

Beyond Precision and Recall – Ensuring 'Aboutness' in Topical Classification Using Confidence Scores

Haystack 2022 Conference April 27, 2022

Kimberly Hoffbauer, Lead News Taxonomist Sophie Lagace, Lead Legal Taxonomist Mark Shewhart, Principal Data Scientist

RELX Overview



RELX is a global provider of information-based analytics and decision tools for professional and business customers. RELX serves customers in more than 180 countries and has offices in about 40 countries. It employs more than 33,000 people over 40% of whom are in North America. www.relx.com



SCIENTIFIC, TECHNICAL & MEDICAL



LexisNexis Legal & Professional is a leading global provider of **legal**, **regulatory** and **business information** and **analytics** that help customers increase productivity, improve decision-making and outcomes, and advance the rule of law around the world.

- We help **lawyers** win cases, manage their work more efficiently, serve their clients better and grow their practices.
- We assist **corporations** in better understanding their markets, monitoring their brands and competition, and in mitigating business risk.
- We collaborate with **universities** to educate students, and we support nation-building with governments and courts by making laws accessible and strengthening legal infrastructures.

We combine information, analytics, and technology to help our customers reach essential insights, make more informed decisions, and achieve better outcomes.



Markets

Law Firms

Corporate

Fed/State/Local Government Law Schools and Universities

Users

Lawyers

Librarians

Media Professionals

Students & Academics

Content

- Cases
- Statutory Codes
- Legislation
- Dockets
- Legal Guidance & Forms
- News
- Company & Financial Info
- Patents/IP
- Public Records
- People Records

*Not a comprehensive list



Confidence Score – A Crucial Requirement

Importance of Taxonomies and Classification

- 1. Taxonomies are a huge part of our product offering: used in pre-filters, post-filters, analytics, and potentially to improve search relevancy.
- 2. Based on customer feedback, we know our users love using classification.

Automated Classification

- 1. Automated classification is critical given the scope of our taxonomies and the volume of our content.
- 2. Our classification is based on automated processes, which results in a wide array of relevancy in results; some very relevant and some more marginal.
- 3. A confidence score mechanism for placing the best results at the top and the more marginal results at the bottom increases the efficacy of our taxonomies.
- 4. Having a confidence score within the data and leveraging that score in search is CRUCIAL to having the very best Taxonomy product.



Confidence Score – A Crucial Requirement

- 1. Our vision: Consistent Confidence Scores Everywhere, Every Topic
- Every document-level topic that gets assigned shall have a score that predicts "aboutness" (aboutness = breadth and depth of discussion).
- 3. It doesn't matter what technology assigns a topic they shall all assign a uniform scale with consistent meanings in 4 defined score range buckets.
 - Very strong, strong, passing, marginal/weak
- 4. Global Benefits of Vision
 - Product functionality and search relevance improvements
 - Search engine boosting algorithms across content types and topic types can be more consistent with uniform scores and score range bucket meanings.
 - Search filtering by minimum score can likewise be consistent.
 - Analytic products can leverage, dynamically customizing views.
- 5. Opens doors for future vision.



Lexis Advance - Topic pre-filter and post-filter

Lexis Advance® Quicklaw® Advanced Search Searc					
l● CA ~ hearsay	,		Civil Evidence 🗸 📿		
Narrow By: Civi	I Evidence $ imes$		Clear \precsim X		
Jurisdiction	Aboriginal Law	Information Technology			
	Administrative Law	Insurance Law			
Content Type	Alternative Dispute Resolution	Intellectual Property Law			
	Banking and Finance Law	International Law			
Legal Topics	Bankruptcy and Insolvency Law	Labour Law and Arbitration			
	Civil Evidence	Landlord and Tenant Law			
Recent &	Civil Procedure	Legal Profession			



