

Noise induced hearing loss: Building an application using the ANGELIC methodology

Latifa Al-Abdulkarim ^a, Katie Atkinson ^{a,*}, Trevor Bench-Capon ^a, Stuart Whittle ^b,
Rob Williams ^b and Catriona Wolfenden ^b

^a *Department of Computer Science, The University of Liverpool, UK*

^b *Weightmans LLP, Liverpool, UK*

Abstract. The ANGELIC methodology was developed to encapsulate knowledge of particular legal domains. In this paper we describe its use to build a full scale practical application intended to be used internally by Weightmans, a large firm of legal practitioners with branches throughout the UK. We describe the domain, Noise Induced Hearing Loss (NIHL), the intended task, the sources used, the stages in development and the resulting application. An assessment of the project and its potential for further development is discussed. The project shows that current academic research using computational models of argument can prove useful to legal practitioners confronted by real legal tasks.

Keywords: Reasoning with legal cases, abstract dialectical frameworks, knowledge elicitation, decision making in personal injury claims

1. Introduction

Although AI and Law has produced much interesting research over the last three decades, [16], there has been disappointingly little take-up from legal practice. One important exception is the approach to moving from written regulations to an executable expert system based on the methods proposed in [31], which has been developed through a series of ever larger companies: Softlaw, Ruleburst, Haley Systems and, currently, Oracle,¹ where it is marketed as *Oracle Policy Automation*. Important reasons why Softlaw and its successors were able to succeed were its well defined methodology and its close integration with the working practices of its customer organisations. There has recently, however, been an unprecedented degree of interest in AI and its potential for supporting legal practice. In addition to books such as [30] and reports such as *Capturing Technological Innovation in Legal Services*² there have been many articles in the legal trade press such as *Legal Business*³ and *Legal Practice Management*.⁴ There are

*Corresponding author. E-mail: K.M.Atkinson@liverpool.ac.uk.

¹<http://www.oracle.com/technetwork/apps-tech/policy-automation/overview/index.html>.

²Law Society (2017). <http://www.lawsociety.org.uk/support-services/research-trends/capturing-technological-innovation-report>.

³*AI and the law tools of tomorrow: A special report*. www.legalbusiness.co.uk/index.php/analysis/4874-ai-and-the-law-tools-of-tomorrow-a-special-report. All websites accessed in September 2017.

⁴*The Future has Landed*. www.legalsupportnetwork.co.uk. The article appeared in the March 2015 edition.

websites such as *Artificial Lawyer*⁵ and that of the International Legal Technology Association.⁶ There have also been UK national radio programmes with editions devoted to AI and Law topics including both *Law in Action*⁷ and *Analysis*⁸ and events, such as panels run by the Law Society of England and Wales⁹ and *Alternative AI for Professional Services*.¹⁰ At the ICAIL 2017 conference there was a well attended workshop on *AI in Legal Practice*.¹¹ We have also seen law firms such as Dentons invest in AI via Nextlaw Labs, their legal technology venture development company.¹² Other examples include TLT, who have partnered with LegalSifter,¹³ and Slaughter and May, who are contributing to the funding of Luminance.¹⁴

The above examples serve to show that the legal profession has never been so interested in, and receptive to, the possibilities of AI for application to their commercial activities. There are, therefore, opportunities which need to be taken. In this paper we describe the use of the ANGELIC (ADF for kNowledGe Encapsulation of Legal Information for Cases) methodology [3], developed to encapsulate knowledge of particular legal domains, to build a useful practical application for internal use by a firm of legal practitioners, to enable mutual exploration of these opportunities. ANGELIC is able to draw on more than 25 years of improved academic understanding of the issues since the publication of [31], the paper which contained the original Softlaw proposal. In particular [31] lacked any hierarchical structure. ANGELIC in contrast emphasises a hierarchical structure and is driven by its documentation, which comprises

- a hierarchical structure of nodes representing issues, intermediate concepts and factors, ascribed to cases on the basis of their particular facts. The structure corresponds to the abstract factor hierarchy of CATO [7]. The structure produced by ANGELIC is always a tree.
- a set of links between these nodes. The links are directed, so that the acceptability (or otherwise) of a parent is determined solely by the acceptability (or otherwise) of its children.
- a set of acceptability conditions. For each node, a series of tests is given, each sufficient to determine whether the node should be accepted or rejected, together with a default in case none of the tests is satisfied. These tests are all exclusively in terms of children of the node in question.

As discussed in Section 3, where a further description of the methodology is given, this documentation can be seen as an Abstract Dialectical Framework (ADF) [23]. This paper is intended to describe an application of the methodology, rather than the methodology itself, which has already been presented in [3]. As such this paper alone is not intended to provide sufficient detail for someone interested in using the methodology for their own applications, or evaluating it. Such interested parties are referred to [3]

⁵<https://www.artificiallawyer.com>.

