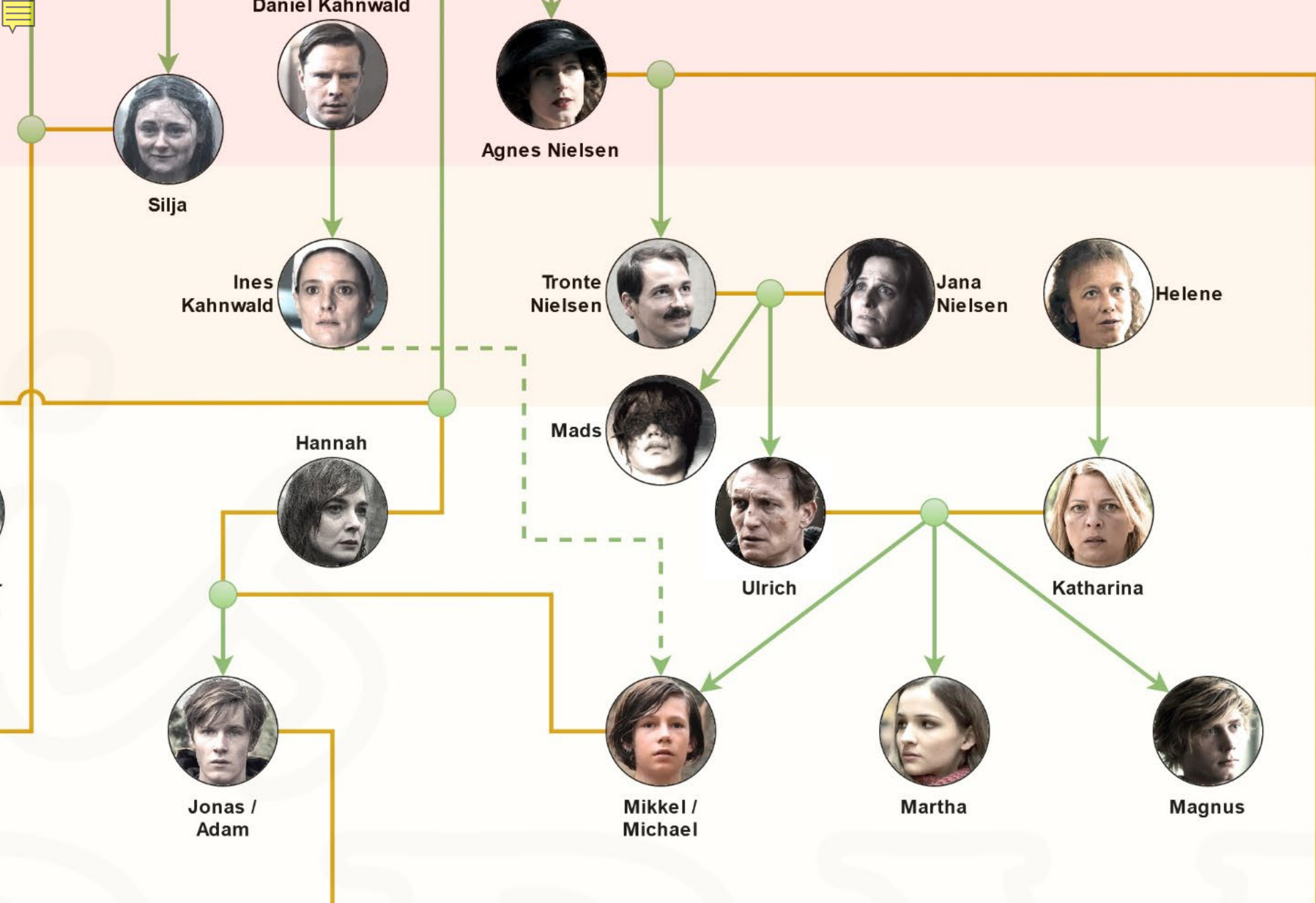
- Consultor de Data & AI da MICROSOFT ISD BRASIL
- Colunista em diversos portais (SQLServerCentral, iMasters)
 - +150 artigos
 - +100 podcasts
 - +050 palestras

O Que São Grafos

- Modelos que lidam com **Relações** complexas entre **Entidades**
- Usados em todo tipo de negócio: social media, engenharia...
- *"6 Graus de Separação"*
- Exemplo:
 - LINKEDIN: classificação de contatos





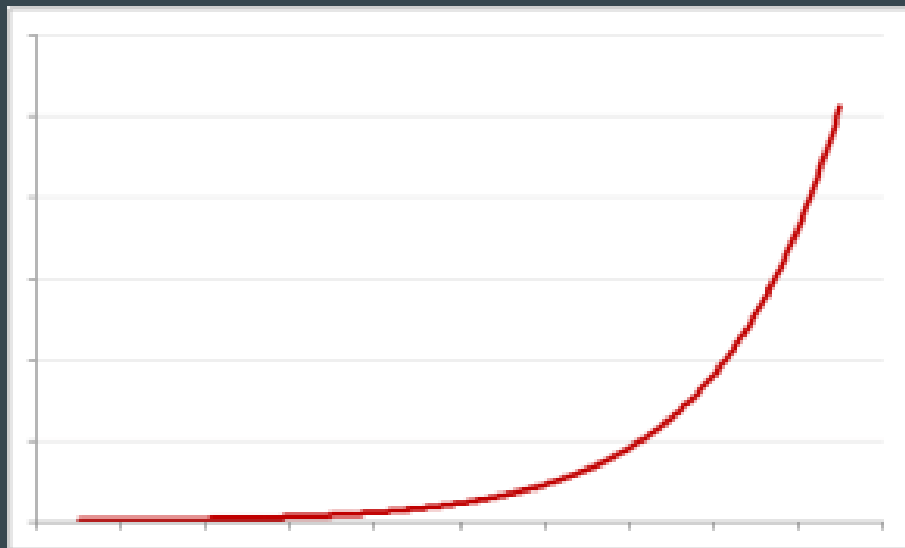


53-1954

1986-1987

Onde Aplicar Modelos de Grafos

- Modelos com relacionamentos complexos
- Quando ENTIDADES se apresentam com grande número de elementos distintos