Lexis Advance - Topic Browse

C Lexis Advance Quicklaw	😚 Browse			English Françai	S	Client: -None- History
Browse						×
Sources >	Canada 🛛	Civil Evidence	> ^	Burden and standard of proof	> ^	Actions for Hearsay rule
Topics >	👯 Australia	Civil Procedure	>	Constitutional issues	>	Get documents
Find a topic Q	France 🥝	Commercial Law	>	Documentary evidence	>	∇ Add topic as a search filter
		Conflict of Laws	>	Exclusionary rules	>	🗘 Create a topic alert
	Hong Kong	Constitutional Law	>	Hearsay rule	>	
	New Zealand	Construction Law	>	Methods of proof	>	
	🗮 United Kingdom 🕜	Contracts	>	Opinion evidence	>	
	United States of	Corporations, Partnership	s >	Parol evidence rule	>	
	America 7	Creditors and Debtors Law	~	Physical evidence	>	
			v /	Privilege	>	
		Criminal Law	>	Scientific evidence	>	
		Damages	>	Weight	>	
		Education Law	>	Witnesses	>	

Nexis SmartIndexing Pre-Filters

Topic Look Up				Browse Subject Terms		
					Enter Search Term	
Index Terms (Topics, Tags)				15 Results for cryptocurrency		
ſ					- 🗌 Business News	
					- 🗌 Economy & Economic Indicators	
			~		- 🗌 Digital Economy	
Browse:	Subject	Industry	Company	People	- 🗌 Electronic Commerce	
					- 🗌 Digital Currency	
					- Cryptocurrency	
					- Cryptocurrency Crimes	
					- 🗌 Financial Technology	
					- 🗌 Digital Currency	
					- Cryptocurrency	
					- 🗌 Cryptocurrency Crimes	

Word Wheel

Index Terms (Topics, Tags)

Cryptology

C LexisNexis

Word Wheel

\$SHOOT Token Seed Sale Continues to Make Wave, IDO & Staking Platform Demo Released

 13 Apr 2022
 MENAFN - Press Releases (English)

 416 words
 1 hit

1irstcoin (FST) Trading 17.1% Lower Over Last Week

13 Apr 2022 Newstex Blogs 695 words 1 hit

View list of similar documents (1)

88mph Price Hits \$19.00 on Exchanges (MPH)

13 Apr 2022 Newstex Blogs 766 words 1 hit

Nexis SmartIndexing Post-Filter

... company's proprietary polymer resin which ensures that etherium® By E-Stone countertops are not only beautiful, but also heat, stain, scratch, mold and mildew-resistant. They can be installed in as little as one day and their beauty will last a lifetime...

Subject

Business News

Company Activities & Management

Economy & Economic Indicato

Government & Public Administration

International Relations & National Security

✓ More

Select multiple

Industry

Banking & Finance Computing & Information Technology Manufacturing Media & Telecommunications

Blockchain Platform As A Service/Low-Code Web3 Dapp NFT Production Tool Updated

 30 Mar 2022
 Newstex Blogs
 Interwire

 708 words
 2 hits

... has been used to develop use-case applications to full production in weeks. The platform works with a variety of blockchains, including Etherium, Enterprise Etherium, Polygon, and Binance Smart Chain. For more information about SettleMint's blockchain...

Introducing Attidude, a collection of cool and unique men proudly living on Etherium blockchain

08 Mar 2022 ENP Newswire

712 words 2 hits

... Introducing Attidude, a collection of cool and unique men proudly living on Etherium blockchain... ... Attidude, a collection of cool and unique men proudly living on Etherium blockchain and hosted on IPFS is set to make its debut in the exciting world of NFTs with its pre-sale date set for 03/22/22 at 9:22 am PDT (or

Advantages of Searching Using Classification

- 1. To easily access results on a given topic
- 2. To add context
 - Hearsay + Topic Civil evidence
- 3. To search for a concept that has many synonyms
 - Topic Driving while intoxicated: also called offences relating to conveyances, impaired driving, DWI, DUI, driving under the influence, drunk driving, etc.
- 4. To improve results when search terms are really common or to disambiguate
 - Topic Negligence (in Tort Law)
 - Topic Competition (in Commercial Law)
 - Topic Will Smith (the actor vs anybody named Will Smith)
 - Topic Apple (the company vs the fruit)
 - Topic Georgia (the country vs the state vs a name)
- 5. To simplify the search by replacing one of its components with a topic:
 - Disciplinary sanction for inappropriate behavior outside of work = Topic Off-duty conduct under Employee Discipline in Employment Law

Legal Topics at LexisNexis – Automated Classification

Machine learning models

-1.7472056953544384e-06 doc length + -0.00325418079297124 * total ops + 0.0022553615191532794 * true ops + 0.022859433969558137 * num unique + -0.015421547337865575 * el num unique + -0.0025221382157179307 * freg count + 0.004099394857401066 * el_freq_count + 0.003957939296 **Boolean rules** 1.6066056927268358e-05 * num_pos + -9.0789039473 0.00013543918274854834 * min_pos + -6.507109814

