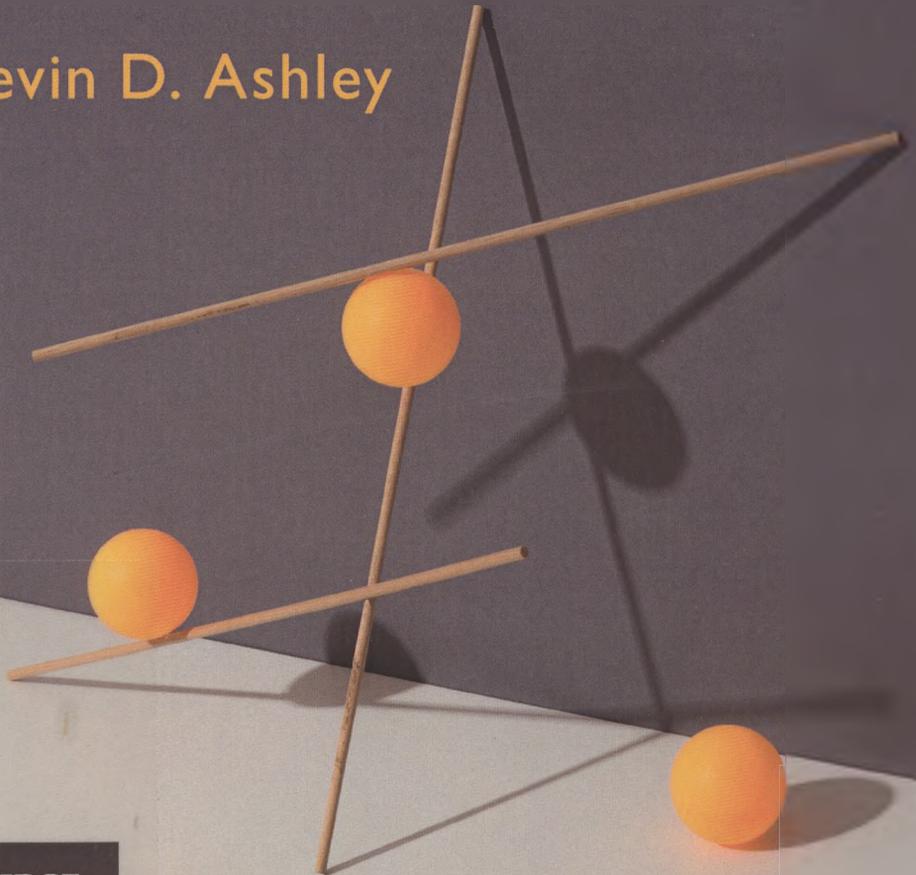


ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS

New Tools for Law Practice
in the Digital Age

Kevin D. Ashley



CAMBRIDGE

ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS

New Tools for Law Practice in the Digital Age

The field of artificial intelligence (AI) and the law is on the cusp of a revolution that began with text analytic programs like IBM's Watson and Debater and the open-source information management architectures on which they are based. Today, new legal applications are beginning to appear, and this book – designed to explain computational processes to non-programmers – describes how they will change the practice of law, specifically by connecting computational models of legal reasoning directly with legal text, generating arguments for and against particular outcomes, predicting outcomes, and explaining these predictions with reasons that legal professionals will be able to evaluate for themselves. These legal apps will support conceptual legal information retrieval and enable cognitive computing, enabling a collaboration between humans and computers in which each performs the kinds of intelligent activities that they can do best. Anyone interested in how AI is changing the practice of law should read this illuminating work.

Dr. Kevin D. Ashley is a Professor of Law and Intelligent Systems at the University of Pittsburgh, Senior Scientist, Learning Research and Development Center, and Adjunct Professor of Computer Science. He received a B.A. from Princeton University, a JD from Harvard Law School, and Ph.D. in computer science from the University of Massachusetts. A visiting scientist at the IBM Thomas J. Watson Research Center, NSF Presidential Young Investigator, and Fellow of the American Association for Artificial Intelligence, he is co-Editor-in-Chief of *Artificial Intelligence and Law* and teaches in the University of Bologna Erasmus Mundus doctoral program in Law, Science, and Technology.

