

Jak to się robi w Ameryce, czyli

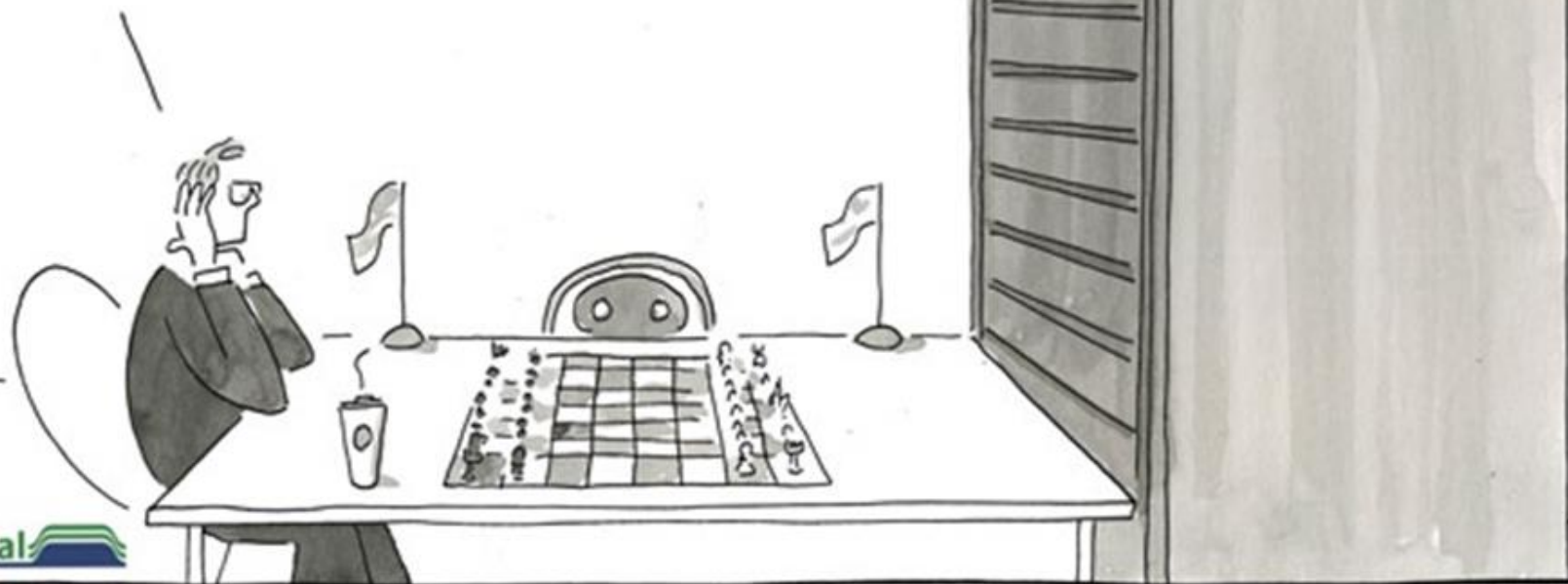
Wykorzystanie technologii w eDiscovery

Julia Barg

MAN VERSUS MACHINE

NOW THAT I HAVE COFFEE,
I'M READY TO CHALLENGE
YOU TO REVIEW 10
TERABYTES OF EMAIL

CHECKMATE



Postępowanie dowodowe w USA

- Strony postępowania odpowiedzialne za zebranie całego materiału dowodowego
- *Discovery* – etap postępowania sądowego lub administracyjnego, po wniesieniu pozwu, w którym każda ze stron może uzyskać dowody od strony przeciwnej, współ-strony, lub stron trzecich za pomocą dowodowych instytucji prawnych
- Wdrożenie *legal hold*
 - Identyfikacja potencjalnie istotnych danych
 - Zabezpieczenie dokumentów (elektronicznych i papierowych)
 - Zebranie zabezpieczonych danych i dokumentów

Procesowanie i przechowywanie danych

▪ Procesowanie

- Przekształcenie danych na jednolity format
- *de-NISTing*
- De-duplikacja dokumentów [*de-duping*]
- Wstępne filtrowanie danych [*pre-filtering*]

