



an IPRO company

Europe Show

27 - 29 September 2021

Bridging the Gap between Legal and Technology



Johannes C. Scholtes

Chief Data Scientist – ZyLAB / iPRO



Tala Jomaa

Business Development German Markets- ZyLAB /
iPRO

Legal Technology & Legal Defensibility



Both directions:

- Teaching legal professionals technology
- Help technicians understand legal requirements

Legal Defensibility

When using technology in the legal domain, it is important to take the following requirements in account:

- Applicable legislation: which legal framework(s) apply? Are there any contradictions? Can technology address these?
- Ethical standards.
- Solid quality control on all automatic and manual actions.
- Understand where there is a risk for bias. Explain and document how this can be prevented.
- Transparency of technology: you need to be able to explain it, also to laymen.
- Robustness of technology: what happens with faulty or noisy input data?
- Reproducibility of technology. Using the same algorithm on the same data should get you the same results, also at a later moment in time
- Documentation: *Chain of Custody en Audit Trails* (who did what when, *inclusion-*, *exclusion-* and error reports)
- Safeguarding forensic integrity of the data and the process
- Is there (international) case law on the use of this technology? Has the technology been challenged in court before?

THIS IS WHAT LEGAL PROFESSIONALS SHOULD ASK QUESTIONS ABOUT!

This is what should be included in a framework dealing with legal defensibility & eDiscovery technology

Examples of Bad Science versus Bad Ethics

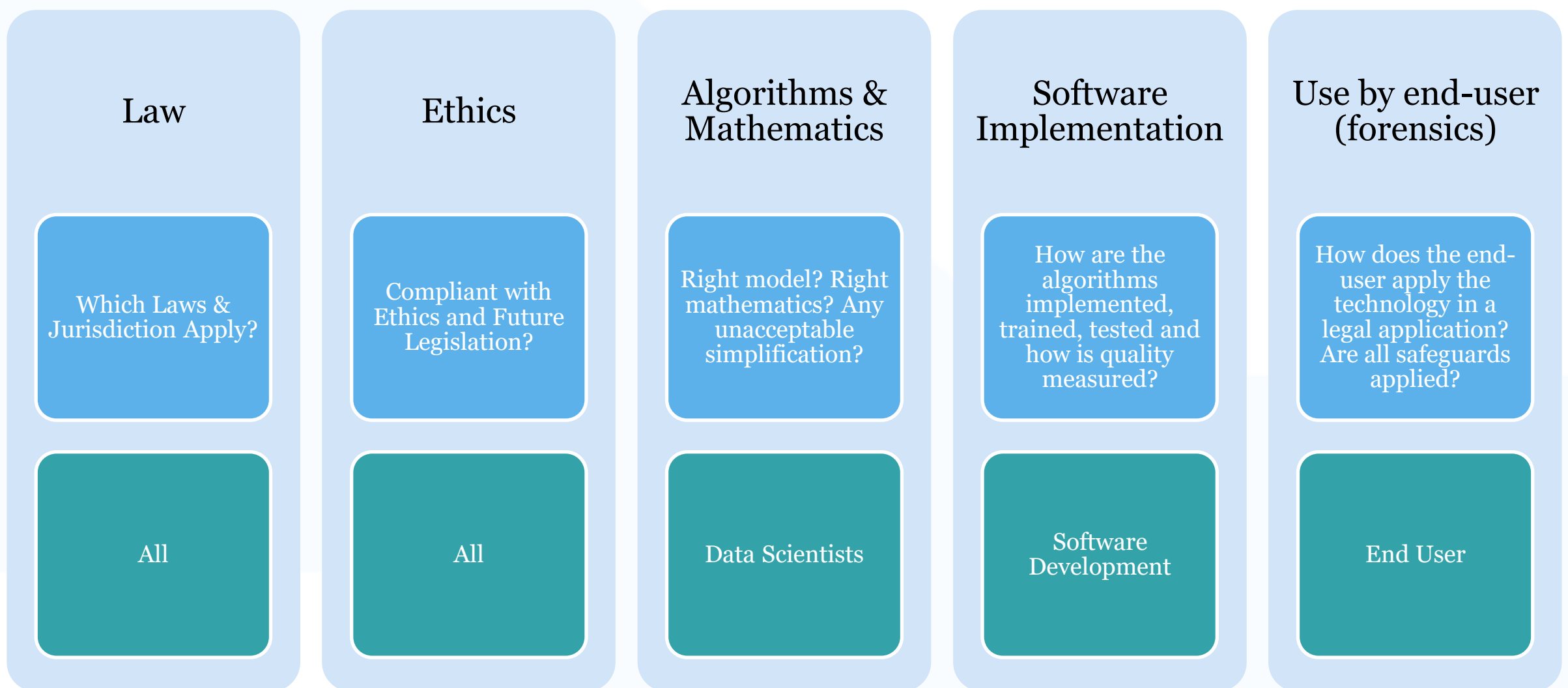
Bad Ethics

- No compliant to applicable legislation.
- Not being prepared for compliance with possible future legislation.
- No accountability for ethical values and standards.
- Prejudice or discriminative bias in data sets for machine learning.
- No transparency of algorithms.

Bad Science

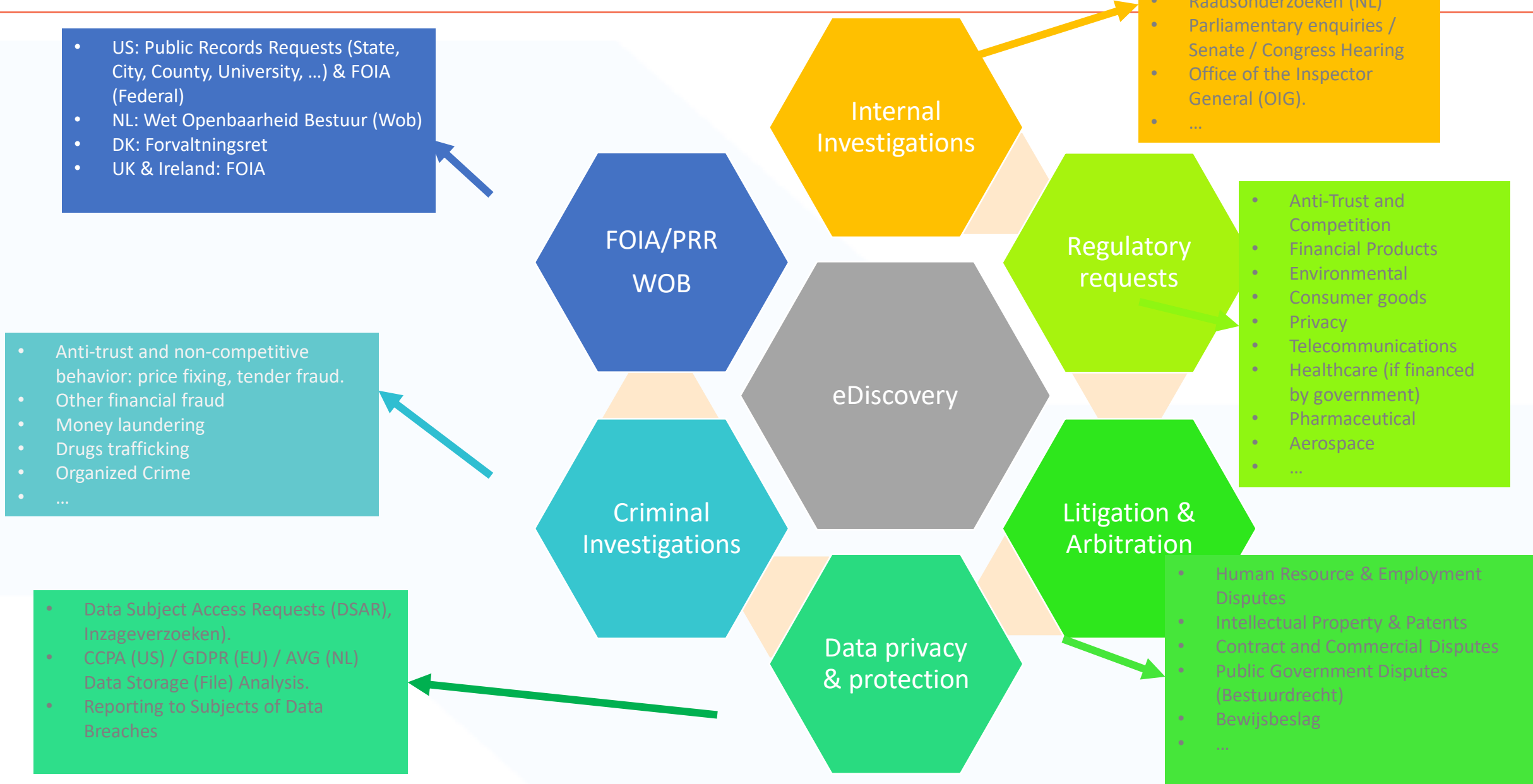
- Not being aware of any simplifications of real-world situations or limitations of applicability in the algorithms or mathematics you use.
- No objective quantitative validation of your algorithms or implementation.
- Using wrong quantitative validation measurements. E.g. using accuracy instead of precision & recall on unbalanced data sets.
- Creation of validation sets by just one individual: risk of human bias.
- Not measuring the disagreement between different individuals when using multiple individuals to create data sets (kappa distance).
- Selection or measurement bias in data sets for machine learning.
- Not testing algorithms against noisy input data.
- Jumping to conclusions: correlation does not imply causation.
- Not using forensic safe-guards (security, chain of custody, checks and balances, ...)
- Marketing ...