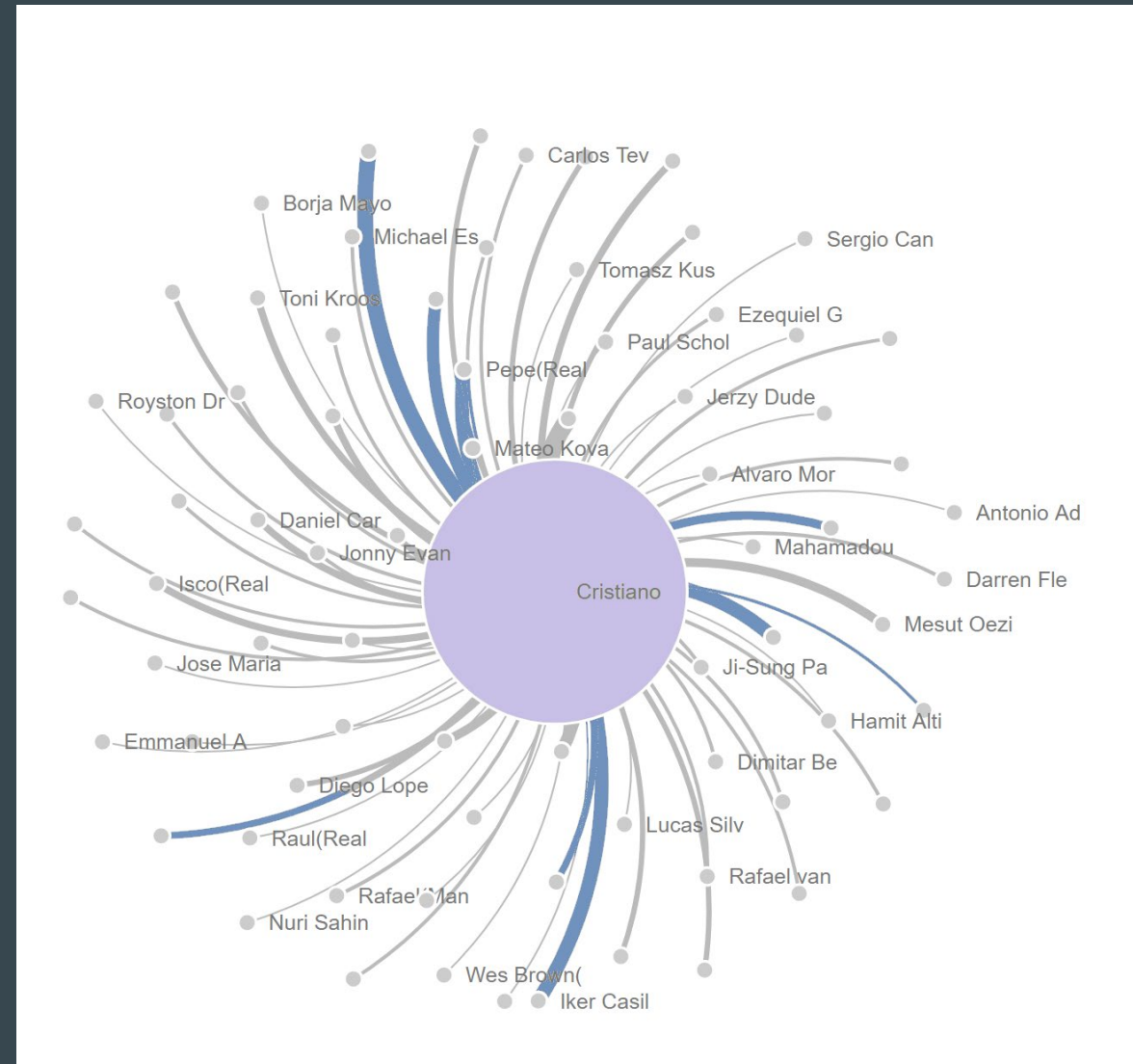
BD de Grafos = BD Relacional?

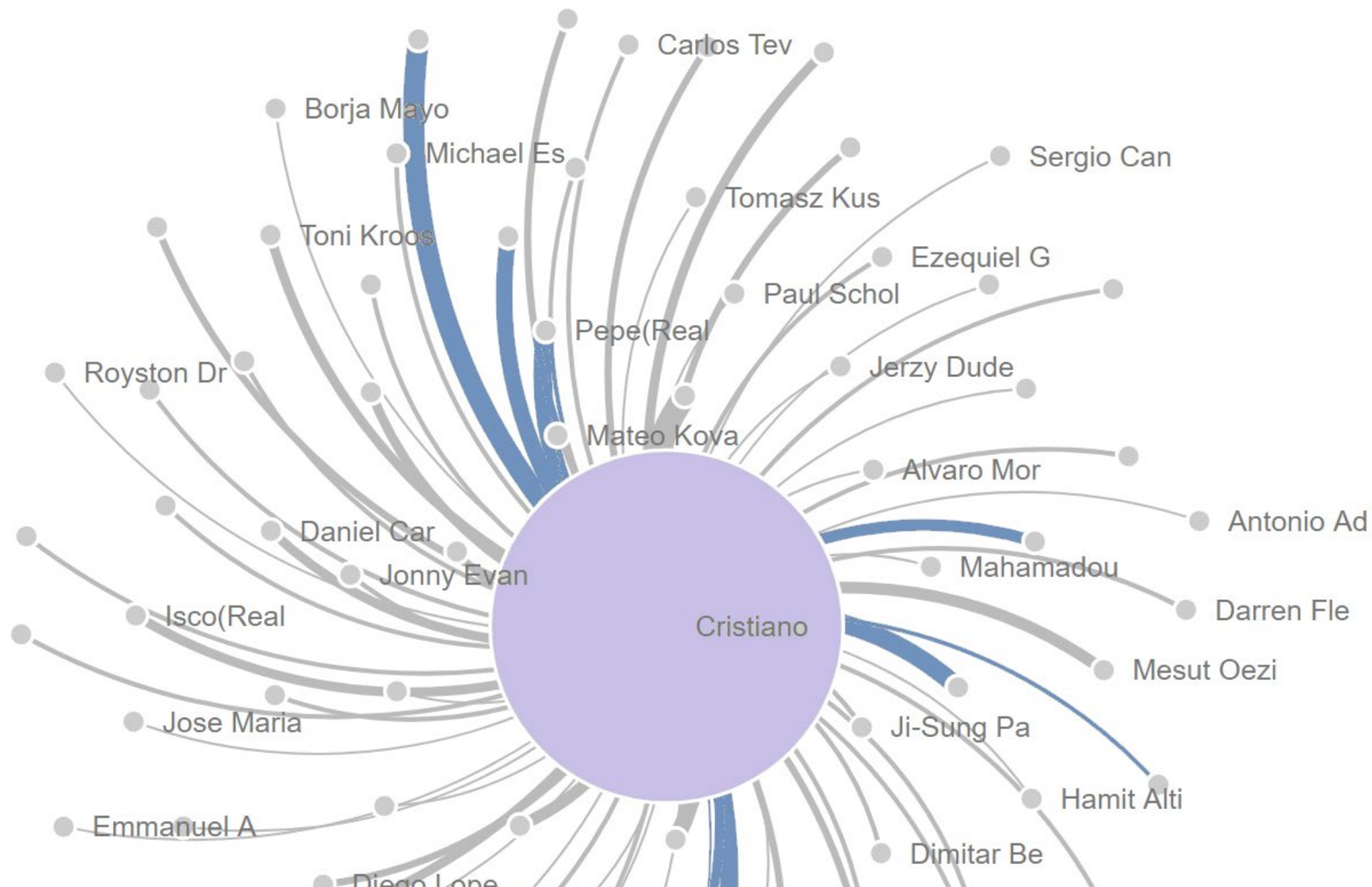
- São bastante diferentes
- **Entidades** no modelo relacional são tabelas e nos grafos cada elemento da tabela é uma entidade.
- Grafos tratam de **Relações**, mas estas podem ser mais complexas e numerosas que as definidas no modelo relacional
- Relacional: raramente passa de 20 relações entre tabelas
- Grafos: uma entidade pode facilmente ter milhares relações



