



an IPRO company

# Europe Show

**27 - 29 September 2021**

Insights for Investigations



**Jelle Oorebeek**

Fraud Investigator & Project Advisor



**Matthijs Luisman**

Service Delivery Manager

# Investigations Golden W's

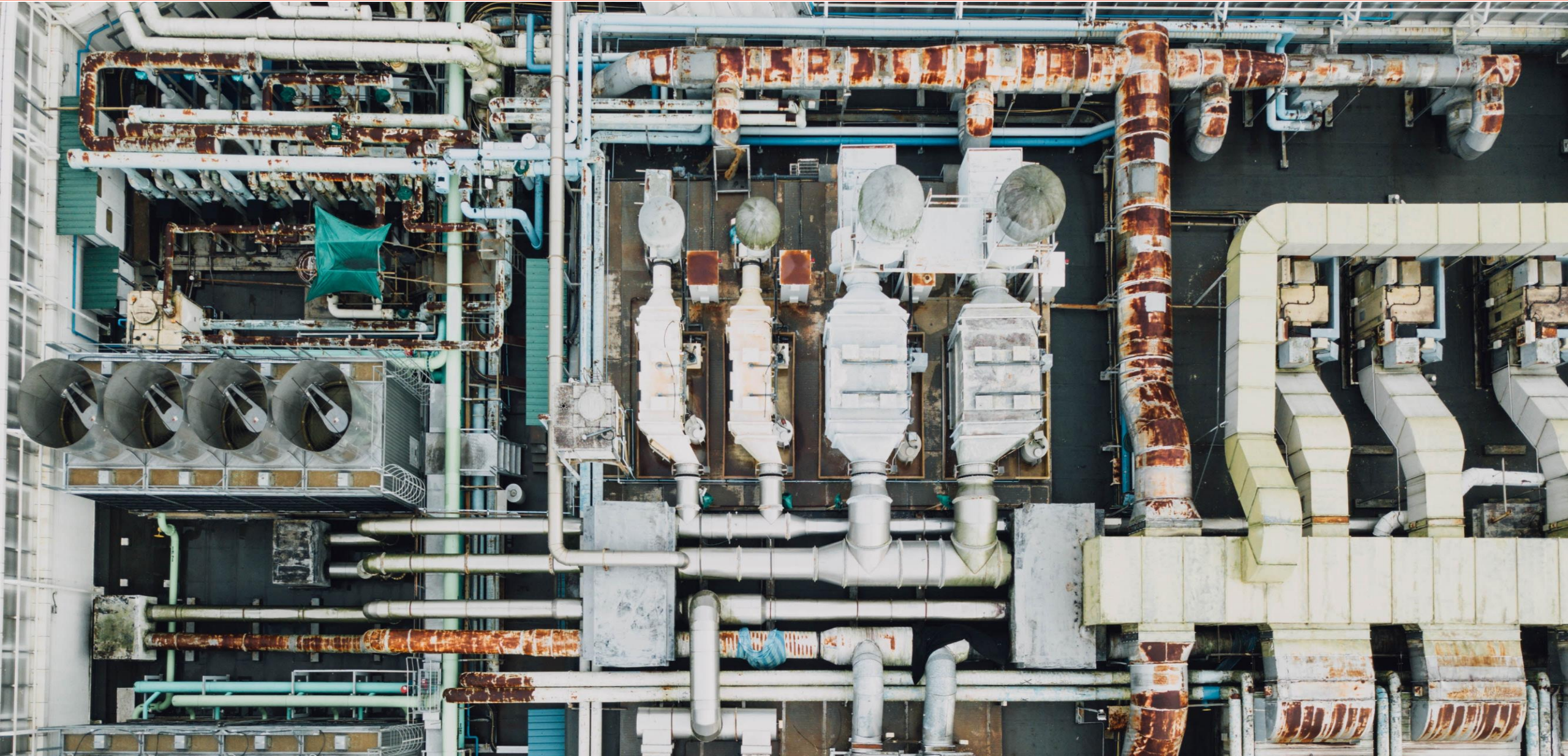
- We need to specify the correct questions that will help us gather the correct data.
- What happened?
- Who is involved?
- When did it happen?
- Where did it happen?
- With what means or methods?
- Why did it happen?
- These are all questions that an investigator will ask themselves when investigating crime. How can technology help answer these questions.