Responsible & Defensible use of Legal Technology



Legal Requirements

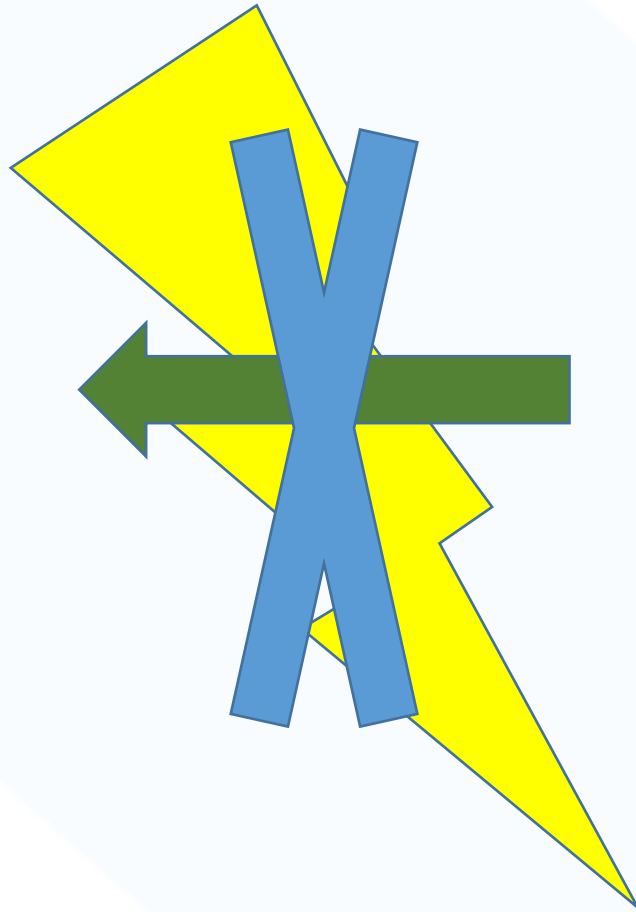
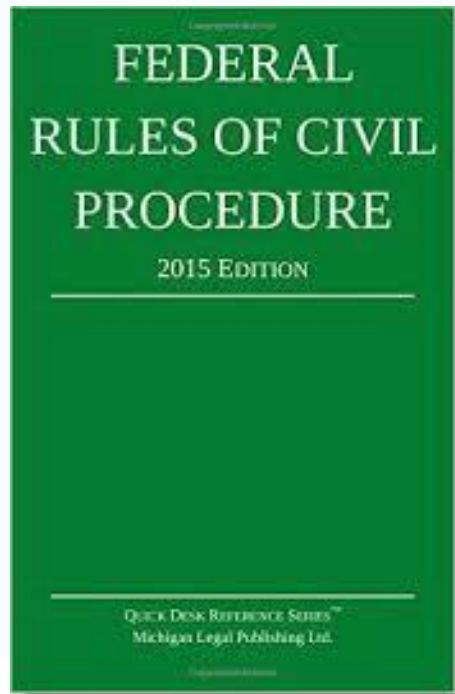
Applicable legislation depends on use case



What do you typically have to deal with?

- Privacy & data protection regulations
- Proportionality
- Subsidiarity
- Additional regulations and case law per use case? Recommended best practices? Mandatory audits for non-compliance?
- Copyrights, patents, specific contractual obligations, ...

Often (international) legislation is in contradiction to each other



Large Controversy on Google Books Project

Google: “we’re not scanning all those books to be read by people. We’re scanning them to be read by Artificial Intelligence”



Outcome: the US District Court’s eventual ruling in November 2013 in Author’s Guild, Inc. v Google, Inc. took a more pragmatic view, holding that Google’s use of the works, including “a type of research referred to as ‘data mining’ or ‘text mining’” was “fair use” under US copyright law: Google Books provides significant public benefits. It advances the progress of the arts and sciences, while maintaining respectful consideration for the rights of authors and other creative individuals, and without adversely impacting the rights of copyright holders. In Europe courts decided differently!

Ross versus West-Law (Thomson Reuters)

- Ross used law books from West-law to extract legal facts (using text-mining).
- Machine learning algorithms were trained with this data
- West-Law accused Ross of copyright violations.
- Interesting fact: most of West Laws data is based on public data (court verdicts) but is enriched (often also with text-mining).
- Ross ceased business and accused West-law of anti-trust behavior and abuse of monopoly positions

Source: <https://www.abajournal.com/news/article/ross-intelligence-to-shut-down-amid-thomson-reuters-lawsuit>

TECHNOLOGY

ROSS Intelligence will shut down amid lawsuit from Thomson Reuters

BY LYLE MORAN

DECEMBER 11, 2020, 11:50 AM CST

Like 54 Share Tweet Share



Image from Shutterstock.com.

ROSS Intelligence, which launched an AI-based legal research platform six years ago, announced Friday that it plans to cease operations early next year because of an ongoing copyright infringement lawsuit that Thomson Reuters brought against the company.

Thomson Reuters filed a [lawsuit](#) with the U.S. District Court for the District of Delaware in May, alleging that ROSS Intelligence had stolen “critical features” of Thomson Reuters’ Westlaw legal research platform to develop its own legal research offering.

“Litigation is expensive—no matter how speculative the claims against you nor how worthy your position,” ROSS Intelligence said in a [statement on the company’s website](#). “With our company ensnared by this legal battle, we have been unable to raise another round of funding to fuel our development and marketing efforts. Our bank account is running out, and we

must cease operations in the new year.”

ROSS Intelligence said as of Monday, it has stopped accepting new customers, and by Jan. 31, the platform will no longer be available.

“Between now and then, our priority is to help our current customers transition to other services,” the company said.

Existing case law on use of certain technology helps!

- Jurisprudence on the use of certain technology in earlier cases:
 - Has it been used?
 - Has it been challenged?
 - Did it survive such challenges?



International Case Law References



ZyLAB technology has been used in the following international cases by the prosecution, the lawyers, chambers and other participants of the following international courts and tribunals:

- ❖ UN International Court of Justice for various cases and archives, including the Nurnberg files. More information: <http://www.icj-cij.org/>.
- ❖ UN International Criminal Tribunal for the former Yugoslavia (UN ICTY). More information: <http://www.icty.org/>.
- ❖ UN International Criminal Tribunal for Rwanda (UN ICTR). More information: <http://www.icttr.org/>.
- ❖ Extraordinary Chambers in the Courts of Cambodia/United Nations Assistance to the Khmer Rouge Trials (ECCC/UNAKRT) <http://www.eccc.gov.kh/> and www.unakrt-online.org
- ❖ UN Special Court on Sierra Leone (UN-SCSL). More information: <http://www.sc-sl.org/>.
- ❖ Serious Crimes Investigations Team/United Nations Integrated Mission in East Timor (UNMIT) www.unmit.org
- ❖ United Nations International Independent Investigations Committee in Lebanon (UN IIIC)
- ❖ UN Tribunal for the Sea (UN ITLOS). More information: <http://www.itlos.org/>.
- ❖ European Union Rule of Law Mission in Kosovo (EULEX-Kosovo). More information: <http://www.eulex-kosovo.eu/>.
- ❖ European Union Anti-Fraud Department: OLAF: [European Anti-Fraud Office | European Commission \(europa.eu\)](http://european-anti-fraud-office.europa.eu)

2012: Case Law on Assisted Review (TAR)

- The Honorable Andrew J. Peck served for 23 years (from February 1995 until his retirement in February 2018) as a United States Magistrate Judge for the Southern District of New York, including a term as Chief Magistrate Judge from 2004 to 2005.
- Judge Peck is recognized internationally for bringing electronic discovery competency to the attention of both the judiciary and bar. He is widely described as one of the first judges to tackle the subject of e-discovery head on. His landmark decision in the 2012 employment class action *Monique Da Silva Moore, et. al. v. Publicis Groupe & MSL Group*, was the first judicial decision approving the use of technology-assisted review (TAR). By 2015, Judge Peck declared in *Rio Tinto v. Valle* that it was black-letter law that if the responding party wished to use TAR, courts would allow it. In the third of his trilogy of TAR cases, *Hyles v. City of New York*, he stated that while he preferred the use of TAR, neither the requesting party nor the court could require a reluctant responding party to use TAR.



Ethical Requirements

Values and Ethics

Ethics or moral philosophy deals with systematizing, defending, and recommending concepts of right and wrong conduct.

Ethics seeks to resolve questions of human morality by defining concepts such as good and evil, right and wrong, virtue and vice, justice and crime.

Centuries of study into moral philosophy. Many different schools and views. There is no single school of thought that settles all moral questions.



Ethics and Computer Science?

Association of Computing Machinery (ACM) defined 7 principles in relation to accountability and transparency of algorithms:

1. Awareness of bias
2. Access and redress to prevent discrimination
3. Accountability
4. Explanation
5. Data Provenance
6. Auditability
7. Validation and Testing



Association for
Computing Machinery

Computers are ruthless... fruit picking game



- Google's AI got "highly aggressive" when competition got stressful in a fruit-picking game.
- Shoot other players at beginning of game.

Scientific & Mathematical Requirements: preventing bad science

Transparency of your technology

Know your algorithms, their mathematics, their assumptions and this their limitations!