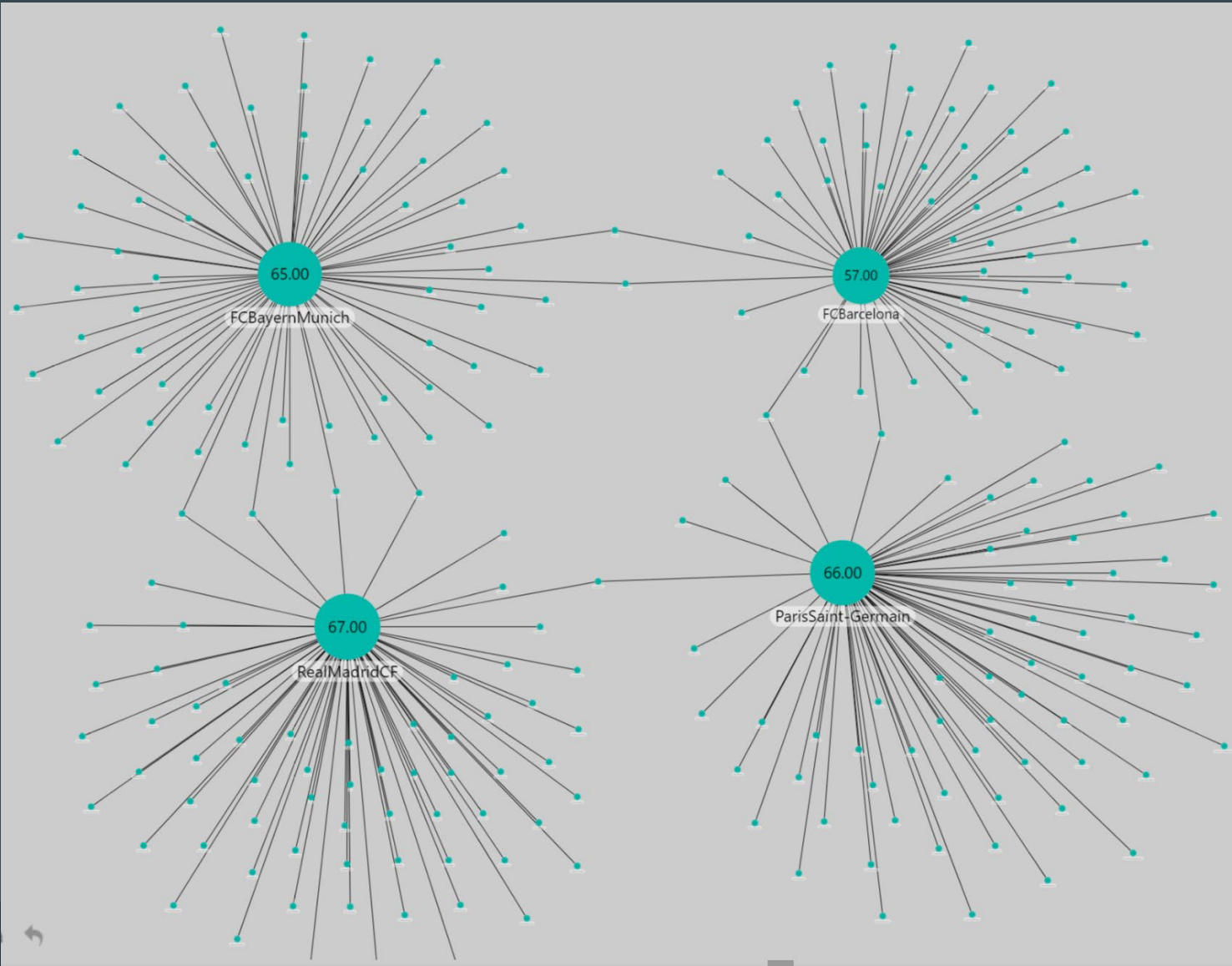
Um Único Grafo

- Número de jogos que cada jogador fez com ou contra Cristiano Ronaldo
- Linhas grossas representam maior número de jogos

