



# Exploration de grands graphes de connaissance avec SemSpect

Vincent Vialard

[vialard@derivo.de](mailto:vialard@derivo.de)

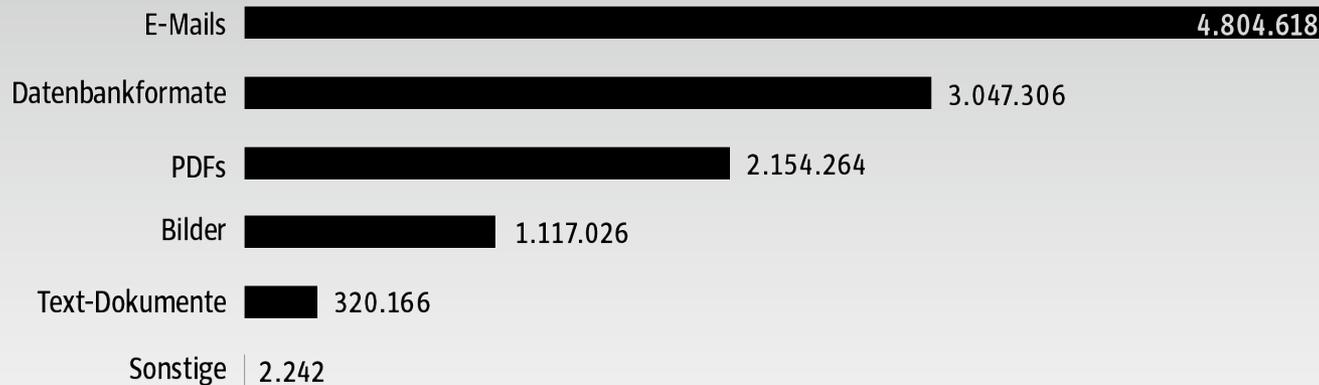
**derivo**

Semantische Systeme

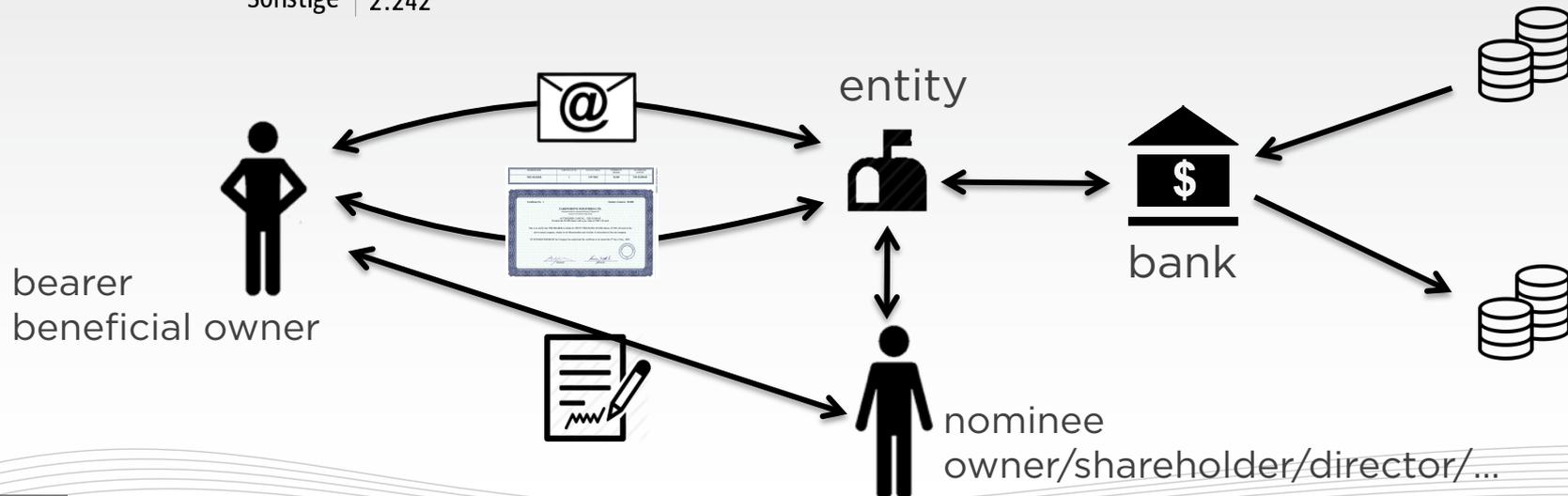
[www.derivo.de](http://www.derivo.de)

