
# Legal Analytics:
# Concepts, Methods, and Markets

*Günther Schefbeck*

Applied Artificial Intelligence Conference 2019
Vienna, 27 May 2019

# Outline


- What is Legal Analytics?
- Back to the basics: concepts and methods
- The growing Legal Analytics market
- Examples for Legal Analytics services
- Examples for Legal Analytics projects in Europe
- Prospects
- Reading recommendations

# A prophet‘s voice


- “In a few years, lawyers will rely more and more on computers to perform many tasks for them. They will not rely on computers simply to do their bookkeeping, filing or other clerical tasks. They will also use them in their research and in the analysis and prediction of judicial decisions. In the latter tasks, they will make use of modern logic and the mathematical theory of probability, at least indirectly.”

# A prophet‘s voice

ManyLaws

- “In a few years, lawyers will rely more and more on computers to perform many tasks for them. They will not rely on computers simply to do their bookkeeping, filing or other clerical tasks. They will also use them in their research and in the analysis and prediction of judicial decisions. In the latter tasks, they will make use of modern logic and the mathematical theory of probability, at least indirectly.”
- *Reed C. Lawlor, What computers can do: analysis and prediction of judicial decisions, American Bar Association Journal 49 (1963), 337-344*

# „Big Data Analytics“

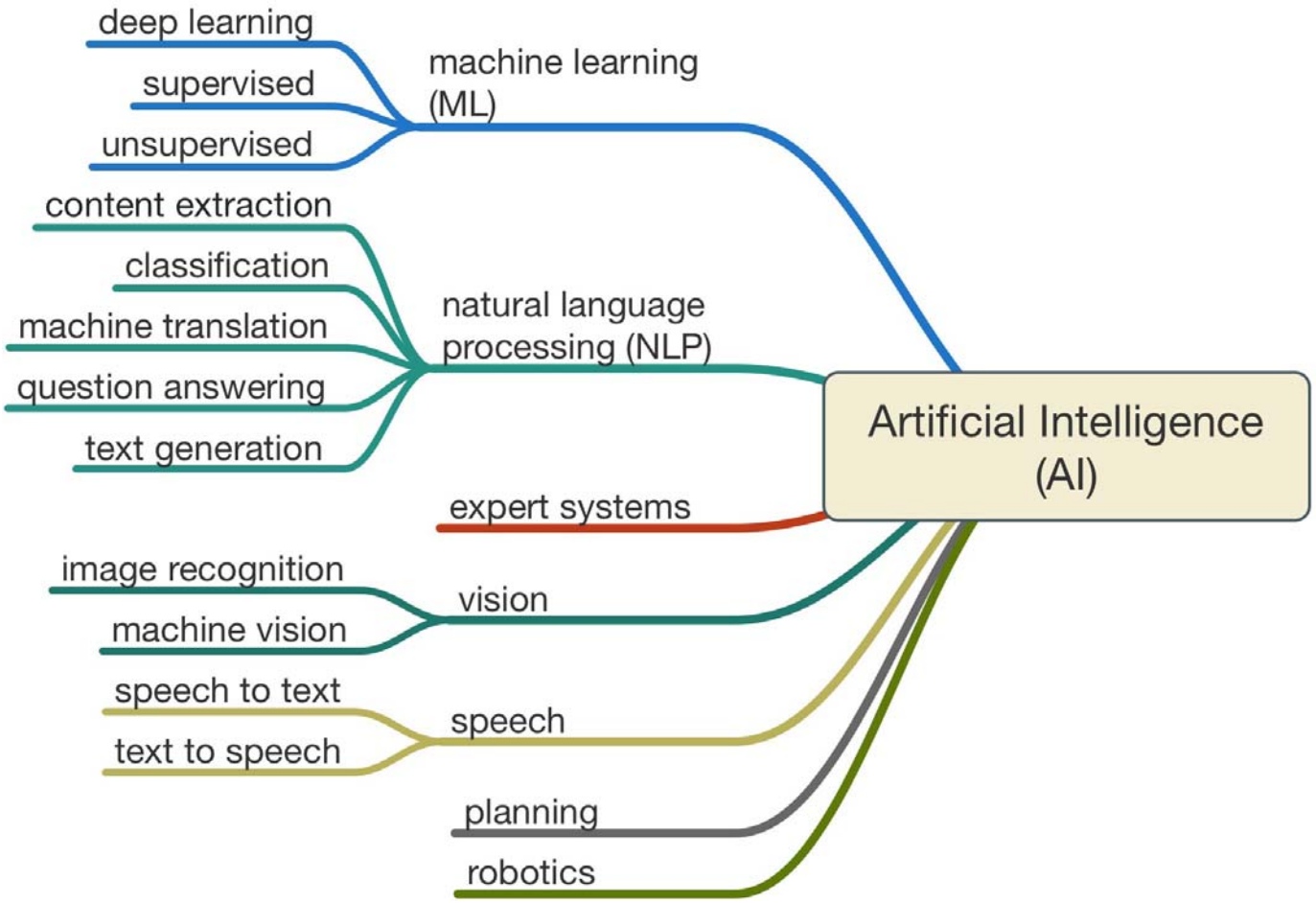

- “Big Data”: mass data (“3 v”: “volume, velocity, variety”)
  - not being kept in relational databases
  - sources: e.g., online behaviour, use of search engines, “Social Media”, electronic communication, use of networked devices (IoT), OGD
  - processed in parallel
- “Big Data Analytics”: machine learning tools used to recognize patterns in mass data and draw conclusions
- “Predictive Analytics”: predictions derived from mass data analyses
- Precondition: quick growth of computing power
- Quick technology change: e.g., MapReduce replaced by Spark
- Areas of application: e.g., product and service marketing, prevention and prosecution of crime and terrorism, prediction of epidemics, influencing elections, credit assessment, “Social Scoring”