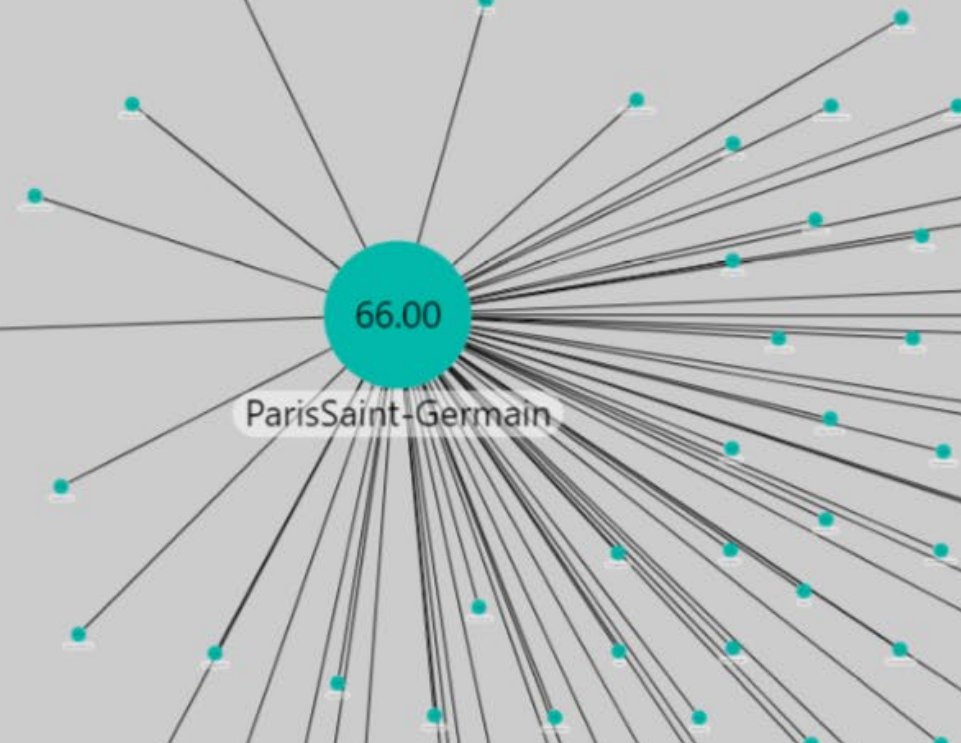
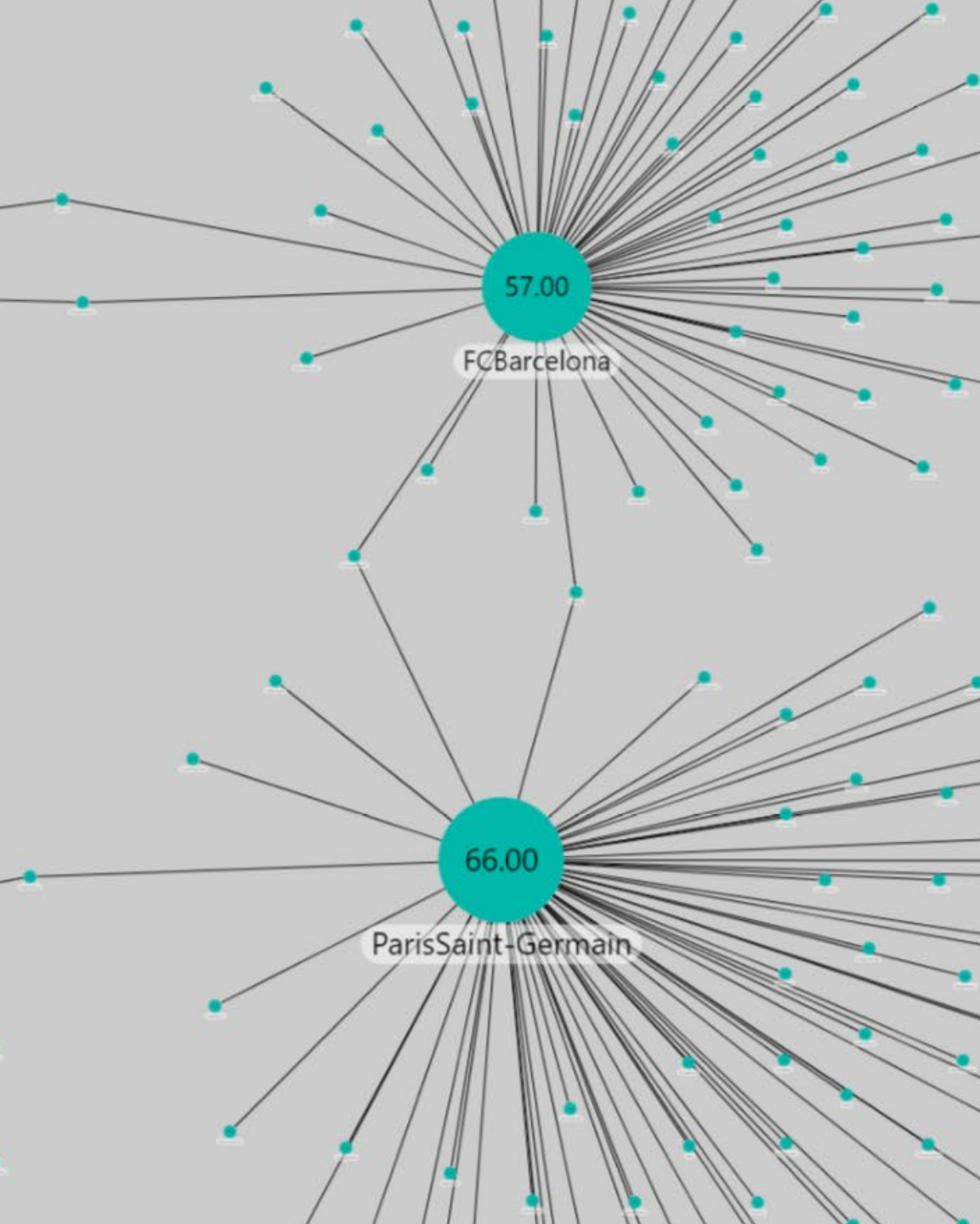
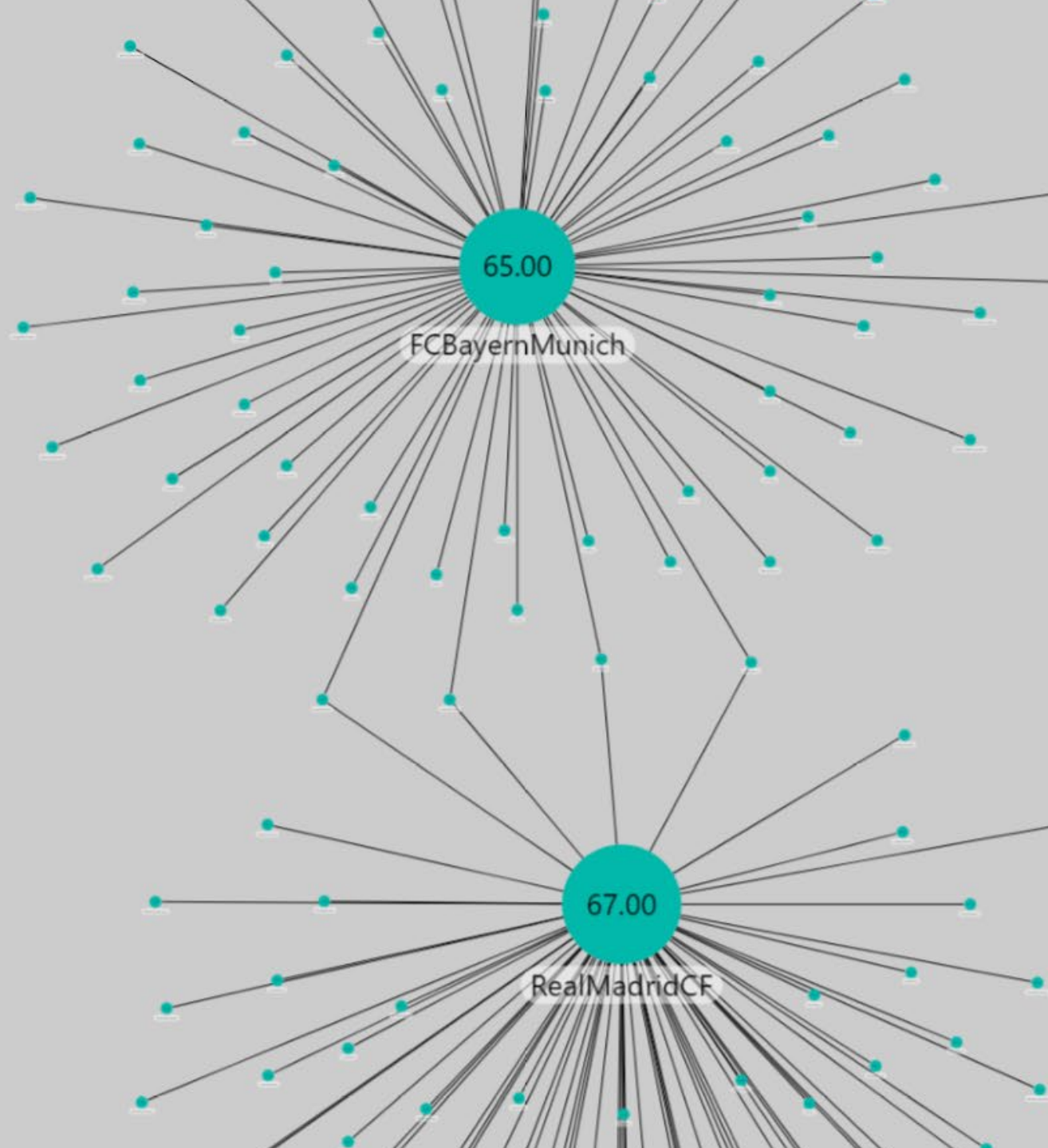