⁶<https://www.iltanet.orgUK>.

⁷*Artificial Intelligence and the Law*. www.bbc.co.uk/programmes/b07dlxmj.

⁸*When Robots Steal Our Jobs*. www.bbc.co.uk/programmes/b0540h85.

⁹The full event of one such panel can be seen on youtube at www.youtube.com/watch?v=8jPB-4Y3jLg. Other youtube videos include Richard Susskind at www.youtube.com/watch?v=xs0iQSyBoDE and Karen Jacks at www.youtube.com/watch?v=v0B5UNWN-eY.

¹⁰<https://www.alternativeai.co.uk/>.

¹¹<https://nms.kcl.ac.uk/icail2017/ailp.php>.

¹²<https://www.dentons.com/en/whats-different-about-dentons/connecting-you-to-talented-lawyers-around-the-globe/news/2017/may/nextlaw-labs-invests-in-next-generation-ai-powered-expertise-finding-platform-profinda>.

¹³<http://www.tltsolicitors.com/insights-and-events/news/tlt-and-legalsifter-combine-ai-and-legal-advice-in-contract-negotiation-solution/>.

¹⁴<http://www.legalweek.com/sites/legalweek/2017/11/29/slaughter-and-may-contributes-to-10m-funding-round-for-ai-startup-luminance/?sreturn=20180008164521>.

which gives a full description of the methodology, its technical details, and its advantages. That paper also discusses its application to three different legal domains, each of which is illustrated by complete examples of the structures, the acceptance conditions for each domain and the executable code which can be produced from the design.

In Section 2 we provide an overview of the law firm for which the application was developed, the domain, and the particular task in that domain at which the application was directed. The section also includes a discussion of why argumentation is central to the task we consider, and hence why we adopt the traditional manual analysis approach for this application, rather than using machine learning or argument mining techniques. Section 3 gives a fuller overview of the ANGELIC methodology, while Section 4 describes the sources used to develop the application. Section 5 discusses the process of capturing and refining the domain knowledge and describes the development of an interface to enable the knowledge to be deployed for the required task. Section 6 provides a discussion of the current and future possibilities for development of the work and Section 7 concludes the paper.¹⁵

2. Application overview

The application was developed for Weightmans LLP. Weightmans LLP is an Alternative Business Structure (ABS) and a top 45 UK law firm with revenue of £95 million which employs 1,300 people across 7 offices. Weightmans is a full service law firm and is highly respected in the public sector, acting for many local, police and fire authorities, and NHS trusts. Weightmans provides strong, diverse commercial services for public sector bodies, large institutions, owner-managed businesses and Public Liability Companies and has a full family and private client service including: wills, tax, probate and residential conveyancing. Weightmans is a proud leading national player in insurance, with a formidable reputation and heritage with one of the largest national defendant litigation solicitor practices and an annual turnover in civil litigation work approaching £60 million. Weightmans deals with motor, liability and other classes of claims for clients from the general insurance industry, other compensators including NHS Resolution (formerly the NHS Litigation Authority), local authorities, and self-insured commercial organisations such as national distribution and logistics companies.

2.1. Noise induced hearing loss

Amongst other things, Weightmans acts for employers and their insurance companies and advises them when they face claims from claimants for Noise Induced Hearing Loss (NIHL) where it is alleged that the hearing loss is attributable to negligence on the part of the claimant's employer(s), or former employer(s), during the period of the claimant's employment. Weightmans receive approximately 5,000 new NIHL instructions annually, the majority of which are at the pre-litigation stage. Liability in law to compensate the claimant cannot be established in more than half of the claims in which Weightmans are asked to advise (clients will not normally involve lawyers unless there is some substantial issue in the claim). When damages are paid, the average figure is around £3,700. Industry wide, the Association of British Insurers (ABI) figures show that 2013 was a peak year for NIHL claims notifications with the number reaching over 85,000 with an estimated cost of over £400 million.¹⁶

¹⁵This paper is a revised and extended version of [5], presented at the Jurix 2017 conference.

¹⁶<https://www.abi.org.uk/globalassets/sitecore/files/documents/consumer-guides/tackling-the-compensation-culture.pdf>.

In each case Weightmans typically represents both the insurer and the insured employer (as their interests are aligned). Weightmans advises whether the claimant (the current or former employee) has a good claim in law and, if appropriate, the likely amount of any settlement (which the insurer will normally meet on behalf of the employer). Weightmans' role is thus to identify potential arguments which the employers or their insurance companies might use to defend or mitigate the claim. Compensators are thus looking to use the explicit statement of domain knowledge produced in this project primarily to improve their ability to settle valid claims and, when appropriate, pay proper and fair compensation in a timely manner, whilst using the application to challenge cases which may have no basis in law or may otherwise be defensible. In such an application it is essential that the *arguments* be identified. Black box pronouncements are of no use: it is the reasons that are needed. Note that the idea is to identify usable arguments, not to replicate the argumentation process. As Branting remarks in [20], “[machine learning] techniques typically are opaque and unable to support the rule-governed discourse needed for persuasive argumentation and justification”.