(Legal)

Damages for psychological injuries

((damages or damage award! or damages award! pre/1 for pre/2 psychological or mental or emotional or nervous or post-traumatic stress or post traumatic stress pre/1 injur! or distress or shock or disorder or upset or trauma or suffering)) OR (caps(Action) pre/1 for damages for pre/1 psychological injur! or emotional injur! or mental distress or emotional distress or nervous shock or post-traumatic stress disorder or post traumatic stress disorder or mental suffering) OR ((((psychological! or mental! or emotional! or nervous or post-traumatic or post traumatic w/3 injury or injuries or injured or injuring or distress or shock or disorder or upset or trauma or suffering) w/7 damages or damage award!)) and not (lpin or bpin or (caps(Physical) pre/1 "and psychological injuries" pre/2 caps(Physical)))) OR (((cognitive impairment! or depression or (change! w/3 personality) or neurosis or neuroses or phobia! or ptsd or anxiety) w/7 damages or damage award!)) OR (damages pre/5 caps(Emotional) pre/2 mental distress)

Definition

Name: CRYPTOCURRENCY

TERM01 = bitcoin w/4 coinTERM01 = bitcoin w/4 coinsTERM01 = bitcoin w/4 cryptoTERM01 = bitcoin w/4 cryptosTERM01 = bitcoin w/4 currenciesTERM01 = bitcoin w/4 currencyTERM01 = bitcoin w/4 dogecoinTERM01 = bitcoin w/4 ethereum TERM01 = bitcoin w/4 litecoinTERM01 = bitcoin w/4 moneroTERM01 = bitcoin w/4 rippleTERM01 = bitcoin w/4 tether TERM01 = bitcoin w/4 xrpTERM01 = crypto-based currenc)) TERM01 = crypto based currenc)) TERM01 = crypto currenc)) TERM01 = crypto-currenc)) TERM01 = crypto digital currenc)) TERM01 = crypto virtual currenc)) TERM01 = cryptocurrenc)) TERM01 = cryptographic (rrenc)) TERM01 = cryptograph (rrenc)

> SmartIndexing rules (News)

Legal Topics at LexisNexis – Measuring Precision

STF tool – 4-point scale

4 = very strong

- 3 = strong
- 2 = passing
- 1 = weak/bad hit

- Major reference precision: percentage of very strong/strong results (ratings 4 and 3) within topic results
- Minor reference precision: percentage of very strong/strong results <u>and</u> passing (ratings of 4,3,2) within topic results

Major = 70% Minor = 80%

Legal Topics at LexisNexis – Automated classification

Confidence Score to Predict Aboutness

• On the topic 'Damages for psychological injuries', the 2 following results have very different relevance

The Law about Compensable Damages for Psychological Injury 40 A review of the case law and the legal literature reveals that a variety of different names have been used to describe mental suffering, which has variously been called nervous shock, mental anguish, mental suffering, emotional trauma , emotional distress, and emotional injury. In this case, the Plaintiffs' witness Dr. Robert Maunder, a psychiatrist, whose evidence I will discuss later, testified about what he described as the " psychological injury	 f The Plaintiff also claims damages for emotional distress. There is no evidence the Defendant's unnecessary replacement of the ball joints, and inadequate repair of the brakes caused the Plaintiff's any emotional distress and that claim is dismissed. 17 Accordingly, there will be judgment in favour of the Plaintiff for the sum of \$1,300.00, together with prejudgment interest in accordance with the Courts of Justice Act. 18 Coursel may make an appearament with the trial second pater to argue
 of the Uninfected Persons. 41 On the motions for summary judgment, whether the psychological injuries as described by Dr. Maunder or as described by the Plaintiffs' witnesses from their own experiences amounted in law to compensable damages was a major point of contention between the parties. Before describing the evidentiary background and reviewing the evidence, it is help and I believe necessary to discuss first the state of the law in Ontario about the providence of the providence of the law in Ontario about the providence of the providence of	- <classify:classitem> -<meta/> <metaitem name="appliedby" value="boolean"></metaitem> <metaitem name="intitle" value="false"></metaitem> <metaitem name="score" value="59"></metaitem> <metaitem name="InternalID" value="LL000CWBJ"></metaitem> <metaitem name="InternalName" value="Psychological injuries
DAM750"></metaitem></classify:classitem>
the nature of compensable damages for psychological injury. This issue is subset of the larger issue of the laws' treatment of what often is called ner - <classify:classitem> -<meta/> <metaitem name="appliedby" value="boolean"></metaitem> <metaitem name="intitle" value="false"></metaitem> <metaitem name="score" value="89"></metaitem> <metaitem name="InternalID" value="LL000CWBJ"></metaitem> <metaitem name="InternalID" value="Psychological is </pre></classify:classitem>	 Confidence score is added to topic in document metadata and predicts how relevant the document is to the topic