▪ Przechowywanie i wyszukiwanie

- 2 rodzaje baz danych w ramach 1 platformy
- dtSearch
- Raport wyszukiwanych haseł [*STRs*] + *persistent highlight set*
- Wyszukiwanie dokumentów

Analiza danych i e-dokumentów: analiza wspomagana technologią

▪ ***Sample-based learning***

- Powtarzalność treści dokumentów wykorzystana do grupowania i kodowania dokumentów
- Bazuje na małej grupie dokumentów zakodowanych przez prawnika
- Kategoryzacja wszystkich dokumentów w bazie danych

▪ ***Active learning***

- Narzędzie uczy się na podstawie decyzji recenzentów dotyczących kodowania
- Binarny system klasyfikacji (istotne vs. nieistotne, responsywne vs. nierespowsywne)
- Dwie metody:
 - *Prioritized review* – identyfikacja dokumentów najistotniejszych dla sprawy
 - *Coverage review* – podział dokumentów na dwie kategorie

Analiza danych i e-dokumentów:

Relativity Analytics

- **Structured** [strukturalna] – analizuje tekst dokumentów aby wskazać podobieństwa i różnice między dokumentami
 - *Email threading*
 - Identyfikacja tekstowa prawie duplikatów
 - Wykrywanie języków
 - Identyfikacja powtarzającej się treści
- **Conceptual** [konceptualna] – analizuje strukturę treści i pojęć w dokumentach, nawet jeśli zawierają różne frazy i słowa kluczowe
 - *Clustering*
 - Wyszukiwanie pojęć
 - Rozszerzenie słów kluczowych
 - Znajdowanie podobnych dokumentów

Analiza danych i e-dokumentów:



&

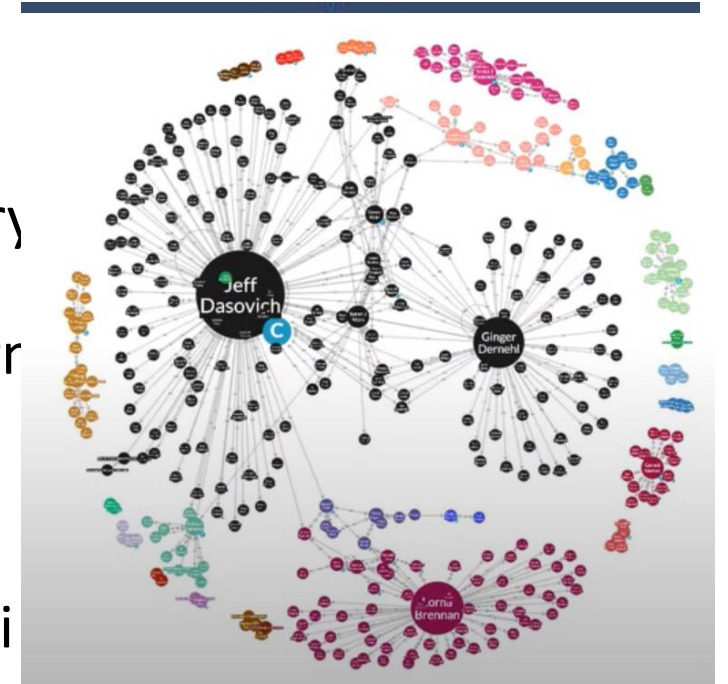


■ Brainspace

- oprogramowanie uczące komputer, inteligentnie wykrywa unikalne frazy w dużych zbiorach danych
- Interaktywna wizualizacja obrazuje uczenie się platformy dostosowującej się do toku myślenia użytkownika

■ NexLP - *story engine*

- Przetwarzanie języka naturalnego, analiza poznawcza i
- Znajduje połączenia między rozbieżnymi i nieustrukturyzowanymi dokumentami
- W połączeniu z danymi strukturalnymi tworzy wszechstronny obraz całego uniwersum danych



Prezentacja dowodów

- Metadane
- Tekst dokumentu
- Przywileje, czyli tajemnica zawodowa, ochrona danych osobowych i tajemnice przedsiębiorstwa
- *Redactions*
- Znakowanie dokumentów



HOW E-DISCOVERY WORKS



JUDGE



COUNSEL



IT