

A Brief History of Al & Law: Key Milestones for Legal IR & NLP

Jack G. Conrad
TR Labs -- Applied Research
45th European Conference on Information Retrieval
ECIR 2023
Thursday, 6 April 2023



Outline

Al & Law: A Brief History of Key Research & Milestones

- Background
- Foundations of the Research Community
- Overview of Field
 - Sketch of 35+ Years of Al and Law
 - The Role of Legal Startups
- From Past to Present
 - Some Comparisons
- A Look to the Future
 - What will next generation capabilities look like?



Al and Law – Foundations of the Research Community

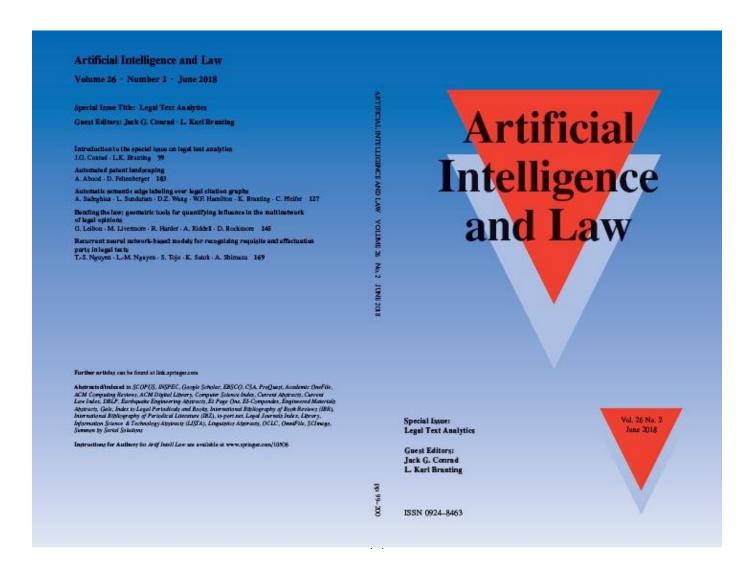
The AI and Law Journal

Established in 1992 Editors

- Kevin L. Ashley
- Trevor Bench-Capon
- Giovanni Sartor

Extensive Editorial Board with specialists covering a broad range of topical research areas.

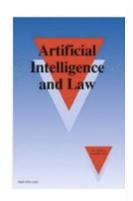
Published by Springer (NL)





The AI and Law Journal

Key Issue – Reviewing 30 years of research papers in the journal (divided into 3 decades)



Open Access | Published: 09 August 2022

Thirty years of artificial intelligence and law: the third decade

Serena Villata, Michal Araszkiewicz, Kevin Ashley, Trevor Bench-Capon ☑, L. Karl Branting, Jack G. Conrad & Adam Wyner

Artificial Intelligence and Law 30, 561–591 (2022) | Cite this article 2778 Accesses | 3 Citations | 9 Altmetric | Metrics

Abstract

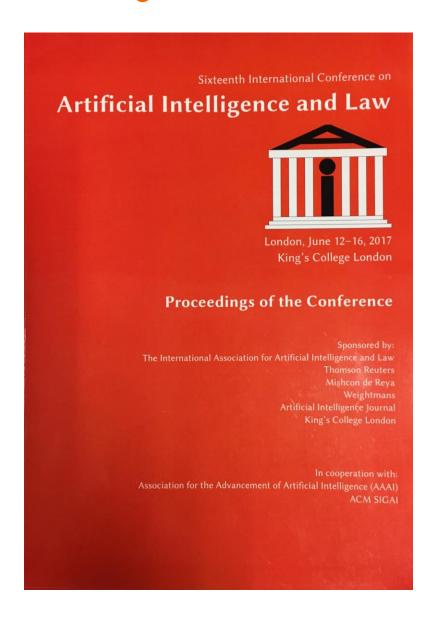
The first issue of *Artificial Intelligence and Law* journal was published in 1992. This paper offers some commentaries on papers drawn from the Journal's third decade. They indicate a major shift within Artificial Intelligence, both generally and in AI and Law: away from symbolic techniques to those based on Machine Learning approaches, especially those based on Natural Language texts rather than feature sets. Eight papers are discussed: two concern the management and use of documents available on the World Wide Web, and six apply machine learning techniques to a variety of legal applications.

