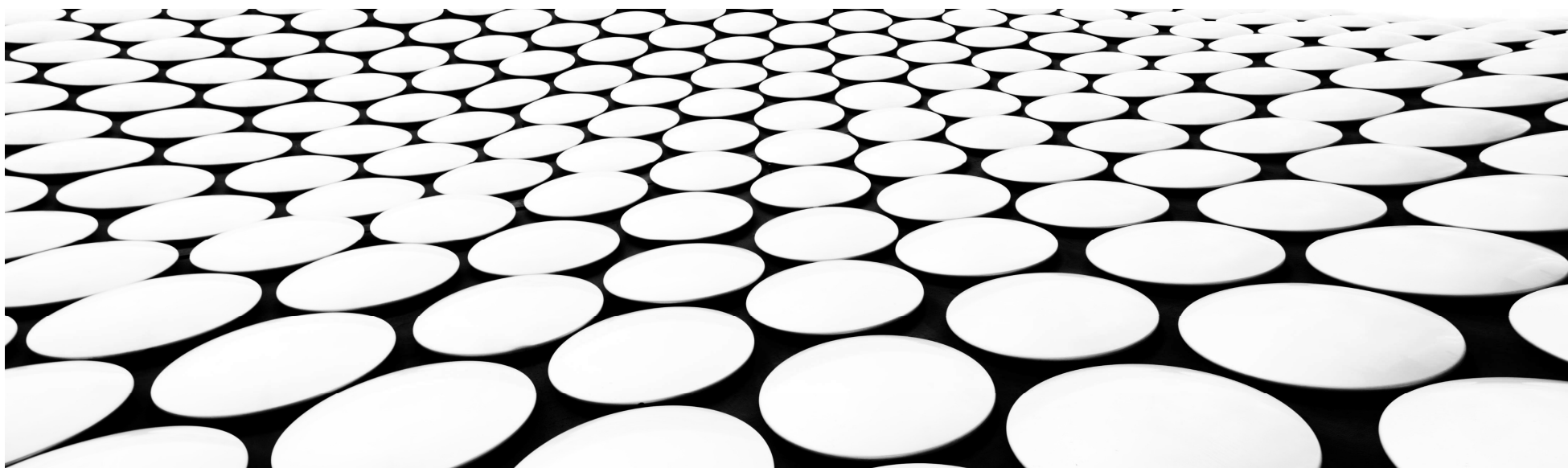

SLIMME ONDERZOEKS TECHNOLOGIE

NFFI | 31 MAART 2021





INTRODUCTIE

Jelle Oorebeek
eDiscovery specialist &
fraude onderzoeker
+31 6 41775649
jelle.oorebeek@zylab.com

meer druk uit verschillende richtingen



- **meer data**
- **globalisering**
- **wetgeving & controles**
- **privacy**
- **veranderende business modellen**
- **veeleisende klanten**
- **boze werknemers**

het resultaat... meer (interne) onderzoeken

- toezichhouders
- rechtszaken & arbitrage
- fraude onderzoeken
- patent, contract & hr disputen
- M&A, herfinanciering en structurering



informatie verzoeken komen in golven



- incident gedreven
- snel handelen
- geen tijd voor goede aanbesteding
- geen tijd voor training
- onvoorspelbare data hoeveelheden



data verovert de wereld

- data volumes groeien exponentieel elk jaar
- gemiddelde data volume per zaak groeit
- aantal zaken met kleinere hoeveelheden data groeit

(voorbeeld. 1000 e-mails < 250MB)

big data



complexe privacy uitdagingen

efficitur ac augue ac, **tristique**. Sed laoreet, orci quis dapibus
unc dui et dolor. **Proin** venenatis mattis leo a suscipit.
vida, neque quam dignissim est, a cony
fringilla, massa
ro. Cras ex erat, maximus id finibus s
quis blandit nisi eleifend
eleif
non scelerisque sem. Morbi a diam
ipsum sit

mensen missen feiten

de beste mensen presteren tot 85%

Gebaseerd op een studie waar advocaten in vergelijking met AI 5 contracten beoordeelden

1%

Niet weten is geen optie

Bron: <https://www.lawgeex.com/resources/aivslawyer/>



het vinden van feiten is technischer

vergeleken met 20 jaar geleden

mankracht is niet langer de oplossing



EDISCOVERY = ZOEKEN EN BEOORDELEN VAN INFORMATIE OP RELEVANTIE

‘Recordboete dreigt voor Facebook’

**Volkskrant, RTL en NOS eisen
meer transparantie over
MH17-ramp**

Waarheidsvinding wordt
steeds lastiger, ook voor de
politie

RvS: Verslagen ministerraad over MH17 hoeven niet openbaar

Celstraffen geëist om fraude bij bouw KPMG-kantoor

**ZEMBLA onthult: recyclingbedrijven
kunstgras overtreden structureel
milieuregels**

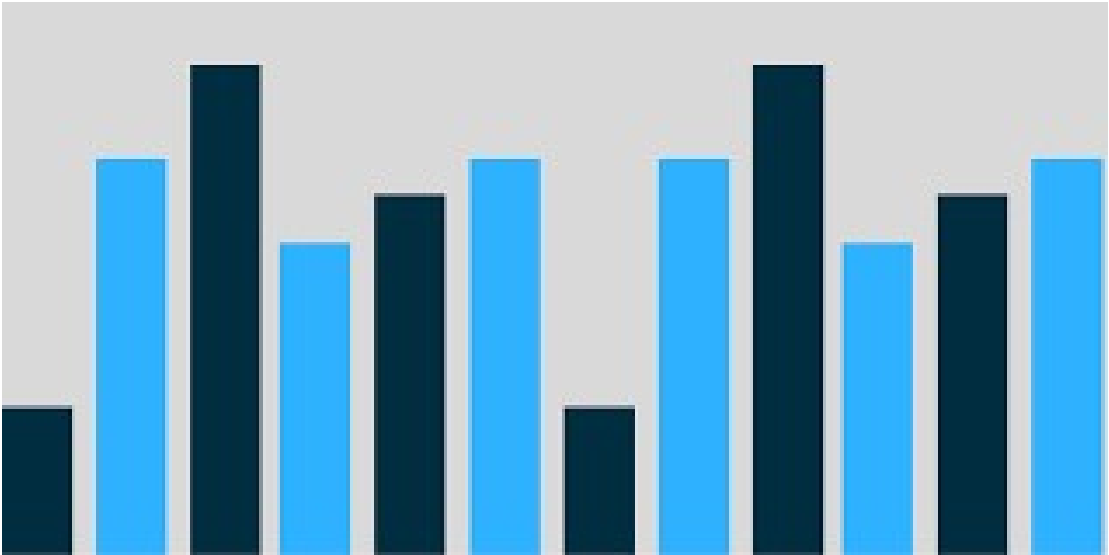
'Ten koste van het milieu wordt hier geld verdiend'

**Accountants Baker Tilly Berk
krijgen bijna miljoen euro boete
voor falend klantenbeleid (FD)**

FTM start met de Shell Papers



POLLS





HET ONDERZOEK PROCES



JURIDISCHE VERDEDIGBAARHEID

Bij het gebruik van technologie binnen het juridische domein is van essentieel belang rekening te houden met:

- Wetgeving
- Ethische aspecten en maatschappelijke normen en waarden
- Solide kwaliteit controle
- Voorkoming van vooringenomenheid (bias)
- Transparantie van technologie (je moet het kunnen uitleggen)
- Robuustheid van technologie (wat als je foute invoer data hebt?)
- Reproduceerbaarheid van technologie (het moet steeds dezelfde uitkomst geven)
- Documentatie: *Chain of Custody en Audit Trails* (wie heeft wat wanneer gedaan, *inclusion-*, *exclusion-* en foutrapporten)
- Forensische integriteit moet geborgd zijn