Relacionamentos entre Grafos



Jogadores que atuaram por cada Time num período de N temporadas



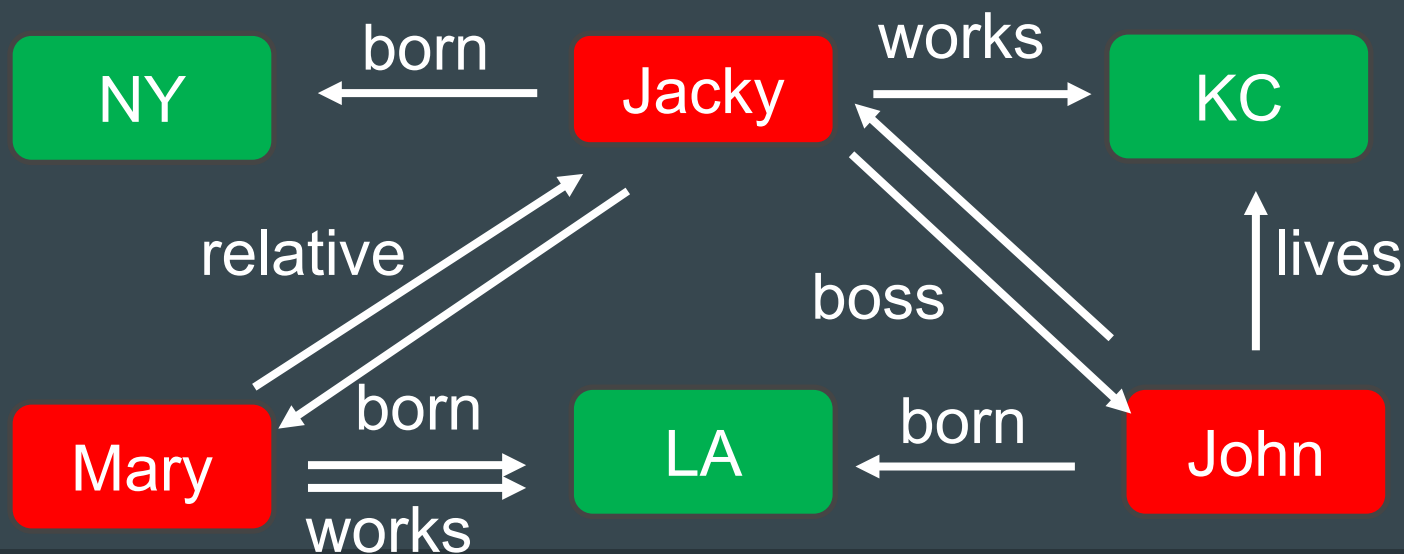
Definição do Modelo

Grafos e T-SQL

- Disponíveis desde 2017 (On-Premises e na nuvem)
- Visão lógica de dados
 - Tabelas de nós (NODE) lidam com ENTIDADES
 - Tabelas de bordas (EDGE) tratam de RELACIONAMENTOS
- Porta de entrada para universo dos grafos
- Carga de dados para modelo de grafos requer cuidado

Exemplo de Modelo de Grafos

- ENTIDADES: Cidades (verde), Pessoas (Vermelho)
- RELACIONAMENTOS: onde nasceu, onde mora, onde trabalha, quem é parente, quem é chefe



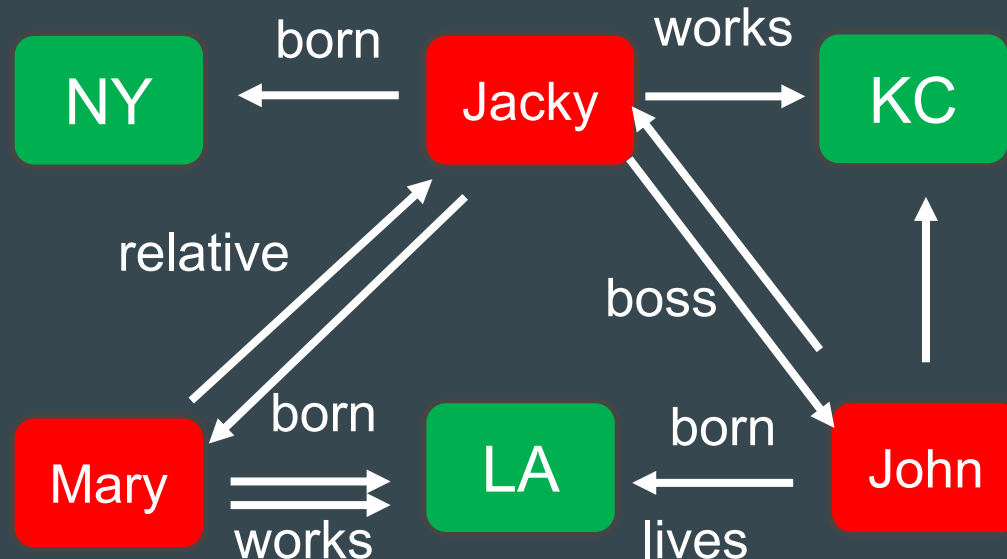
Definição do Modelo de Grafos

1. **NODES** (i.e. Entities)

