



# Aplicaciones de la Analítica de relaciones

Alonso Guevara

# Agenda

- LLMs
- RAG
- Grafos
- Analítica de grafos
- Comunidades y Clústeres
- Centrality
- Graph RAG
- Casos de uso
  - Ucrania y Rusia
  - Tech Against Trafficking
  - Corrupción en procuraduría
  - Análisis de info textual
  - Archivo Nacional

# Agenda

- Casos de uso
  - Ucrania y Rusia
  - Tech Against Trafficking
  - Corrupción en procuraduría
  - Análisis de info textual
  - Archivo Nacional



- “El auténtico genio consiste en la capacidad para evaluar información incierta, aleatoria y contradictoria”
- -Winston Churchill



- “True genius resides in the capacity for evaluation of uncertain, hazardous, and conflicting information.”
- -Winston Churchill

- Modelo de inteligencia artificial diseñado para comprender y generar texto en lenguaje natural a una escala considerable.
- Pueden generar texto coherente y relevante en respuesta a preguntas o instrucciones dadas.



- Los LLMs pueden ser pre-entrenados en grandes conjuntos de datos y luego ajustados
- El uso de LLMs plantea preguntas importantes sobre la privacidad de los datos, el sesgo algorítmico y la generación de contenido potencialmente dañino o engañoso.



# RAG

- Retriever-Augmented Generation
- Retriever: Este componente se encarga de recuperar información relevante de una gran base de datos o corpus de texto en respuesta a una consulta dada
- Generator: Una vez que el Retriever ha identificado los documentos relevantes, el componente de generación toma esa información y la utiliza para generar una respuesta coherente y relevante.

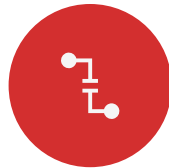




# Grafos



Un grafo es una estructura matemática que consta de un conjunto de nodos y un conjunto de aristas que conectan estos nodos entre sí.

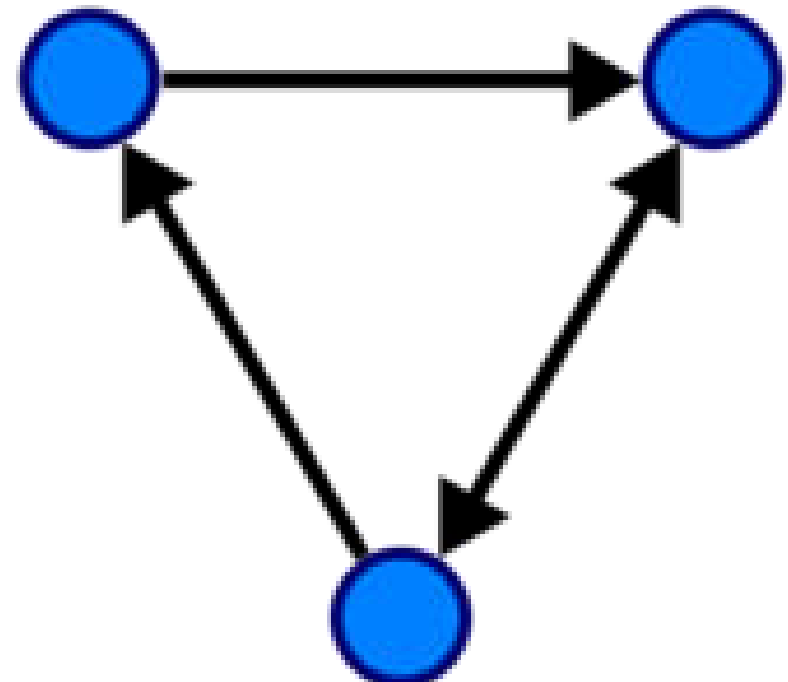


En un grafo, los nodos representan entidades individuales, mientras que las aristas representan las relaciones o conexiones entre estas entidades.