The Al and Law Journal

Key Issue – Reviewing 25 years of ICAIL research papers (3-4 papers per year)



ICAIL: Proceedings of the Int'l Conference on Al and Law – Latest Issue 2021



- From the biennial conference
- Print + Digital versions
- Accessible from the ACM Digital Library
- Full length, short and demo papers
- Introduction from PC Chairs
- Listings of workshops
- Hosted by IAAIL.org (more anon)



ICAIL: Proceedings of the Int'l Conference on AI and Law – 18 + counting

ICAIL 1997

The Sixth International Conference on Artificial Intelligence and Law

Location: University of Melbourne Law School

Program Chair: L. Karl Branting, University of Wyoming

Conference Chair: John Zeleznikow (La Trobe University, Australia) and Dan Hunter (University of

Melbourne, Australia)

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail97.html

ICAIL 1995

The Fifth International Conference on Artificial Intelligence and Law

Location: University of Maryland, College Park, MD

Program Chair: Trevor Bench-Capon, Department of Computer Science, University of Liverpool

Conference Chair: L. Thorne McCarty, Department of Computer Science, Rutgers, The State

University of New Jersey

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail95.html

ICAIL 1993

The Fourth International Conference on Artificial Intelligence and Law

Location: Vrije Universiteit, Amsterdam

Program Chair: Kevin Ashley, School of Law, University of Pittsburgh

Conference Chair: Ania Oskamp, Open Universiteit, the Netherlands

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail93.html

ICAIL 1991

The Third International Conference on Artificial Intelligence and Law

Location: St Catherine's College, University of Oxford

Program Chair: Marek Sergot, Department of Computing, Imperial College London

Conference Chair: Richard Susskind, Masons, UK

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail91.html

ICAIL 1989

The Second International Conference on Artificial Intelligence and Law

Location: University of British Columbia, Vancouver

Program Chair: Edwina L. Rissland, Department of Computer Science, University of Massachusetts

Conference Chairs: J. C. Smith and Robert T. Franson (University of British Columbia)

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail89.html

ICAIL 1987

The First International Conference on Artificial Intelligence and Law

Location: Northeastern University, Boston, MA

Program Chair: L. Thorne McCarty, Department of Computer Science, Rutgers, The State University

of New Jersey

Conference Chair: Carole D. Hafner, Northeastern University, Boston

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail87.html

ICAIL 2009

The Twelveth International Conference on Artificial Intelligence and Law

Location: Autonomous University of Barcelona

Program Chair: Carole D. Hafner, College of Computer and Information Science, Northeastern

University, Boston

Conference Chair: Pompeu Casanovas, Institute of Law and Technology, Autonomous University of

Barcelona

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2009.html

ICAIL 2007

The Eleventh International Conference on Artificial Intelligence and Law

Location: Stanford Law School, Stanford, California

Program Chair: Radboud Winkels, Leibniz Center for Law, University of Amsterdam, The Netherlands

Conference Chair: Anne Gardner, Atherton, California

Website: http://www.iaail.org/past-icail-conferences/icail2007/index.html

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2007.html

ICAIL 2005

The Tenth International Conference on Artificial Intelligence and Law

Location: University of Bologna

Program Chair: Anne Gardner, Atherton, California
Conference Chair: Giovanni Sartor, University of Bologna

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2005.html

ICAIL 2003

The Ninth International Conference on Artificial Intelligence and Law

Location: University of Edinburgh

Program Chair: Giovanni Sartor, University of Bologna

Conference Chairs: John Zeleznikow (The Edinburgh Law School, University of Edinburgh) and Lillian

Edwards (The Edinburgh Law School, University of Edinburgh)

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2003.html

ICAIL 2001

The Eighth International Conference on Artificial Intelligence and Law

Location: Washington University, St. Louis, MO

Program Chair: Henry Prakken, Department of Information and Computing Sciences, Utrecht

University, The Netherlands

Conference Chair: Ronald P. Loui, Department of Computer Science, Washington University, Saint

Louis, MO

Website: http://www.iaail.org/past-icail-conferences/icail2001/index.html

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2001.html

ICAIL 199

The Seventh International Conference on Artificial Intelligence and Law

Location: University of Oslo, Norway

Program Chair: Thomas F. Gordon, GMD, Germany

Conference Chair: Jon Bing, NRCCL (Norwegian Research Center for Computers and Law) and

Andrew J. I. Jones, University of Oslo

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail1999.html

ICAIL 2021

The Eighteenth International Conference on Artificial Intelligence and Law

Location: University of São Paulo, Brazil

Program Chair: Adam Zachary Wyner (Swansea University, United Kingdom)
Conference Chair: Juliano Maranhão (University of São Paulo, Brazil)