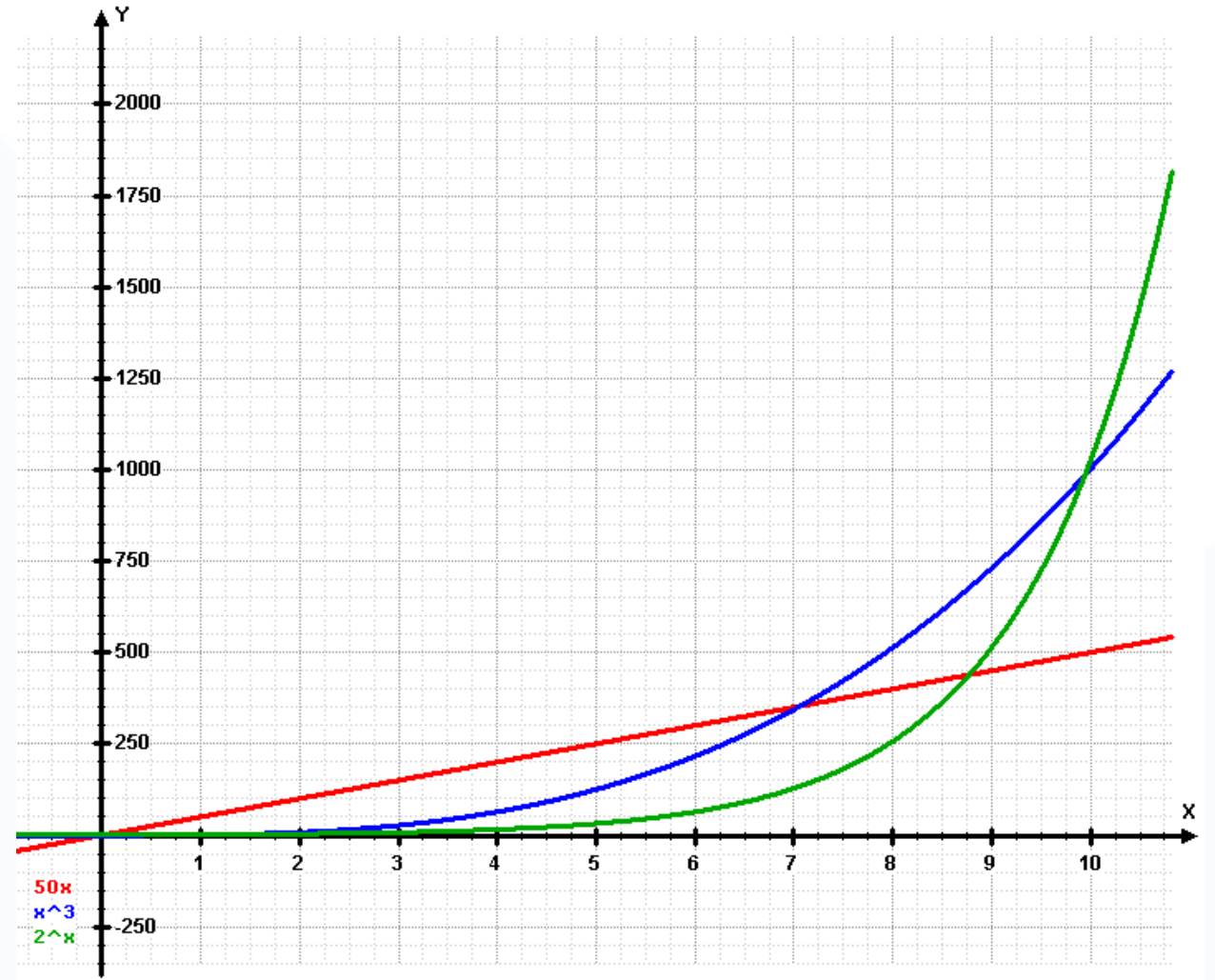
Also:

- Cultural invariance
- Temporal Invariance (other standards or values over time)
- Lacking domain expertise
- Internet based research: only popular data, not good broad un-biased spreading,. No provenance as well.
- Disagreeing judges when building data sets. Not measuring Kappa distances.
- Cherry picking in results
- Just stupid errors
- Simply overlooking exceptions (errors and omissions)
- Long tail of exceptions and rare occurrences

Can you explain the workings to laymen?

Predicting COVID Outbreaks

- Use the correct mathematical model
- Linear or exponential?

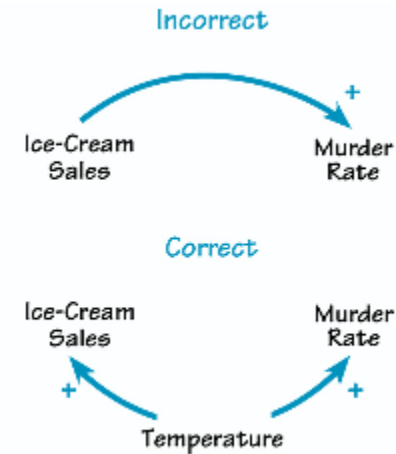


Primum non Nocere*: dealing with false positives

- Text mining and data mining would seem to provide a great deal of potential for inadvertent or deliberate misinterpretation of data.

Example: mining data extracted from a range of travel-related transaction databases (e.g., airline passenger manifests, railway ticket websites, etc.), to attempt to identify suspicious patterns of travel behavior that can be linked to named individuals. Support & Confidence: *"correlation does not imply causation"*

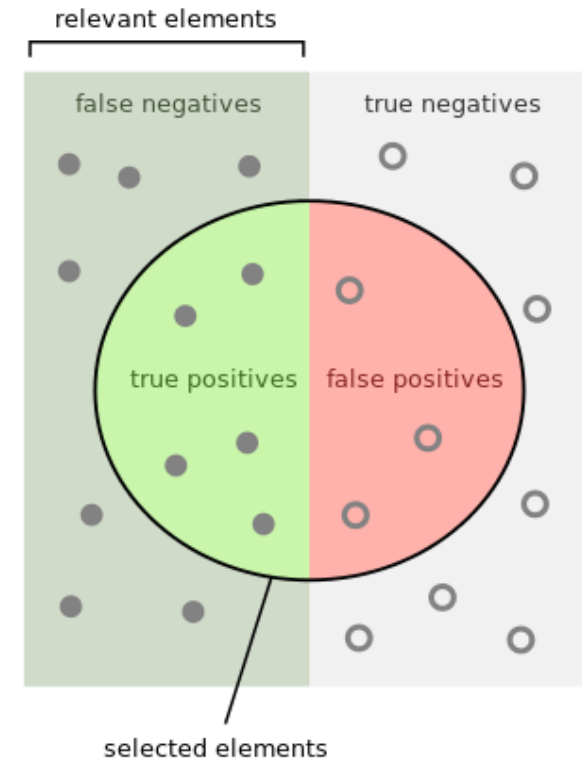
- Named entity recognition: error-prone, highly contextual sensitive, often well-informed guesswork: 90% precision on 100.000 patters, still leaves us 10.000 WRONG conclusions!
- Complex relation extraction: often no more than 60% f1 values, so even lower precision and recall!



* first, do not harm

Use the Correct Quantitative Measurements

- Lack of precision leads to noise, too many false hits, too much work to review, which yields **high cost of review**.
- Lack of recall leads to missing relevant documents which yields **risk**.



How many selected items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

Should we instead use the accuracy measure for evaluation?

- Given a query, an engine classifies each doc as “Relevant” or “Nonrelevant”
- The **accuracy** of an engine: the fraction of these classifications that are correct
 - $(tp + tn) / (tp + fp + fn + tn)$
- **Accuracy** is a commonly used evaluation measure in machine learning classification work
- Why is this not a very useful evaluation measure in IR?

High accuracy does not mean you find something useful

- How to build a 99.9999% accurate search engine on a low budget....
- $\text{Accuracy} = (\text{tp} + \text{tn}) / (\text{tp} + \text{fp} + \text{fn} + \text{tn})$

snoogle.com

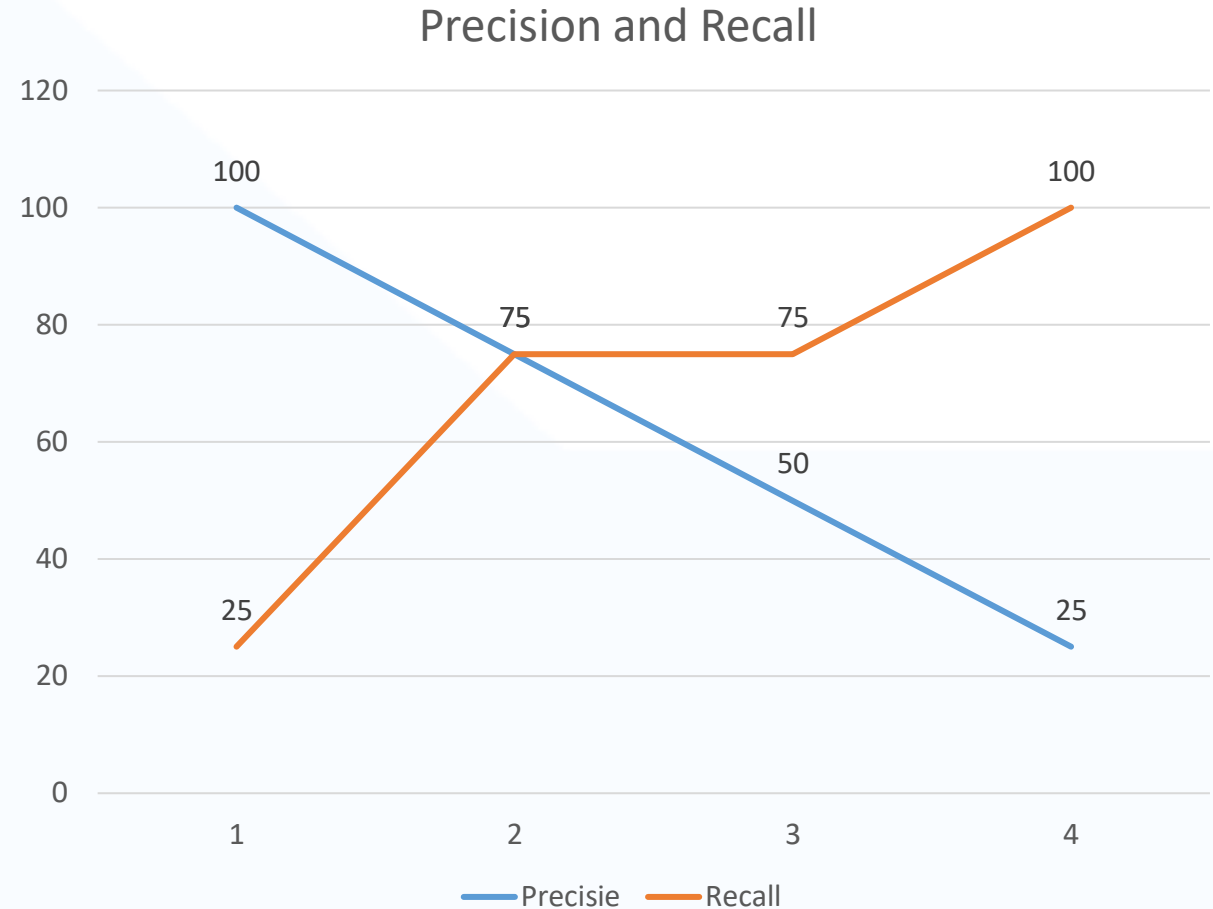
Search for:

0 matching results found.

- Suppose we have 100.000 docs, we find 1 of them, there are 1000 relevant. Then $\text{accuracy} = 1 + 99000 / 1 + 0 + 999 + 99000 = 99\%$. But we missed 999 documents!