# Grafos

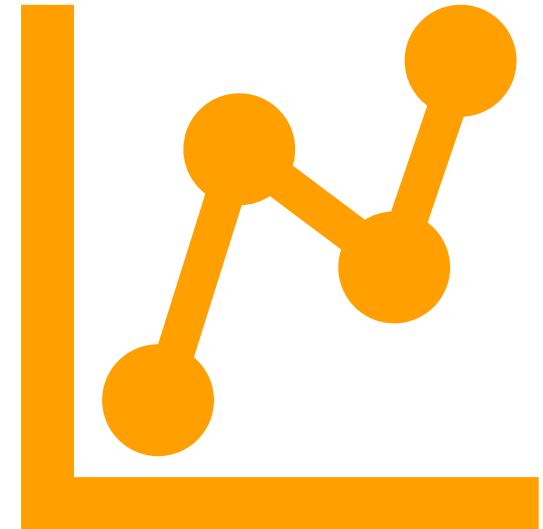
- Herramienta poderosa para análisis y visualización de datos
- Usos comunes:
  - Redes
  - Estructuras organizacionales
  - Búsqueda y recorridos
  - Bases de datos



This Photo by Unknown Author is licensed under [CC BY-SA](#)

# Analítica de grafos

- Campo interdisciplinario que se centra en el análisis y la interpretación de datos estructurados en forma de grafos.
- Se basa en técnicas y algoritmos específicos diseñados para extraer información significativa de las relaciones y estructuras presentes en los grafos.



# Comunidades y Clústers

- Proceso mediante el cual se identifican conjuntos de nodos en el grafo que están más densamente conectados entre sí que con el resto de los nodos en el grafo.
- Estos conjuntos de nodos, conocidos como comunidades, representan subgrupos o clústers dentro del grafo que exhiben una fuerte cohesión interna y una conectividad relativamente baja con los nodos fuera de la comunidad.

# Centralidad

- La centralidad en un grafo es una medida que evalúa la importancia o influencia de un nodo en la red en función de su posición y sus conexiones con otros nodos. En otras palabras, la centralidad de un nodo indica cuán "central" o relevante es dentro del grafo.
- **Centralidad de grado:** Mide la importancia de un nodo en función del número de conexiones que tiene con otros nodos en el grafo. Los nodos con un alto grado de centralidad de grado están más conectados en la red.

# Graph RAG

- GraphRAG es una técnica para buscar información basada en una consulta del usuario y proporcionar los resultados como referencia para que un AI genere una respuesta.



# Graph RAG

- La mayoría de los enfoques de RAG utilizan la similitud de vectores como técnica de búsqueda.
- Graph RAG utiliza grafos de conocimiento generados por LLMs para proporcionar mejoras sustanciales en el rendimiento de preguntas y respuestas al realizar análisis de documentos de información compleja. Esto se basa en nuestra investigación reciente, que señala el poder del aumento de sugerencias al realizar descubrimientos en conjuntos de datos privados.

Query: "What is Novorossiya?"

<b>Baseline RAG</b>	<b>GraphRAG</b>
<p>The term "Novorossiya" is not mentioned in the provided context. However, to provide some information, Novorossiya, or New Russia, is a historical term referring to a region in the south of Ukraine that was conquered by the Russian Empire in the 18th century. In recent years, the term has been revived by Russian-backed separatists in Eastern Ukraine to refer to the territories they claim, but it is not officially recognized as a separate entity by international bodies.</p>	<p>Novorossiya, also known as New Russia, is a term with historical and contemporary significance. Historically, it referred to a region in the southern part of the Russian Empire that is now part of Ukraine. In the contemporary context, it has been used by pro-Russian separatists to refer to certain areas of Ukraine [Data: Entities (6494, 912, 6493)].</p> <p>The term Novorossiya has been used in the context of a political movement in Ukraine, particularly by Russian nationalists. This movement has been associated with efforts to destabilize Ukraine and create a so-called "Novorossiya" [Data: Entities (912); Relationships (15211, 15206)]. The term has also been used to denote a separatist state in Ukraine, known as the Federal Republic of Novorossiya (FRN) [Data: Entities (7690)].</p>



Query: "What has Novorossiya done?"