- Epistemological implication: focus on correlation instead of causality

# „Legal Analytics“


- “Legal Analytics”: applying “Big Data Analytics” methods on legal mass data
- Term: used since 2013 by “LexMachina” (LexisNexis service)
- Issue: data significantly smaller than, e.g., in commercial or medical applications
- Advantage: data better structured than usual mass data, partly even semantically annotated (e.g., EUROVOC)
- Areas of application: e.g., legal research (e.g., implicit references, cases of precedence), recognition of decision patterns (administrative authorities, courts, judges), prediction of decisions, assistance in legal writing, classification, evaluation, or generation of legal documents (e.g., contracts)

# AI concepts and methods

ManyLaws




*(Michael Mills, 2016)*

# „Legal Analytics“ concepts and methods


- Legal information usually represented in texts (general and individual norms, accompanying information like explanatory notes or decision comments)
- Methods of Natural Language Processing (NLP), e.g.
  - “N-grams”: fragmentation of text into contiguous sequences of (n) items (e.g., words)
  - “Bag of Words”: representation of a document as a collection (multiset) of terms (words or n-grams) ignoring the sequential order of the terms in the document
  - “Word (and/or phrase) embeddings”: mapping of words/phrases to a continuous vector space allowing to calculate distances/similarities
  - “Topics”: clustering together semantically similar n-grams, suggesting that similar words appear in similar contexts
- Machine Learning (ML)
  - Unsupervised: inferring groupings of unlabeled instances based on their content
  - Supervised: inferring a classification model from labeled training data

# Legal Analytics services


- Legal-domain specific
  - Conducting legal research
  - E-discovery (e.g., cases of precedence, case-specific evidence)
  - Predicting litigation outcome/duration, comparing litigation strategies
  - Legal document creation and review
  - Comparing firm performance/experience
  - Self-help legal resources (e.g., chatbots)
- Generic
  - Administrative support
  - Performing due diligence
  - Compliance management
  - Cybersecurity, data privacy