Precision & recall: reverse proportional

- Increase precision: AND, W/5, NOT
- Increase recall: OR, *, ?, Thesaurus, fuzzy.
- Both (up to certain level): Quorum search



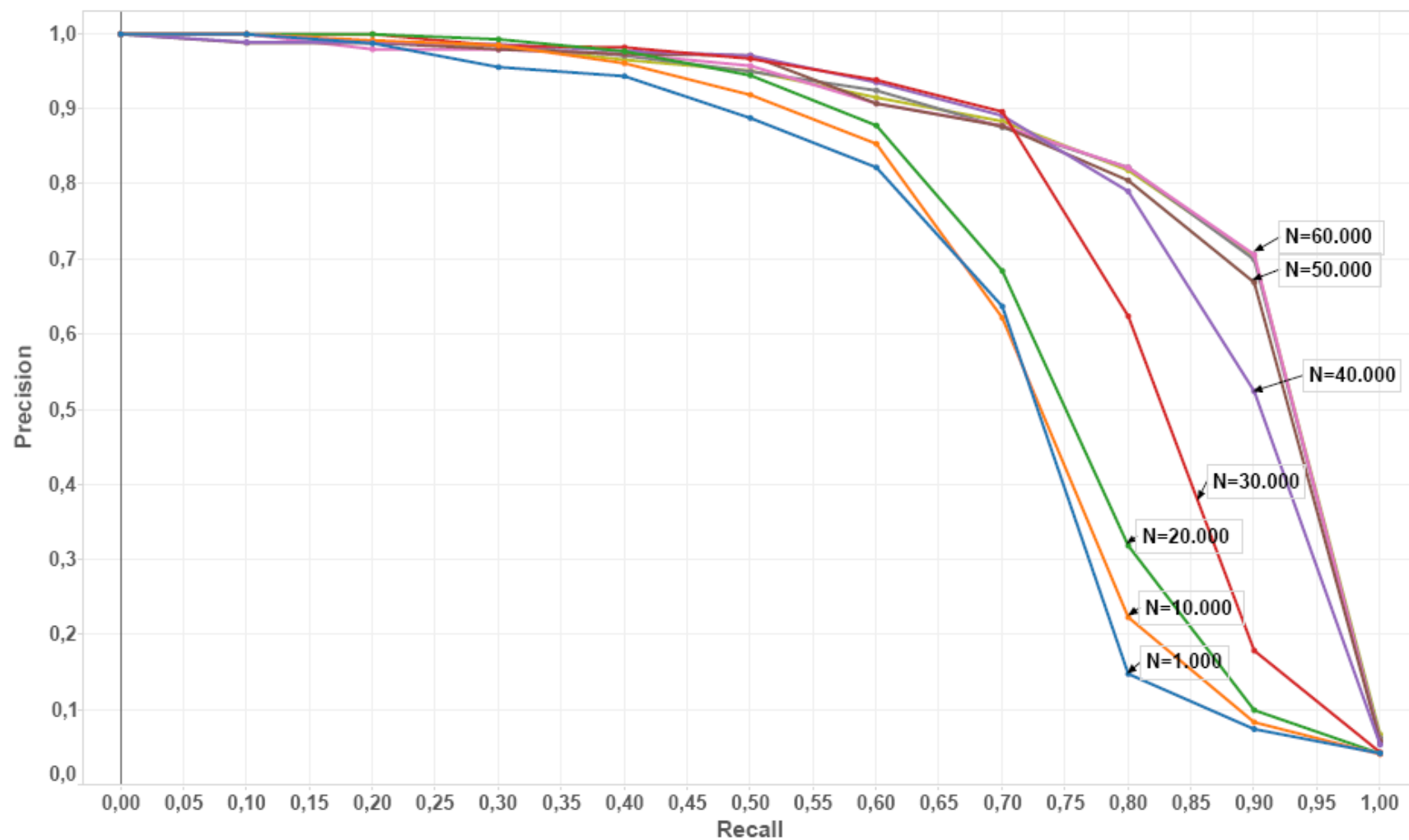


Figure 8.2: Precision-Recall curves of classifiers of various iterations on the evaluation set - Simulation of CAL protocol on the whole RCV1 corpus with topic code GV10

A combined measure: F

- Combined measure that assesses precision/recall tradeoff is **F measure**:

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

Human Performance

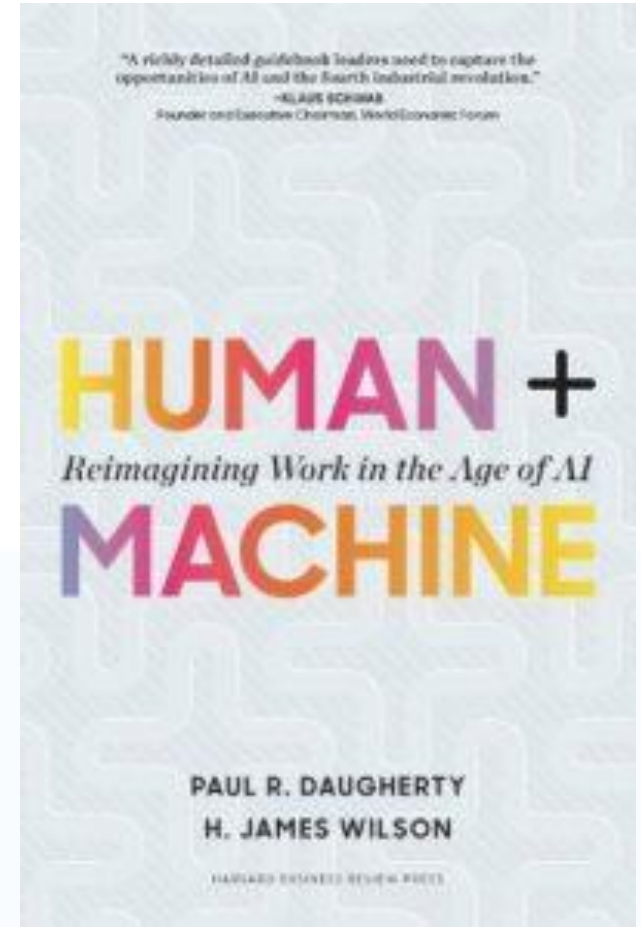
- When both precision and recall are over 80%, human performance is approached.
- This applies to the best humans.
- It can be argued that values over 80% are often subject to different interpretations and discussions.



Human, machine or Machine & Human?

What we observe in the real world:

1. (Good) human f-1 values are 60-80%
2. The best algorithms typically reach 80-90%
3. By first having a computer do the heavy lifting and initial selection and then have human beings make the final call: 95+%!

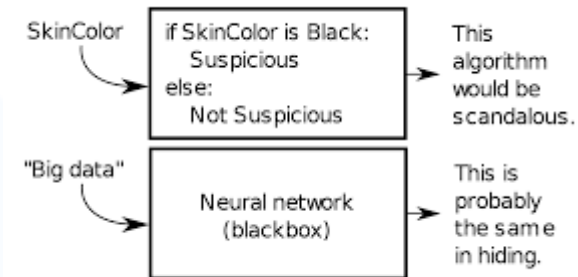
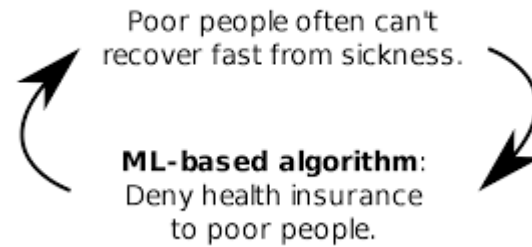


Implementation Requirements

There are different types of Bias

- Sample bias: the data used to train model does not accurately represent the real-world. Aka “balance” in training data. *Train car to drive only on day time images and not on night images.*
- Prejudice bias: training data is influenced by cultural stereotypes. *Images of people at work. Men are coding software, women are in kitchen cooking. Aka selection bias or confirmation bias*
- Measurement bias: systematic value distortion due to issues with the device used to measure features. *Using special filters on camera when taking pictures. Images without these filters will have lower recognition.*
- Algorithm bias: wrong algorithm. Wrong bias-variance balance. *Linear one that cannot deal with non-linear data.*

Examples of Prejudice Bias



COOKING

ROLE	VALUE
AGENT	▶ WOMAN
FOOD	▶ PASTA
HEAT	▶ STOVE
TOOL	▶ SPATULA
PLACE	▶ KITCHEN



COOKING

ROLE	VALUE
AGENT	▶ WOMAN
FOOD	▶ FRUIT
HEAT	▶ —
TOOL	▶ KNIFE
PLACE	▶ KITCHEN



COOKING

ROLE	VALUE
AGENT	▶ WOMAN
FOOD	▶ MEAT
HEAT	▶ GRILL
TOOL	▶ TONGS
PLACE	▶ OUTSIDE



COOKING

ROLE	VALUE
AGENT	▶ WOMAN
FOOD	▶ VEGETABLES
HEAT	▶ STOVE
TOOL	▶ TONGS
PLACE	▶ KITCHEN



COOKING

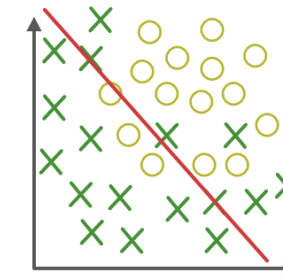
ROLE	VALUE
AGENT	▶ MAN
FOOD	▶ —
HEAT	▶ STOVE
TOOL	▶ SPATULA
PLACE	▶ KITCHEN

Machine Learning: Variance and Bias

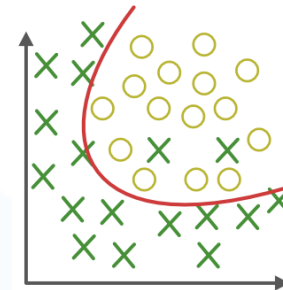
- Bias (error): The bias error is an *error from erroneous assumptions* in the learning algorithm. High bias can cause an algorithm to miss the relevant relations between features and target outputs (high bias leads to under fitting).
- Variance: The variance is an *error from sensitivity to small fluctuations in the training set*. High variance can cause an algorithm to model the random noise in the training data, rather than the intended outputs (high variance leads to overfitting)

Machine Learning

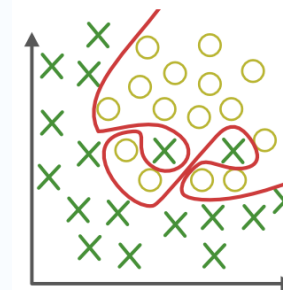
- Underfitting: occurs when a statistical model or machine learning algorithm cannot adequately capture the underlying structure of the data. It occurs when the model or algorithm does not fit the data enough. Caused by low variance but high bias.
- Overfitting: “the production of an analysis that corresponds too closely or exactly to a particular set of data, and may therefore fail to fit additional data or predict future observations reliably”. Caused by high variance and low bias. Green line on right.