<b>Baseline RAG</b>	<b>GraphRAG</b>
<p>The text does not provide specific information on what Novorossiya has done.</p>	<p>Novorossiya, a political movement in Ukraine, has been involved in a series of destructive activities, particularly targeting various entities in Ukraine [Entities (6494, 912)]. The movement has been linked to plans to destroy properties of several Ukrainian entities, including Rosen, the Odessa Canning Factory, the Odessa Regional Radio Television Transmission Center, and the National Television Company of Ukraine [Relationships (15207, 15208, 15209, 15210)]. In addition to these, Novorossiya has also been associated with plans to destroy properties of Roshen, the Odessa Conservatory Factory, the Odesa Regional Radio Television Transmission Center, and the National Telecommunications Company of Ukraine [Relationships (8329, 8330, 8331, 8332)]. Furthermore, Novorossiya has been implicated in plans to blow up ATMs of PrivatBank, the largest state bank in Ukraine [Relationships (8335, 8335)].</p> <p>The Office of the General Prosecutor in Ukraine has reported on the creation of Novorossiya, indicating the government's awareness and potential concern over the activities of this movement [Relationships (8264, 8264)]. The destructive activities of Novorossiya have had significant impacts on the targeted entities and have likely contributed to the ongoing tensions and conflicts in the region. In summary, Novorossiya has been involved in a series of destructive activities targeting various entities in Ukraine. These activities have had significant impacts on the targeted entities and have likely contributed to the ongoing tensions and conflicts in the region. The Ukrainian government, through the Office of the General Prosecutor, has acknowledged the existence and activities of Novorossiya, indicating a level of concern over the movement's actions.</p>

Baseline RAG	GraphRAG
<ol style="list-style-type: none"><li>1. Improving the quality of life in cities and urban development.</li><li>2. The meeting between Vladimir Putin and the President of the United Arab Emirates, Mohammed bin Zayed Al Nahyan.</li><li>3. The National Rating of the Investment Climate in the Regions of Russia.</li><li>4. The state of the Russian economy and the growth of domestic brands.</li><li>5. The discussion on the meaning of life and the role of science in understanding it.</li></ol>	<p>The data presents a wide range of themes, but the top five most prevalent themes can be identified as follows:</p> <ol style="list-style-type: none"><li><b>1. Conflict and Military Activity:</b> A significant portion of the data revolves around the ongoing conflict in Ukraine, with numerous reports detailing military activities, occupations, and attacks by various entities. This theme includes the actions of the Ukrainian military, Russian forces, and various other military units and entities [Data: Reports (513, 241, 151, 493, 343, +more)].</li><li><b>2. Political and Governmental Entities:</b> Many reports focus on political figures and governmental entities, highlighting their roles, relationships, and actions in the context of the current situation. This includes entities like the Kremlin, the President of Ukraine, the US State Department, and various regional administrations [Data: Reports (98, 493, 265, 129, 196, +more)].</li><li><b>3. Infrastructure and Environmental Concerns:</b> Several reports highlight potential environmental impacts and infrastructure threats. These include potential environmental damage due to industrial activities, flooding crises, and threats to infrastructure due to military conflicts [Data: Reports (308, 154, 201, 348, 216, +more)].</li><li><b>4. Community Analysis and Threat Assessment:</b> Several reports provide detailed analyses of specific communities, often centered around a particular location or entity. These analyses typically discuss the key entities within the community, their relationships, and any associated threats. Many reports also include assessments of potential threats associated with the entities or situations they discuss [Data: Reports (323, 596, 518, 583, 179, +more)].</li><li><b>5. Health, Humanitarian Concerns, and Emergency Situations:</b> Some reports focus on health and humanitarian issues, such as the Shereshevsky-Turner Syndrome community and the potential risk of intestinal diseases in the Odessa region. The data also covers emergency situations, including natural disasters, accidents, and attacks [Data: Reports (204, 149, 372, 154, 532, +more)]. These themes provide a comprehensive overview of the current situation, highlighting the complexity and interconnectedness of the issues at hand.</li></ol>

# Casos de uso



# Ucrania y Rusia

- Análisis de ciber ataques
- Estudio de orígenes y destinos de información falsa y mal intencionada



# Tech Against Trafficking

- La analítica de relaciones puede ser utilizada para combatir la trata de personas.
- Identificación de posibles víctimas y alerta automática
- Análisis de herramientas contra la trata de personas.