2.2. Need for explicit arguments

It is this need to find the arguments that are available in the particular case under consideration which militates against currently fashionable systems which exploit the use of the large amounts of data now available, typically using some form of Machine Learning. Consider, for example, [6] which received a good deal of attention in the UK national media in 2017.¹⁷ It was designed to predict whether cases heard before the European Court of Human Rights would result in violations or non-violations of three Articles of the European Convention on Human Rights. It was stated in that paper that “our models can predict the court’s decisions with a strong accuracy (79% on average)”. They used a dataset comprising 250 cases related to Article 3, 80 to Article 6 and 254 to Article 8, balanced between violation and non-violation cases. Although the resulting model was intended to provide a way of predicting outcomes in pending cases *before* they are decided, availability dictated that they used the published decisions, written *after* the decision had been made, to produce their models. Since these decisions reflect the outcome as well as the facts, this may raise questions and potential concerns about the methodology,¹⁸ in that the decisions may contain indications of the outcome not in the documents presented during the trial, but that is not pertinent here. The 79% success rate reported does not represent strong enough accuracy for our purposes, since it is doubtful whether the kind of application wanted by Weightmans could tolerate an error rate in excess of 20% for the decision support task we are focussing on. Again that is not really the issue here. What does matter is that the system offers only a binary *violation* or *non-violation* answer. This is not what is needed for our application: Weightmans’ clients will expect some justification for why the claimant’s arguments are valid, or, if they are worth challenging, the exact reason for the challenge needs to be identified. It should be possible to attempt to discover the features being used to make the predictions: over 25 years ago rules were being extracted from a neural net trained on extensive cases data [13]. The task in [13] was to identify the 12 relevant features from cases containing 64 features. The system achieved over 98% accuracy in classifying cases (note that the input data was structured rather than natural language, accounting for the high success rate), but this was achieved without considering 4 of the 12 relevant factors, demonstrating that successful prediction does not prove that the model discovered is correct or complete.

¹⁷Artificial intelligence ‘judge’ developed by UCL computer scientists, <https://www.theguardian.com/technology/2016/oct/24/artificial-intelligence-judge-university-college-london-computer-scientists>.

¹⁸A fact recognised in the paper if not in the press release.

No extraction of rules is attempted in [6], but they do offer “the 20 most frequent words, listed in order of their SVM [Support Vector Machine] weight”. These do not, however, look immediately promising: the list for topic 23 of article 6 predicting violation, for example, is:

court, applicant, article, judgment, case, law, proceeding, application, government, convention, time, article convention, January, human, lodged, domestic, February, September, relevant, represented

One doubts whether this would be the list expected by a legal expert.

2.3. Problems with uncritical use of past decisions

A major problem with looking at the past decisions to build a model for use with future cases, is that the past decisions may contain all sorts of, perhaps unconscious, biases and prejudices. The need to look critically at the training data is shown by Microsoft AI chatbot, Tay, a machine learning project,¹⁹ designed for human engagement. As it learns from previous examples, some of its responses are inappropriate and indicative of the types of interactions some people are having with it. Indeed, so racist and misogynist were its learned responses that it was necessary to withdraw it very quickly. Although less extreme than so-called social media, any corpus is likely to embody some human foibles. For example [25] states:

Our results indicate that text corpora contain recoverable and accurate imprints of our historic biases, whether morally neutral as towards insects or flowers, problematic as towards race or gender, or even simply veridical, reflecting the status quo distribution of gender with respect to careers or first names.

This also applies to legal decisions which very often reflect the attitudes and assumptions prevalent at the time they were written. Provided that one is aware of, or even trying to uncover, such biases, this can be useful (e.g. [28], reported in *The Guardian* newspaper as “Judges are more lenient after taking a break, study finds”²⁰). In the newspaper article, Jonathon Levav, one of the co-authors of [28], is reported as saying

You are anywhere between two and six times as likely to be released if you’re one of the first three prisoners considered versus the last three prisoners considered.

Time of day is also found to affect asylum decisions made by the US Department of Homeland Security in [29]. A prediction system may need to know such information in order to be accurate and such prediction-based systems can be used for a variety of purposes. However, a decision support system, in which explicit justification for a decision is given primacy, needs to uncover and remove such biases, not encourage and reflect them. Another extensive study, this time of US Supreme Court decisions (covering more than 240,000 justice votes and 28,000 cases outcomes over nearly two centuries (1816–2015)), is [32], which achieved an accuracy of around 70%.²¹

¹⁹Microsoft ‘deeply sorry’ for racist and sexist tweets by AI chatbot, <https://www.theguardian.com/technology/2016/mar/26/microsoft-deeply-sorry-for-offensive-tweets-by-ai-chatbot>.