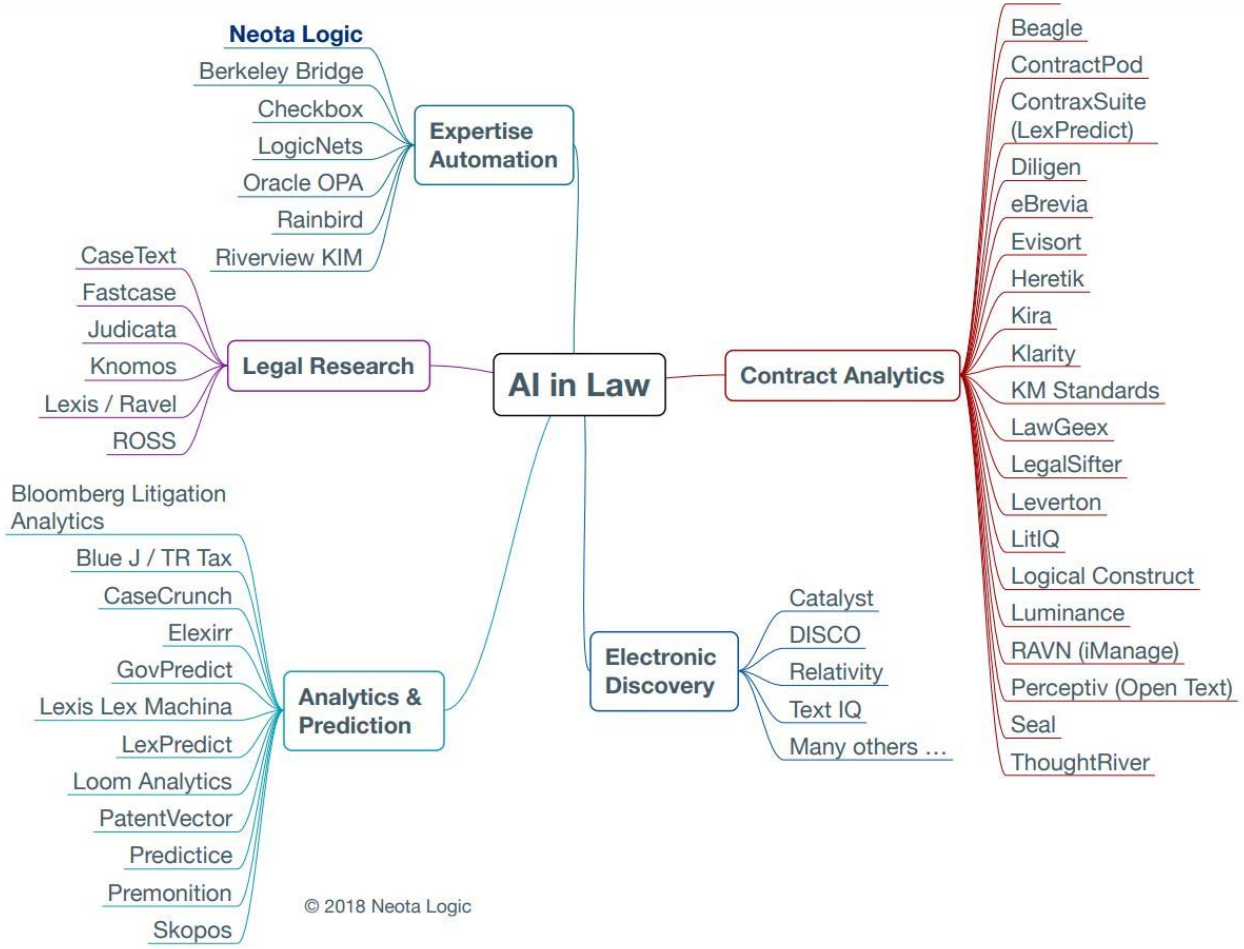
# Issues and opportunities

ManyLaws

- Legal analytics issues
  - Lack of regulatory framework
  - Liability issues
- New business opportunities for law firms
  - IP protection (AI patent litigation)
  - Law enforcement access to data
  - Violation of privacy by (AI) algorithms