# How technology assists in investigations





# Data processing





# Investigations Golden W's and how technology can solve them

- We need to specify the correct questions that will help us gather the correct data.
- *What happened?*
- *Who is involved? **\_persons\_** **\_organizations\_***
- *When did it happen? **\_times\_** **\_dates\_***
- *Where did it happen? **\_localities\_***
- *With what means or methods? **\_currency\_** **\_financial information\_***
- *Why did it happen? **\_sentiment\_***
- *These are all questions that an investigator will ask themselves when investigating crime. How can technology help answer these questions.*

# What techniques are being used...

## Search

We've posted all 112 of them below...

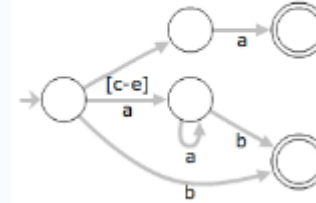
- Qaddafi, Muammar
- Al-Gathafi, Muammar
- al-Qadhafi, Muammar
- Al Qathafi, Mu'ammar
- Al Qathafi, Muammar
- El Gaddafi, Moamar
- El Kadhafi, Moammar

## Pattern Recognition

Pattern (edit) : `a*b|[cde]a` Input:

Instructions: (show)

**Nondeterministic Finite-State Automaton**



NL99 BANK 0123 4567 89

## Text-Mining

Automatically find names  
of people, places, products,  
and organizations in text  
across many languages.


# pros and cons...

## Search

- **NEED TO KNOW** what you are looking for
- “out of vocabulary words” will be missed (*low recall*)
- A lot of noise is generated (*low precision*)

## Pattern Recognition ([0-9]{3}-[0-9]{2}-[0-9]{4})

- (High recall)
- Hardly any noise is generated (*high precision*)



12/20/2010 MEMBER NO.: 400015290  
SSN: 254-55-0999

## Text-Mining

- **NO NEED TO KNOW** what you are looking for
- “out of vocabulary words” will automatically be found (*high recall*)
- No noise is generated (*high precision*)