²⁰<https://www.theguardian.com/law/2011/apr/11/judges-lenient-break>.

²¹This study used features based on the Supreme Court Database (Spaeth HJ, Epstein L, Martin AD, Segal JA, Ruger TJ, Benesh SC. 2016 Supreme Court Database, Version 2016 Legacy Release v01. (SCDB_Legacy_01) <http://Supremecourtdatabase.org>.) rather than unprocessed natural language. As the authors remark, this database makes available “more than two hundred years of high-quality, expertly-coded data on the Court’s behavior”. Note that the research is into the *behaviour* of the Court, not an attempt to discover laws, The lower accuracy may be explained by the greater difficulty of the task, given the much more extensive range of topics and the much longer period examined.

The techniques might, however, be usefully applied to the task of discovering which base-level factors are present in a case, in the manner of SMILE [10]. Here prejudices are less likely to have an impact, and importantly changes in the law (see below) tend to affect the attribution (as opposed to the significance) of base level factors only to a limited extent. A success rate of 80% is, however, unlikely to be good enough for many practical applications. If 20% of base level factors are misattributed, there is likely to be a significant impact on the arguments available. Given, however, the specificity of the target classification, especially if, like [10], that a separate model is developed for each factor, and the relatively short period involved, the accuracy may well be capable of improvement to acceptable levels. Such a system could then serve as a front end to the system we describe in this paper.²²

2.4. Argument mining

Another currently popular form of AI based on large corpora is *argument mining* (e.g. [33]). The idea here is to extract arguments from the corpus. Again this can prove interesting, especially if the arguments so extracted are subject to an informed critique, but it is not appropriate to the current task. For one thing, the arguments that will be extracted are as likely to be bad, unsound, arguments, as good ones. But even more significant is that our experts *already know* all the arguments that can be used before the courts: what they want is a system that can help less expert users to apply them accurately and consistently to particular cases, and to assess which are likely to succeed in particular cases. The task is not about finding arguments in general, but determining which are applicable to particular cases.

2.5. Change and development in the law

Another serious problem, which faces every corpus-based legal application, is the fact that law changes frequently. This makes it different from other domains. The rules of chess have not changed significantly since 1886 when Steinitz became the first world champion, and so games from that era can still provide instruction for players today. Although fashions change, and theory improves understanding of various positions, several opening variations introduced by Steinitz remain playable today. Thus the games collected in a corpus remain relevant and usable. Similarly the human body does not change (although our understanding of it and the treatment of diseases does) and so we need not worry about, for example, x-rays becoming obsolete. Also such corpora can be built on an international basis whereas law is peculiar to particular jurisdictions, and so large corpora are more difficult to assemble in law. In practice, the number of landmark legal cases required to understand a domain using traditional analysis techniques is relatively small: HYPO [8] used fewer than thirty cases, and the wild animal possession cases have produced a good deal of insight from an investigation using fewer than ten [11]. In law, unlike data mining, rules get their authority from their source, not from the number of supporting instances.

A major problem with law, however, is that it does change, and AI systems directed towards law need to take this into account. This has long been recognized: the authors of [21] describe how very soon after their group had expended a good deal of time and effort in building a detailed model of a piece of legislation, there was a substantial change in the law, casting doubt on the whole of their model. Such a change also renders any cases in a corpus related to the legislation dubious: since one cannot tell what is and what is not affected, one cannot use any of them, and must rebuild the corpus from scratch. Even if such wholesale changes in legislation are not so frequent, there is a constant stream of small updates

²²There is some commercial software available to support such a task, e.g. Kira Systems (<https://kirasystems.com/how-it-works>).

and modifications which can cast doubt on previous cases and their decisions. At least changes resulting from amendments of the legislation can be precisely dated, but changes can also occur through case law, without legislative amendment. Detecting such changes was the topic of [38] and [19]. Shifts in case law are required to adapt to changing social conditions and changing attitudes and values. Judges very often have the discretion to respond to these changes, so that their decisions may differ from those that would have been made, under the same legislation, in previous years. For example, UK Unemployment Benefit was originally conceived of as a six day benefit, since when it was introduced in the late forties Saturday was normally seen as a work day, although there was an exception for those whose customary working week was Monday–Friday. Since then, however, a five day week has become the norm, and the exception is now the standard case. There has been no change in the legislation, but the decisions now reflect this social change. Similarly as Justice Marshall remarked in the context of a 1972 capital punishment case²³ “*stare decisis* must give way to changing values”.²⁴ Sometimes it is a matter of actual change, overturning decisions, but often it is refinement: as [14] shows in a discussion of the automobile exception to the US 4th Amendment. While the original case from the Prohibition era²⁵ remains the basis of the exception, over time other issues have become to be seen as relevant, most notably privacy.