Under-fitting
(too simple to explain the variance)

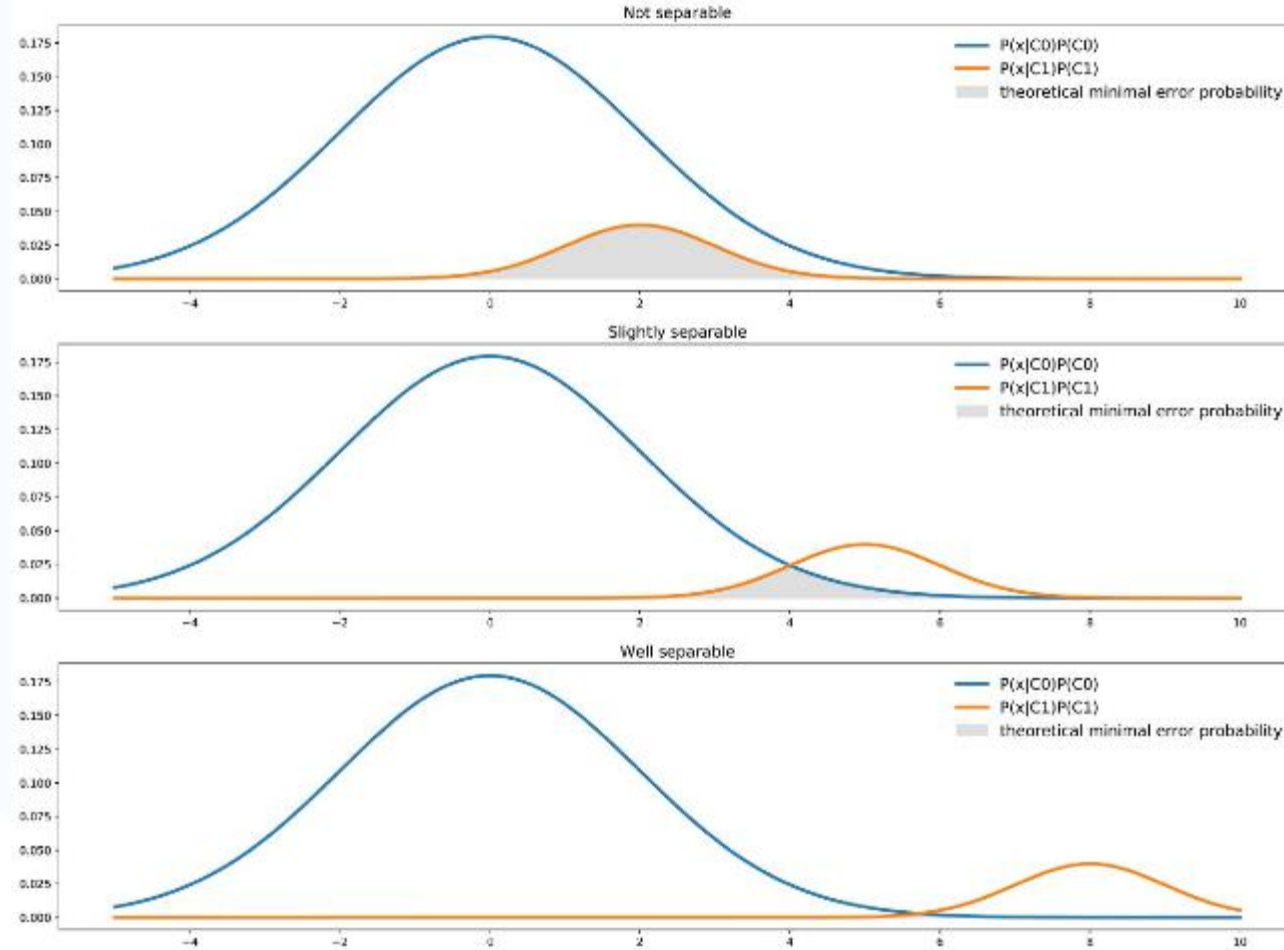


Appropriate-fitting

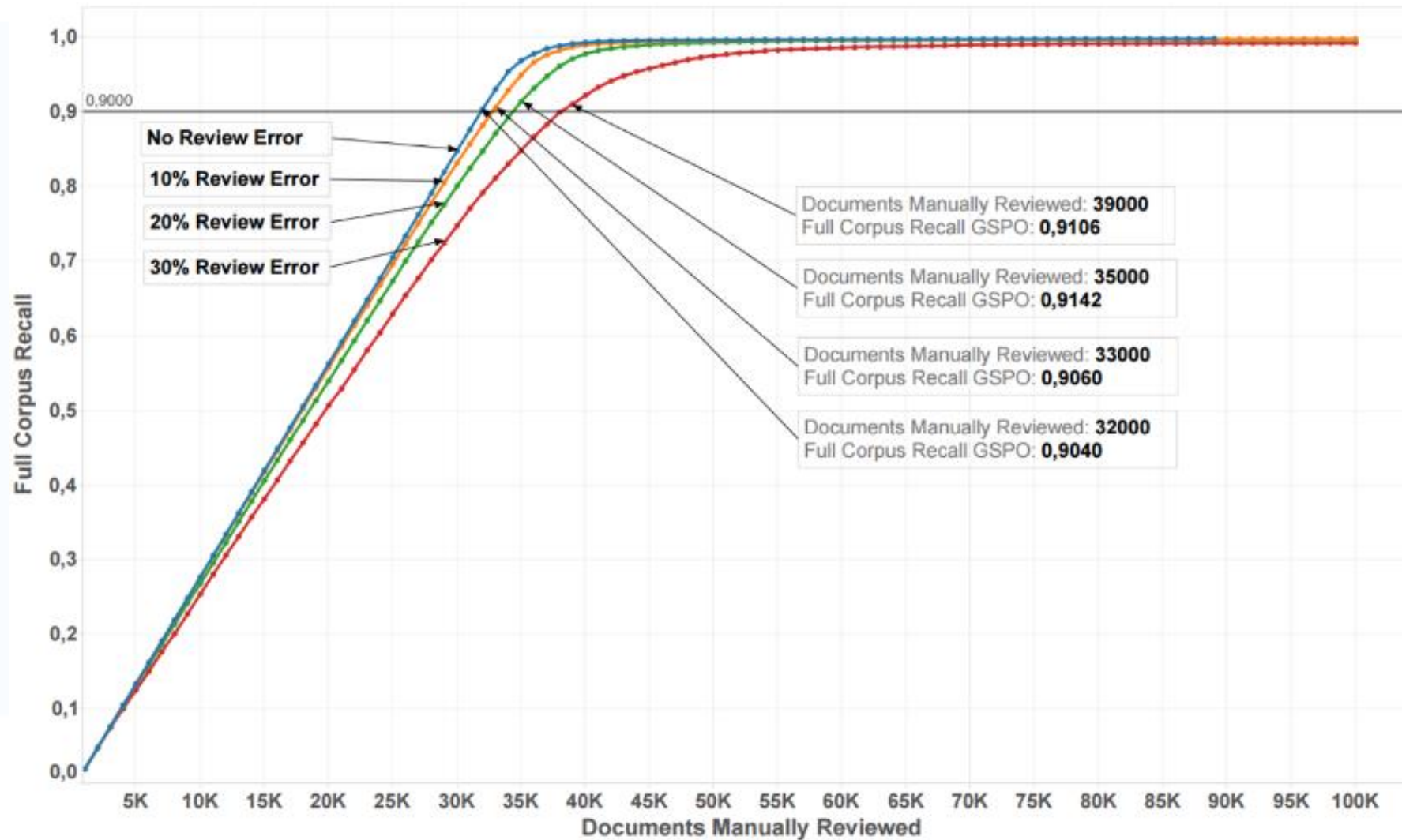


Over-fitting
(forcefitting--too good to be true)

Imbalanced Training Data



Robustness of your algorithms: test it extensively on noisy and incorrect data!



Reproducibility of your technology

- Using the same algorithm on the same data should get you the same results, also at a later moment in time, regardless from where you run the algorithm.

→ Question: *Is the use of Google as a legal search engine defensible when it is used to identify relevant case law?*

Usage (Forensic) Requirements

Legally Defensible Users: Reporting, Sampling & Audit logs

The screenshot displays the ZYLAB eDiscovery Demo interface. The top navigation bar is blue and contains the ZYLAB logo, a hamburger menu, and user information: "EN", "Matter eDiscovery Demo - DOJ vs Enron (Open another)", and "johannes.scholtes@zylab.com (Logout)".