Confidence Score - Output

- Score from 50 to 99 2 categories
 - 50-84: Minor references
 - 85-99: Major references
- Ability to force higher confidence scores in certain scenarios

Implementing Confidence Score

- 1. Classification develop mechanism to produce score and stamp it in the data
- 2. Implementation in search
 - a) Major reference filtering in search
 - b) Boost documents with highest scores in results list
 - c) Major reference filtering in topic alerts

Boolean Based Confidence Score

Boolean rules used to develop machine learning model

Damages for psychological injuries

((damages or damage award! or damages award! pre/1 for pre/2 psychological or mental or emotional or nervous or post-traumatic stress or post traumatic stress pre/1 injur! or distress or shock or disorder or upset or trauma or suffering)) OR (caps(Action) pre/1 for damages for pre/1 psychological injur! or emotional injur! or mental distress or emotional distress or nervous shock or post-traumatic stress disorder or post traumatic stress disorder or mental suffering) OR ((((psychological! or mental! or emotional! or nervous or post-traumatic or post traumatic w/3 injury or injuries or injured or injuring or distress or shock or disorder or upset or trauma or suffering) w/7 damages or damage award!)) and not (lpin or bpin or (caps(Physical) pre/1 "and psychological injuries" pre/2 caps(Physical))) OR (((cognitive impairment! or depression or (change! w/3 personality) or neurosis or neuroses or phobia! or ptsd or anxiety) w/7 damages or damage award!)) OR (damages pre/5 caps(Emotional) pre/2 mental distress) OR (caps(Damage) pre/1 awards pre/5 caps(Psychological or Mental) pre/1 injuries or shock or effects) OR (caps(Damages) pre/10 caps(Psychological or Mental or Emotional) pre/1 injuries or distress) OR (caps(Damages) pre/10 caps(Psychological injuries" pre/25 caps(Psychological) pre/1 injuries) OR (caps(Physical) pre/1 "and psychological injuries" pre/25 caps(Psychological) pre/1 injuries) OR (caps(Physical) pre/1 "and psychological injuries" pre/25 caps(Psychological) pre/1 injuries) OR (caps(Physical) pre/1 "and psychological injuries" pre/10 caps(Psychological) pre/1 injuries)

Multiple Search engine features

Matched Term frequency & Variety (incl. freq limits)

Term Frequency / document length; term density

Location-specific hits in the document

Etc.

Boolean Based Confidence Score - Accuracy Metrics

- Confidence score good at identifying very strong results and weak results:
 - **Scores of 92 and up:** 85% were indeed major references
 - **<u>Scores of 50-59</u>**: 79% were indeed more marginal results
- Applying confidence score in search and boosting high scores translates into moving the strongest references to the top and the weakest references to the bottom
- Significant impact in relevance of top results for topic only searches.
- For searches where a topic is added as a filter to user keywords, weight attributed to topic score will need to be balanced with user keywords
- Applying in search has no negative consequences we are not changing the number of results, just the <u>order in which they appear</u>

Major References (Scores 85-99)

- The SmartIndexing rules (concept definitions) contain a threshold plus terms and phrases in different word concept buckets.
- The buckets have different weights and frequency limits.
- The frequency limits indicate the maximum amount a word concept bucket can contribute to the overall score that is compared to the threshold.
- The threshold is the cumulative score across word concept buckets that needs to be met to assign the topic as a MAJOR REFERENCE.