HOME

MAP

CLEAR

111 Tools Selected

TOOL CATEGORY



TRAFFICKING TYPE



TARGET SECTOR



TARGET USERS



TECHNOLOGY



KEY PHRASES



**1343 Actionline Hotline**  
Hotline

The 1343 Actionline is a 24/7 hotline facility that responds to emergency or crisis calls from victims of human trafficking and their families.

**Analyze**  
Software

Analyze is a full suite for digital investigations into hotlinks (unique identifiers) of known child sexual abuse material (CSAM) found on the

**Apprise App**  
Mobile application

Apprise is a smartphone app that helps NGOs and authorities identify potential victims of exploitation/trafficking.

**Artic**  
Mobile

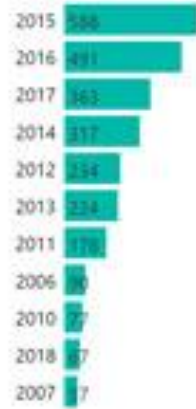
Artic is a mobile application that helps NGOs and authorities identify potential victims of exploitation/trafficking.

# CTDC Global Dataset on Victims of Trafficking

**Privacy resolution (10)** the minimum group size detectable in synthetic/aggregate data  
**Estimated counts:** from synthetic data that reflects the sensitive data at the given resolution  
**Actual counts:** from aggregate data rounded down to the closest multiple of the resolution



## Registration



## Gender



## Age



## Citizenship



## Comparison for Selections

● Actual ● Estimated



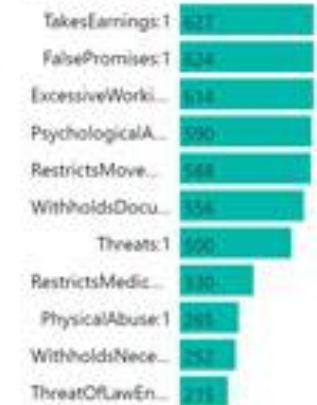
## Exploitation



## TraffickingType



## MeansOfControl



## Recruiter



## Compare

All



# Corrupción en Procuraduría

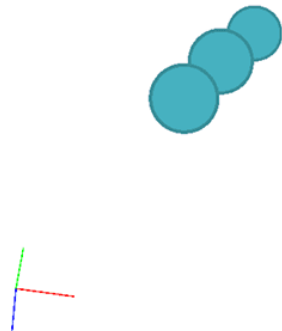
- Una solución que utiliza la incrustación dinámica de gráficos multipartitos con USE (Unfolded Spectral Embedding) para crear representaciones vectoriales de cada empresa que incorpora información de su actividad de adquisiciones históricas, estructura de propiedad y otras fuentes.



- El resultado es un "motor de transparencia" que busca detectar fuentes de relación no descubiertas y luego las desglosa en explicaciones visuales que pueden ser exploradas y evaluadas en el contexto de Microsoft Power BI.
- Para muchos tipos de análisis anticorrupción, no es necesario conocer exactamente quién controla una empresa en particular, sino si un grupo dado de empresas tiene una relación lo suficientemente fuerte como para sugerir posibles no independencias, lo que podría motivar investigaciones más profundas sobre posible coordinación, colusión o control beneficioso común.

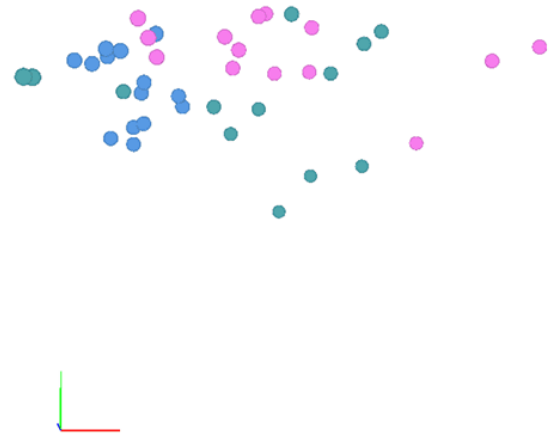
**Dynamic multipartite graph embedding** allows us to combine and compare the structure and activity of a company over time. Companies with *similar behavior* are placed in *similar positions*, in ways that may indicate common ownership or control. The relatedness of any pair of companies can be computed from the distance between their positions in the embedded space.