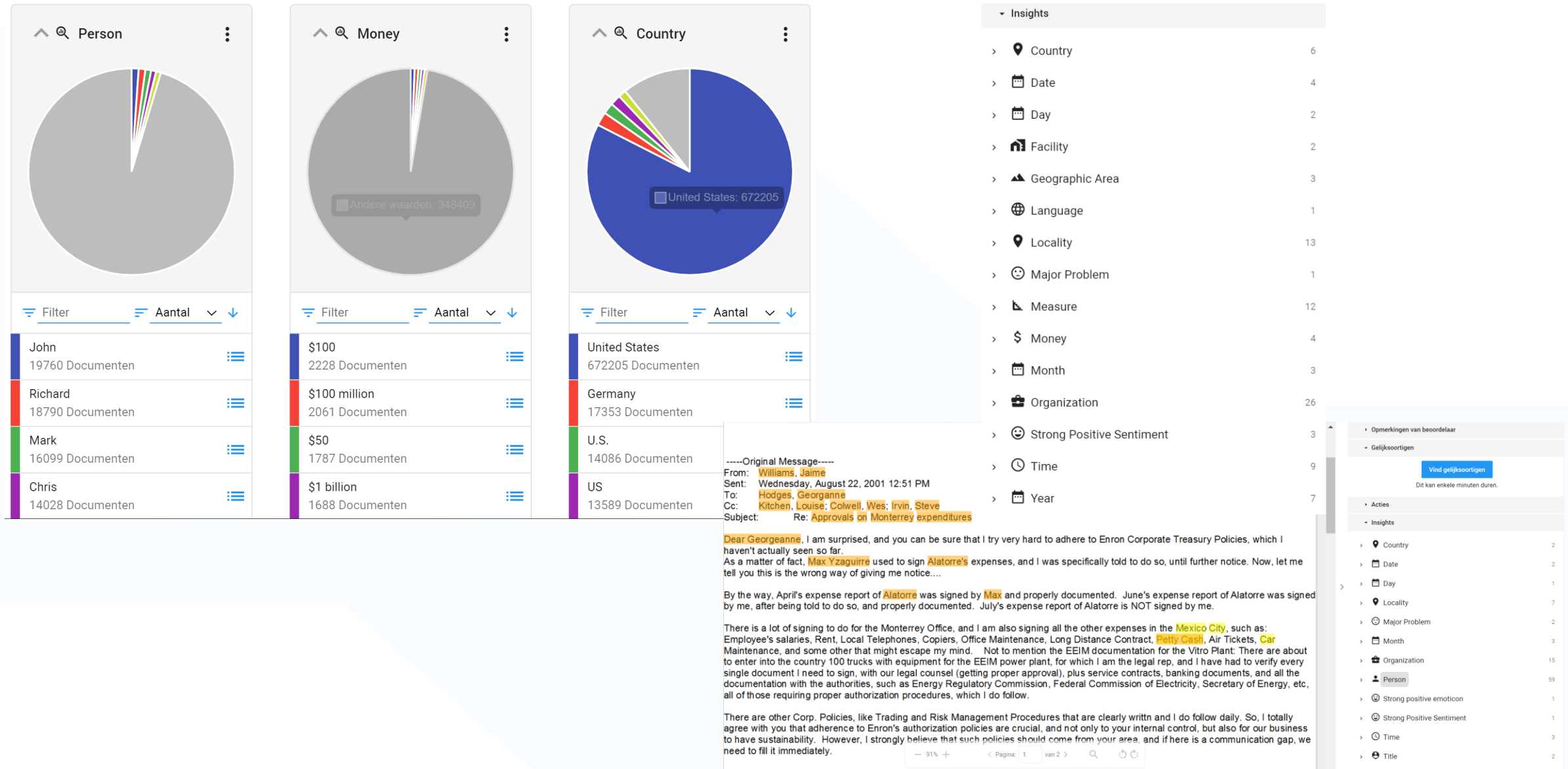




# Insights

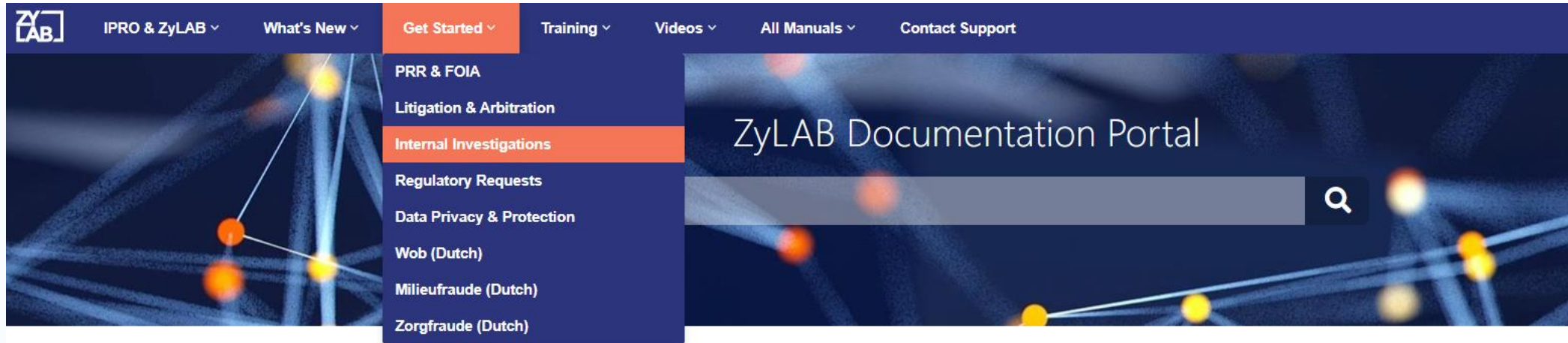
Automated Entity extraction

# Entity extraction & conceptual searching





**Insights as  
starting point for  
Data Analysis**



## What's New



### ZyLAB ONE SaaS

Attention: New features are rolled out gradually. This allows for a smoother transition and a better experience for our users.



### Product Launch Videos

Watch our videos on our latest releases (ZyLAB ONE 6.0 On-Premise and Insights), explained by our Product Directors.



### ZyLAB ONE Legal Hold


Ease the burden of legal hold in a defensible way and improve your ability to identify critical data.



### ZyLAB ONE Insights






Extract relevant information (entities), such as real world objects (persons, locations, organizations, products, etc.), emotions, sentiments and references between them.








What's New ▾Get Started ▾Training ▾Videos ▾All Manuals ▾Contact Support

Internal Investigations / Use Case / Searching for evidence



 Welcome

 Video

 Use Case

- Log In
- + The Select Matter dashboard
  - New internal investigation
  - Upload Data
- + Configuration dashboard
  - How to Navigate a Matter
  - Searching for evidence**
    - + Batch search
    - + Topic Modeling
      - Browse using Facets
      - Search queries

## Searching for evidence

When you are ready to start your investigation and search for evidence, there are multiple ways to start. Different search options are listed the amount of data, the amount of investigators on the case, and how much information is known about the case, certain search options v in finding relevant documents.