# Legal Analytics service providers

ManyLaws

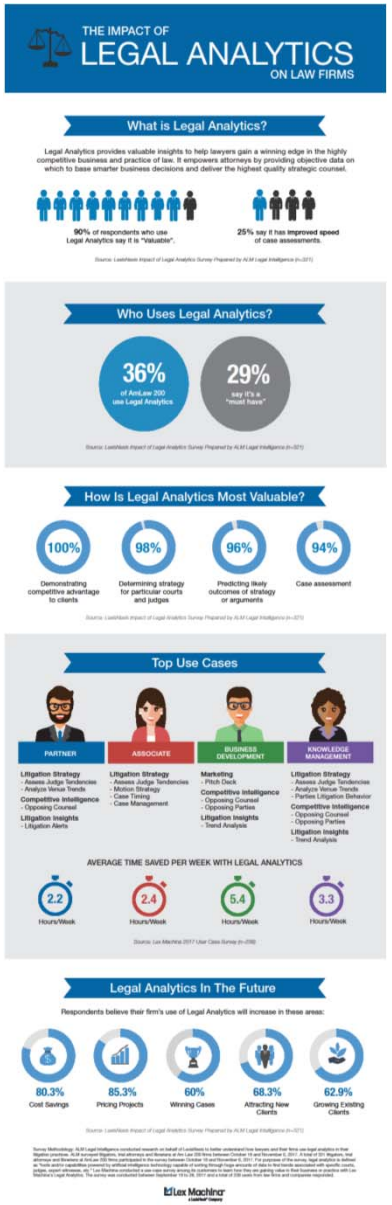


# Examples for Legal Analytics services

ManyLaws

- „The good lawyer knows the law, the better lawyer knows the judge“
  - Ravel Law packages: „Case Analytics“, „Court Analytics“, „Judge Analytics“
  - LexMachina Motion Kickstarter: evaluating granted vs. denied motions (per judge)
  - Premonition: drawing lawyer/judge matrix (to expose “which lawyers win the most before which judge”)
- „Heaven helps those who help themselves“
  - Neota Logic: providing fact- and context-specific answers to legal (e.g., compliance) questions
  - Seal Software: discovering and classifying all of a company‘s existing contracts
  - Argopoint: focusing on legal department management