SmartIndexing Technology Confidence Scores

Passing and more marginal references (Scores 50-84)

- Each word concept bucket can only contribute up to its frequency limit.
- The frequency limits on the word concept buckets tell us something about how ambiguous the terms in that bucket are with respect to the topic.
- Terms in buckets whose frequency limits are less than the threshold cannot by themselves, no matter how frequent, cause the threshold to be met.
- Such word concept buckets require other word concepts present to meet the threshold – this is a disambiguation effect.

- Sum the frequency limits of all word concept buckets for which <u>you found</u> <u>terms</u> in the document.
- If that sum meets the threshold, then you have a PASSING REFERENCE assuming it is not already a MAJOR REFERENCE

SmartIndexing Confidence Score – News & Business

The Federalist Society for Law and Public Policy Studies holds a virtual discussion, beginning at noon, on "stablecoins," focusing on cryptocurrency.

The Washington Daybook

April 14, 2022

Copyright 2022 Federal Information and News Dispatch Inc. All Rights Reserved

Section: DISCUSSION; ||CRYPTOCURRENCY/OUTLOOK|| Finance

Length: 66 words

Body

TIME: Advisory

PARTICIPANTS: <u>Paul Jossey</u> →, principal attorney at <u>Jossey PLLC</u> → and adjunct fellow at the Competitive Enterprise Institute; former Commodity Futures Trading Commission Chairman Timothy Massad, research fellow at Harvard University's Kennedy School of Government; and John Berlau, senior fellow and director of finance policy at the Competitive Enterprise Institute

SmartIndexing Confidence Score – News & Business

The Federalist Society for Law and Public Policy Studies holds a virtual discussion, beginning at noon, on "stablecoins," focusing on cryptocurrency.

The Washington Daybook

Classification

Language: ENGLISH

Publication-Type: Newswire

Journal Code: CNG

Subject: CRYPTOCURRENCY (90%); DIGITAL CURRENCY (90%); PUBLIC POLICY (90%); BANKING & FINANCE REGULATION (88%); FUTURES (88%); CALENDARS (86%); SECURITIES TRADING (72%)

Organization: FEDERALIST SOCIETY FOR LAW & PUBLIC POLICY STUDIES (84%); HARVARD UNIVERSITY (83%); COMMODITY FUTURES TRADING COMMISSION (83%)

Industry: CRYPTOCURRENCY (90%); DIGITAL CURRENCY (90%); BANKING & FINANCE REGULATION (88%); FUTURES (88%); SECURITIES TRADING (72%)

Load-Date: April 6, 2022

Index Terms (Topics, Tags)

Enter a topic or industry term

Browse: Subject Industry Company People Geography Relevancy: Major Terms Only

Confidence Score Boosting Algorithm Development

- Algorithm to boost results with higher scores
- For following topic search scenarios:
 - Topic filter + user keywords
 - Topic only search
- Impact on *topic only* searches should show significant relevance improvements

 only lever is boost on score:
 - Increase number of strong and very strong documents at the top
- Impact on topic + user keywords will be more nuanced user keywords come into play
- Various considerations:
 - Weight to give to score compared to other levers
 - Weight to give to score compared to user keywords
 - Higher boost for higher scores
- Measuring relevance gains via Human Relevance Testing, for 2 search scenarios above