PROBLEMEN VAN EEN ONDERZOEK



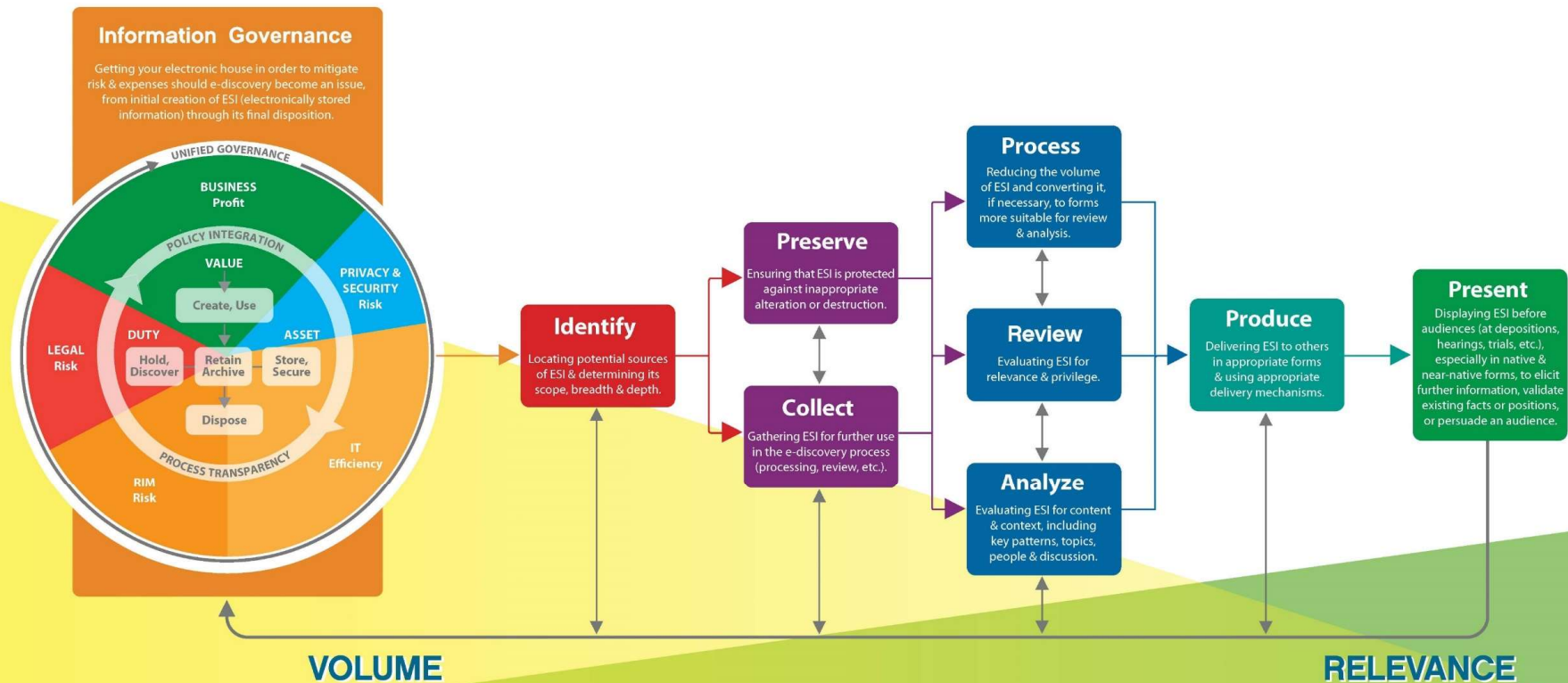
- De opdrachtgever
- De onderzoeker
- Het proces
 - Informatie staat verspreid
 - Structuur aanbrengen in data is lastig
 - Bewijs zoeken duurt lang
 - Falsificatie, ontcrachten van veronderstellingen
 - Irrelevante data en persoonlijke data
 - Risico om bewijs te missen
 - Geen audit log

DE 7 GOUDEN W'S

- De volgende vragen helpen ons de juiste data te verzamelen.
 - Wat is er gebeurd?
 - Wie is er betrokken?
 - Wanneer is het gebeurd?
 - Waar is het gebeurd?
 - Hoe of op welke wijze is het gebeurd?
 - Waarom is het gebeurd?
- Deze vragen zal een onderzoeker zichzelf stellen tijdens het onderzoek. Maar hoe kan technologie ons helpen deze vragen te beantwoorden?

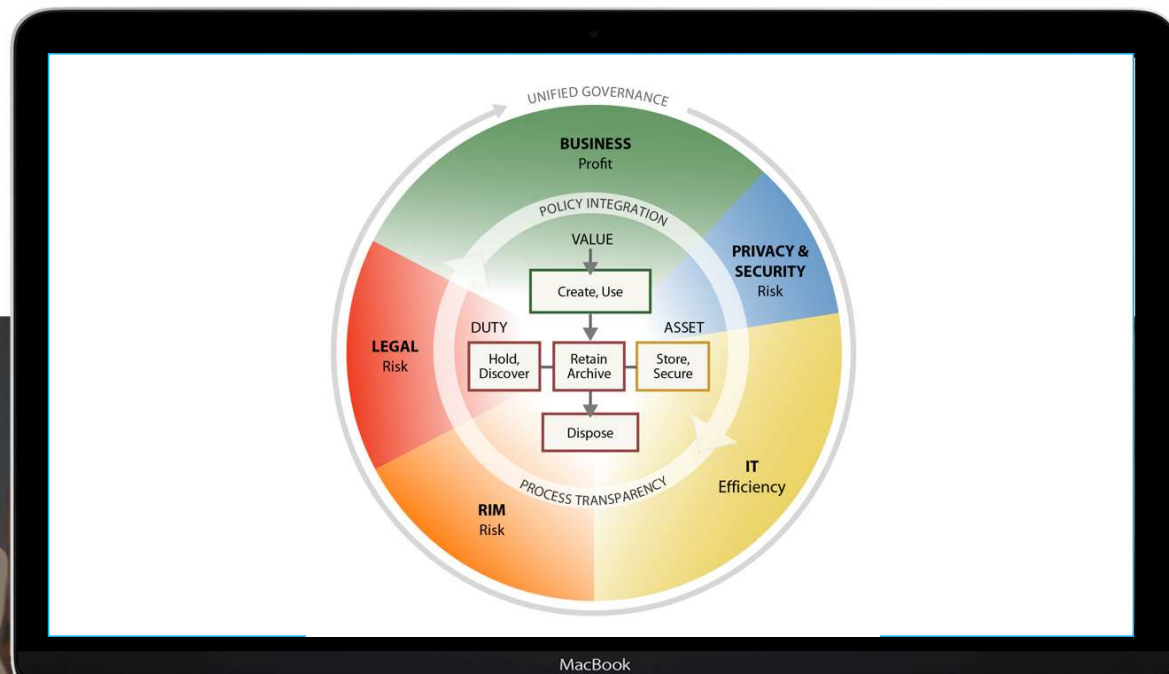
Electronic Discovery Reference Model

Standards, Guidelines, and Practical Resources for Legal Professionals and E-Discovery Practitioners