The main content area is titled "eDiscovery Demo - DOJ vs Enron" with a link "Open Another Matter". Below the title, it says "Please select one of the following options to proceed:". A grid of 12 options is presented, each with an icon and a label:

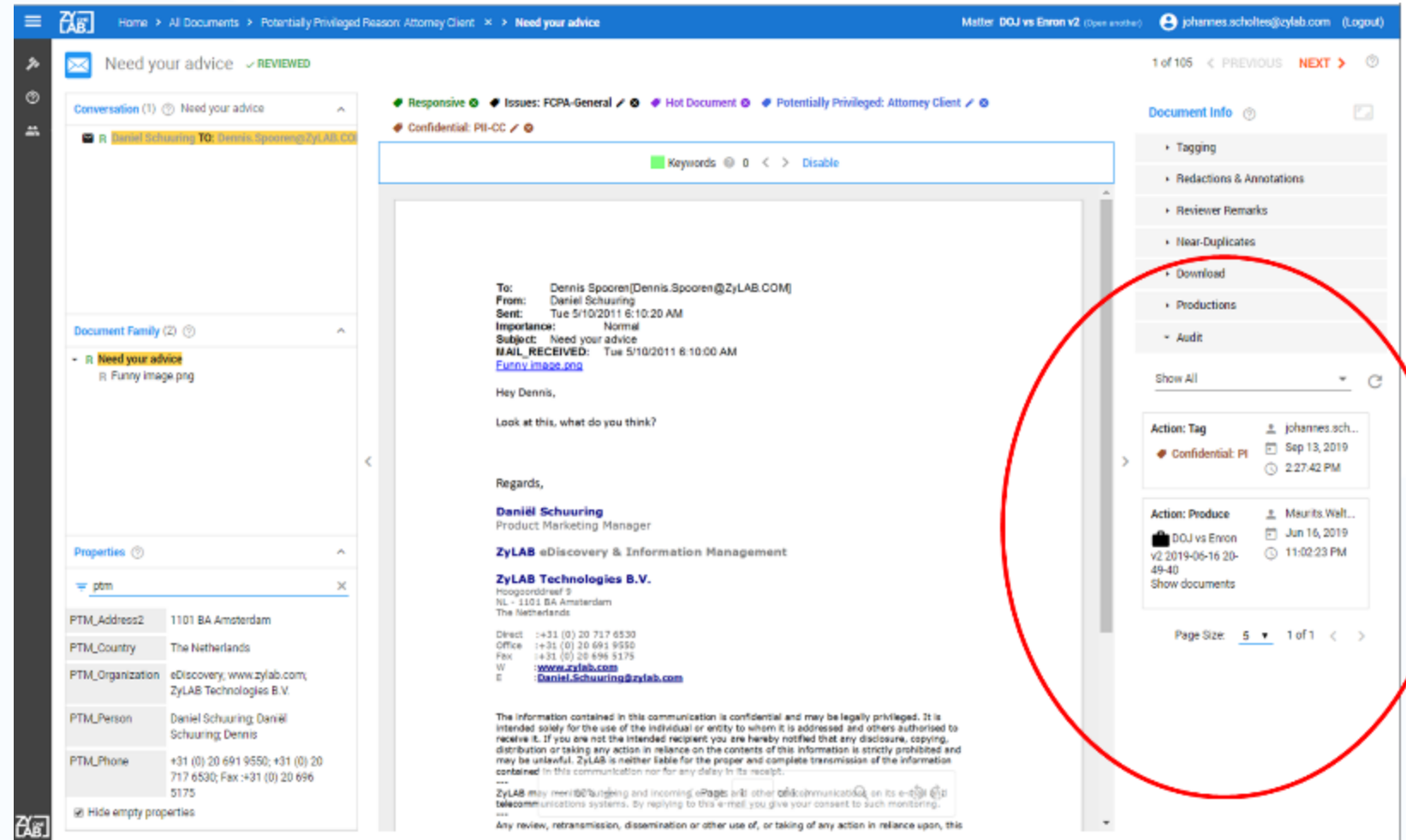
- Continue Review (blue arrow icon)
- Browse (yellow grid icon)
- Upload (blue cloud with arrow icon)
- Overview (blue grid icon)
- Batch Search (magnifying glass icon)
- Sampling Overview (blue magnifying glass over document icon)
- Production Overview (brown briefcase icon)
- Audit Log (blue circular arrow icon)
- Review Batches (black person icon)
- Reports (orange grid icon)
- Configuration (green gear icon)
- Assisted Review (blue line graph icon)

Three red circles are drawn around the "Reports", "Sampling Overview", and "Audit Log" options, highlighting them as key features for legal defensibility.

The bottom of the interface shows a URL bar with the address: <https://sb01.myzylab.com/ediscovery/legalreview/htmlapp/#/matter/17/reports>.

Documentation: *Chain of Custody* (Who did what when?)

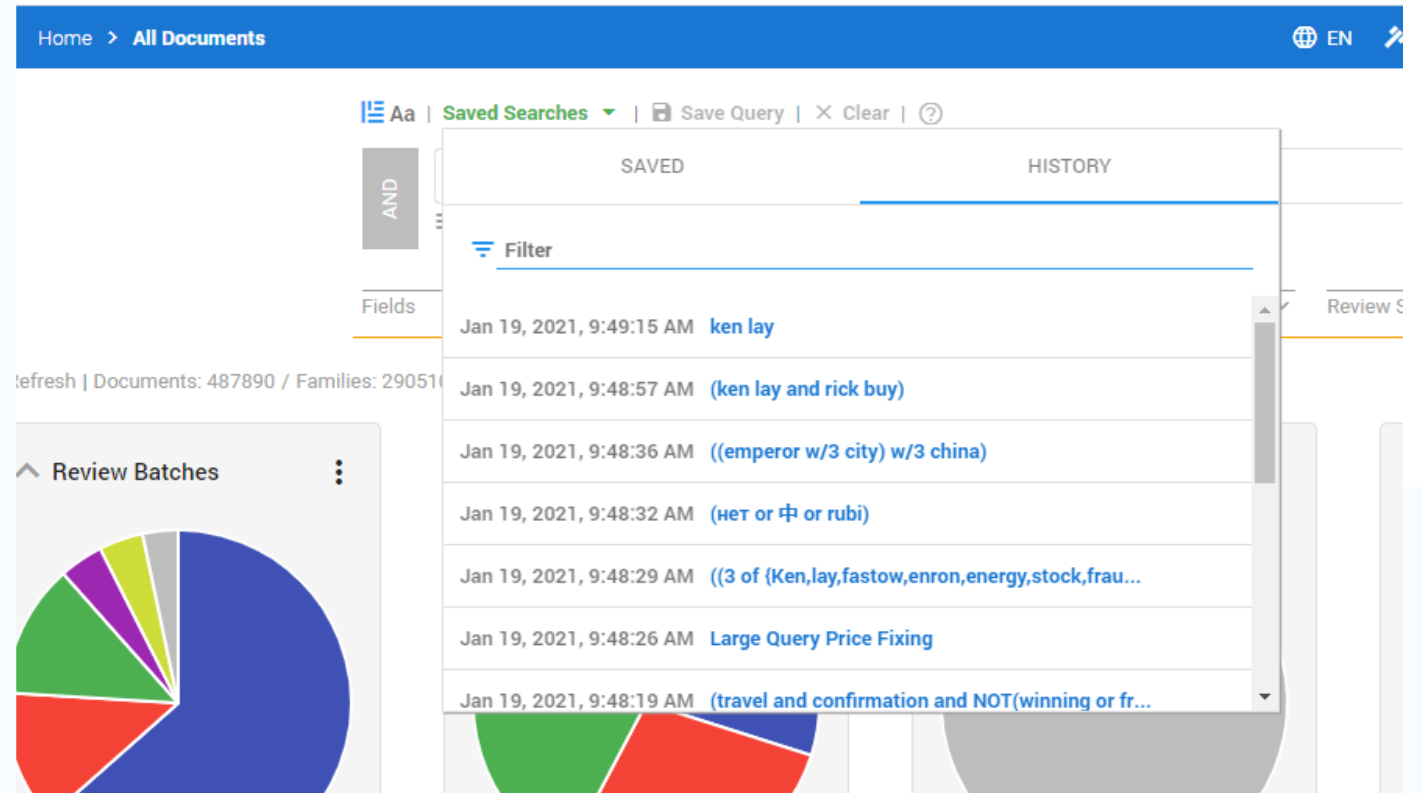
- Search logs
- Audit logs
- Reporting
- Inclusion reports
- Exclusion reports
- ...



The screenshot displays the ZyLAB eDiscovery & Information Management interface. The main window shows an email document titled "Need your advice" with a status of "REVIEWED". The email is from Daniel Schuuring to Dennis Spooren. The interface includes a left sidebar with navigation options like "Conversation", "Document Family", and "Properties". The right sidebar, titled "Document Info", is circled in red and contains sections for "Tagging", "Redactions & Annotations", "Reviewer Remarks", "Near-Duplicates", "Download", "Productions", and "Audit". The "Action: Tag" section shows a tag "Confidential: PI" applied by Johannes Sch... on Sep 13, 2019. The "Action: Produce" section shows a production of "DOJ vs Enron v2 2019-06-16 20-49-40" by Maunits Walt... on Jun 16, 2019. The interface also includes a search bar, a list of documents, and a footer with legal disclaimers.