Combining company behavior over all periods



Colors represent clusters of similar companies.

Comparing company behavior across time periods



Drag to rotate, shift-drag to pan, scroll to zoom.

Colors represent the same company over time.

Maximum relatedness

1.00

Average relatedness

1.00

Similar company count

Search  [Clear All](#)

X Company::8098 X Company::8296

X Company::8298

Company::8153	38
Company::7936	28
Company::8020	28
Company::8179	25
Company::7942	24
Company::8042	24
Company::8150	24
Company::8232	24
Company::7898	23
Company::8206	23

Select to calculate N-company relatedness.

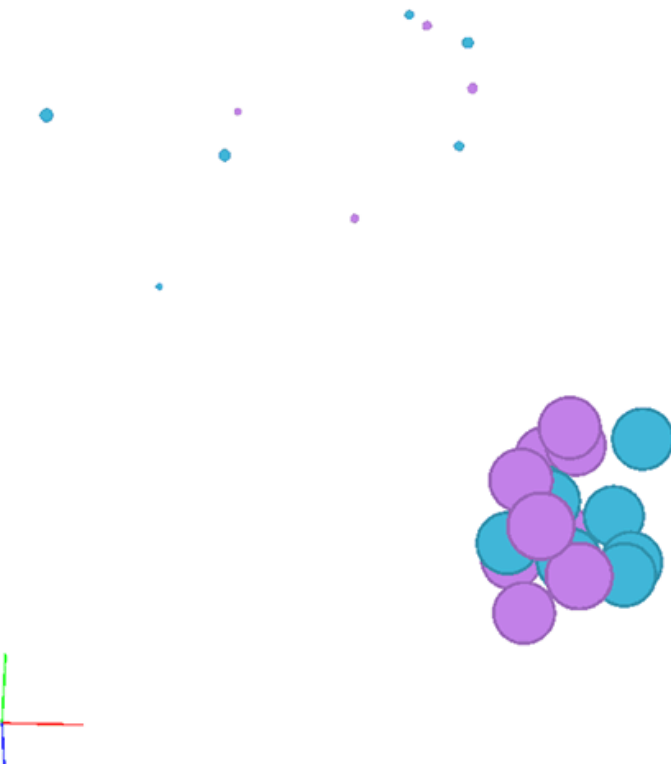
**Dynamic multipartite graph embedding** allows us to combine and compare the structure and activity of a company over time. Companies may exhibit similar behavior at the same time (*synchronous similarity*) or different times (*asynchronous similarity*). Common ownership or control of two companies could give rise to either form, depending on each company's periods of activity.

Combining company behavior over all periods



Colors represent clusters of similar companies.

Comparing company behavior across time periods



Colors represent the same company over time.

Search for company

Synchronous similarity

Edge	SyncSimilarity
Company::8216 -- Company::8274	0.01

Asynchronous similarity

Edge	AsyncSimilarity
Company::8216 -- Company::8274	0.99
Company::8096 -- Company::8203	0.71
Company::8123 -- Company::8245	0.68
Company::8010 -- Company::8196	0.68

Select to show synchronicity relationship.

**Statistical measures of similarity** can be obtained from the vector-based representations of companies in the embedded space. Companies may exhibit similar behavior at the same time (*synchronous similarity*) or different times (*asynchronous similarity*). Common ownership or control of two companies could give rise to either form, depending on each company's periods of activity.

Search for company

Overall similarity



Synchronous similarity

[Clear All](#)