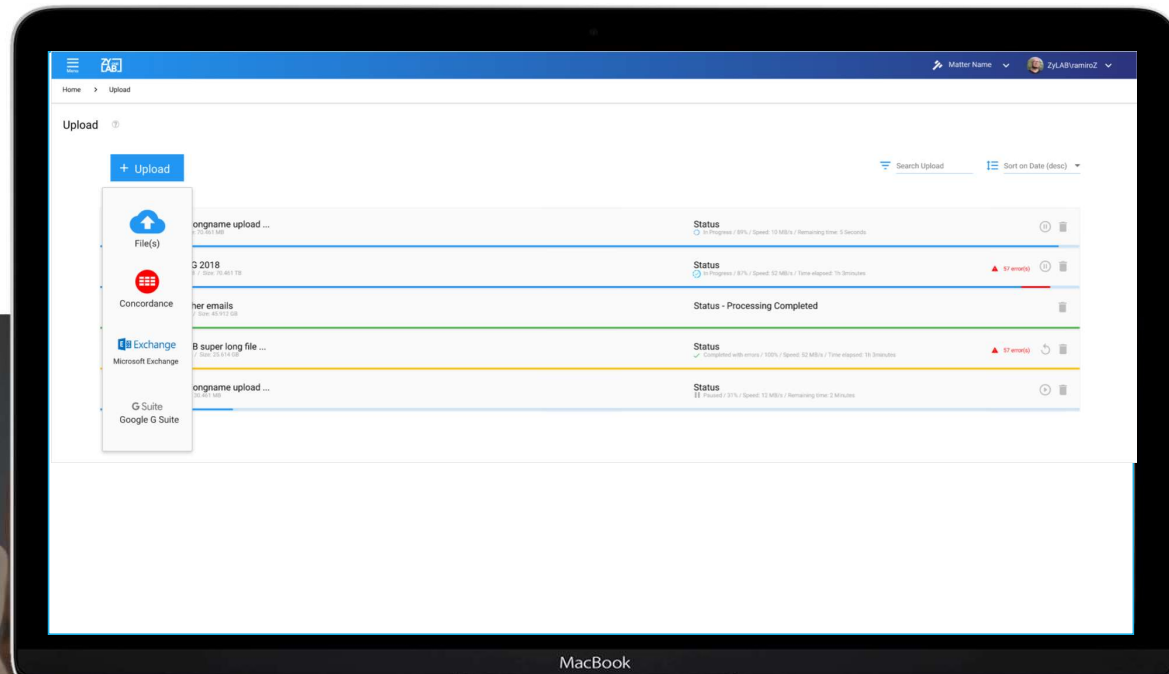
Information Governance

Identificeer data bronnen en eigenaar



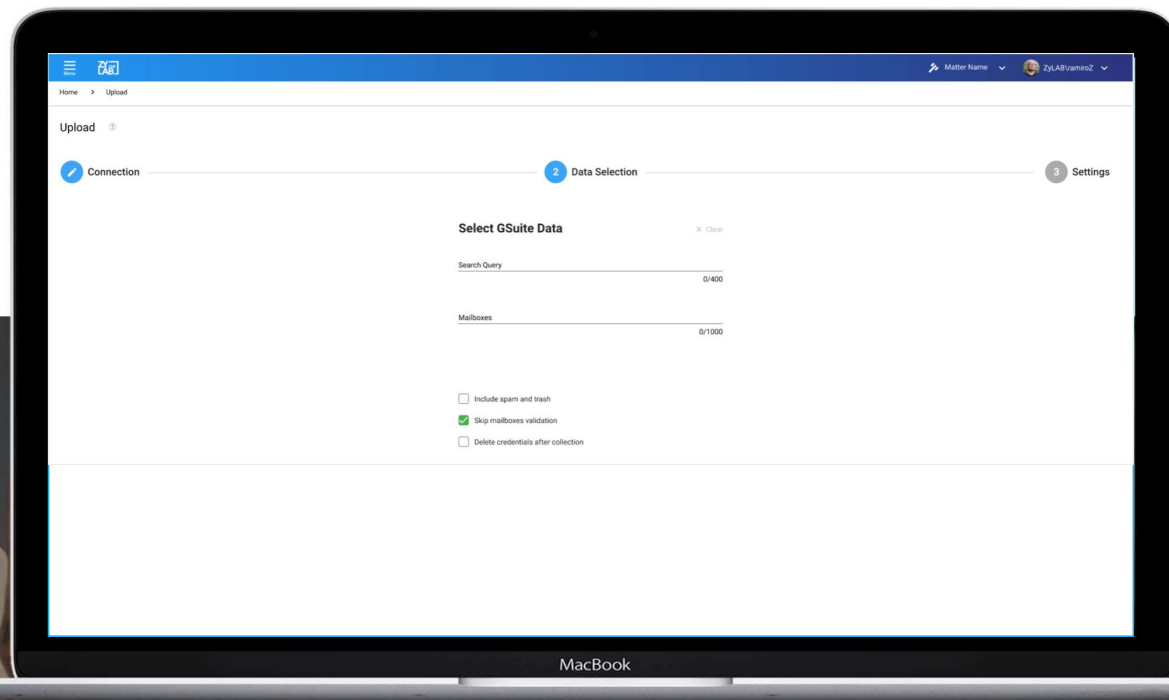
Verzamel de data

Upload de data naar het onderzoeks platform



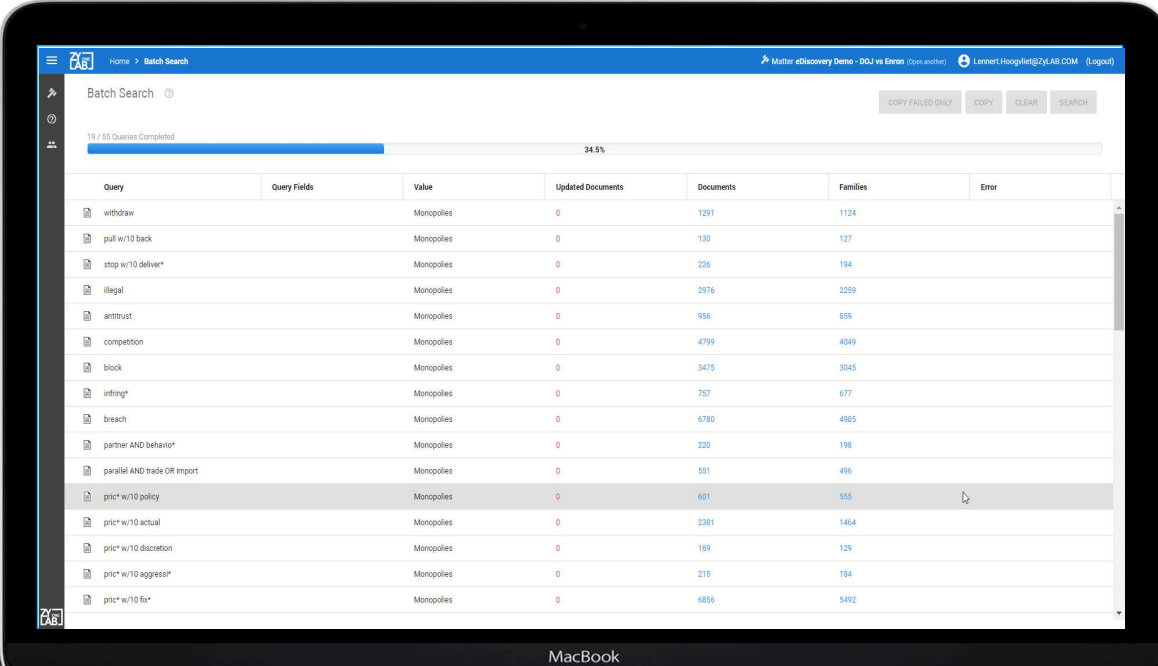
Filter de data

Verwijder duplicaten en filter relevante informatie



Bepaal de scope

Definieer de zoekvragen of gebruik de ingebouwde analyse functies



Batch Search

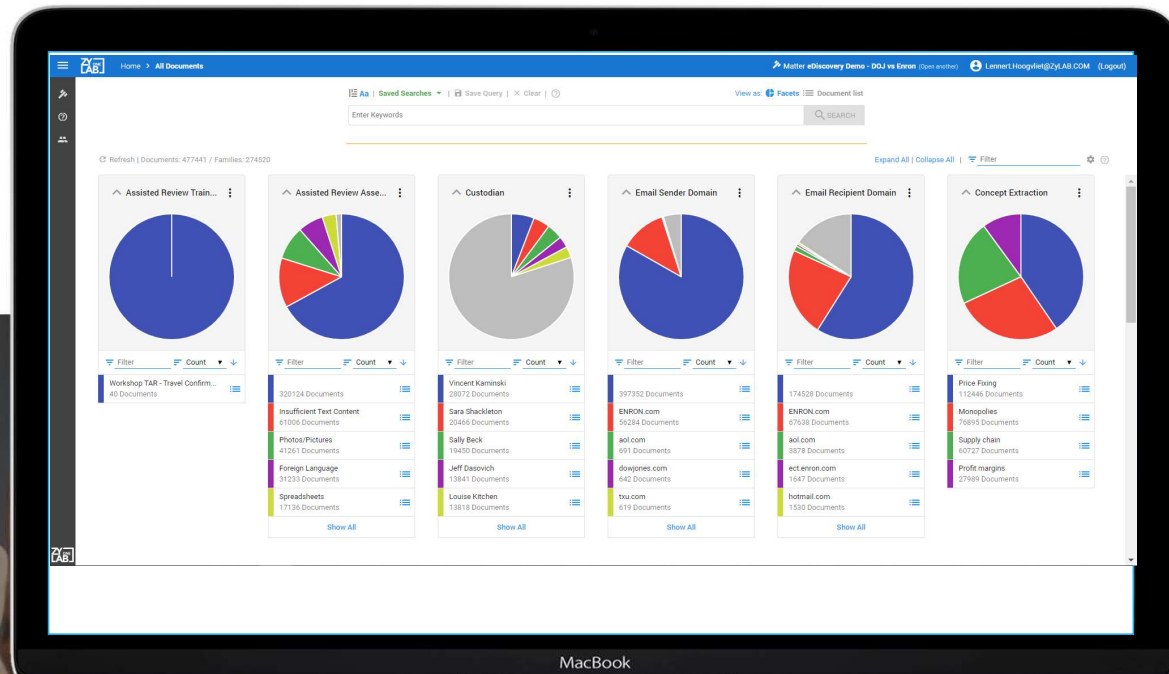
19 / 55 Queries Completed

34.5%

Query	Query Fields	Value	Updated Documents	Documents	Families	Error
withdraw		Monopolies	0	1291	1124	
pull w/10 back		Monopolies	0	130	127	
stop w/10 deliver*		Monopolies	0	225	194	
illegal		Monopolies	0	2976	2259	
antitrust		Monopolies	0	956	859	
competition		Monopolies	0	4799	4049	
block		Monopolies	0	3475	3045	
infring*		Monopolies	0	757	677	
breach		Monopolies	0	6780	4985	
partner AND behavior*		Monopolies	0	220	198	
parallel AND trade OR import		Monopolies	0	551	496	
prtc* w/10 policy		Monopolies	0	601	555	