- a) **People** (Jacky, John, Mary)
- b) **Cities** (NY, LA, KC)

2. **EDGES** (i.e. Relationships)

- a) Born in
- b) Works in
- c) Lives in
- d) Relative
- e) Boss



Implementação do Modelo de Grafos

- Importante:

- **BORDAS** são definidas com uma **DIREÇÃO**

Person WAS BORN IN City

Person WORKS IN City

Person1 IS A RELATIVE Person2

Person2 IS A RELATIVE Person1

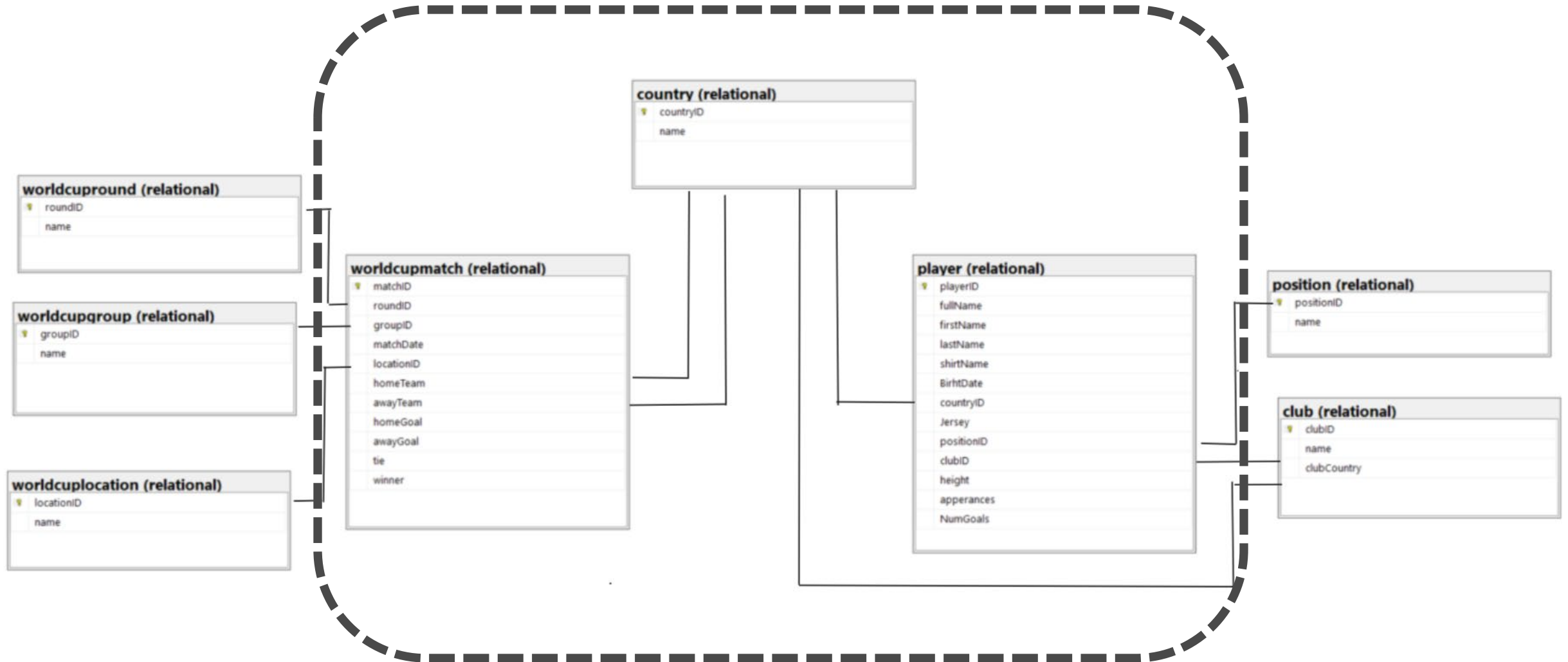
Rel. Bidirecionais:
tratar no **modelo** ou
então na **consulta**

Mão na Massa

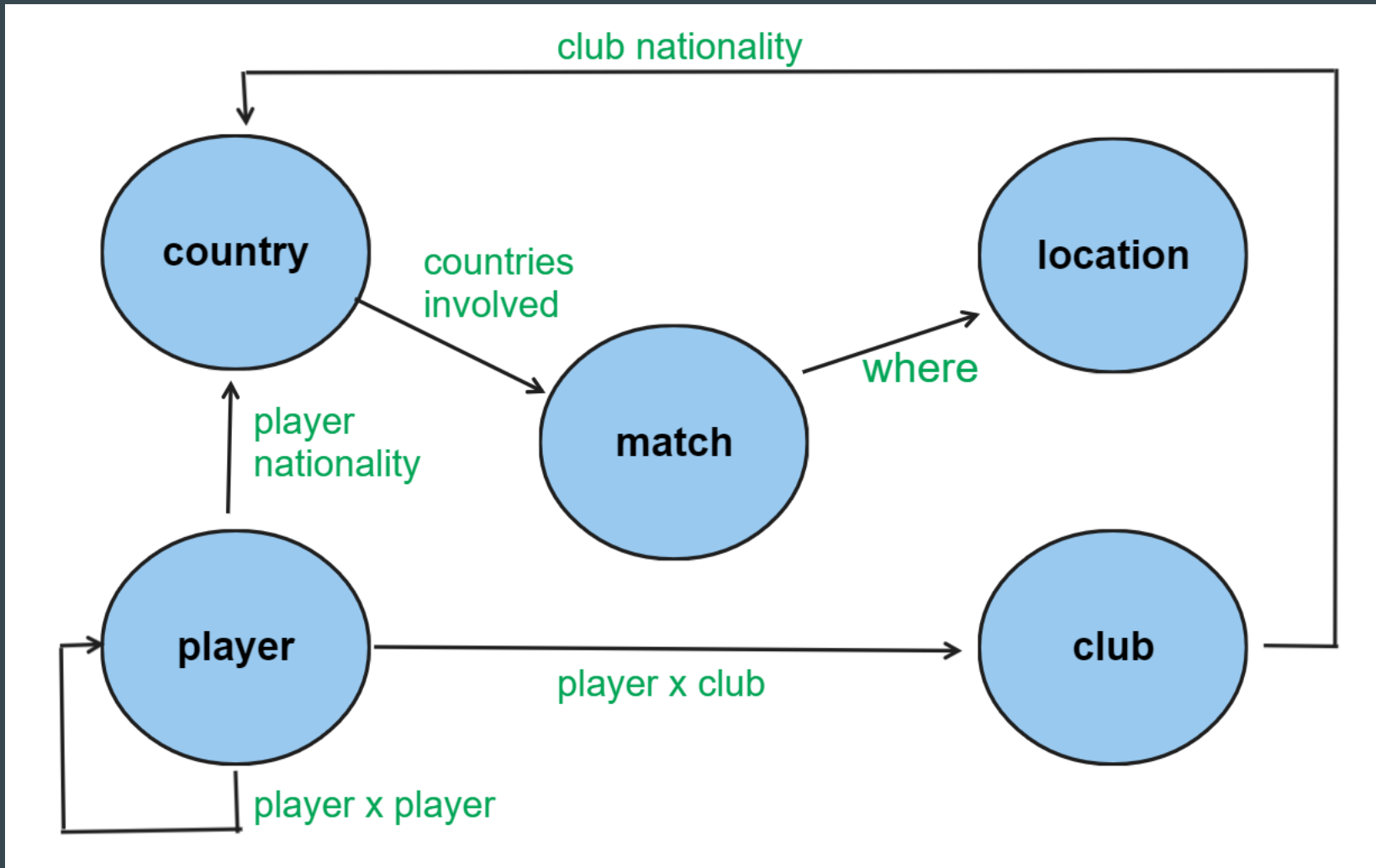
- Dataset: **Copa do Mundo 2022** (FIFA)
- 32 equipes
- 831 atletas
- 296 clubes