Asynchronous similarity

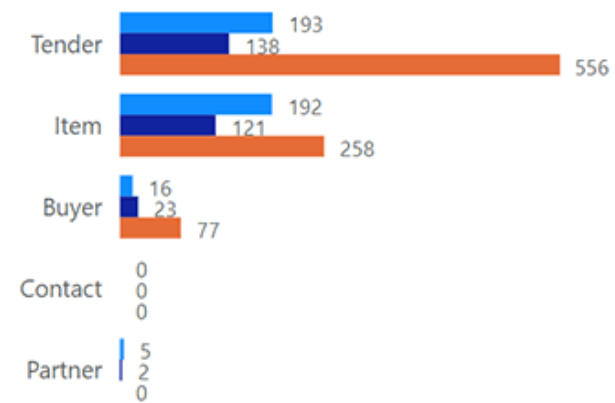


CommonContact  False  True

CommonPartner  False  True

Similarity by type

● Left Only ● Right Only ● Shared



Left Company

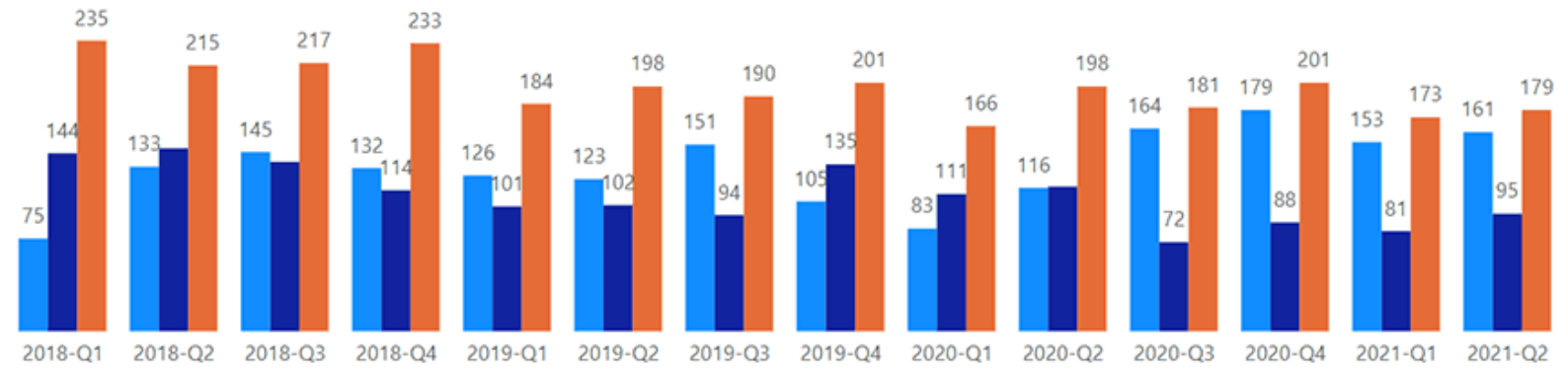
- Linked value
- Buyer::1000
- Buyer::1001
- Buyer::1002
- Buyer::1003
- Buyer::1004
- Buyer::1005
- Buyer::1006
- Buyer::1007
- Buyer::1008

Right Company

- Linked value
- Buyer::1000
- Buyer::1001
- Buyer::1002
- Buyer::1003
- Buyer::1004
- Buyer::1005
- Buyer::1006
- Buyer::1007
- Buyer::1008



Similarity by time



**Statistical measures of similarity** can be obtained from the vector-based representations of companies in the embedded space. Companies may exhibit similar behavior at the same time (*synchronous similarity*) or different times (*asynchronous similarity*). Common ownership or control of two companies could give rise to either form, depending on each company's periods of activity.

Search for company

Overall similarity



Synchronous similarity



Asynchronous similarity

[Clear All](#)

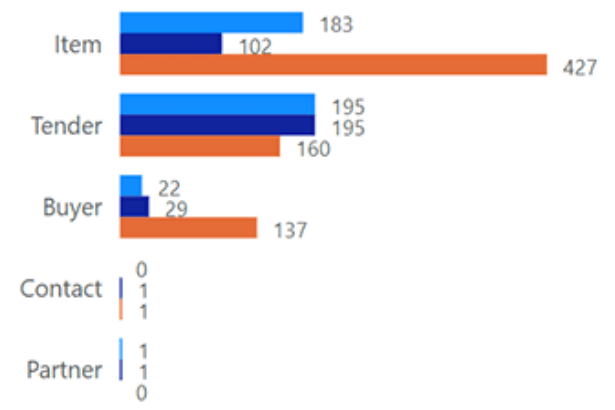


**CommonContact**  
 False  
 True

**CommonPartner**  
 False  
 True

Similarity by type

● Left Only ● Right Only ● Shared



Left Company

Linked value

- Buyer::1000
- Buyer::1001
- Buyer::1002
- Buyer::1003
- Buyer::1004
- Buyer::1005
- Buyer::1006
- Buyer::1007
- Buyer::1008