Artificial Intelligence and Legal Analytics

NEW TOOLS FOR LAW PRACTICE IN
THE DIGITAL AGE

KEVIN D. ASHLEY

University of Pittsburgh School of Law



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314-321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06-04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107171503

DOI: 10.1017/9781316761380

© Kevin D. Ashley 2017

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2017

6th printing 2019

Printed in the United Kingdom by TJ International Ltd. Padstow, Cornwall

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data

ISBN 978-1-107-17150-3 Hardback

ISBN 978-1-316-62281-0 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

For Alida, forever

Contents

<i>List of illustrations</i>	<i>page xv</i>
<i>List of tables</i>	<i>xxi</i>
<i>Acknowledgments</i>	<i>xxiii</i>

PART I COMPUTATIONAL MODELS OF LEGAL REASONING	1
1 Introducing AI & Law and Its Role in Future Legal Practice	3
1.1. Introduction	3
1.2. AI & Law and the Promise of Text Analytics	4
1.3. New Paradigms for Intelligent Technology in Legal Practice	6
1.3.1. Former Paradigm: Legal Expert Systems	8
1.3.2. Alternative Paradigms: Argument Retrieval and Cognitive Computing	11
1.3.3. Toward the New Legal Apps	14
1.4. What Watson Can and Cannot Do	14
1.4.1. IBM's Watson	15
1.4.2. Question Answering vs. Reasoning	18
1.4.3. IBM's Debater Program	23
1.4.4. Text Analytic Tools for Legal Question Answering	27
1.4.5. Sources for Text Analytic Tools	30
1.5. A Guide to This Book	31
1.5.1. Part I: Computational Models of Legal Reasoning	32
1.5.2. Part II: Legal Text Analytics	33
1.5.3. Part III: Connecting Computational Reasoning Models and Legal Texts	34
1.6. Implications of Text Analytics for Students	35

2	Modeling Statutory Reasoning	38
2.1.	Introduction	38
2.2.	Complexities of Modeling Statutory Reasoning	39
2.2.1.	Semantic Ambiguity and Vagueness	39
2.2.2.	Syntactic Ambiguity	40
2.3.	Applying Statutory Legal Rules Deductively	42
2.3.1.	Running a Normalized Version on a Computer	43
2.3.2.	Predicate Logic	44
2.3.3.	Syntactic Ambiguity as Design Constraint	45
2.3.4.	The BNA Program	47
2.3.5.	Some Problems of Translating Statutes into Programs	48
2.4.	The Complexity of Statutory Interpretation and the Need for Arguments	52
2.4.1.	A Stepwise Process of Statutory Interpretation	53
2.4.2.	Other Sources of Legal Indeterminacy	54
2.5.	Management Systems for Business Rules and Processes	56
2.5.1.	Business Process Expert Systems	56
2.5.2.	Automating Business Process Compliance	60
2.5.3.	Requirements for a Process Compliance Language	62
2.5.4.	Connecting Legal Rules and Business Processes	65
2.5.5.	Example of Business Process Compliance Modeling	68
2.6.	Representing Statutory Networks	70
3	Modeling Case-based Legal Reasoning	73
3.1.	Introduction	73
3.2.	Relationship of Legal Concepts and Cases	74
3.2.1.	The Legal Process	74
3.2.2.	The Legal Process Illustrated	75
3.2.3.	Role of Legal Concepts	76
3.3.	Three Computational Models of Legal Concepts and Cases	77
3.3.1.	Prototypes and Deformations	78
3.3.2.	Dimensions and Legal Factors	81
3.3.3.	Exemplar-based Explanations	93
3.4.	Teleological Models of Case-based Legal Reasoning	97
3.5.	An Approach to Modeling Teleological Reasoning	100
3.5.1.	Teleology in Theory Construction	101
3.6.	Design Constraints for Cognitive Computing with Case-based Models of Legal Reasoning	104
4	Models for Predicting Legal Outcomes	107
4.1.	Introduction	107
4.2.	A Nearest Neighbor Approach to Automated Legal Prediction	108