The recognition that law is always changing, so that new decisions may overturn old ones, and that past decisions may require reinterpretation, means that methodologies for building legal knowledge based systems have been designed with the maintainability of the knowledge as a chief concern [17,27]. The development of ontologies in AI and Law was, in part, motivated by this need for maintainability [41]. As discussed in [3], the ANGELIC methodology is no exception: the principled modularisation which is a key feature is designed to facilitate maintenance. Responding to change using ANGELIC is discussed in detail in [2].

2.6. Knowledge elicitation versus machine learning

Thus, given the difficulties with corpora containing dubious, misleading and superseded cases, the nature of the task and the availability of genuine experts, the knowledge for our application will need to be elicited using traditional techniques and deployed in a program not dissimilar from a “good old fashioned” expert system. This seems to meet the current task requirements, which are to support and improve decision making in terms of speed, quality and consistency. The novelty resides in the use of a newly developed methodology, which uses the target structure to drive the elicitation process, and provides an effective form in which the knowledge is captured and recorded: unlike Softlaw it does not restrict itself to encoding written rules, but draws on other forms of documentation and expert knowledge, which may include specific experiences such as previous dealings with a particular site and common sense knowledge, and it structures this knowledge into a set of highly cohesive modules. In this way, the knowledge acquired using ANGELIC is not restricted to specific items of law, or to particular precedents, but can capture the wider negligence principles that experts distill from the most pertinent decisions. Since the knowledge encapsulated is a superset of what is produced in the CATO system [7], it could, were the task teaching law students to distinguish cases, equally well be deployed in that style

²³Furman v. Georgia, 408 U.S. 238 (1972).

²⁴Katz et al. [32] noted that court outcomes are potentially influenced by a variety of dynamics, including public opinion, inter-branch conflict, changing membership and shifting views of the Justices, and changing judicial norms and procedures. They remark that “The classic adage ‘past performance does not necessarily predict future results’ is very much applicable.”

²⁵Carroll v. United States, 267 U.S. 132 (1925).

of program. Note too that the analysis producing the knowledge, as in CATO, is performed by a human analyst and then applied to cases: the knowledge is not derived from the cases, nor is it a machine learning system.

Machine Learning techniques may, however, be able to provide support for moving from natural language to the base level factors required as input to our model. Currently this task relies on the judgement of the users and, as noted in our concluding remarks, some support would be highly beneficial. Different tasks for automation support may well lend themselves to different techniques. A very recent discussion of the limitations of deep learning in general can be found in [34]. Marcus lists ten problems with deep learning, some of which are very applicable to law and related to the above discussion, including that it works best when there are “thousands, millions or even billions of training examples”, a situation unlikely to be found in a particular domain of law; its lack of transparency; its inability to distinguish between causation and correlation, which is linked to the inclusion of biases and prejudices; that it presumes a largely stable world, which is certainly not the case with law; that, although it gives generally good results, it cannot be trusted, especially where the examples show small but significant deviations from the standard case (as often happens in law); and that deep learning models are hard to reuse and to engineer with. Marcus, however, remains excited about deep learning provided its limitations are recognised and the current hype does not lead to it being seen as a panacea: it should be seen “not as a universal solvent, but simply as one tool among many, a power screwdriver in a world in which we also need hammers, wrenches, and pliers, not to mention chisels and drills, voltmeters, logic probes, and oscilloscopes.” Marcus argues for hybrid models, where deep learning and good old fashioned AI are used to supply appropriate parts of a hybrid system, much as we have suggested above. Discussions of what kinds of legal task are most suitable for various AI approaches can be found in [20] and [9].

3. Methodology overview

The ANGELIC methodology builds on traditional AI and Law techniques for reasoning with cases in the manner of HYPO [8] and CATO [7] and draws on recent developments in argumentation, in particular Abstract Dialectical Frameworks (ADFs), introduced in [24] and revisited in [23], and ASPIC+ [36]. The software engineering advantages of the approach are extensively discussed in [3].

Formally ADFs, as described in [23], form a three tuple: a set of nodes, a set of directed links joining pairs of nodes (a *parent* and its *children*), and a set of acceptance conditions. The nodes represent statements²⁶ which, in this context relate to the issues, intermediate factors and base level factors found in CATO’s factor hierarchies, as described in [7]. The links show which nodes are used to determine the acceptability (or otherwise) of any particular node, so that the acceptability of a parent node is determined solely by its children. As explained in [3], it is this feature which provides the strong modularisation that is important from a software engineering point of view, and which is lacking in many previous rule based systems.²⁷

For example, early rule based systems such as the British Nationality Act (BNA) system [40] noted difficulties resulting from structural features of legislation, such as cross reference, which led to unforeseen interactions between different parts of the knowledge base. The BNA was relatively small and, because

²⁶In [23] the statements are three valued (true, false and undecided), but this has subsequently been generalised [22] to arbitrarily truth valued statements.