FIFA 2022 – Modelo Relacional



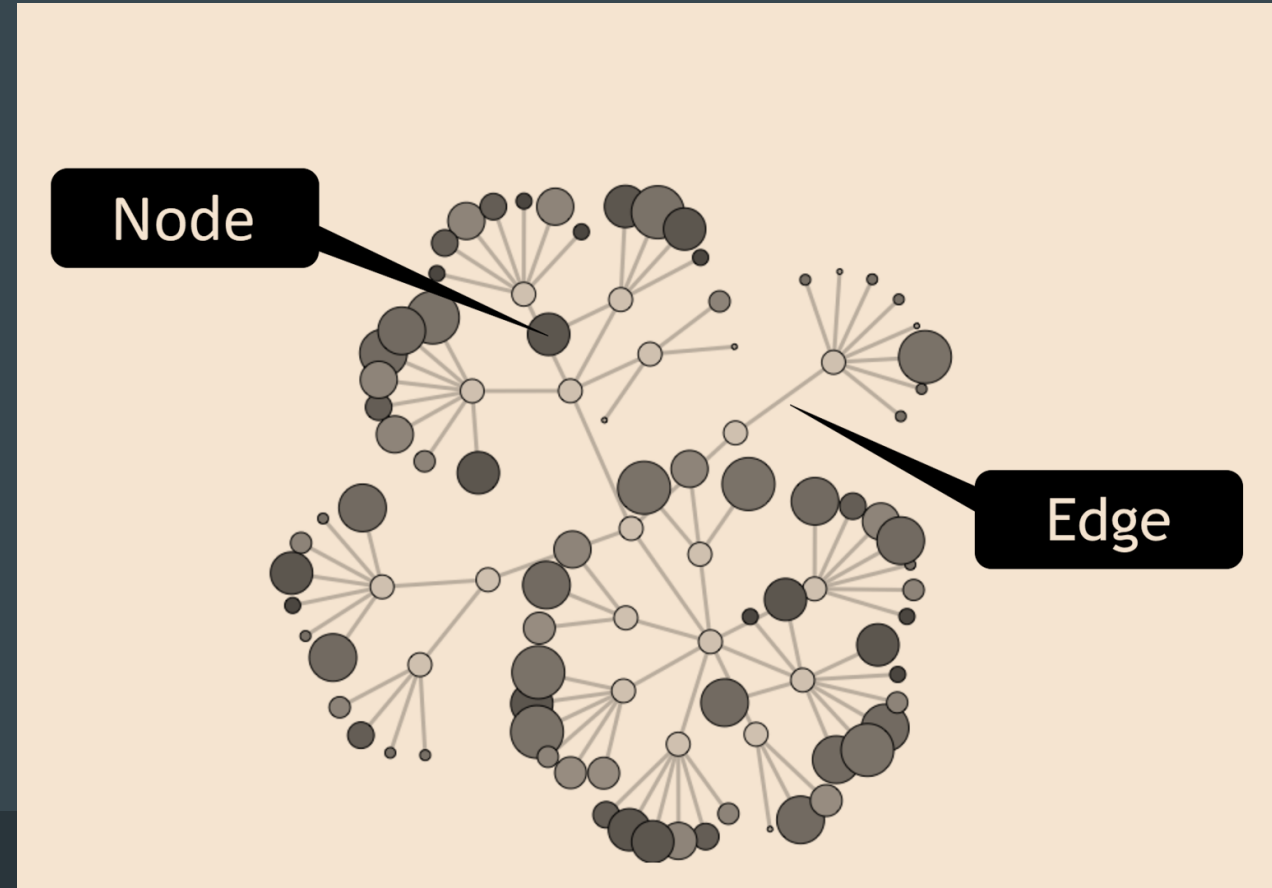
FIFA 2022 – Modelo de Grafos



Implementação do Modelo

Grafos e T/SQL

- Criando tabelas de nós e bordas no Azure SQL DB/SQL Server 2025
- Carga de dados
- Consultas



DDL para Grafos

- Tabelas de **Nós** usam `CREATE TABLE (...) AS NODE`
- **Bordas** usam `CREATE TABLE [(.)] AS EDGE`
- Os nós criam automaticamente uma coluna `$node_id`

```
{"type":"node","schema":"node","table":"club","id":0}
```



Demo

Carga de Dados para Tabelas de Grafos

- INSERT normal para tabelas de **NÓS**
- INSERT nas **BORDAS** deve trazer os **\$node_id** das tabelas envolvidas mais os atributos adicionais

```
INSERT INTO edge.clubXcountry
SELECT NCL.$node_id , NCO.$node_id
FROM node.club NCL
INNER JOIN relational.club RC ON NCL.ID = RC.ID
INNER JOIN node.country NCO ON CO.ID = RC.ID
GO
```



Demo

Análises com Grafos

Consultando Dados de Grafos

- Principal diferença:
Não existe JOIN.