prtc* w/10 actual		Monopolies	0	2381	1464	
prtc* w/10 discretion		Monopolies	0	169	129	
prtc* w/10 aggressi*		Monopolies	0	215	184	
prtc* w/10 fir*		Monopolies	0	6856	5492	

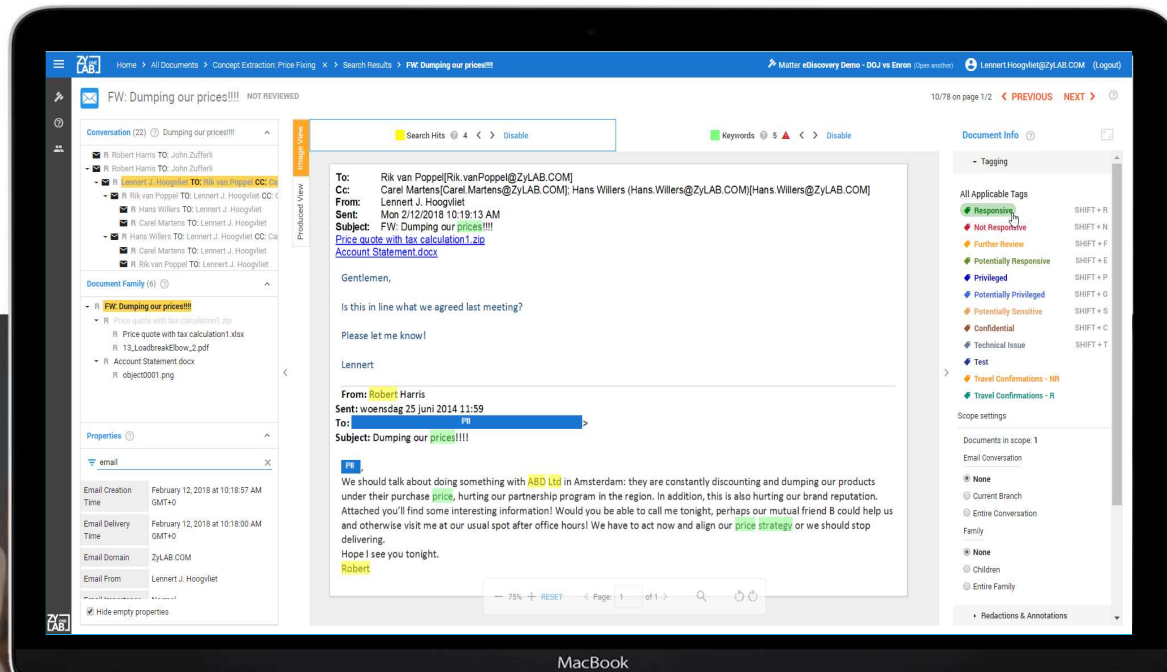
Analyseer de data

Creëer inzicht in de data door het gebruik van Insights



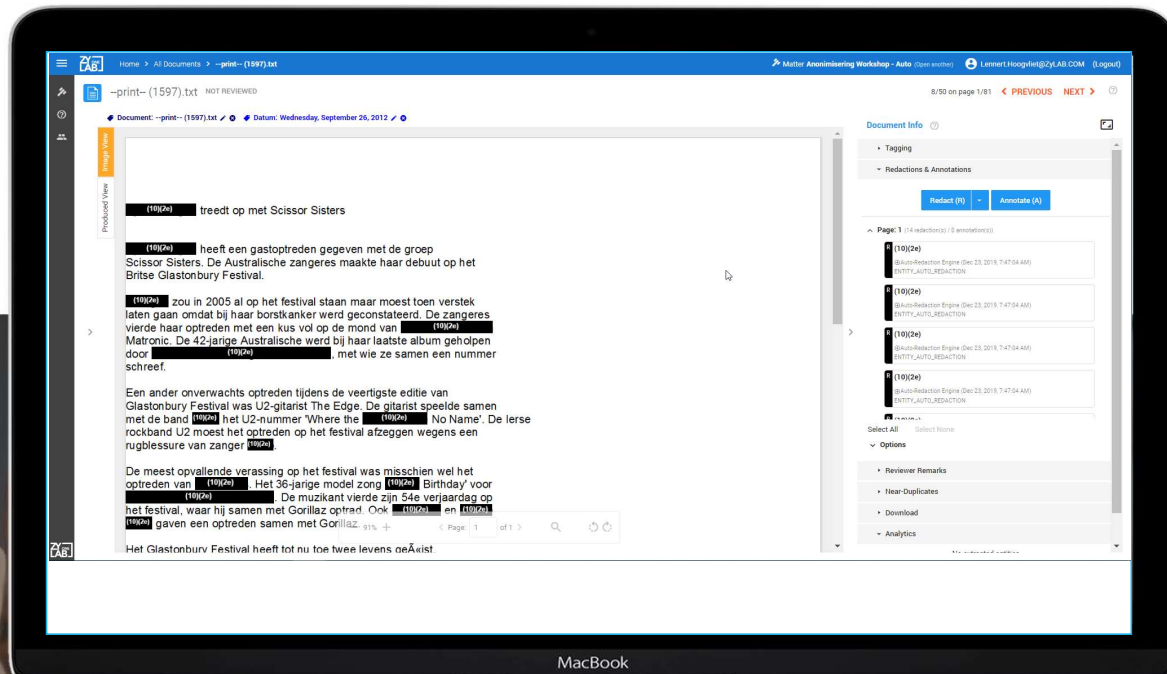
Beoordeel

Beoordeel, label en verwerk de data (documenten)



Anonimiseer

Lak privacy en gevoelige informatie (automatisch af)



Rapporteer de feiten

Produceer de feiten en rapporteer middels industrie standaard

The image shows a laptop screen displaying a 'Production Overview' dashboard. The dashboard is titled 'Production Overview' and features a search bar and a 'Show All' dropdown menu. Below the search bar, there are six panels, each representing a completed eDiscovery job. Each panel includes a green checkmark icon, the job title 'eDiscovery Demo - DOJ vs Enron ...', a 'Completed' status, and various statistics such as file counts and sizes. The statistics for each panel are as follows:

Panel	Created on	Completed on	Bates Range	Images	Native	Text	Placeholders
1	Oct 30, 2019, 12:15:40 PM	Oct 30, 2019, 12:20:37 PM	000000000-0000000006	7	0	0	0
2	Oct 22, 2019, 11:58:37 AM	Oct 22, 2019, 12:07:11 PM	000000000-0000000008	3	0	0	0
3	Oct 15, 2019, 5:31:35 PM	Oct 15, 2019, 5:39:40 PM	000000000-0000000013	14	0	0	0
4	Jun 5, 2019, 4:52:34 PM	Jun 5, 2019, 4:59:55 PM	000000000-0000000025	2	0	0	1
5	Jun 5, 2019, 4:23:32 PM	Jun 5, 2019, 4:37:36 PM	PRE-000000000-SUF-PRE-0000000000...	1	0	1	0
6	Jun 5, 2019, 2:46:48 PM	Jun 5, 2019, 2:54:31 PM	000000000-0000000003	0	4	0	0

ZYLAB ONE

ONDERZOEK TECHNOLOGIE



ZYLAB ONE EN DE CLOUD

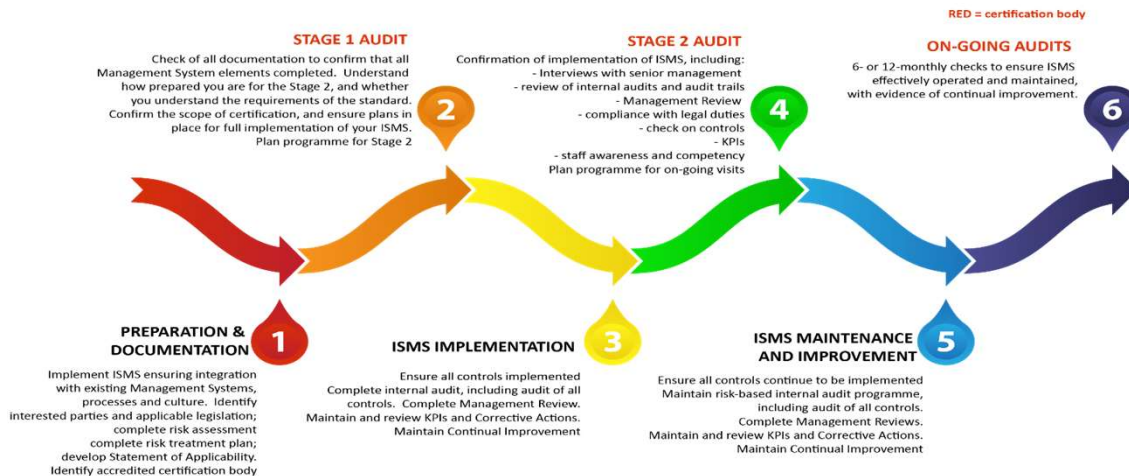


ISO/IEC 27001:13 & SOC 2 COMPLIANCE STATEMENT

ZyLAB has developed and maintained an Information Security Policy which has been in operation since 2008.



Since 2017, the ZyLAB policies, procedures and way of working have been formalized according to the ISO/IEC 27000 series. In November 2018, ZyLAB obtained the certification proving that it has successfully implemented an information security management system in accordance with ISO/IEC 27001:2013.

ISO 27001 CERTIFICATION PROCESS

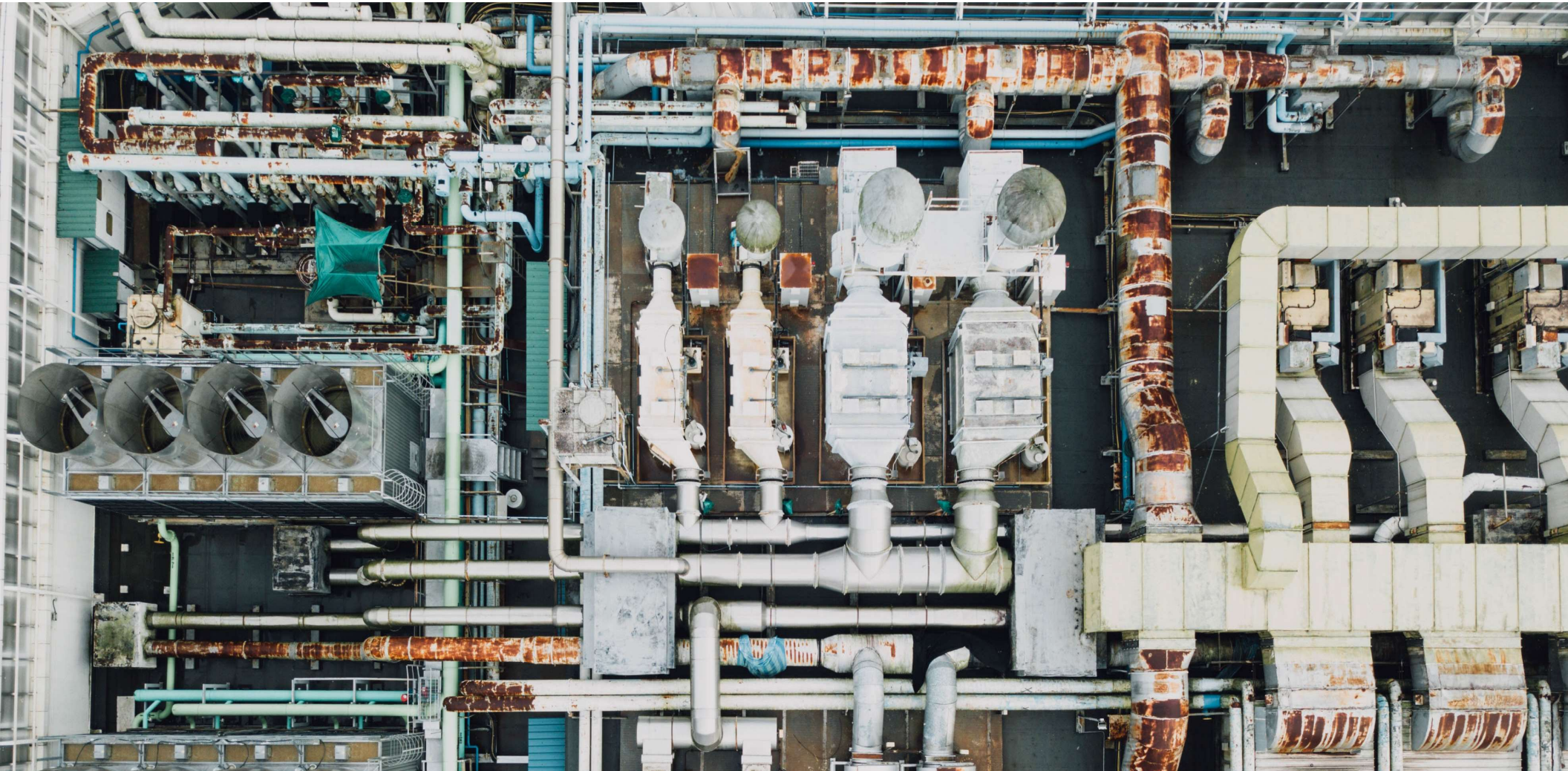




UITVOEREN ONDERZOEK

- 
- Data verwerking
 - Automatischh classificeren
 - Conceptueel zoeken
- 

DATA PROCESSING



DE 7 GOUDEN W'S

- De volgende vragen helpen ons de juiste data te verzamelen.
 - Wat is er gebeurd?
 - Wie is er betrokken? **_personen_**
 - Wanneer is het gebeurd? **_tijd_ _datum_**
 - Waar is het gebeurd? **_plaatsen_ _landen_**
 - Hoe of op welke wijze is het gebeurd? **_bedragen_ _rekeningnummers_ _creditcardnummers_**
 - Waarom is het gebeurd? **_negatief sentiment_**
- Deze vragen zal een onderzoeker zichzelf stellen tijdens het onderzoek. Maar hoe kan technologie ons helpen deze vragen te beantwoorden?

WELKE TECHNIEKEN ZIJN BESCHIKBAAR...

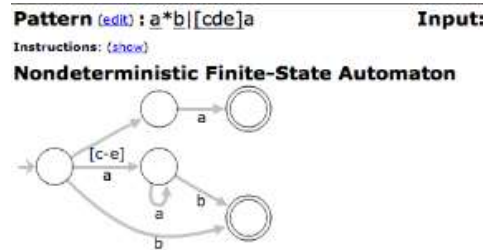
Zoeken

Patroon herkenning

Text-Mining

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

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



NL99 BANK 0123 4567 89

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

VOOR EN NADELEN...