Right Company

Linked value

- Buyer::1000
- Buyer::1001
- Buyer::1002
- Buyer::1003
- Buyer::1004
- Buyer::1005
- Buyer::1006
- Buyer::1007
- Buyer::1008

0.47

0.29

401

725

328

Similarity

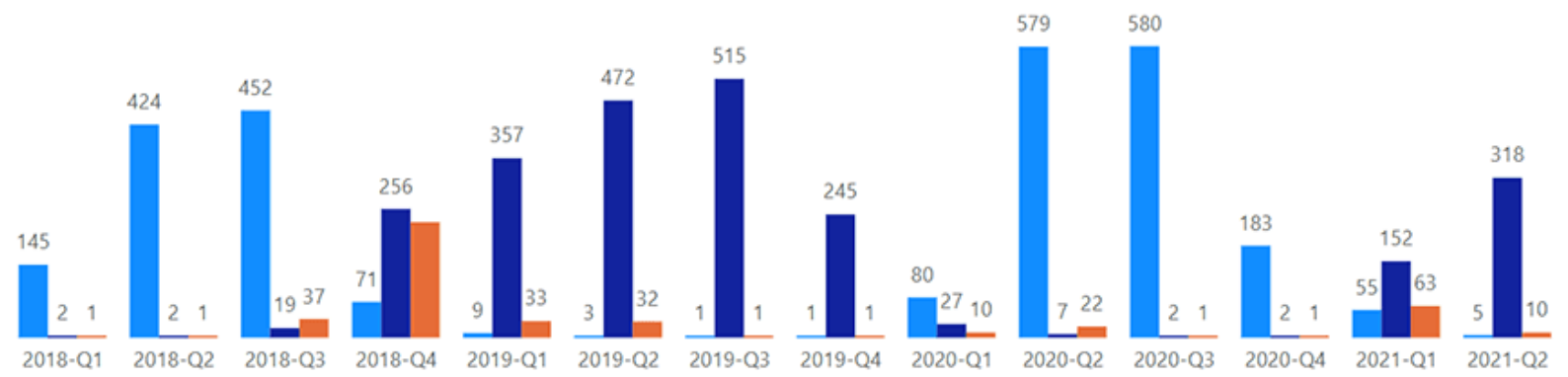
Synchronization

Left Company Only

Shared

Right Company Only

Similarity by time



**Clusters of related entities** can be extracted from variations in the density of the dynamic multipartite graph embedding. The cluster network (center) shows common partners, common contact details, and inferred company-company links (in red). The evidence graph (right) shows observed similarity by party and period for inferred company-company links (red).

### 1. Company

Company::8153	38
Company::7936	28
Company::8020	28
Company::8179	25
Company::7942	24
Company::8042	24
Company::8150	24
Company::8232	24
Company::7898	23