- [Batch search](#)

A batch search enables you to execute multiple queries at once. This is most useful in an internal investigation when you want to scope of your data. By executing a list of queries that include keywords and queries that are suspect and could lead to evidence, hits displayed will give an indication of the scope. In this way, you can quickly narrow down on possibly relevant documents and start your investigation.
- [Topic Modeling](#)

When you are not sure which information can be found in the collected data, you can use topic modeling to quickly get an overview narrow down on certain topics of interest and start searching manually, or you can select a particular topic and use **Assisted Review** documents that have similar contents.
- [Browsing](#)

By browsing through facets, you can quickly narrow down your scope and start searching in specific places in the data. This is useful have a clear idea of which information you want to zoom in on.
- [Search queries](#)

By using search queries, you can search rather specifically for a certain topic, person, time period, or all three of them at once. Unfortunately sometimes be a bit of a hit or miss, which is why its recommended to use search queries once you've narrowed down the data. This is accomplished by using batch search or topic modeling, but also by browsing through facets.
- [Assisted Review](#)

Assisted Review enables you to use artificial intelligence to search for relevant documents. By starting with a set of documents, topic, a search query, or relevant documents that have been found already, the computer will search for documents that have similar. All you must do is review them.











**Note:** Assisted Review can also be used check if you might have missed evidence once you conclude your investigation. By using documents that contain evidence to start Assisted Review, the computer will search for documents similar to those. In this way all relevant documents in the data have been found.

# Using Insights for Data Analysis

- Determine your Investigation Topics
  - Patterns
    - Queries
    - Regex
    - Entity detection
    - Entity Search
    - All of above combined
- Data Analysis (Statistics)
  - What is happening and how often
  - Which persons are involved and at what level
- Example: Communication Analysis
  - Who is talking to Whom about what?

# Data Analysis: Communication\*

*Investigation focus: 2 persons and the internal email communication*

Communication Analysis					
Sum of Documents	Column Labels 				
	 Board	 General Management	 Local Management	 Department A	Grand Total
Row Labels 					
 One-to-one Person A	12597	26559	2781	1682	43619
 One-to-many Person A	15001	63493	15114	4275	97883
 One-to-one Person B	171	8864	247	291	9573
 One-to-many Person B	763	25757	1035	1204	28759
Grand Total	28532	124673	19177	7452	179834


\* Data Analysis services provided ZyLAB eDiscovery Consultants





# Data Analysis: Communication\*


*What topics can be identified in the email communication?*





**Communication Analysis**

Sum of Documents

Column Labels 

 Board  General Management  Local Management  Department A Grand Total

Row Labels 



 One-to-one Person A	12597	26559	2781	1682	43619
HRM	1006	1478	129	69	2682
Communication	3067	4267	774	186	8294
Services	1372	3076	285	195	4928
Fraud	310	385	37	36	768
Organization	3462	6378	920	492	11252
Finance	3380	10975	636	704	15695
 One-to-many Person A	15001	63493	15114	4275	97883
 One-to-one Person B	171	8864	247	291	9573
 One-to-many Person B	763	25757	1035	1204	28759
Grand Total	28532	124673	19177	7452	179834

\* Data Analysis services provided ZyLAB eDiscovery Consultants



# Data Analysis: Communication\*

*What topics have been discussed 1-1 to whom?*  
*Which persons should be interviewed and about what?*

Communication Analysis										
Sum of Documents		Column Labels 				Board Total	General Management	Local Management	Department A	Grand Total
		Board								
		Person C   Person G   Person L   Person M								
Row Labels 										
One-to-one Person A		10093	1213	683	608	12597	26559	2781	1682	43619
HRM		786	141	40	39	1006	1478	129	69	2682
Communication		2403	320	237	107	3067	4267	774	186	8294
Services		1079	109	72	112	1372	3076	285	195	4928
Fraud		263	10	19	18	310	385	37	36	768
Organization		2750	333	219	160	3462	6378	920	492	11252
Finance		2812	300	96	172	3380	10975	636	704	15695
One-to-many Person A		7777	3370	1607	2247	15001	63493	15114	4275	97883
One-to-one Person B		3	141	15	12	171	8864	247	291	9573
One-to-many Person B		12	602	92	57	763	25757	1035	1204	28759
Grand Total		17885	5326	2397	2924	28532	124673	19177	7452	179834

\* Data Analysis services provided ZyLAB eDiscovery Consultants



For more information:

[\*\*Docs.zylab.com\*\*](https://docs.zylab.com)

*ZyLAB One What's new, Use cases  
and documentation*

# ZY LAB] Q&A



**Thank you**