

Al to automatically detect sensitive documents

Using advanced machine learning techniques, Text IQ's Deep Concept Analysis (DCA) aids in the identification and review of sensitive documents.

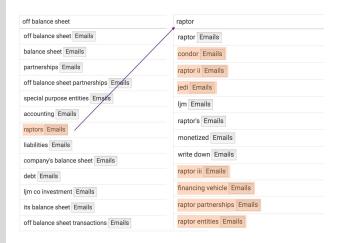
How it works

The engine identifies documents within a data set that contain words and phrases that are of *similar meaning* to known words and phrases. Customers utilize DCA to search for documents related to 1) the use of embarrassing language, 2) detection of code words, and/or 3) documents triggering privacy concerns (i.e. PII/PHI). After review, customers can export all relevant document IDs to flag and/or withhold from production.

All examples are pulled from Text IQ's Deep Concept Analysis running against the publicly available Enron data set

Offensive Language Detector Searching "idiot" returns the word "chicks" and identifies a highly offensive company email idiot Emails I was a stud in NYC but I pretty much get about on average phone numbers a night and at least hot chicks that say that they want to go home with me every night I goodt I love the buyside I have i'm really Emails bankers calling me everyday with opportunities and they pretty much cater to my every whim you know golfing events lavish dinners a night out clubbing-free guys I work with are also all chill! I live in the same app-full tubby Emails muttered Emails got Emails of Korea to work and when we go out What can I say live is season tix Emails CHUNG is KING of his domain here in Seoul chicks Emails dammit Emails come out and visit my ass ASAP I'll show you guys an unbelievable fw fw Emails time My contact info is below Oh by the way someone's gotta start fedexing me boxes of domes I brought out about but I think I'll run out of them by Saturday blonde Emails anyway Emails female news anchor who Emails Peter Chung The Carlyle Group huh Emails 0 nonchalantly Emails **©** his dad Emails **@** joke Emails **33** just kidding Emails 0 exclaimed Emails girlfriend Emails Ø redneck Emails

Code Word Detector



Searching for "off balance sheet" returns the code word "Raptor"; searching "Raptor" returns multiple other code words