# Legal Analytics impact




*(ALM Legal Intelligence Survey 2017)*

# Market effect of Legal Analytics

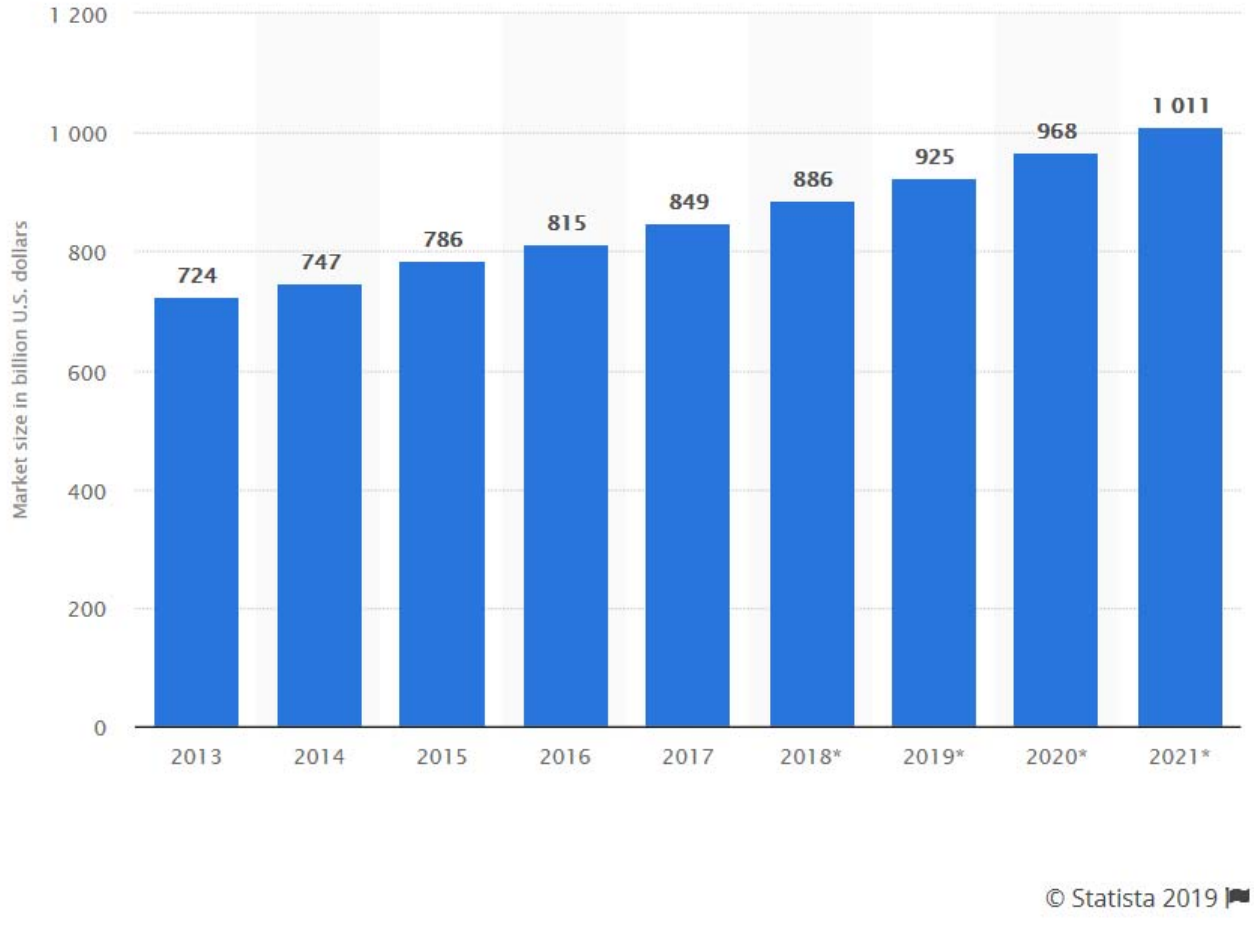
ManyLaws

| Market Problem | AI Solution | Lawyer Impact | Examples |
|---|---|---|---|
| **The volume costs of lawyer effort** | Machine-learning programs perform lawyer tasks faster and therefore much less expensively. | Law firms can no longer sell lawyer hours spent on these tasks, reducing inventory. | E-discovery, contract analysis, legal research |
| **The scarcity costs of lawyer expertise** | Expert systems tap into legal knowledge databases to answer questions and forecast outcomes. | Lawyers' expertise becomes less exclusive, increasing supply and lowering price. | Expert applications, outcome prediction |

*(Jordan Furlong, Law21, 2018)*

# Global legal services market

ManyLaws



# Legal Analytics market

ManyLaws

- Legal Analytics market expected to grow from 451 mio. USD in 2017 to 1859 mio. USD in 2022, at a compound annual growth rate of 32,7 % *(Marketsandmarkets report 2017)*
- Market dominated by „big players“ like LexisNexis (LexMachina) and Thomson Reuters (Westlaw Edge), but examples like Ravel Law show there are opportunities for start-ups
- Market dominated by U.S. based enterprises, but with growing European market new opportunities for European enterprises
- Example: 2021.AI (Denmark) offering, e.g., ML optimized legal search engine, case matching, prediction of claim acceptance and probabilty of winning cases, contract sorting

# European projects


- Academic proofs of concept, e.g., *Nikolaos Aletras et al., Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective, PeerJ Computer Science 2 (2016)*: analysis of ECHR decisions by applying NLP methods (n-grams and topics used to train a machine learning algorithm), achieved prediction accuracy 79 %
- National projects, e.g., LexDatafication (CIRSFID, University of Bologna): processing Italian legal data for semantic analysis by migrating them to AKN and annotating them
- European projects, e.g., Lynx: developing services for enterprises assisting them in complying with European and national norms

# Prospects

- Sarcasm indicating uncertainty about future of legal profession: e.g., „Siri, Esq.“, „Alexa, guilty or not guilty?“
- Will lawyers in large part be “replaced by advanced systems, or by less costly workers supported by technology or standard processes, or by lay people armed with online self-help tools” *(Richard Susskind, The End of Lawyers?, Oxford 2008, 2)*?
- Pioneering China:
  - First Internet Court established in Hangzhou in 2017
  - Primary competence for e-commerce cases
  - Digital proceedings
  - Three procedural modes: synchronous, diachronous, and autonomous (AI based)
  - 60 % reduction in time consumption, 50 % reduction in duration of proceedings

# Reading recommendations


- AI basics: *Viktor Mayer-Schönberger/Kenneth Cukier, Big Data: A Revolution That Will Transform How We Live, Work, and Think (London 2014)*
- AI risks: *Cathy O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (New York 2016)*
- Practical Legal Analytics: *Ed Walters (ed.), Data-Driven Law: Data Analytics and the New Legal Services (Boca Raton 2019)*
- Theoretical Legal Analytics: *Kevin D. Ashley, Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age (Cambridge 2017)*

ManyLaws

# Thank you for your attention!

More information:

*guenther.schefbeck@parlament.gv.at*