# Panama Papers Leak (ICIJ)

- 2,6 téraoctets de données
- 11,5 millions de documents

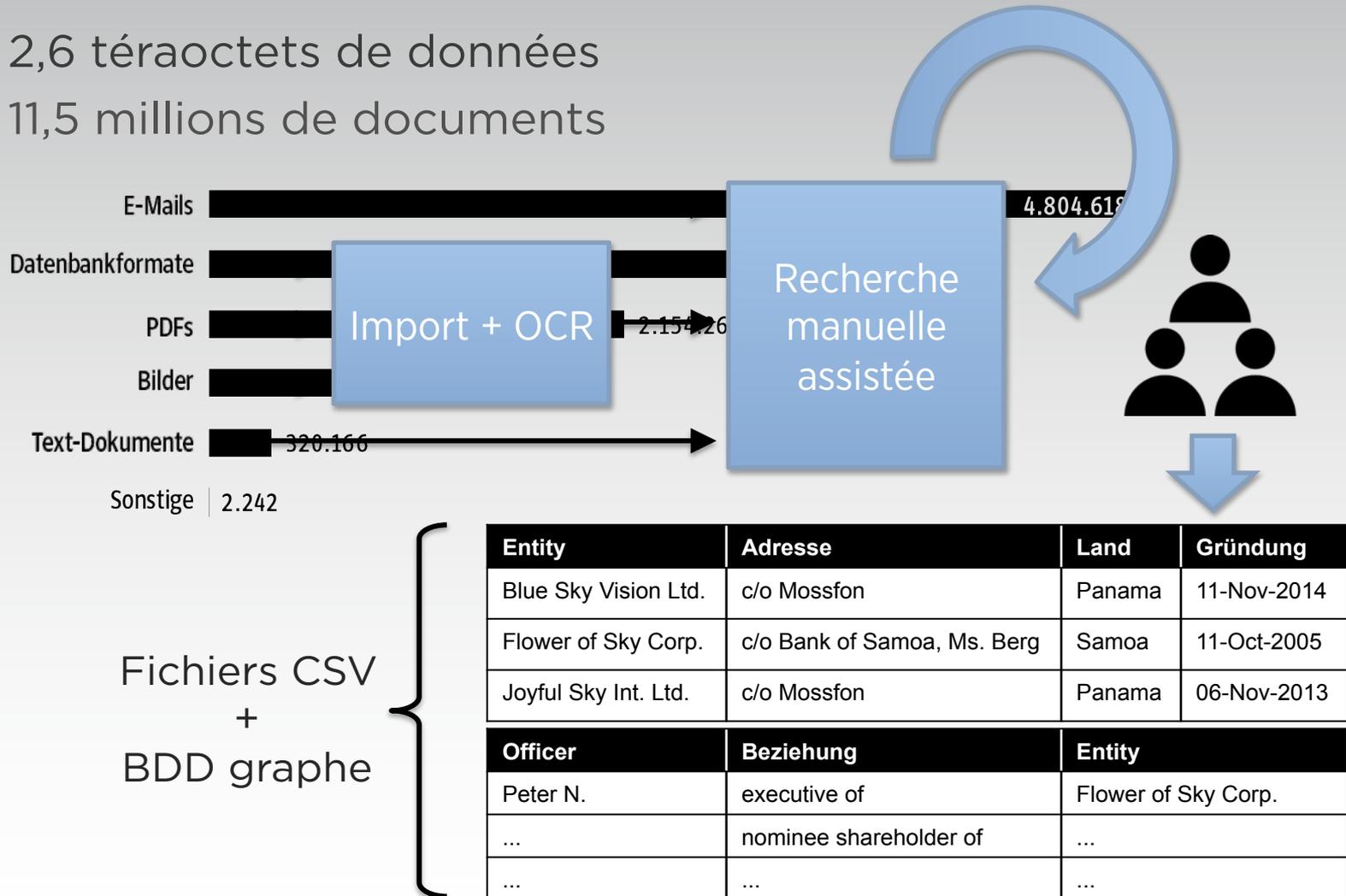


Quelle: SZ



# Données publiées par l'ICIJ

- 2,6 téraoctets de données
- 11,5 millions de documents

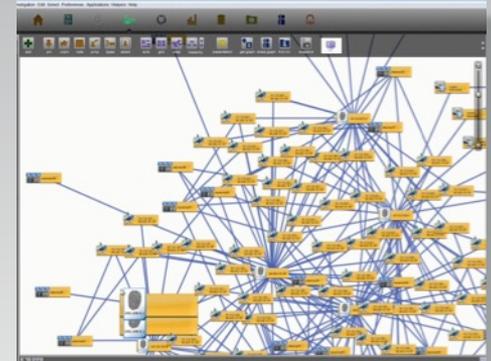




# Approches classiques

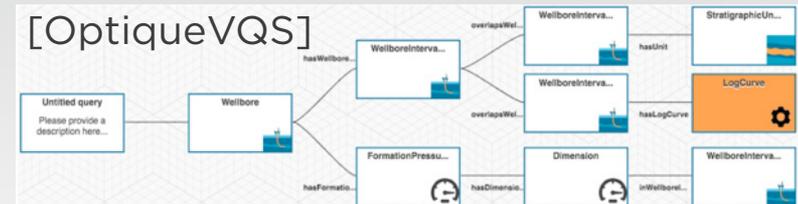
- Visualisation du graphe

[Palantir Gotham]