Zoeken

- **MOET WETEN WAT JE ZOEKT**
- “*synoniemen*” worden niet gevonden (*lage recall*)
- Veel ruis (*lage precisie*)

Patroon Herkenning([0-9]{3}-[0-9]{2}-[0-9]{4})

- (Hoge recall)
- Weinig ruis(*hoge precisie*)

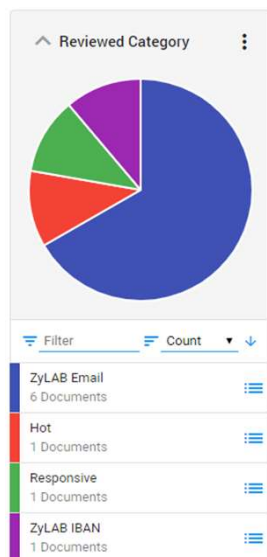
12/20/2010 MEMBER NO.: 400015290
SSN: 254-55-0999

Text-Mining

- **HOEFT NIET TE WETEN WAT JE ZOEKT**
- “*synoniemen*” worden automatisch gevonden(*hoge recall*)
- Geen ruis (*hoge precisie*)

AUTO CLASSIFICATIE EN VISUALISATIE

- Auto classificatie regels
- Gebaseerd op zoekvraag of entiteit extractie



↓ ZyLAB IBAN

Name: ZyLAB IBAN 10 / 128 Intake Only Continuous

Aa | Saved Searches | Save Query | Clear

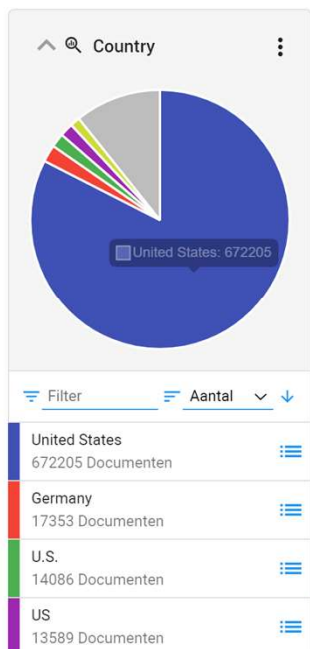
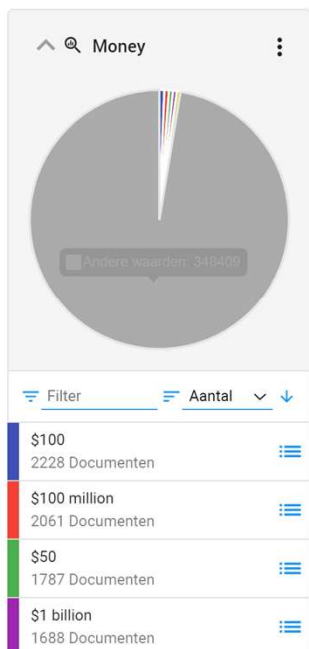
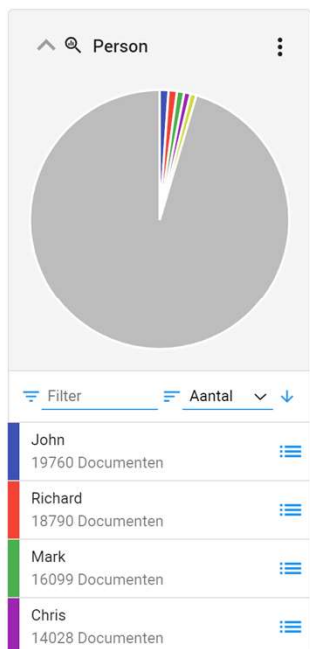
AND [a-z]{2}[0-9]{2}[a-z0-9]{4}[0-9]{7}[a-z0-9]{0,16}

Fields Tags R

When Matches

Reviewed Category = ZyLAB IBAN

ENTITEIT EXTRACTIE EN CONCEPTUEEL ZOEKEN



---Original Message---
 om: Williams, Jaime
 nt: Wednesday, August 22, 2001 12:51 PM
 x: Hodges, Georgianne
 : Kitchen, Louise; Colwell, Wes; Irvin, Steve
 bject: Re: Approvals on Monterrey expenditures

ar Georgianne, I am surprised, and you can be sure that I try very hard to adhere to Enron Corporate Treasury Policies, which I haven't actually seen so far.
 ; a matter of fact, Max Yzaguirre used to sign Alatorre's expenses, and I was specifically told to do so, until further notice. Now, let me I you this is the wrong way of giving me notice....

By the way, April's expense report of Alatorre was signed by Max and properly documented. June's expense report of Alatorre was signed by me, after being told to do so, and properly documented. July's expense report of Alatorre is NOT signed by me.

There is a lot of signing to do for the Monterrey Office, and I am also signing all the other expenses in the Mexico City, such as: Employee's salaries, Rent, Local Telephones, Copiers, Office Maintenance, Long Distance Contract, Ferry Tickets, Air Tickets, Car Maintenance, and some other that might escape my mind. Not to mention the EEIM documentation for the Vitro Plant. There are about to enter into the country 100 trucks with equipment for the EEIM power plant, for which I am the legal rep, and I have had to verify every single document I need to sign, with our legal counsel (getting proper approval), plus service contracts, banking documents, and all the documentation with the authorities, such as Energy Regulatory Commission, Federal Commission of Electricity, Secretary of Energy, etc, all of those requiring proper authorization procedures, which I do follow.

There are other Corp. Policies, like Trading and Risk Management Procedures that are clearly written and I do follow daily. So, I totally agree with you that adherence to Enron's authorization policies are crucial, and not only to your internal control, but also for our business to have sustainability. However, I strongly believe that such policies should come from your area, and if there is a communication gap, we need to fill it immediately.

Opmerkingen van beoordeelaar

Gelijksortingen

Vind gelijksoortigen

Dit kan enkele minuten duren.

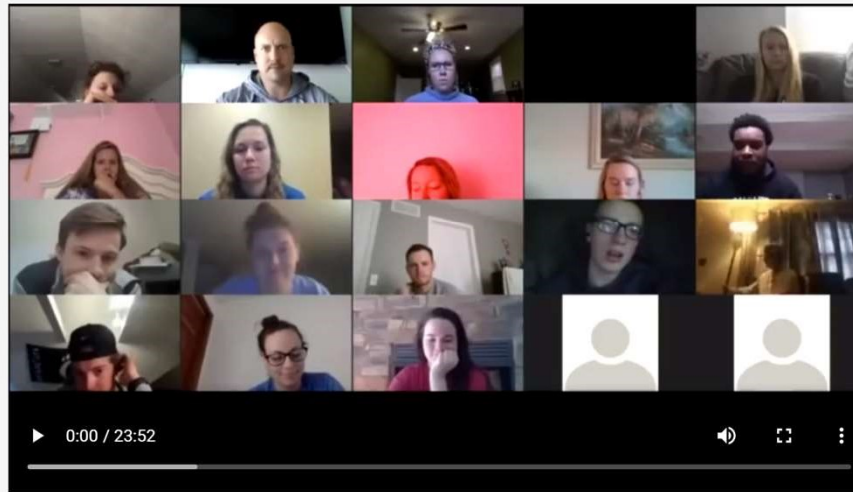
Acties

Insights

- Country 2
- Date 2
- Day 1
- Locality 7
- Major Problem 2
- Month
- Organization
- Person 59
- Strong positive emotion 1
- Strong Positive Sentiment 1
- Time 3
- Title 2

AUDIO & VIDEO ONDERSTEUNING

Search Hits 4 < > Disable



Show Alternatives

Playback rate: 1x

SPEAKER 1: Try to do um. We'll try to do just like a **half an hour meeting** today I don't want to another half an hour on Friday we'll see how it works what I'm Gonna do as I'm you know than. I'm going to lay a raid the power points so that when you download the power points on canvas I am basically it'll be just me kind of walking through those power points never read it to Greg so these sessions what does go over kind her if you go questions and just the mean main saying gobble times a week. All right but It's not like. We're all done anything a successor kind of a hanging out by a computer is anyway were fans are well I'll go bodies good. Erm hobnob our Ward I in hope you stay in safe right um. I know they're It's It's a weird time I must see you know you guessed so as the kind of gold Mount some. Um. Let's try this out good summer it starts share my screen here in we're Gonna and go Lamb in here in and see what you guys um what she can do right tell you that thumbs up from some people you Ian Gaga are Arsenal he real um who Arthur's that your arse out there good stuff. All all right some sure in a screen

SPEAKER 2: in

SPEAKER 1: end once once Suzanne now good Sue erm see if he goes can't see the um the power Ah that's Gonna come up soon here

SPEAKER 2: cool.