²⁷For a full description of CATO, which remains central to discussions of reasoning with legal cases in AI and Law, see [15]. For a full discussion of how CATO was reconstructed using ANGELIC see [3].

new, was an unamended piece of legislation, but the difficulties were exacerbated when the techniques was applied to larger pieces of legislation which had been subject to frequent amendment [18]. The need to cope with changes in the law [21] led to suggestions for structuring the knowledge base, such as hierarchical formalisations [39] and so-called isomorphic representations [17], and methodologies for building maintainable Knowledge Based Systems [27]. Ontologies represented another attempt to provide the needed structure [41]. In recent computational argumentation, in contrast, since the logic is not affected by the structure, less attention has been paid to the structure and modularisation of the knowledge base. For example, although ASPIC+ [35] partitions its knowledge bases into axioms and contingent premises, this is for logical rather than software engineering reasons. In ANGELIC, the structure provides both the advantages of a hierarchical structure, and a fine grained, domain relevant, partitioning of the knowledge base.

The acceptance conditions for a node state how precisely its children relate to that node. In ANGELIC the acceptance conditions for non-leaf nodes are a set of individually sufficient and jointly necessary conditions for the parent to be accepted or rejected (although there is no reason why different, polyvalent, truth values should not be assigned, so that the statement types of [4] could be accommodated if desired). Thus the acceptance conditions can be seen as an ordered series of tests, each of which leads to acceptance or rejection of the node if satisfied, with the default to cover situations where none of the tests are satisfied. The tests are applied in order, thus prioritising them: once a satisfied test is found, the remaining tests are not applied. For leaf nodes, acceptance and rejection is determined by the user, on the basis of the facts of the particular case being considered. Essentially the methodology results in an ADF, the nodes and links of which correspond to the factor hierarchy of CATO [7]. The acceptance conditions for nodes with children contain a prioritised set of sufficient conditions for acceptance and rejection and a default. The base level factors which make up the leaf nodes are assigned a value by the user, and do not use defaults. Collectively, the acceptance conditions can be seen as a knowledge base as required by the ASPIC+ framework [35], but distributed into a number of tightly coherent and loosely coupled modules to conform with best software engineering practice [37].²⁸ Thus the acceptance conditions are used to generate arguments, and the ADF structure to guide their deployment.

The methodology is supported by tools [1] developed in parallel with, and informed by, this project intended to guide the knowledge acquisition, visualise the information, record information about the nodes such as provenance, and to generate prototype code to enable expert validation, and support refinement and enhancement. Once the knowledge is considered acceptable, a user interface is developed, in conjunction with those who will use the system in practice, to facilitate the input of the information needed for particular cases and present the results needed to support a particular task. The interface is tailored to the task. Different interfaces would enable the knowledge to be reused for a variety of different applications, including retrieval and prediction.

4. Sources

Several sources were supplied by Weightmans and used to provide the knowledge of the Noise Induced Hearing Loss domain to which the ANGELIC methodology was applied.

- **Experts:** Weightmans made available domain experts to introduce the domain, provide specific documents and to comment on and discuss the developing representation. Weightmans covers disease-related cases nationwide out of five offices. There are 130 specialist disease claims handlers dealing

²⁸For a full discussion of the relationship between ANGELIC and ASPIC+ see [12].

with live and legacy market claims from conditions ranging from Asbestos to Q Fever. Jim Byard, Partner, was the lead domain expert and is recognised as an industry leader. Jim specialises in occupational disease cases with particular interest in respiratory disease, work related upper limb disorders and noise induced hearing loss claims. Jim has been involved in much high profile litigation, defending class actions brought by claimants and represents the lead appellant in the Notts and Derbyshire Textile Litigation (*Baker v Quantum Clothing*) which came before the UK Supreme Court in November 2010. Jim has been part of the Civil Justice Council Working Party on NIHL which has recently published its final report, *Fixed Costs in Noise Induced Hearing Loss Claims*.²⁹

- **Documents:** The documents included a thirty-five page information document (*the Information Document*), which is an extract from a longer disease guide produced by Weightmans for use by their clients and lawyers; an eighteen question *check list* produced by Weightmans to train and guide their employees in the questions they should be asking; and a number of *anonymised example cases* illustrating different aspects of the domain. The Information Document and check list are useful reference tools for all levels of Weightmans employees and contain intellectual property confidential to Weightmans.
- **Users:** Potential users of the system were made available to assist in building and refining the interface to ensure it reflected the practicalities of file handling, as opposed to being only an abstract or theoretical tool.