4.3.	Introduction to Supervised Machine Learning	109
4.3.1.	Machine Learning Algorithms: Decision Trees	110
4.4.	Predicting Supreme Court Outcomes	111
4.4.1.	Features for Predicting Supreme Court Outcomes	112
4.4.2.	Applying Supervised Machine Learning to SCOTUS Data	112
4.4.3.	Evaluating the Machine Learning Method	113
4.4.4.	Machine Learning Evaluation Measures and Results	114
4.5.	Predicting Outcomes with Case-based Arguments	114
4.5.1.	Prediction with CATO	115
4.5.2.	Issue-based Prediction	115
4.5.3.	IBP's Prediction Algorithm	117
4.5.4.	Evaluating IBP's Predictions	119
4.6.	Prediction with Underlying Values	121
4.7.	Prediction based on Litigation Participants and Behavior	123
4.8.	Prediction in Cognitive Computing	125
5	Computational Models of Legal Argument	127
5.1.	Introduction	127
5.1.1.	Advantages of CMLAs	128
5.2.	The Carneades Argument Model	129
5.3.	An Extended Example of a CMLA in Action	131
5.3.1.	Family Law Example with Carneades	132
5.3.2.	Arguing with Defeasible Legal Rules	134
5.3.3.	Integrating Arguing with Cases and Rules	135
5.4.	Computational Model of Abstract Argumentation	139
5.5.	How CMLAs Compute Winners and Losers	141
5.5.1.	Resolving Conflicting Arguments about Facts	142
5.5.2.	Resolving Conflicting Arguments about Values	143
5.5.3.	Resolving Conflicting Arguments about Legal Rules	144
5.6.	How Practical are Computational Models of Legal Argument?	144
5.6.1.	Role of Proof Standards in CMLAs	145
5.6.2.	Integrating Probabilistic Reasoning into CMLAs	147
5.7.	Value Judgment-based Argumentative Prediction Model	149
5.7.1.	VJAP Domain Model	150
5.7.2.	VJAP Values Underlying Trade Secret Regulation	151
5.7.3.	VJAP Argument Schemes	154
5.7.4.	VJAP's Argument-based Predictions	156
5.7.5.	VJAP Program Evaluation	158
5.8.	Computational Model of Evidentiary Legal Argument	160
5.9.	Computational Models of Legal Argument as a Bridge	164

PART II	LEGAL TEXT ANALYTICS	169
6	Representing Legal Concepts in Ontologies and Type Systems	171
6.1.	Introduction	171
6.2.	Ontology Basics	172
6.3.	Sample Legal Ontologies	174
6.3.1.	The e-Court Ontology	174
6.3.2.	van Kralingen's Frame-based Ontology	176
6.4.	Constructing Legal Ontologies	178
6.5.	Ontological Support for Statutory Reasoning	181
6.6.	Ontological Support for Legal Argumentation	185
6.6.1.	A Target Application for Legal Argument Ontology	185
6.6.2.	An Ontology for the Argument Microworld	190
6.6.3.	Limits for Automating Legal Argumentation through Ontological Support	198
6.6.4.	Ontological Support for Cognitive Computing in Legal Argumentation	201
6.7.	Type Systems for Text Analytics	202
6.7.1.	Defining a Type System	202
6.7.2.	Type System Example: DeepQA	203
6.8.	LUIMA: A Legal UIMA Type System	204
6.9.	LUIMA Annotations can Support Conceptual Legal Information Retrieval	208
7	Making Legal Information Retrieval Smarter	210
7.1.	Introduction	210
7.2.	Current Legal Information Retrieval Services	211
7.3.	An Example of Using Commercial Legal IR Systems	212
7.4.	How Legal IR Systems Work	214
7.5.	IR Relevance Measures	216
7.5.1.	Boolean Relevance Measure	216
7.5.2.	Vector Space Approach to Relevance	217
7.5.3.	Probabilistic Model of Relevance	218
7.6.	Assessing Legal IR Systems	221
7.7.	Recent Developments in Legal IR Systems	223
7.8.	Comparing Legal IR and CMLAs	226
7.9.	Improving Legal IR with AI & Law Approaches	226
7.9.1.	Integrating Legal Ontologies and IR	227
7.9.2.	Integrating Legal IR and AI & Law Relevance Measures	227