- O que se usa:
Operador **Match** na cláusula **WHERE**

Ex:

```
WHERE MATCH (t1-(t2)->t3)
```



Demo

VIX25

www.dataes.com.br 

Observe a DIREÇÃO

- Operador MATCH usa ao menos 1 nó e 1 borda
- Mas o MATCH define a direção da busca de dados

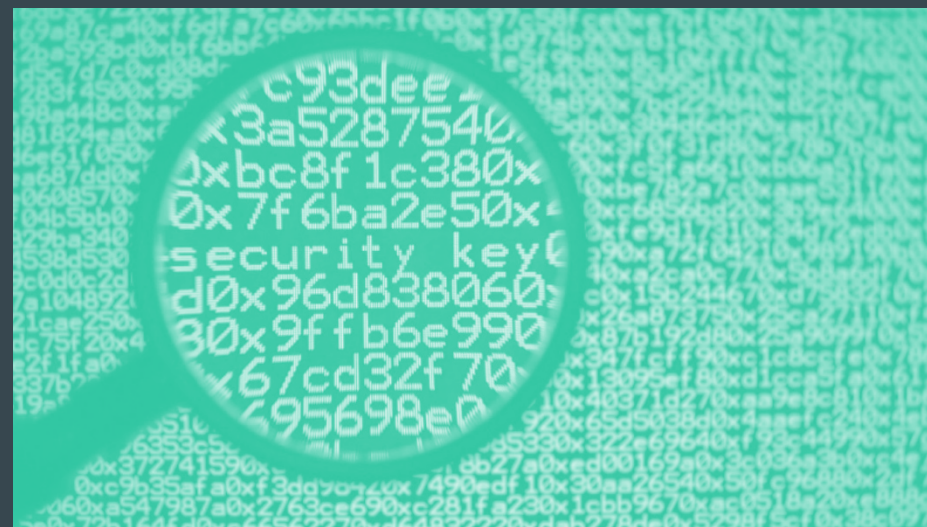
```
SELECT .....  
FROM nodes.Player P1, edges.PlayersInvolved PIN, nodes.Game M  
WHERE MATCH (t1-(t2)->t3)
```

É idêntico a

```
SELECT .....  
FROM nodes.Player P1, edges.PlayersInvolved PIN, nodes.Game M  
WHERE MATCH (t3 <-(t2)-t1)
```

Relacionamentos Duplos

- Exemplo: COUNTRY vs MATCH
- Partida: Mandante vs Visitante



- `WHERE MATCH (C1<-(MC1)-M-(MC2)-> C2)`
- `AND MC1.teamStatus = 'home'`
- `AND MC2.teamStatus = 'away'`

`AND M.SEASON = '2009/2010'`



Demo

VIX25

Função SHORTEST_PATH

- Quando não se sabe a “distância” do relacionamento
- “Amigo de um amigo de um amigo”
- Relacional é excessivamente complexo nesses casos

```
WHERE MATCH(SHORTEST_PATH(no1(-(borda)->no2)+))  
AND no1.Player = 'Player1'
```



http://bit.ly/sql_shortest_path

VIX25

www.dataes.com.br 

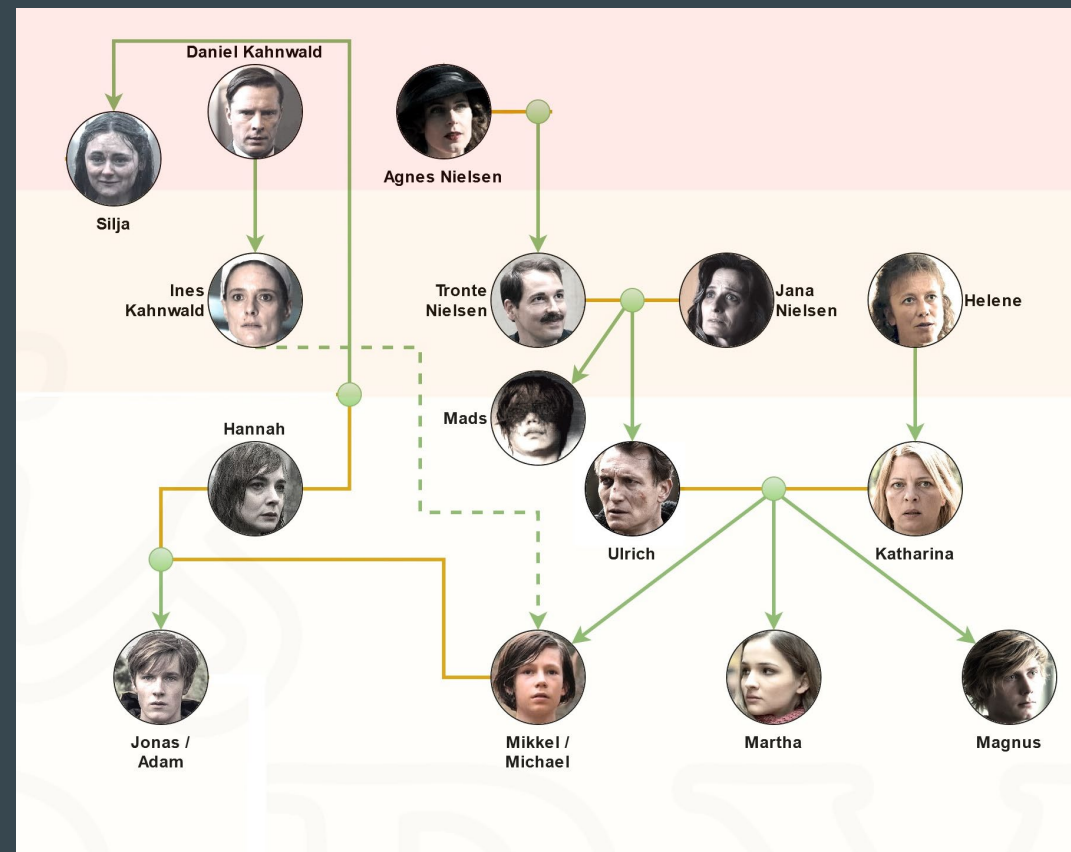
Menor Caminho: Relacional vs Grafo



Demo

Comentários Finais

- Grafos são muito úteis
- Sintaxe TSQL é fácil de lidar
- Carga de dados não é elementar
- Escolha relacionamentos relevantes
- POWER BI oferece recursos para apresentação dos grafos





Leituras Sugeridas

1. Bate-bola com Grafos e SQL Server Parte 1 – Definindo o Modelo – bit.ly/sqlgraph1
2. Bate-bola com Grafos e SQL Server Parte 2 – Implementando o Modelo – bit.ly/sqlgraph2
3. Bate-bola com Grafos e SQL Server Parte 3 – Visualizando Dados – bit.ly/sqlgraph3



Perguntas

Obrigado!



<https://www.linkedin.com/in/wagner-crivelini/>