Website: https://icail.lawgorithm.com.br/

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2021.html

ICAIL 2019

The Seventeenth International Conference on Artificial Intelligence and Law

Location: Cyberjustice Laboratory, University of Montreal, Canada Program Chair: Floris Bex (Utrecht University & Tilburg University)

Conference Chair: Karim Benyekhlef (University of Montreal)

 $\textbf{Website}: \ \textbf{http://www.iaail.org/?q=article/icail-2019-17th-international-conference...}$

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2019.html

ICAIL 2017

The Sixteenth International Conference on Artificial Intelligence and Law

Location: Strand Campus of King's College London

Program Chair: Guido Governatori, Data61, CSIRO, Australia

Conference Chair: Jeroen Keepens, King's College London, UK

Website: https://nms.kcl.ac.uk/icail2017/

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2017.html

ICAIL 2015

The Fiftheen International Conference on Artificial Intelligence and Law

Location: University of San Diego School of Law

Program Chair: Katie Atkinson, University of Liverpool, UK

Conference Chair: Ted Sichelman, University of San Diego School of Law

Website: http://sites.sandiego.edu/icail/

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2015.html

ICAIL 2013

The Fourtheen International Conference on Artificial Intelligence and Law

Location: National Research Council (CNR) of Italy, Rome

Program Chair: Bart Verheij, University of Groningen, The Netherlands

Conference Chair: Enrico Francesconi, ITTIG-CNR, Italy

Website: http://icail2013.ittig.cnr.it

Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail2013.html

ICAIL 2011

The Thirteenth International Conference on Artificial Intelligence and Law

Location: University of Pittsburgh School of Law, Pittsburgh, PA

Program Chair: Tom van Engers, Leibniz Center for Law, University of Amsterdam, The Netherlands

Conference Chair: Kevin Ashley, University of Pittsburgh School of Law

Website: http://www.law.pitt.edu/events/2011/06/icail-2011-the-thirteenth-interna...
Bibliography: http://www.informatik.uni-trier.de/~ley/db/conf/icail/icail/2011.html



Does IAAIL Support Regional Conferences?

Europe – each November or December

Originated in the Netherlands



Asia - each autumn

Typically hosted in Japan

JURISIN

Sixteenth International Workshop on Juris-informatics (JURISIN 2022)

COLIEE – Competition on Legal Information Extraction and Entailment

- Held every year
 - With ICAIL or with JURISIN, depending on year
 - Generally ~5 tasks involving case or statute retrieval and entailment, legal question answering, etc.
 - Tasks can be using Japanese or English (translation) texts



ICAIL: Workshops covering latest trends in Al and Law

List of Workshops from ICAIL 2019



<u>Tutorial - Introduction to AI & Law</u>

Tutorial - Defeasible Logic for Normative Reasoning

Doctoral Consortium

2nd Workshop on AI in Legal Practice

<u>1st Workshop on Al vs. Intelligent Assistance for Legal Professionals</u>

in the Digital Workplace (LegalAIIA)

<u>LegRegSW 2019: A Workshop on a Legislation and Regulation</u>

<u>Semantic Web - a shared corpus task</u>

1st Workshop on Artificial Intelligence and the Administrative State

ASAIL 2019: Third Workshop on Automated Detection, Extraction and

Analysis of Semantic Information in Legal Texts

<u>Competition on Legal Information Extraction and Entailment (COLIEE</u>

<u> 2019)</u>

2nd MWAIL: Multilingual Workshop on AI & Law research

<u>Tutorial – Automated Deduction with Legal Texts</u>



What Organization Hosts ICAIL?

IAAIL.org: International Association for Artificial Intelligence and Law



ICAIL 2023 - 19TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND LAW - CALL FOR PAPERS, DEMONSTRATIONS, WORKSHOPS AND TUTORIALS

Submitted by iaail on Fri, 07/29/2022 - 14:02

June 19-23, 2023

Hosted by:

Universidade do Minho, Braga, Portugal

Universidade do Minho Law School with JUSGOV – Research Center for Justice and Governance, Centro Algoritmi and LASI – Laboratório Associado de Sistemas Inteligentes

Conference Website: https://icail2023.di.uminho.pt

Tags: Call for Papers

READ MORE

JURIX 2022 - CALL FOR WORKSHOP AND TUTORIALS OF THE 35TH INTERNATIONAL CONFERENCE ON LEGAL KNOWLEDGE AND INFORMATION SYSTEMS

Submitted by iaail on Wed, 07/27/2022 - 10:14

Besides the standard conference track and the Doctoral Consortium, JURIX 2022 will host a number of workshops and tutorials related to the theme of the conference.