7.9.3. Augmenting Legal IR Relevance Assessment with Citation Networks	230
7.9.4. Detecting Concept Change	232
7.10. Conclusion	233
8 Machine Learning with Legal Texts	234
8.1. Introduction	234
8.2. Applying Machine Learning to Textual Data	234
8.3. A Basic Setup for Applying ML to Legal Texts	236
8.4. Machine Learning for e-Discovery	239
8.4.1. Litigators' Hypotheses in e-Discovery	240
8.4.2. Predictive Coding Process	241
8.4.3. Assessing Predictive Coding Effectiveness	243
8.4.4. Other Open Issues in Predictive Coding	246
8.4.5. Unsupervised Machine Learning from Text	247
8.5. Applying ML to Legal Case Texts in the History Project	248
8.5.1. History Project System Architecture	249
8.5.2. ML Algorithms: Support Vector Machines	251
8.5.3. History Project SVM	252
8.6. Machine Learning of Case Structures	253
8.7. Applying ML to Statutory Texts	254
8.7.1. Statutory Analysis	254
8.7.2. An Interactive ML Tool for Statutory Analysis	255
8.8. Toward Cognitive Computing Legal Apps	257
9 Extracting Information from Statutory and Regulatory Texts	259
9.1. Introduction	259
9.2. Research Overview Regarding Extracting Information from Statutory Texts	260
9.3. Automatically Extracting Functional Information from Statutory Provisions	262
9.3.1. Machine Learning to Extract Functional Types of Provisions	263
9.3.2. Text Classification Rules to Extract Functional Information	265
9.4. ML vs. KE for Statutory Information Extraction	266
9.5. Extracting Logical Rules from Statutes and Regulations	268
9.6. Extracting Requirements for Compliant Product Designs	270
9.6.1. Implementing Compliance with Extracted Regulations	272
9.6.2. Semiautomated Approaches to Improving Human Annotation for Compliance	272
9.7. Extracting Functional Information to Compare Regulations	275
9.7.1. Machine Learning for Constructing Statutory Networks	276
9.7.2. Applying an ML Algorithm for Statutory Texts	278

9.7.3.	Evaluating the ML Algorithm on Statutory Texts and Dealing with Sparse Training Data	280
9.7.4.	Applying LUIMA to Enrich Statutory Text Representation	282
9.8.	Conclusion	283
10	Extracting Argument-Related Information from Legal Case Texts	285
10.1.	Introduction	285
10.2.	Argument-Related Information in Legal Cases	286
10.3.	Extracting Legal Argument Claims	287
10.3.1.	Machine Learning to Classify Sentences as Propositions, Premises, and Conclusions	287
10.3.2.	Text Representation	288
10.3.3.	Applying Statistical Learning Algorithms	289
10.3.4.	Argument Grammar for Discourse Tree Structure	291
10.3.5.	Identifying Instances of Argument Schemes	293
10.4.	Extracting Argument-Related Legal Factors	294
10.4.1.	Three Representations for Learning from Text	294
10.4.2.	How Well Did SMILE Work?	297
10.4.3.	Annotating Factor Components	298
10.5.	Extracting Findings of Fact and Cited Legal Rules	299
10.5.1.	Applying the LUIMA Type System	299
10.5.2.	Preparing Gold Standard Cases	300
10.5.3.	LUIMA-Annotate	301
10.5.4.	Evaluating LUIMA-Annotate	304
10.6.	Annotation of Training Data	305
10.6.1.	Annotation in IBM Debater	306
10.6.2.	Annotation Protocols	308
10.6.3.	Computer-Supported Annotation Environments	308
PART III CONNECTING COMPUTATIONAL REASONING MODELS AND LEGAL TEXTS		311
11	Conceptual Legal Information Retrieval for Cognitive Computing	313
11.1.	Introduction	313
11.2.	State of the Art in Conceptual Legal IR	315
11.3.	LUIMA Architecture	316
11.3.1.	LUIMA-Search	316
11.3.2.	Reranking Documents with LUIMA-Rerank	320
11.4.	An Experiment to Evaluate LUIMA	321
11.4.1.	Evaluation Metrics	323
11.4.2.	LUIMA vs. CLIR	324