ZyLAB ONE samenvatting

Audio & Video ondersteuning

☰ ZYLAB
Matter Name | ZYLAB\devtest2

Home > All documents > Create Production

(15:30) English Import26505 - Forensic16.wav
🔍 | 🗑️ | ✅ Reviewed

← PREVIOUS
NEXT →

Conversation

Document Family

Properties

Accessed

Appointment End

Appointment Start

Assigned To

BinaryContentHash

Concepts Hits

Created

DedupSHA5BinHash

Deletion Status

Document Set Id

Responsive

Show Alternatives
⏮️ ⏪ ⏩ ⏭
Playback rate: 1x

UNKNOWN SPEAKER: Yeah yeah. It's a term what it he give you the wrong number and Here's looking at it 1 night. 1 2. Sorry. fine. Yeah means you got Long Road. Yeah. Hello. Why. so that means 30 time there's a naked number. Law you tell you at a part of numbers are you sure. Yeah but I just reduced Stott's Hydro Ken and now my requirements requirements number. Wow are went to the 2 3 she's gone. Yeah think that Besides wrong. When I need to take that money. No no no you right. 2 law. Under this. 9. Ok so positive number as long can a numbers are you sure Yes. Now. so what did you say is that boy me what it says imbalance is you taking the resource imbalance it OK OK thanks for coming home oh no I am I130K I've got AI got a resource imbalance divided by. Number of days times 24 times A199 pence times -1. Well is the fine I see 8 years. That sign is right it depends which are you looking at it because the embarrassing recess a mountain light 25 actually moved sure long because you take care are you loads of you might have thing you resources the Hydro on the phone matter G99 want to harsh required. 32. Hydro engineer. 25. Ok so what are you resource imbalance shows a number 10 please. Excellent. to mackerel. Ok so that means we. Access by law in Hydro now it makes sense to me I am I got it. Yeah thank you for explaining my sheet. Sorry fixed up. No no this is what we have you when we talk about. This is the exactly 10 and we need 10 at the tournament and you can extend as the are now obviously well having sex for size was poor budget purposes. For those purposes I made some assumptions that while. They may not. You know. Track. I figured they were as good as any you could make and that is that. Ah. oh requirements. On a monthly basis we'll be. Right but the his life to numbers in just added the because back. Peak get a hotel. And. on the energy the Hydro are foam energy years contractual marriage Helena that's good number. On the Hydro I'd use last juice. Petrol. and the bats properly. I Dunno I mean that. Oh up It's Gonna look like but right now January a lot of. Course the thing though they might they might come back and say oh by the way It's better the year A1 It's 9 point it could always have and but certainly below the last property was 30 listen when Hello work December 2000 McGuire Megawatt hour. We didn't like yeah be better off better off. Looking at the at the passion is degree maybe we do see you by 530%. January it looks like It's a reduction what but the we can be part of a more Conservative she need to bylaw 100 and 65 but John over buy film out please how much was it as a. um I tell you what I'm Gonna do. I'm Go in to the. Revise the Hydro for 30 worry. And. March down to. No access. We'll start from past. And. we're that takes years 5 numbers were concerned and then I'll look forward to see how we need to just Presley which me you'll get back. Later. bye. Bye. pass. Able to get 100 megawatt every year for the next 6 you. Well done no Ah no no you ready to pull the trigger man oh no. I am here we go to my. God in our subpoena. Looks like we're pretty close got to go ahead we just been for out exactly what you need oh well hopefully not the next you know nobody into this week oh Tomorrow or Friday you pull the trigger I just Wanna. You know obviously feel comfortable would you comfortable that you know these prices I mean obviously you don't have a Crystal Ball you know that is don't crap out yeah. Fundamental. error the wild card those can be in the West Hiya Jo are thanks but some of that this day. I can say with all. 3. Fair amount of that this is very bullish picture I take. And comes out why It's because to me gas is where oh where did the true supply demand balance comes coming through the 4th. All these other things. Yeah political whereas grabber a what people think supply demand balance is a are currently without a you know dollar to get our catch really short the picture of what worth where we're at and you are California just messed up yeah but TV few things belt the month what from 100 and 50 to 300. When the weather ahead There'll be time for the weather doesn't had and price is get the Prize but if you say you a normal life they wanted a powder and a. think that's high jump thing er no 1 knows What's Gonna happen there's there's look at the last 40 50 years this chart their coats it buy a colour buy a mobile the lower than average. Bunches oh well and It's not fair I would say not a guaranteed that it like that but I think. Everything points to and really bullish puncture um at sea. So. Rolling blackouts yeah I didn't know I'd heard about a little while ago San Francisco you got no the real tragedies killed. Traffic lights go now oh yeah and now. The more sophisticated Hi Tech manufacturing and the just headquarters at company headquarters in silicon Valley bad Rolling blackouts you don't like electricity the how did you ever used and be cared that damn thing to work perfect so much promise. Well no just Freaks me out of time has come out and say that. We're not going to be rest of the country not going to be impacted the you know how old is got. The work was problem and. The for the rest of the company of you OK heard rest of the country. Where the Hell are these guys later. Well earned. Nasdaq, NASDAQ index is probably 580% and Cal. Ahead. head you said by NASDAQ and say the rest of the clutch it doesn't get affected well you know majority of the company if that would represent the NASDAQ is in California going to have an impact of the prices rise by step substantial figure. You've heard the I feel a cop that than and know working with you in saying that this is call a good thing to do then we did that. Take a look at the Excel spreadsheet I like it a little bit I think It's not only from the standpoint of like it out that price for it and just right now looking for to say that er It's bullish but also because are

Document Info

Tagging

All Applicable Tags

- Confidential SHIFT+C
- French SHIFT+F
- German SHIFT+G
- Italian SHIFT+I
- South-African SHIFT+S
- Responsive SHIFT+R
- Not Responsive SHIFT+N
- Dutch SHIFT+D
- Archived SHIFT+A
- Danish SHIFT+D
- Turkish SHIFT+T
- Russian SHIFT+R

Scope Settings

Documents in scope: 1

Email Conversation

None

Current Branch

Entire Conversation

Redaction & Annotations

Reviewer Remarks

Near Duplicates

Download

Productions

Audit

Collection & Processing

■ Collection & Connectivity

- ✓ File shares
- ✓ MS Exchange
- ✓ G-Suite
 - ✓ Google Docs
 - ✓ Gmail
- ✓ Office 365
 - ✓ Outlook 365
 - ✓ SharePoint
 - ✓ Drop box

■ Processing

- ✓ Support for +700 file formats
- ✓ Support for +400 languages
- ✓ Mime-type detection
- ✓ Deduplication
- ✓ Hashing
- ✓ Single-instance storage
- ✓ Filtering
- ✓ Multi-directional OCR
- ✓ Customizable processing rules

(AUTO) Data classification

- **Rules based**

- ✓ Full text search
- ✓ Pattern Search (REGEX)
- ✓ Language detection
- ✓ Data enrichment (FA)
 - ✓ Property extraction
 - ✓ Full-text search
 - ✓ Regular expression
 - ✓ Meta-data

Auto Classification



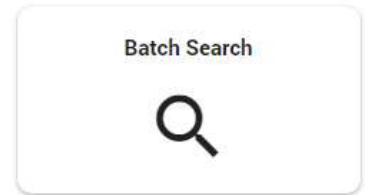
Review & QC

- Search
 - ✓ Full-text search
 - ✓ Pattern Search
 - ✓ Meta data search
- Data visualization
 - ✓ Facets
 - ✓ Timelines
- Duplicate detection
 - ✓ Exact duplicates
 - ✓ Near duplicate
- Tagging
 - ✓ Manual tagging
 - ✓ Bulk tagging
 - ✓ Tagging scope
- ✓ Redaction
- ✓ Annotation
- ✓ Dual monitor support