Workshops provide an informal setting where participants have the opportunity to discuss specific

ICAIL 2023 WILL BE HELD IN BRAGA, PORTUGAL

ICAIL 2021 | G...

The next International Conference on AI and Law (ICAIL 2023) will be organized by the University of Minho, Braga, Portugal

FEATURED VIDEOS



ICAIL 2021 - Great Opening

LATEST PHOTOS



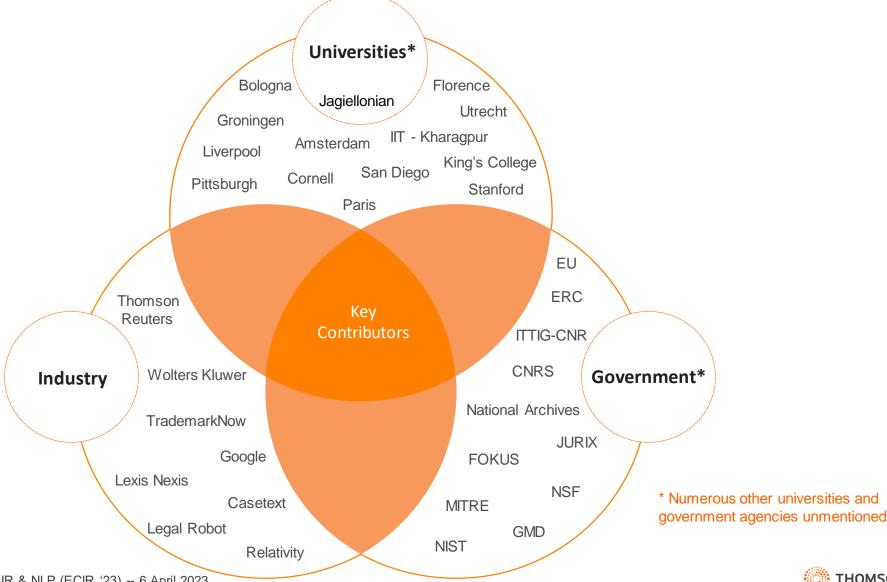






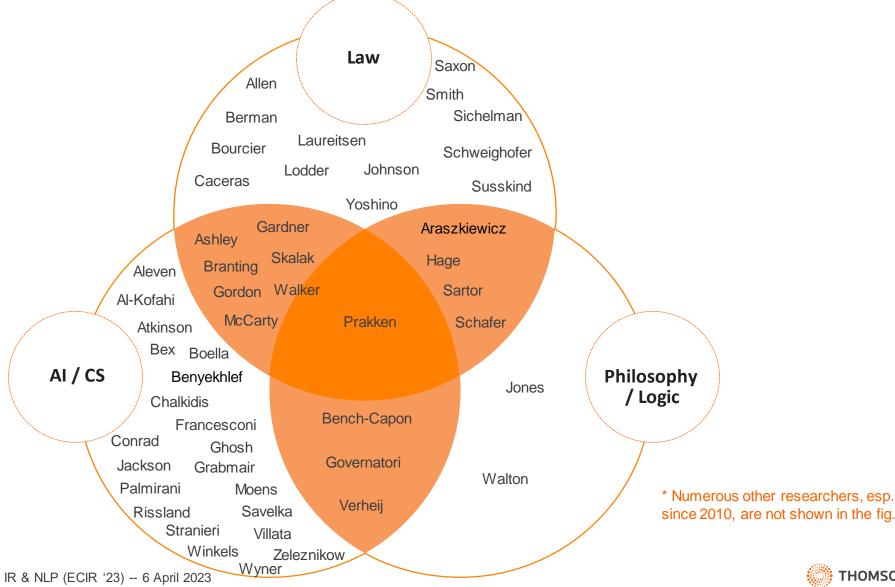
The Community and the Conference (IAAIL and ICAIL)