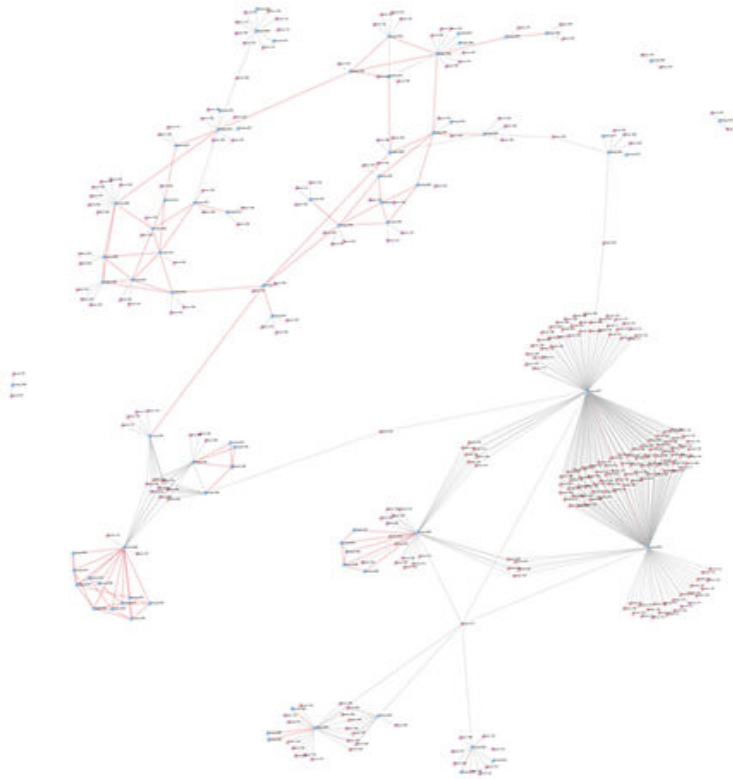
### 2. Cluster

Clear All

5	2486
8	1405
4	979
9	918
6	744
1	657
2	548
7	446
3	417
0	110

### 3. Load company cluster

### 4. View company cluster / Select company



### 5. Left company

 Company::7978

### 6. Right company

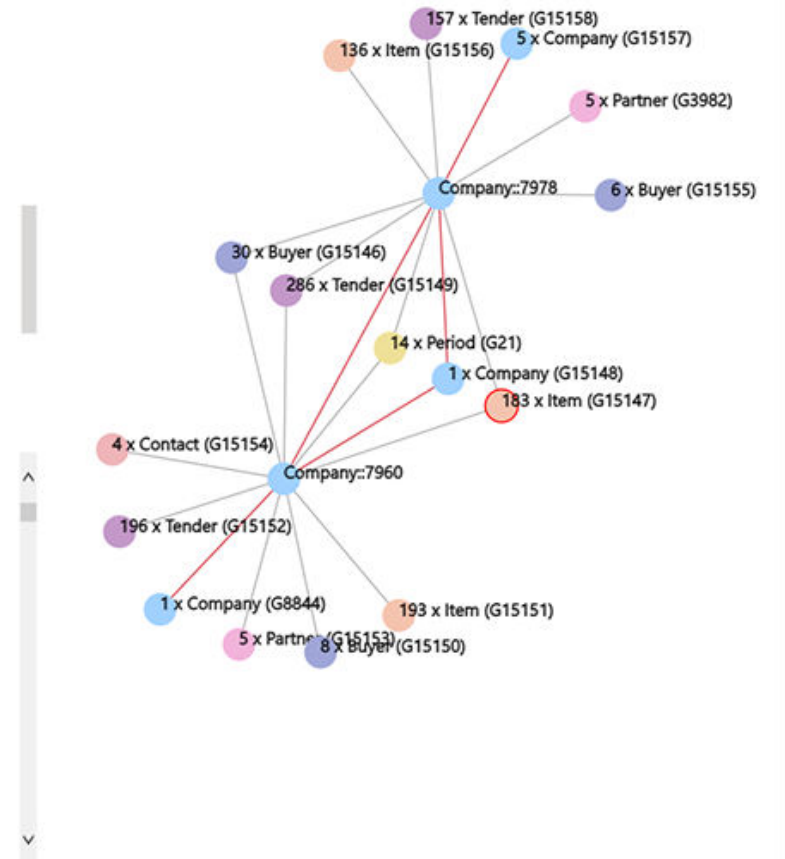
- Company::7932
- Company::7960
- Company::7978
- Company::7995

### 8. Attribute group

Value

- Item::1003
- Item::1004
- Item::1010
- Item::1012
- Item::1016
- Item::1018
- Item::1020
- Item::1021

### 7. View evidence graph / Select attribute group





# Análisis de info sin estructura

- Permite encontrar estructura y patrones en datos que carecen de esto
- Ingerir miles de archivos y estudios que para un humano o varios sería imposible.
- Aplicar Graph RAG para descubrimiento de información



# Archivo Nacional

- La analítica de relaciones puede ser utilizada para analizar datos históricos.
- La analítica de relaciones puede ser utilizada para identificar patrones y relaciones en los datos históricos.
- La analítica de relaciones puede ser utilizada para mejorar la comprensión de la historia.



# Gracias!

- Preguntas??