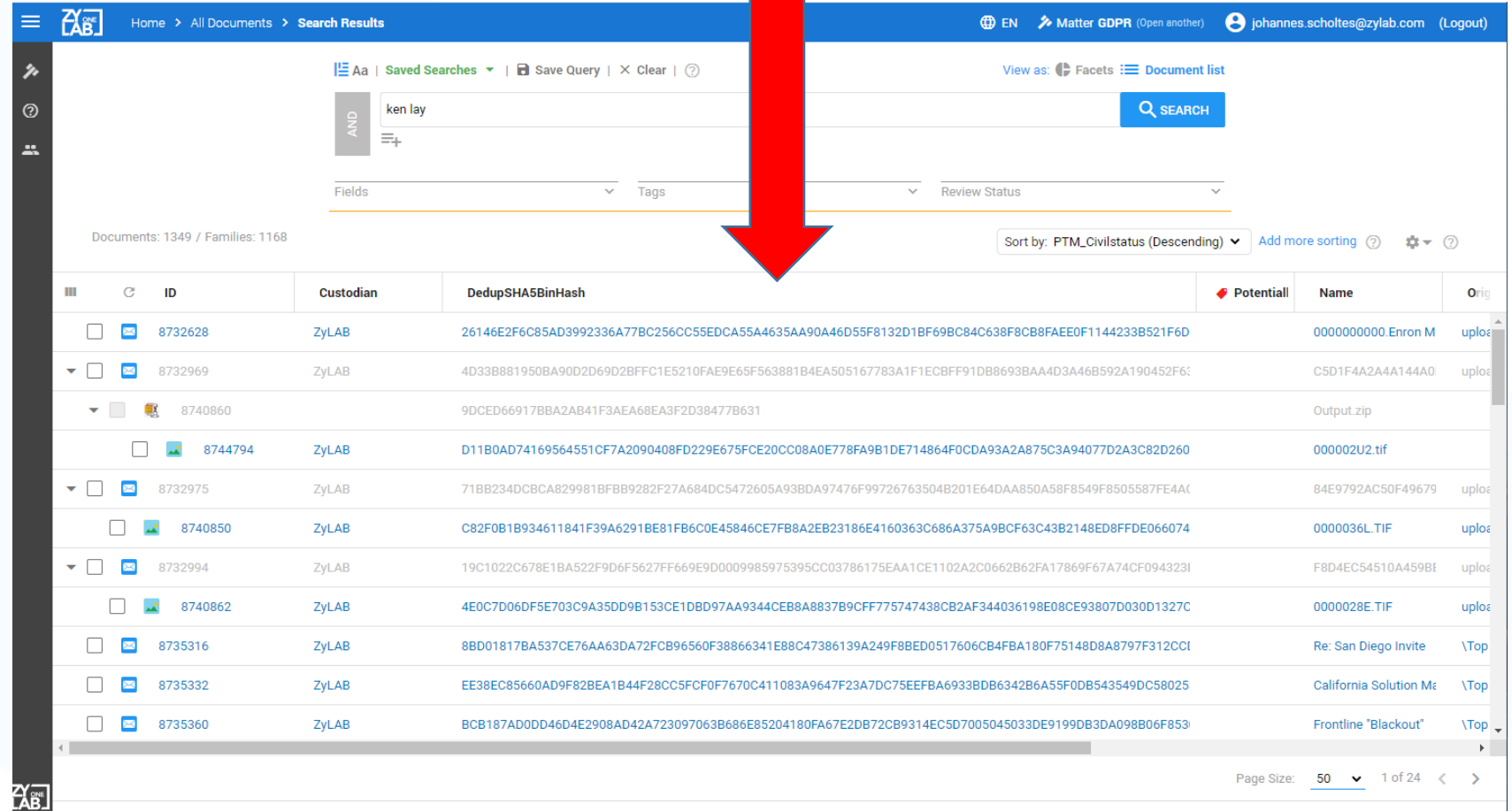
Search logs: evidence and falsification

- Keep track of previous queries and results.
- Document that one searches both for evidence, but also for potential falsifications.
- Avoid tunnel vision.



Forensic Integrity

- Hashing at collection
- Multiple purpose:
 - De-duplication
 - Preservation of forensic integrity by uniquely storing evidence of non-changed content of file



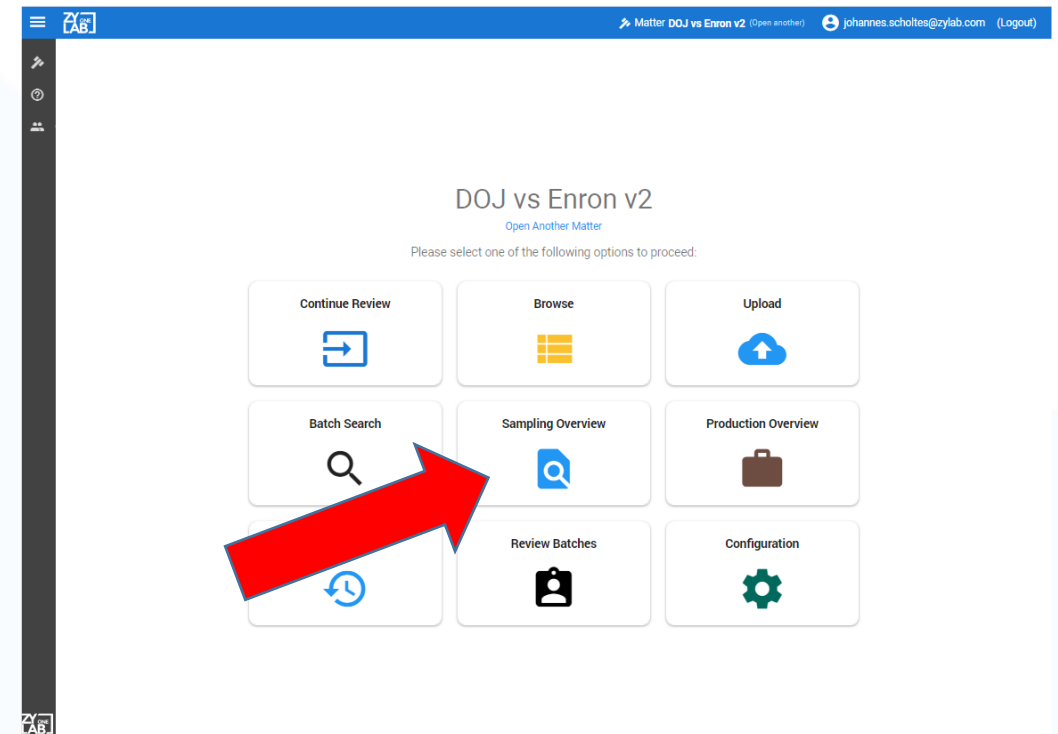
The screenshot displays the ZyLAB Search Results interface. The top navigation bar includes the ZyLAB logo, 'Home > All Documents > Search Results', and user information for 'johannes.scholtes@zylab.com'. The search bar contains the query 'ken lay'. Below the search bar, there are filters for 'Fields', 'Tags', and 'Review Status'. The results are sorted by 'PTM_Civilstatus (Descending)'. The table lists documents with columns for 'ID', 'Custodian', 'DedupSHA5BinHash', 'Potential', 'Name', and 'Orig'. A large red arrow points from the top of the slide to the 'DedupSHA5BinHash' column.

ID	Custodian	DedupSHA5BinHash	Potential	Name	Orig
8732628	ZyLAB	26146E2F6C85AD3992336A77BC256CC55EDCA55A4635AA90A46D55F8132D1BF69BC84C638F8CB8FAEE0F11442338521F6D		0000000000.Enron M	uploa
8732969	ZyLAB	4D33B881950BA90D2D69D2BFFC1E5210FAE9E65F563881B4EA505167783A1F1ECBFF91DB8693BAA4D3A46B592A190452F6C		C5D1F4A2A4A144A0	uploa
8740860		9DCED66917BBA2AB41F3AEA68EA3F2D38477B631		Output.zip	
8744794	ZyLAB	D11B0AD74169564551CF7A2090408FD229E675FCE20CC08A0E778FA9B1DE714864F0CDA93A2A875C3A94077D2A3C82D260		000002U2.tif	
8732975	ZyLAB	71BB234DCBCA829981BFB89282F27A684DC5472605A93BDA97476F99726763504B201E64DAA850A58F8549F8505587FE4AC		84E9792AC50F49679	uploa
8740850	ZyLAB	C82F0B1B934611841F39A6291BE81FB6C0E45864CE7FB8A2EB3186E4160363C686A375A9BCF63C43B2148ED8FFDE066074		0000036L.TIF	uploa
8732994	ZyLAB	19C1022C678E1BA522F9D6F5627FF669E9D0009985975395CC03786175EAA1CE1102A2C0662B62FA17869F67A74CF094323I		F8D4EC54510A4598I	uploa
8740862	ZyLAB	4E0C7D06DF5E703C9A35DD9B153CE1DBD97AA9344CEB8A8837B9CFF775747438CB2AF344036198E08CE93807D030D1327C		0000028E.TIF	uploa
8735316	ZyLAB	8BD01817BA537CE76AA63DA72FCB96560F38866341E88C47386139A249F8BED0517606CB4FBA180F75148D8A8797F312CCI		Re: San Diego Invite	\Top
8735332	ZyLAB	EE38EC85660AD9F82BEA1B44F28CC5FCF0F7670C411083A9647F23A7DC75EEFBA69338DB6342B6A55F0DB543549DC58025		California Solution Me	\Top
8735360	ZyLAB	BCB187AD0DD46D4E2908AD42A723097063B686E85204180FA67E2DB72CB9314EC5D7005045033DE9199DB3DA098B06F853		Frontline "Blackout"	\Top

Page Size: 50 1 of 24

Quality Control by using sampling


- **Test** the results of all **automatic processes** always manually on a random sample.
- This is the most reliable way to explain laymen the working of a complex algorithm and gain trust.
- Alternative are expert witnesses.



Using sampling

1. Whenever using any kind of automatic (AI) process
2. Select the documents on which the process has been applied
3. Create a random sample of X % (subject of negotiations and risk), typically $X = 2-5\%$ on large data sets (>1000) to $10\%-50\%$ on small data sets. There are statistical models that can provide guidance on exact measurements.
4. Review this set and keep track of maximum obtainable quality.
5. When the quality drops below a certain threshold: go back to #1 and adjust the parameters of your automatic process. Address most common errors.
6. Continue until the quality random sample is above the set quality threshold.

How large should your sample be?

			Sample size calculator	
What margin of error can you accept? <small>5% is a common choice</small>	<input type="text" value="5"/> %	The margin of error is the amount of error that you can tolerate. If 90% of respondents answer yes, while 10% answer no, you may be able to tolerate a larger amount of error than if the respondents are split 50-50 or 45-55. Lower margin of error requires a larger sample size.		
What confidence level do you need? <small>Typical choices are 90%, 95%, or 99%</small>	<input type="text" value="95"/> %	The confidence level is the amount of uncertainty you can tolerate. Suppose that you have 20 yes-no questions in your survey. With a confidence level of 95%, you would expect that for one of the questions (1 in 20), the percentage of people who answer yes would be more than the margin of error away from the true answer. The true answer is the percentage you would get if you exhaustively interviewed everyone. Higher confidence level requires a larger sample size.		
What is the population size? <small>If you don't know, use 20000</small>	<input type="text" value="20000"/>	How many people are there to choose your random sample from? The sample size doesn't change much for populations larger than 20,000.		
What is the response distribution? <small>Leave this as 50%</small>	<input type="text" value="50"/> %	For each question, what do you expect the results will be? If the sample is skewed highly one way or the other, the population probably is, too. If you don't know, use 50%, which gives the largest sample size. See below under More information if this is confusing.		
Your recommended sample size is	377	This is the minimum recommended size of your survey. If you create a sample of this many people and get responses from everyone, you're more likely to get a correct answer than you would from a large sample where only a small percentage of the sample responds to your survey.		