Demographics

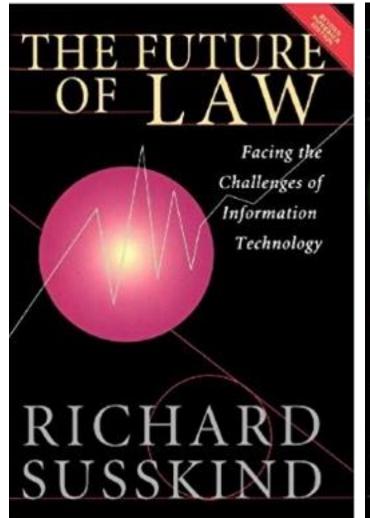


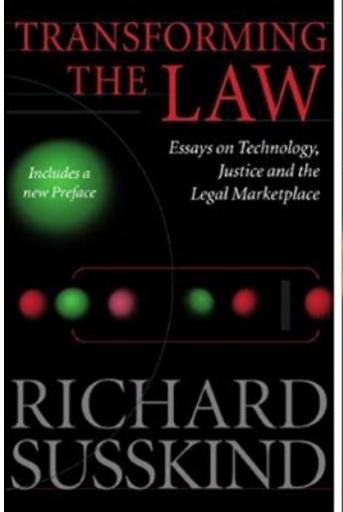
The Community and the Conference (IAAIL and ICAIL)

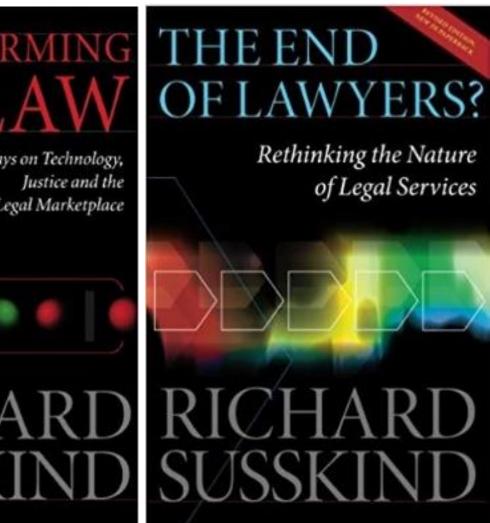
Demographics



Richard Susskind and Al & Law









- Legal Services
 Visionary
- Author, Lecturer
- Advisor to Lord Chief Justice
- AI & Law
 Consultant

ONLINE COURTS AND THE FUTURE OF JUSTICE RICHARD SUSSKIND

2019

THOMSON REUTERS®

1996 2000 2010

Noteworthy Early AI & Law Events

Among the earliest newsworthy activities in the community (1971 – 75)

TAXMAN Dr. L. Thorne McCarty

- Work on project began while author was a Law and Computer fellow at Stanford Law School (1971-73), and supported by a grant from the IBM Corporation.
- Additional work was undertaken during a workshop at Stanford Law School in 1975
- Computer time for the project was provided by the Stanford Artificial Intelligence Laboratory
- The work introduced some novel ways of computer-based semantic processing developed in the field of Artificial Intelligence
- McCarty describes a computer program that uses techniques that enables a computer to apply certain fact situation concepts in the area of corporate reorganization taxation
- In his Harvard Law Review article, he considers both the limitations of the program and the means by which it may be feasibly improved
- He further argues that if developed further, by means of increasingly sophisticated techniques, it may prove of practical use to a lawyer

VOLUME 90

MARCH 1977

NUMBER 5

HARVARD LAW REVIEW

REFLECTIONS ON TAXMAN: AN EXPERIMENT IN ARTIFICIAL INTELLIGENCE AND LEGAL REASONING †

L. Thorne McCarty *

After introducing some basic techniques of semantic information processing developed in the field of artificial intelligence, Professor McCarty describes a computer program using such techniques which enables a computer to apply to certain fact situations concepts of the area of the taxation of corporate reorganizations. He considers both the present limitations of the program and ways in which it might feasibly be improved, and argues that if developed further by means of increasingly sophisticated techniques, it may prove of practical use to tax lawyers. The development of such computer models also has a dual theoretical purpose. It clarifies the structure of the area of law modeled, and the investigation of the limits of such models provides the foundation for a theory of the nature of legal concepts.