Each of these sources played an invaluable role at various stages of the knowledge representation process, each making useful and complimentary contributions by providing different perspectives on the domain.

5. Representing the knowledge

Following an introductory discussion of the domain between the experts and the knowledge engineers, it was agreed that the workflow of the knowledge representation process could be seen as comprising a series of steps. The party responsible for performing each step is shown in brackets.

1. Analyse the available documents and identify the nodes and the links between them (knowledge engineers).
2. Organise the components into an ADF (knowledge engineers).
3. Define the acceptance conditions for these components (knowledge engineers).
4. Review the initial ADF (domain experts).
5. Refine the ADF to accommodate the changes requested by the domain experts (knowledge engineers). The review and refine steps are iterated until the ADF is signed off by both parties.
6. Extract a Prolog program from the acceptance conditions (knowledge engineers).
7. Run the program on the example cases to confirm the structure can generate the arguments in those cases (knowledge engineers).
8. Identify any necessary modifications (domain experts and knowledge engineers).
9. Specify a task-orientated GUI (end users and programmers).

Key stages are further described in the following subsections.

²⁹<https://www.judiciary.gov.uk/publications/establishing-fixed-costs-and-better-procedures-for-noise-claims/>.

5.1. Document analysis

The initial discussion with the domain experts provided an excellent orientation in the domain and the key issues. These issues included the fact that claims were time limited, and so had to be made within three years of the claimant becoming aware of the hearing loss. Both actual awareness, usually the date of an examination, and constructive awareness (the date on which the claimant should have been aware that there was a problem) need to be considered. Then there is a question of the nature of the hearing loss: there are many reasons why hearing deteriorates, and only some of them can be attributed to exposure to noise. Then there is the possibility of contributory negligence: there is a Code of Practice with which the employers should have complied, and it is also possible that the employee was in part to blame, by not wearing the ear defenders provided, for example.

Next, the Information Document was used to identify the components that would appear in the ADF, putting some flesh on the skeleton that emerged from the initial discussion. The Information Document provides summaries of the main definitions, the development of the legal domain rules, the assessment of general damages for noise induced hearing loss cases, and Judicial College Guidelines for the assessment of general damages. Other medical conditions related to hearing loss are listed and described.

At this stage components had been identified, and where these were elaborated in terms of the conditions that were associated with them, links between these components could be identified. For example hearing loss can be sensorineural, but can also be attributed to a number of other factors: natural loss through aging, loss accompanying cardio-vascular problems, infections, certain drugs, etc. Only sensorineural loss can be noise induced, and so hearing loss arising from the other factors cannot be compensated. The Information Document gives an indication of the various different kinds of hearing loss, and then further information of what may cause the various kinds of loss.

At the end of this phase we have a number of concepts, some of which are elaborated in terms of less abstract concepts, and some potential links. The next step is to organise these concepts in an appropriate structure.

The check list document (see the list of source documents in Section 4) was kept back to be used after the concepts had been organised into a hierarchy, to determine whether the hierarchy bottomed out in sensible base-level factors. The check list comprised a set of 18 questions and a “traffic light” system indicating their effect on the claim. The idea was to associate base level factors with the answers to these questions. For example, question 1 asks whether the exposure ceased more than three years before the letter of claim: if it did not, the claim is *ipso facto* within limitation and other kinds of defence must be considered.

Similarly the cases were not used to build the initial ADF but were held back to provide a means of working through the ADF to check that the arguments deployed in those cases could be recovered from the ADF.

5.2. Component organisation

The main goal now is to move from unstructured information gathered from the documentation to structured information. The main issues had been identified in the initial discussion and the document analysis. These were used to identify and cluster the relevant intermediate predicates from the documents. These nodes were further expanded as necessary to produce further intermediate predicates and possible base level factors. The checklist was then used to identify, and where necessary add, base level factors. The documents from the sample of particular cases were used to provide examples of possible

facts, and the effect these facts had on decisions. The result was a factor hierarchy diagram where the root shows the question to be answered, while the leaves show facts obtainable from the sample cases. All this was recorded in a table that described the factors in the domain and their related children (immediate descendants).³⁰ Here we will consider one particular node from the NIHL application, *breach of duty*:

Factor: Breach of duty

Description: The employer did not follow the code of practice in some respect.

Children: Risk assessments were undertaken; employee was told of risks; methods to reduce noise were applied, protection zones were identified, there was health surveillance, training.

The six children of *breach of duty* listed above are the main things required of an employer under the code of practice, and so provide a list of the ways in which a breach of duty might have occurred. They may be further elaborated: for example noise reduction includes measures such as shielding the machinery and providing appropriate ear protection.