<http://www.raosoft.com/samplesize.html>

Select document for Quality control

Home > All Documents > Search Results

EN Matter DOJ vs Enron v2 (Open another) johannes.scholtes@zylab.com (Logout)

View as: Facets Document list

ken lay

SEARCH

AND OR NOT W/S P/5 ~2 () [0-9a-z]{3} 2 OF (query1, query2, query3)

Fields Tags Review Status

Documents: 3636 / Families: 3486 35 document(s) selected Reset Sort by: Default

	Name	Custodian	Email From	Email To	Email Subject	PTM_Person
<input type="checkbox"/>	Organizational Announcement	Albert Meyers	mbx_OfficeChairman@	DL-GA-all_enron_world	Organizational Announ	
<input type="checkbox"/>	Message from Ken Lay	Albert Meyers	mbx_OfficeChairman@	DL-GA-all_enron_world	Message from Ken Lay	bert meyers; Meyers
<input checked="" type="checkbox"/>		Albert Meyers	Theresa.Villeggiante@	John.Anderson@ENRO		Anderson; Anderson Jc
<input type="checkbox"/>	powerdaily.02.01.2002.pdf	Albert Meyers				Aquila; Davis; Edwards;
<input type="checkbox"/>	Ken Lay Resigns from Board	Albert Meyers	mbx_annceenron@ENR	DL-GA-all_enron_world	Ken Lay Resigns from t	
<input type="checkbox"/>	Organizational Announcement	Andrea Ring	mbx_OfficeChairman@	DL-GA-all_enron_world	Organizational Announ	Andrea; ENA; Greg Whi
<input type="checkbox"/>	Message from Ken Lay	Andrea Ring	mbx_OfficeChairman@	DL-GA-all_enron_world	Message from Ken Lay	bert meyers; Meyers
<input type="checkbox"/>	Ken Lay Resigns from Board	Andrea Ring	mbx_annceenron@ENR	DL-GA-all_enron_world	Ken Lay Resigns from t	kam keiser; Keiser; Ken
<input type="checkbox"/>	Holiday Party - Canceled	Andrea Ring	IMCEANOTES-Ken+20L	IMCEANOTES-All+20Er	Holiday Party - Cancele	
<input type="checkbox"/>	Organizational Announcement	Andrew Lewis	mbx_OfficeChairman@	DL-GA-all_enron_world	Organizational Announ	Andrea; ENA; Greg Whi
<input type="checkbox"/>	Message from Ken Lay	Andrew Lewis	mbx_OfficeChairman@	DL-GA-all_enron_world	Message from Ken Lay	bert meyers; Meyers
<input type="checkbox"/>	Ken Lay Resigns from Board	Andrew Lewis	mbx_annceenron@ENR	DL-GA-all_enron_world	Ken Lay Resigns from t	
<input type="checkbox"/>	Ken Lay and Jeff Skillin on CNNfn	Andrew Lewis	Public Relations	All Enron Houston	Ken Lay and Jeff Skillin	

Page Size: 100 1 of 35



Aa Saved Searches Save Query X Clear ?

View as: Facets Document list

ken lay SEARCH

Fields Tags Review Status

Documents: 3636 / Families: 3486 35 document(s) selected Reset

Sort by: Default

	Name	Custodian	Email From	Email To	Email Subj
<input type="checkbox"/>	Organizational Announcement	Albert Meyers	mbx_OfficeChairman@	DL-GA-all_enron_world	Organizati
<input type="checkbox"/>	Message from Ken Lay	Albert Meyers	mbx_OfficeChairman@	DL-GA-all_enron_world	Message fr
<input type="checkbox"/>		Albert Meyers	Theresa.Villeggiante@l	John Anderson@ENRO	
<input type="checkbox"/>	powerdaily.02.01.2002.pdf	Albert Meyers			
<input type="checkbox"/>	Ken Lay Resigns from Board	Albert Meyers	mbx_anncenron@ENR	DL-GA-all_enron_world	Ken Lay Res
<input type="checkbox"/>	Organizational Announcement	Andrea Ring	mbx_OfficeChairman@	DL-GA-all_enron_world	Organizati
<input type="checkbox"/>	Message from Ken Lay	Andrea Ring	mbx_OfficeChairman@	DL-GA-all	
<input type="checkbox"/>	Ken Lay Resigns from Board	Andrea Ring	mbx_anncenron@ENR	DL-GA-all	
<input type="checkbox"/>	Holiday Party - Canceled	Andrea Ring	IMCEANOTES-Ken+20L	IMCEANOTES-All+20Er	Holiday Par
<input type="checkbox"/>	Organizational Announcement	Andrew Lewis	mbx_OfficeChairman@	DL-GA-all_enron_world	Organizati
<input type="checkbox"/>	Message from Ken Lay	Andrew Lewis	mbx_OfficeChairman@	DL-GA-all_enron_world	Message from Ken Lay bert meyers; Meyers
<input type="checkbox"/>	Ken Lay Resigns from Board	Andrew Lewis	mbx_anncenron@ENR	DL-GA-all_enron_world	Ken Lay Resigns from f
<input type="checkbox"/>	Ken Lay and Jeff Skillin on CNNfn	Andrew Lewis	Public Relations	All Enron Houston	Ken Lay and Jeff Skillin

- Columns
- List Options
- Views
- Bulk Tagging
- Review Batches
- Batch Downloads
- Productions
- Sampling
- Reports
- Open Connected View
- Delete Selection

- Create for Selection
- Show All Samples

ZY
LAB

EN

Matter DOJ vs Enron v2 (Open another)

johannes.scholtes@zylab.com (Logout)

Sampling Wizard ?

1 Sampling

2 Preview

3 Apply

Sample name *

Random QC Sample

16 / 400

Sample description

Sample to check quality of automatic redaction

46 / 1000

Sampling Size:

10

%

4 of 35 documents

Sampling split

☒ Documents

☐ Families

Cancel

Next

✓ Sampling

2 Preview

3 Apply

✉ FW: Yahoo! Finance Story - Yahoo - Enron debt sinks on cre... 3 NOT REVIEWED

Abort

1 of 4

Incorrect

Correct

Conversation (2) ? Yahoo! Fina... ^

✉ [Unknown message]

✉ R Mark.Smith@ENRON.com TO

Image View

Produced View

Search Hits 0 < > Disable

Keywords 3 ▲ < > Disable

Document Info ?



Properties ? ^

Filter

Accessed April 26, 2018 at
8:34:15 AM
GMT+0

Analytics_State No

Assisted Review Travel
Training Batches Confirmation -
Travel
Confirmation

Bates Numbers [0000000046 -
0000000046]

☒ Hide empty properties

To: Black Troy [E-mail]; Breslau Craig [E-mail]
Cc: Zipper Andy [E-mail]; May Larry [E-mail]
From: Smith Mark
Sent: Fri 10/26/2001 11:06:39 PM
Importance: Low
Subject: FW: Yahoo! Finance Story - Yahoo - Enron debt sinks on credit crunch fears
MAIL_RECEIVED: Fri 10/26/2001 11:06:39 PM

^
-----Original Message-----
From: Yahoo! Finance [E-mail]
Sent: Friday, October 26, 2001 3:05 PM
To: Smith, Mark
Subject: Yahoo! Finance Story - Yahoo - Enron debt sinks on credit crunch fears

Mark Smith ([E-mail]) has sent you a news article

Personal message:

Yahoo - Enron debt sinks on credit crunch fears
<http://biz.yahoo.com/rf/011026/n26531689_1.html>

Yahoo! Finance <<http://finance.yahoo.com/?u>> Finance Home <<http://finance.yahoo.com>>
<<http://help.yahoo.com/help/fin/fnews>>

Reuters <<http://www.reuters.com>>

75% + RESET < Page: 1 of 1 >



Sampling

Statistic

Correct 0.00% (0 documents)

Incorrect 0.00% (0 documents)

To Do 100.00% (4 documents)

Best Case 100.00%
Estimate

▶ Reviewer Remarks

▶ Actions

▶ Productions



Sampling

2 Preview

3 Only

NYC Article.docx NOT REVIEWED

Properties ?

Filter

Accessed June 25, 2018 at 11:34:22 AM GMT+0

Analytics_State No

Contains Redaction Yes

Created June 25, 2018 at 11:34:22 AM GMT+0

Custodian John Doe

Deletion Status None

DocProperty_Authors Brenda Dodd

DocProperty_Source Microsoft Office Word

Document Author Brenda Dodd

Document Date June 25, 2018 at 11:34:35 AM GMT+0

☒ Hide empty properties

Image View

Produced View

Search Hits 0 < > Disable

Keywords 0 < > Disable

Abort

4 of 4

Incorrect

Correct

Document Info ?

Sampling

Statistic

Correct 50.00% (2 documents)

Incorrect 25.00% (1 documents)

To Do 25.00% (1 documents)

Best Case Estimate 75.00%

Reviewer Remarks

Actions

How running an \$837,000 office pool destroyed this man's life

Updated Mar 18, 2015; Posted Mar 16, 2015
2.9k shares

By [Steve Politi](#)

E-mail

Columnist, NJ Advance Media for NJ.com

They were everywhere, 11 cops in full body armor, tearing apart his modest three-bedroom condo. And all John Boveri could do was watch.

They knocked down a closet shelf. They ripped down a wall inside the laundry room. They even used a box cutter to shred a bedroom pillow. Finally, when the search didn't produce the evidence they expected, one of them demanded answers.

"Where is the cash?" he asked Boveri, who said he was handcuffed and seated in a recliner in his living room. "Where are the betting slips?"

Boveri tried, once again, to explain. He was not a bookmaker. He was not hiding anything. He was a football pool operator who was so open about his business, and so convinced he was doing it legally, that he listed his home address in Parlin on the pool's website.

But this wasn't any ordinary office pool. This one, he told investigators and NJ Advance Media, was worth \$837,000 in 2009. This one had 8,000 entries from around the globe -- men and women sending in \$100 for a chance to win that huge pot. The players included prominent sports broadcasters, New Jersey state troopers, dozens of lawyers and -- according to Boveri -- the agents for Tiger Woods and other PGA Tour golfers.

It had grown from a few dozen Wall-Street-brokers in 1996 to a potential life-changer for whomever could string together a season's worth of NFL winners. But on that September day in 2010, as the cops from the Monmouth County Prosecutor's Office carted box after box from his

Page 1 of 2

ZY LAB] Q&A

**ZY
LAB** Thank you