Noteworthy Early AI & Law Events

Among the earliest newsworthy activities in the community (1990 – 93)

WIN (Westlaw is Natural) **Dr. Howard Turtle**

- Worked in the Center for Intelligent Information Retrieval (CIIR) at UMass - Amherst while a PhD grad student under Prof. Bruce Croft
- Designed & implemented the *Inquery* natural language search engine as part of his PhD program
- Using Bayesian inference nets and tf.idf term & phase scoring, produced the first natural language search engine that could scale
- Was hired by West Publishing Co. (Westlaw) upon graduation.
- At West, led a secret project to productionize Inquery, making it arguably the world's the first industrial natural language search engine, which was known as 'WIN'
- Launched in 1992, in a pre-web world, it became available to the 1 million attorneys in the U.S. by way of West's own proprietary network.
- Received Online magazine's "Product of the Year" award in 1992.
- Inquery has been used for comparative IR experiments at CIIR over the three decades since Dr. Turtle's tenure at UMass.
- Related reference: https://dl.acm.org/doi/10.5555/188490.188559



You can use WIN in virtually all WESTLAW databases.

And go instantly to the most relevant portion of each

document. Even display cases in chronological order

but still retain their relevancy ranking.

first from West Publishing. Dedicated

More ways to win

to bringing America's lawyers faster

and easier ways to find the law.

FOR A NO-OBLIGATION DEMONSTRATION OR MORE INFORMATION, CALL 1-800-328-9352.

WIN is protected by U.S. Patent No. 5,265,065

Noteworthy Past Al & Law Events

Among the newsworthy activities within the community (2006 –)

TREC (Text Retrieval Conference) for Legal (E-Discovery)

Jason Baron JD, Dr. Doug Oard, Dr. David Lewis, Gordon Cormack, et al.

- Sponsored by the National Institute for Standards and Technology (NIST) in US
- Goal is to evaluate the ability of IR techniques meet the needs of the legal profession.
- It is especially focused on tools and methods to help retrieve electronic business records, principally for use as evidence in civil litigation, a.k.a., "e-Discovery."
- Each year's 'task' involves three phases that include planning, experimentation, and reporting.
- Prior tasks have addressed e-Discovery needs and performance against the Philip Morris tobacco settlement collection, end-to-end system evaluation, an interactive task, a learning task, a total recall task, and others.
- As a research community, TREC largely relies on volunteers to conduct its trials.
- Following the research task for any given year is an invitation-only conference for participants.
- Related references:
 - https://trec-legal.umiacs.umd.edu/
 - http://users.umiacs.umd.edu/~oard/desi7/#History



TREC Legal Track

About | 2012 | 2011 | 2010 | 2009 | 2008 | 2007 | 2006 | Papers | Other



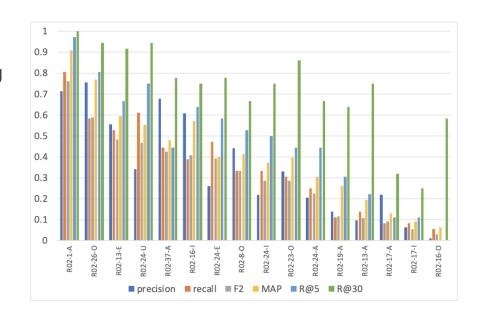
Noteworthy Past Al & Law Events

Among the newsworthy activities within the community (2014 –)

COLIEE (Competition on Legal Information Extraction/Entailment)

Randy Goebel, Ken Satoh, Juliano Rabelo, Yoshinobu Kano, et al.

- The most significant and longest running legal NLP competition in existence.
- Has historically been run against Japanese legal data, though with an accompanying translated English language collection.
- Participating teams can decide which collections they wish to work with.
- There are generally up to five tasks run per year, which have included tasks such as:
 - Task 1: Information retrieval for caselaw
 - Task 2: Confirmation of entailment relation between an existing case and an unseen one
 - Task 3: Information retrieval task for statutes
 - Task 4: Entailment/question answering task on retrieved civil code statutes
 - Task 5: Entailment/question answering task without retrieved civil code statutes
- The competition takes place each year. Participation is open to any group.
- Participants are invited to attend a workshop to report on their results after each competition. It is co-located with ICAIL in odd years and with JURISIN in even years.
- Related reference: https://sites.ualberta.ca/~rabelo/COLIEE2023/



Statute retrieval task: Averages of precision, recall, F2, MAP, R 5, R 10, and R 30 for non-easy questions with a single relevant article.