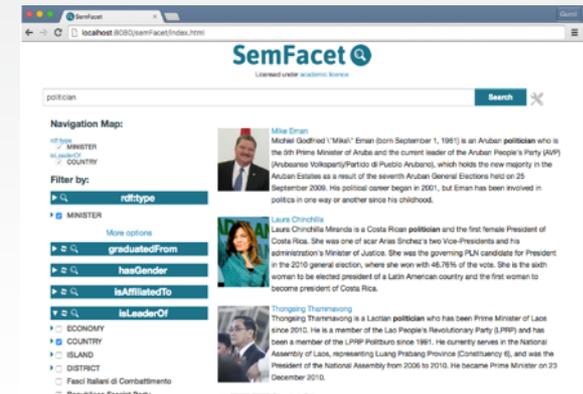


- Interfaces de requêtes (textuelle, graphique)

[OptiqueVQS]



- Navigation par facettes (SemFacet, Sparklis...)



# Exploration sur le site de l'ICJI

**ICJI**

**Jérôme Cahuzac**  
Budget minister at the Ministry of the Economy, Finance and External Trade (2012-2013); Deputy, National Assembly of France (1997-2002, 2007-2012)

[Read the story](#) [Explore the data](#)

**MONFORT CAPITAL PARTNERS JLT** registered  
Shareholder → **CERMAN GROUP LIMITED** Beneficial owner  
Beneficiary → Jérôme Cahuzac  
Beneficiary → Mr. Jerome Andre CAHUZAC  
Shareholder → **TALWAY INTERNATIONAL CORP.** registered address  
85 avenue de Breteuil<br/>75007 Paris<br/>France

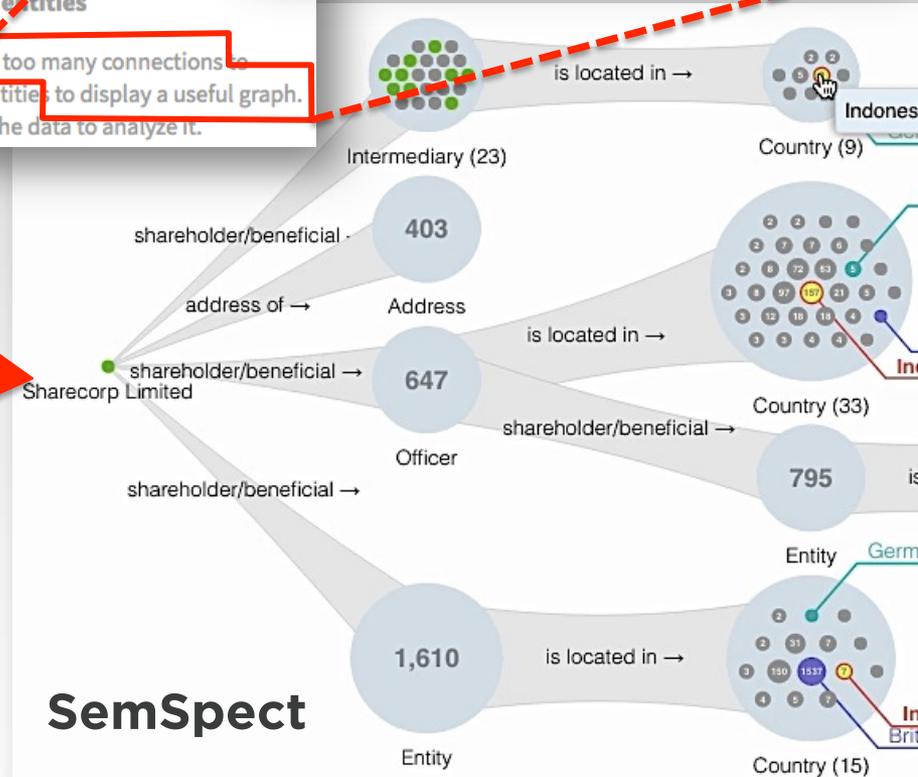
**ICJI**

## Sharecorp Limited

Connected to **23 intermediaries**  
Connected to **403 addresses**  
Connected to **624 officers**  
Connected to **1631 entities**

**ⓘ** This selection has too many connections to other officers and entities to display a useful graph. You may [download](#) the data to analyze it.

too many connections to display a useful graph.