Search interface showing a complex query: `((abd ltd and robert) OR [A-Z]{3}[0-9]{3}) OR (account~1 AND (>16000)) OR ([0-9]{3}-[0-9]{2}-[0-9]{4}) w/2 Social Security* or Soc Sec or SSN* or SSNS or SSID)`

Filters: Custodian = John Doe

Source filters: Source = Laptop, Source = John Doe Desktop, Source = Desktop, Source = Email



Responsive Hot Document

Search Hits 4 < > Disable

From: Robert Harris

Sent: woensdag 25 juni 2014 11:59

To: PII PII PII

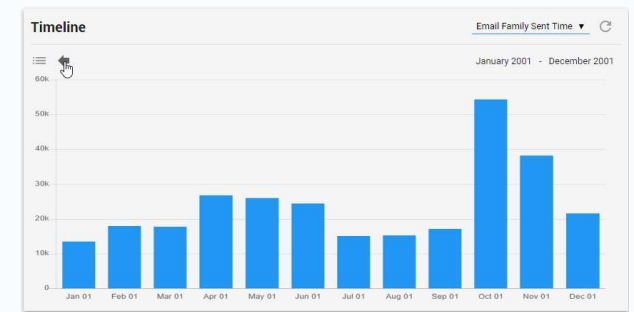
Subject: Dumping our prices!!!!

John,

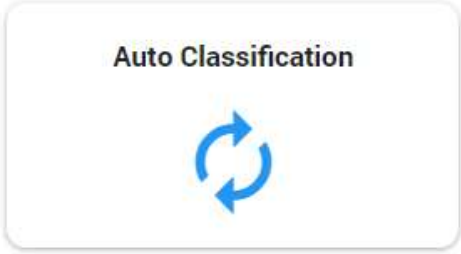
We should talk about doing something with **ABD Ltd** in Amsterdam: they are constantly discounting and dumping our products under their purchase price, hurting our partnership program in the region. In addition, this is also hurting our brand reputation. Attached you'll find some interesting information! Would you be able to call me tonight, perhaps our mutual friend B could help us and otherwise visit me at our usual spot after office hours! We have to act now and align our price strategy or we should stop delivering.

Hope I see you tonight.

Robert



Machine learning (TAR)

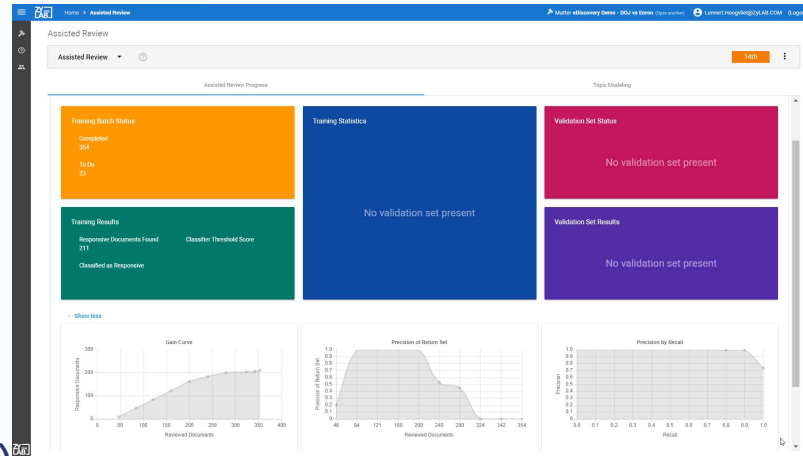


Supervised

- ✓ Assisted document

Unsupervised

- ✓ Topic modeling
- ✓ Professional Text Mining (NLP)
 - ✓ Entity extraction
 - ✓ Fact extraction



PII detection & redaction

- PII - detection

- Pattern Search (REGEX)

- ✓ BSN/SSN
- ✓ Passport

- Professional Text Mining

 - Entity extraction

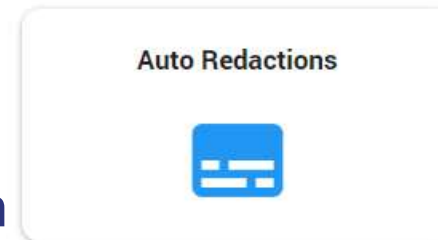
 - ✓ Persons
 - ✓ Organizations
 - ✓ Address
 - ✓ Email
 - ✓ Etc.

- Redaction

- Manual redaction

- Auto-redaction

 - Search based (Full text + Patterns)
 - Entity redaction
 - ✓ Persons
 - ✓ Organizations
 - ✓ Address
 - ✓ Email
 - ✓ Etc.



Hierbij stuur ik u de gegevens zoals eerder afgesproken met uw **entity** **entity** had mij ook verteld om dit zo snel mogelijk door te sturen. Ben benieuwd naar project XYZ.

entity

Adres: **entity**, **entity**

entity

ZYLAB ONE API

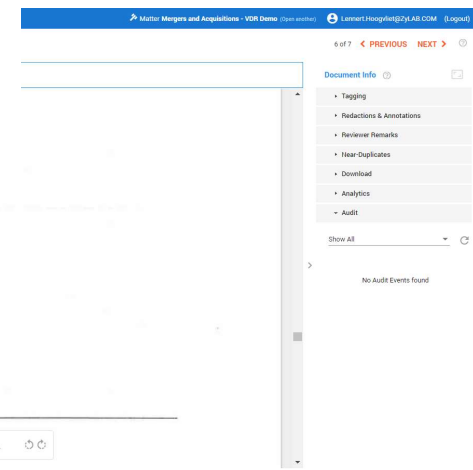
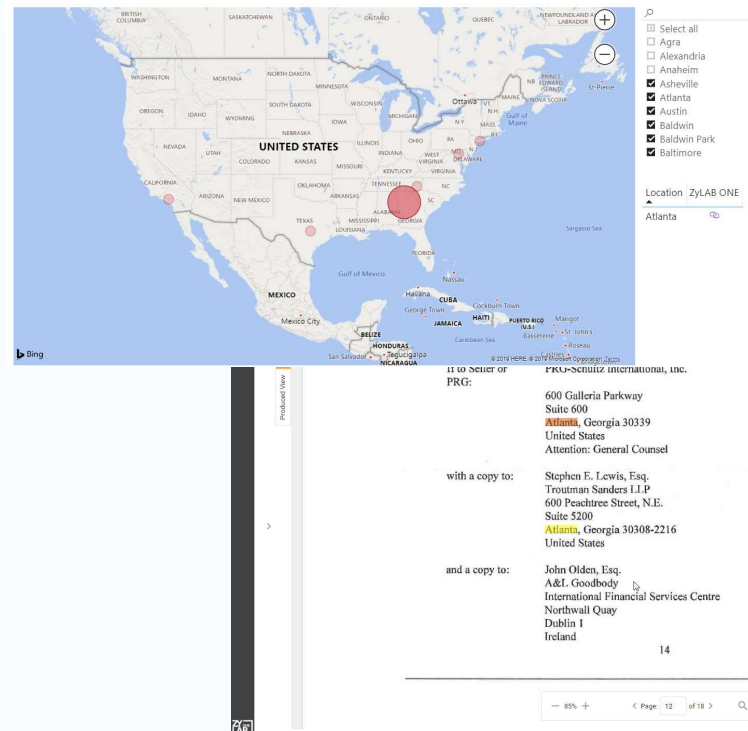
■ Productions

- ✓ Native
- ✓ Images
- ✓ Text
- ✓ Load files

■ Public API

- ✓ Ingestion
- ✓ Search
- ✓ Edit
- ✓ Delete
- ✓ Download
- ✓ Processing framework
 - 10 Custom Collectors
 - 10 Custom Processors
 - 10 Custom Indexers

Geo Mapping based on extracted entities from Zylab ONE IES (integration with Zylab ONE)



ZY
LAB.]