Noteworthy Past Al & Law Events

Among the newsworthy activities within the community (2021)

Optimizing the Role of Training

L. Zheng, N. Guha, B. Anderson, P. Henderson, D. Ho

- Observe that it remains unclear when researchers should rely on domain-specific resource-intensive pre-training
- Legal applications have shown few examples of substantive gains when domain pre-training has been used
- Hypothesize that current results come from legal tasks that are too easy, fail to meet conditions when legal pre-training can help
- Present a new dataset [CaseHOLD (Table 2)] consisting of 53K+ multiple choice questions to identify the relevant holding in a case
- Claimed to be a representative and difficult challenge
- Subsequently measure performance NLP tasks gains on CaseHOLD and other datasets
- Show that legal domain pre-training with custom vocab. outperforms BERT pre-trained on a general corpus
- Show domain pre-training may be warranted when task is similar to the pre-training corpus
- Reference: https://dl.acm.org/doi/abs/10.1145/3462757.3466088

When Does Pretraining Help? Assessing Self-Supervised Learning for Law and the CaseHOLD Dataset of 53,000+ Legal Holdings

Lucia Zheng* zlucia@stanford.edu Stanford University Stanford, California, USA Neel Guha* nguha@stanford.edu Stanford University Stanford, California, USA

Brandon R. Anderson banderson@law.stanford.edu Stanford University Stanford, California, USA

Peter Henderson phend@stanford.edu Stanford University Stanford, California, USA

ABSTRACT

While self-supervised learning has made rapid advances in natural language processing, it remains unclear when researchers should engage in resource-intensive domain-specific pretraining (domain pretraining). The law, puzzlingly, has yielded few documented instances of substantial gains to domain pretraining in spite of the fact that legal language is widely seen to be unique. We hypothesize that these existing results stem from the fact that existing legal NLP tasks are too easy and fail to meet conditions for when domain pretraining can help. To address this, we first present CaseHOLD (Case Holdings On Legal Decisions), a new dataset comprised of over 53,000+ multiple choice questions to identify the relevant

Daniel E. Ho

dho@law.stanford.edu Stanford University Stanford, California, USA

CCS CONCEPTS

 Applied computing → Law; • Computing methodologies → Natural language processing; Neural networks.

KEYWORDS

law, natural language processing, pretraining, benchmark dataset

ACM Reference Format:

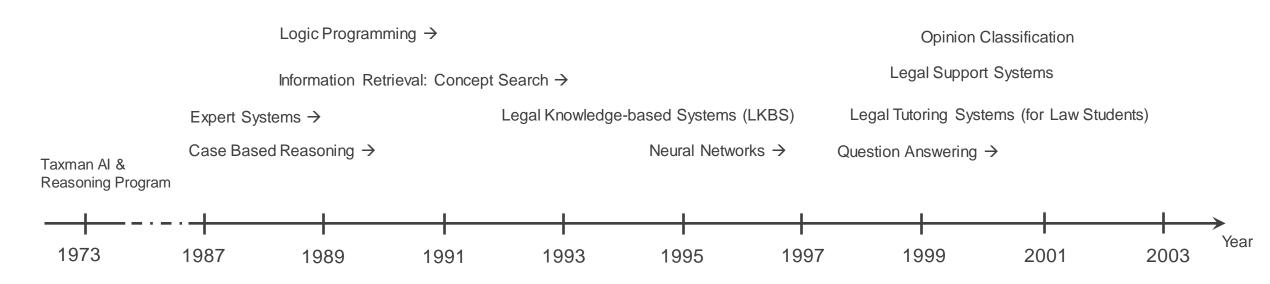
Lucia Zheng, Neel Guha, Brandon R. Anderson, Peter Henderson, and Daniel E. Ho. 2021. When Does Pretraining Help? Assessing Self-Supervised Learning for Law and the CaseHOLD Dataset of 53,000+ Legal Holdings. In Eighteenth International Conference for Artificial Intelligence and Law (ICAIL '21),



35+ Years of AI and Law

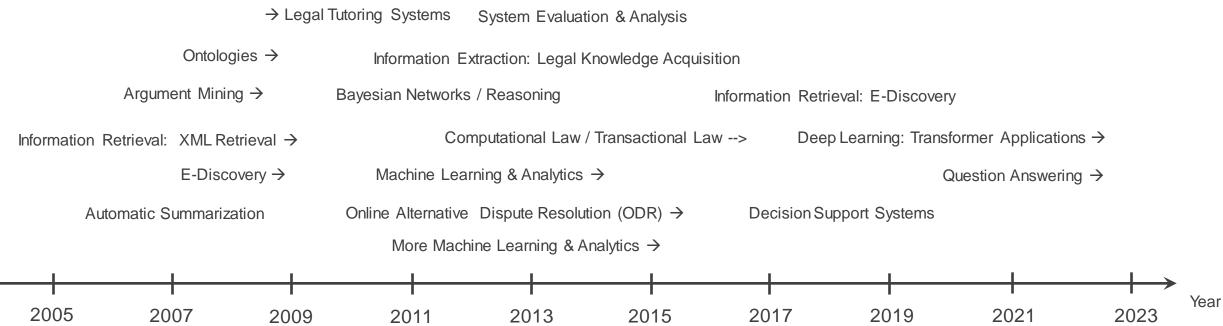


Al & Law across 4+ decades (1985-2003)



Al & Law across 4+ decades (2005-2023)

Legal Expert Assistants



Al & Law: Beyond Foundations

What's been the impact of the community on the field?