11.5. Continuing to Transform Legal IR into AR	327
11.5.1. Connecting LARCCS and Legal IR Systems	328
11.5.2. Querying for Cases with Extended Argument-Related Information	329
11.5.3. New Legal Annotation Types	332
11.5.4. Prospects for Annotating Expanded Legal Types	336
11.5.5. Eliciting Users' Argument Needs	339
11.6. Conceptual Information Retrieval from Statutes	342
11.6.1. A Type System for Statutes	343
11.6.2. Network Techniques for Conceptual Legal IR	345
11.6.3. Conceptual Legal IR with Statutory Network Diagrams	346
11.7. Conclusion	349
12. Cognitive Computing Legal Apps	350
12.1. Introduction	350
12.2. New Legal Apps on the Market	351
12.2.1. Ross	351
12.2.2. Lex Machina	353
12.2.3. Ravel	353
12.3. Bridging Legal Texts and Computational Models	354
12.4. Cognitive Computing Apps for Testing Legal Hypotheses	354
12.4.1. A Paradigm for CCLAs: Legal Hypothesis-Testing	355
12.4.2. Targeted Legal Hypotheses	357
12.4.3. Operationalizing Hypotheses	359
12.4.4. Interpreting Hypotheses	361
12.5. Challenges for Cognitive Computing Legal Apps	367
12.5.1. Challenges: Automatically Annotating Legal Argument-Related Information	368
12.5.2. Challenges: Manual Annotation of Training Instances	373
12.5.3. Challenges: Query-Interface Design	379
12.6. Detecting Opportunities for New Hypotheses and Arguments	381
12.7. What to Do Next?	384
12.8. Conclusion	390
<i>Glossary</i>	393
<i>Bibliography</i>	403
<i>Index</i>	421

ARTIFICIAL INTELLIGENCE AND LEGAL REASONING

Biblioteka Główna
Akademii Sztuki Wojennej
26705/III (CB)



03-026705-000-0

The field of artificial intelligence (AI) and the law is on the cusp of a revolution that began with text analytic programs like IBM's Watson and Duxbury, and open-source information management architectures on which they were built. Today, new legal applications are beginning to appear, and this book seeks to explain computational processes to non-programmers – describing how they will change the practice of law, specifically by connecting computational models of legal reasoning directly with legal text, generating arguments for and against particular outcomes, predicting outcomes, and explaining these predictions with reasons that legal professionals will be able to evaluate for themselves. These legal apps will support conceptual legal information retrieval and enable cognitive computing, enabling a collaboration between humans and computers in which each performs the kinds of intelligent activities that they can do best. Anyone interested in how AI is changing the practice of law should read this illuminating work.

Copy.
004.056

DR. KEVIN D. ASHLEY is a Professor of Law and Intelligent Systems at the University of Pittsburgh, Senior Scientist, Learning Research and Development Center, and Adjunct Professor of Computer Science. He received a B.A. from Princeton University, a JD from Harvard Law School, and a Ph.D. in computer science from the University of Massachusetts. A visiting scientist at the IBM Thomas J. Watson Research Center, NSF Presidential Young Investigator, and Fellow of the American Association for Artificial Intelligence, he is co-Editor-in-Chief of *Artificial Intelligence and Law* and teaches in the University of Bologna Erasmus Mundus doctoral program in Law, Science, and Technology.

CAMBRIDGE
UNIVERSITY PRESS
www.cambridge.org

ISBN 978-1-316-62281-0

9 "781316"622810">