Results for Topic Criminal Injury Compensation

	title Vithout boost		Show / Hide Answer	
1.	<u>R. v. Eze, [2022] O.J. No. 379 [search fra</u> Score: 17337 <u>Show / Hide Entry</u>		title V With Boost	<u>Show / Hide Answer</u>
2.	[2022] O.J. No. 379 2022 ONSC 277 <u>R. v. Northwest Territories (Department o</u> Pricing: subscription Price: Not In Re	1.	Skerget and Criminal Injuries Compensation Board et al., 16 O.R. Pricing: subscription Price: Not In Response Score: 1337	. <u>(2d) 447 [search fragment] [xpath]</u> how / Hide Entry
3.	[2014] N.W.T.J. No. 54 2014 CSHG pa Her Majesty the Queen v. Zebedee et al. [xpath] Pricing: subscription Price: No	2.	Sweet v. The Attorney General of Ontario et al. [Indexed as: Swe (3d) 420 [search fragment] [xpath] Pricing: subscription Price Hide Entry	et v. Ontario (Attorney General)], 106 O.R. : Not In Response Score: 1312 Show /
4.	<u>81 O.R. (3d) 583</u> <u>[2006] O.J. No. 2624</u> <u>R. v. Fairweather, [2014] O.J. No. 3012 [s</u> Response Score: 16583 <u>Show / Hide</u>	3.	<u>106 O.R. (3d) 420</u> <u>[2011] O.J. No. 2144</u> 2011 ONSC 2650 111 2011 CarswellOnt 3161 <u>Masakeyash, on behalf of Skunk v. Criminal Injuries Compensation</u> <u>Injuries Compensation Board</u>], <u>80 O.R. (3d) 467</u> [<u>search fragmer</u> In Response Score: 1303 Show / Hide Entry	335 D.L.R. (4th) 249 278 O.A.C. on Board [Indexed as: Masakeyash v. Criminal at] [xpath] Pricing: subscription Price: Not
5.	[2014] O.J. No. 3012 2014 ONSC 363 R. v. Fisher, [2019] B.C.J. No. 106 [searc Response Score: 16583 Show / Hide	4.	80 O.R. (3d) 467 [2006] O.J. No. 1429 267 D.L.R. (4th) 258 Ahmed v. Criminal Injuries Compensation Board, 90 O.R. (3d) 47 subscription Price: Not In Response Score: 1290 Show / H	208 O.A.C. 364 2006 CanLII 11440 5 [<u>search fragment</u>] [<u>xpath</u>] Pricing: ide <u>Entry</u>
6.	[2019] B.C.J. No. 106 R. v. Adams, [2009] A.J. No. 266 [search Score: 15829 Show / Hide Entry	5.	<u>90 O.R. (3d) 475</u> <u>[2008] O.J. No. 2478</u> 238 O.A.C. 98 2007 <u>Fregeau and Criminal Injuries Compensation Board</u> , <u>[1973] 2 O.F</u> <u>subscription Price: Not In Response</u> Score: 1290 <u>Show / H</u>	3 CanLII 30297 R. <u>182</u> [<u>search fragment</u>] [<u>xpath</u>] Pricing: ide Entry
7.	[2009] A.J. No. 266 2009 ABQB 160 R. v. Bigcharles, [2015] B.C.J. No. 1487 [Response Score: 15829 Show / Hide	6.	[1973] 2 O.R. 182 [1973] O.J. No. 1888 33 D.L.R. (3d) 278 Manson and Criminal Injuries Compensation Board Indexed as: M Compensation Board) (Div. Ct.), 68 O.R. (2d) 222 [search fragme Not In Response Score: 1290 Show / Hide Entry	 <u>lanson v. Ontario (Criminal Injuries</u> ent] [xpath] Pricing: subscription Price:
8.	[2015] B.C.J. No. 1487 2015 BCSC 12 R. v. Carter, [1990] O.J. No. 3140 [search Response Score: 15829 Show / Hide	7.	<u>68 O.R. (2d) 222</u> <u>[1989] O.J. No. 90</u> 32 O.A.C. 236 Saez-Larrazabal v. <u>Criminal Injuries Compensation Board</u> Gonzale Compensation Board Saez v. Criminal Injuries Compensation Boa Board [Indexed as: Saez-Larrazabal v. Criminal Injuries Compensity	es-Delacaridad v. Criminal Injuries rd Ramos v. Criminal Injuries Compensation sation Board], 111 O.R. (3d) 205 [search
			fragment] [xpath] Pricing: subscription Price: Not In Respon	se Score: 1265 Show / Hide Entry

- Our vision: Consistent Confidence Scores Everywhere, Every Topic
- Automated classification will produce variety of results from very strong to more marginal
- Precision is important but so is recall
- Optimize accuracy via combination of indexing rules and tools
- Important to track and measure topic precision
- Confidence score is crucial to make sure documents with most 'aboutness' appear at the top of the list of results
- No loss of recall
- Enabler for future opportunities: related content via topics, increase use of topics in product functionality, analytics, etc.

Thank you! Merci!

Authors:

Kimberly Hoffbauer: <u>kimberly.hoffbauer@lexisnexis.com</u> Lead News & Business Taxonomist

Sophie Lagace: <u>sophie.lagace@lexisnexis.ca</u> Lead Legal Taxonomist, LexisNexis

Mark Shewhart: Mark.Shewhart.3@lexisnexis.com Principal Data Scientist, LexisNexis