# L'approche SemSpect

grouper les objets par catégorie

7,116

Officer

shareholder/beneficial ←

11,618

Entity in Russia

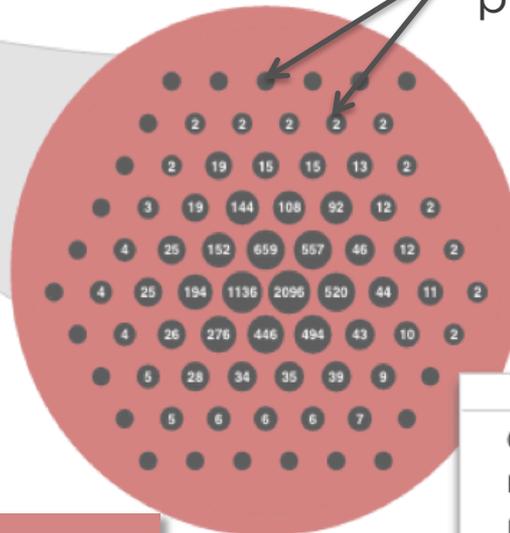
address of →

objets (avec nombre de prédécesseurs)

catégories d'objets en relation

agréger les relations entre les groupes

guider l'exploration par les données



Address (76)

Reachable categories

- Country (5) >
- Entity (43,777) >
- Intermediary (3) >
- Officer (1,004) >

type de relation

address of ← (1,004)

# DEMO

The screenshot displays the SemSpect web application interface. The main area shows a network graph with nodes representing 'Entity in Germany (197)', 'Address (17)', and 'Officer (181)'. Relationships are labeled 'address of →' and 'shareholder/beneficial ←'. A large 'VIDEO' watermark is overlaid on the graph.

**Categories**

- Address (151,592)
- Country (208)
- Entity (495,038)
- Entity in British Virgin Islands ( )
- Entity in France (304)
- Entity in Germany (197)**
- Entity in Hong Kong (51,295)
- Entity in Switzerland (38,077)
- Entity in USA (6,254)
- Entity served by Mossack Fon
- active entities (101,732)
- active entity in british virgin isl
- Intermediary (24,177)
- Officer (370,822)
- Service Provider (3)

**Entity in Germany**

This group contains objects belonging to the category 'Entity in Germany'.

**Source**

Panama Papers	173
Offshore Leaks	24

**Entity status**

Defaulted	75
Dissolved	42
Trash company	22
Dead	21
Inactivated	15
Active Entity	13
Bad debt account	4
Changed agent	3
Resigned as agent	1

**Company Type**

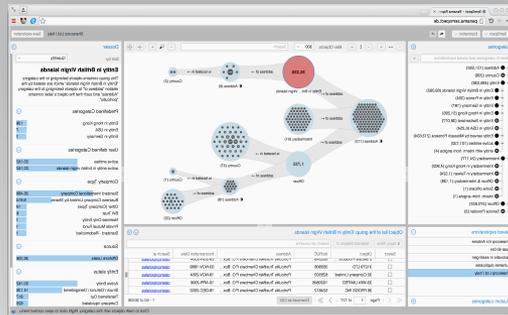
Business Company Limited b...	14
Standard Company under IB...	8
Standard International Comp...	1

**Object list of the group: Officer**

Select	Object	#Predeces...	IJCI ID
<input type="checkbox"/>	ALAIN KORAL HORN	1	5DEF7A6D8D58501E4E7A5A6...
<input type="checkbox"/>	ALEXANDER CHERNOMORDIK	1	862C70979F2A243EC203FE18...
<input type="checkbox"/>	ALEX RODRIGUES ROSA IFANGER	1	AFR97947E13B994D7115R47E

Page 1 of 4 | Apply filters | Download as CSV | Displaying objects 1 - 50 of

# Architecture de SemSpect



**interface SemSpect**  
(html5 /javascript)

**serveur SemSpect**  
conversion des requêtes  
et agrégation des résultats  
(RESTlet java)

modèle + données

**OWL / RDF**

**Neo4J, Stardog,  
Tinkerpop, SQL...**



**Hermit,  
Konklude,**

démonstrateur

**GraphScale**  
gestion BDD,  
abstraction/index,  
matérialisation  
(java)

# Retours d'expérience

- **Pilote installé à Hambourg**  
Investigations
- **Automatismes industriels**  
Contrôle qualité
- **Edition**  
Requêtes paramétrées
- **Musées**  
Contrôle qualité



**Merci**

Plusieurs démos sont accessibles en ligne sur  
**[www.semspect.de](http://www.semspect.de)**

[vialard@derivo.de](mailto:vialard@derivo.de)

**derivo**

Semantische Systeme

[www.derivo.de](http://www.derivo.de)