Via IAAIL.org sponsored

- Conferences
- Workshops
- Proceedings
- Journal



- ICAIL and its workshops have showcased commercial successes, start-ups, luminaries
 - E.g., WIN (first NL vers. of Westlaw), Lex Machina (Stanford start-up now part of Lexis Nexis), seminal E-Discovery works in series of 7 'DESI' workshops
- But a large amount of activity in the AI Legal domain in pockets outside IAAIL
 - E.g., CodeX (Stanford's Center for Legal Informatics) and especially in the startup space



The Role of Legal Start-ups

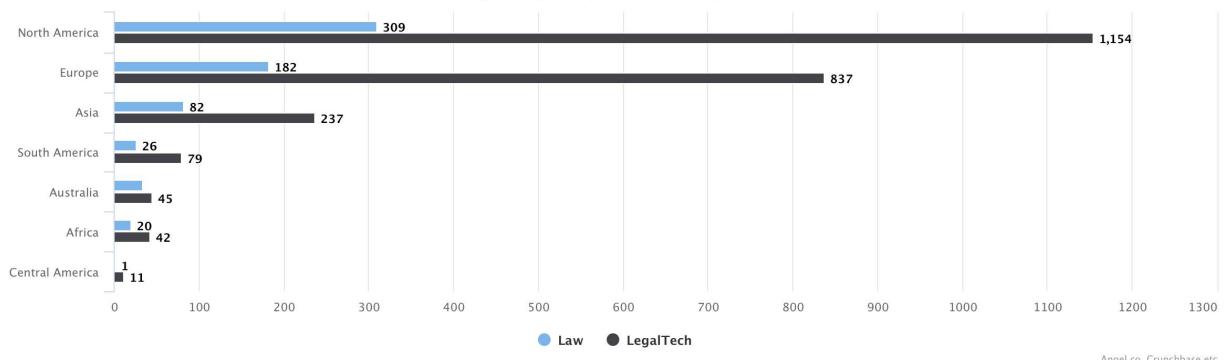


Legal Startups

By Continent

Legal Startups per Continent

Legal vs Legal Tech per continent starting 2010



Angel.co, Crunchbase etc



Legal Startups

By Year

Legal Startups Total by Year





Legal Startups – The Rest of the Story

https://www.legalcomplex.com/blog/2015/12/29/legal-startups-charts/

- Fill gaps in on-going research
- Tend to be highly focused
 - Address specific practitioner needs / problems
 - E.g., distribution of motions in the treatment of a course of litigation
- Often begin with a single jurisdiction or a specific legal practice area
 - E.g., California or IP
- Base next steps on past performance
 - Tied to likelihood of obtaining added funding
 - Trade-offs between expanded coverage and accuracy / value of results





From Past to Present



Al and Law – the Progression

Foundations

- Grand visions (like Al as a whole)
- Much focus on
 - Legal knowledge representation
 - Formal & computational models of reasoning, based largely on inference and argumentation
 - Emphasis also on legal interpretation, decision making
 - Some models developed with domain experts
- Early systems rule-based, e.g., 'TaxMan'
- Applications often focused on narrow subdomains, e.g., trademark infringement

Today and into the Future

- Increasingly data-driven
 - Harness large & ever-growing data repositories
- Approaches to legal analysis has changed
 - It's now more empirical, corpora-based
- Underlying models statistical in nature
 - Semantic, citation-based, or temporal patterns
 - Data-centric techniques enable new algorithmic approaches to legal problem solving
- Analysis continues to leverage core NLP, data mining, and machine learning
 - Deep Learning / Neural Networks making increasingly significant progress
 - LLMs currently taking the world by storm





A Brief History of Al & Law: Key Milestones for Legal IR & NLP

Thank you.

Jack G. Conrad
TR Labs -- Applied Research
45th European Conference on Information Retrieval
ECIR 2023
Thursday, 6 April 2023