The final version of the ADF for NIHL contains 3 issues, 20 intermediate nodes and 14 base level factors, with 39 links. For comparison, the ADF equivalent of CATO given in [3] contained 5 issues, 11 intermediate nodes, and 26 base level factors with 48 links. Thus CATO is larger, but NIHL has more internal structure. The nodes in the tabular presentation of the ADF are annotated to show their provenance (which of the documents they originate from, and the particular section or sections in which they are defined or explained), and whether they relate to particular questions in the checklist document. Although the NIHL ADF contains commercially confidential material, which means that it cannot be shown in this paper, there are several examples of both the diagrammatic and tabular presentations of domain ADFs in [3].

5.3. Defining acceptance conditions

Once the nodes had been identified, acceptance conditions providing sufficient conditions for acceptance and rejection of the non-leaf nodes in terms of their children were provided. These were then ordered by priority and a default provided. The particular example cases provided were used to confirm that the arguments used in them could be recovered from the ADF. Continuing the *breach of duty* example:

Factor: Breach of duty

Acceptance conditions:

1. Employee was not told of risks through the provision of education and training.
2. There were no measures taken to reduce noise.
3. Protection zones were not identified.
4. There was no health surveillance.
5. There was no risk assessment.
6. No relevant training was offered.

Each of these is a sufficient condition to identify a breach of duty. If none of them apply to the case, we can assume, as a default, that there was no breach of duty, and so add rejection of the node as the default.

³⁰Complete examples of both diagrammatic and tabular presentations can be found in [3].

5.4. Review and refinement

After this stage, the analysts and domain experts met to discuss and revise the initial ADF. The final ADF produced using the ANGELIC methodology is acyclic and so can be interpreted using grounded semantics.

Once a final ADF had been agreed, a Prolog program was produced from the acceptance conditions. The program was designed to suggest whether, given a particular set of case facts, there might be a plausible defence against the claim.

5.5. Program implementation

The program is implemented using Prolog, following the process described in detail in [3]. The program was created by writing a Prolog procedure for each node, starting with the leaf nodes and following the links to their parents until the root is reached. The Prolog procedure rewrites each test in the acceptance conditions as a clause and so determines the acceptability of each node in terms of its children. As in [3], which provides detailed examples, this required re-stating the tests using the appropriate syntax. Some reporting was added to indicate whether or not the node is satisfied, and through which condition. Finally some control was added to call the procedure to determine the next node, and to maintain a list of accepted factors. We do not give the example rules or program output here for reasons of commercial sensitivity, but their form is identical to that produced from the Trade Secrets domain ADF in [3]. The ADF acts as a design document, greatly easing the task of the programmer when compared with earlier systems such as [40] and [18] which presented the programmer with raw legislation to rewrite as Prolog code. The closeness between Prolog procedures and expressions of the acceptance conditions, each condition mapping to a clause within the Prolog procedure, makes the implementation quick, easy and transparent. The process of moving from acceptance conditions to Prolog code is essentially a mechanical rewriting into a template (supplying the reporting and control) and so is highly amenable to automation. Automated generation of the Prolog program from the ADF is planned as part of the development of the ANGELIC environment [1]. The program operates by:

- Instantiating the base level factors using the case facts;
- Working up the tree, deciding first the intermediate factors, then the issues, and finally giving the overall outcome. Non-leaf nodes are represented as heads of clauses, and each acceptance condition forms the body of a clause for the corresponding head, determining acceptance or rejection, with the set of clauses for the head completed by a default [26], so that, taken as a set, the acceptance conditions together provide necessary conditions for the acceptance or rejection of the head. The program reports the status of the node and the particular condition which led to this status before moving to the next node. Leaf nodes will be entered by the user as facts.

The program provides a very transparent output (see [3] for examples of output for the US Trade Secret domain) that identifies precisely the path up the hierarchy and hence where any divergences from the expected outcomes occur. The program has been tested on a range of cases (additional to those originally supplied) identified by their base factors to evaluate the output and help the analysts and experts in detecting any errors or potential improvements, carrying with it the factors established thus far.

5.6. Program refinement

Both the initial ADF and the program were, again, shown to and discussed with the domain experts, who suggested corrections and enhancements. The corrections varied: some suggestions were made

about considering missing information from the document, modifying the interpretation of existing acceptance conditions, or adding base factors or new parents to base facts. No changes were related to the main issues or intermediate predicates. As explained in [2], responding to changes in ANGELIC can be easily controlled since although the changes affect nodes individually, because of the modularisation achieved by the ADF, they can be made within a single node, and the programmer can be confident that they will not have ramifications in the rest of the structure. Refinement was an iterative process which was repeated until an ADF acceptable to the domain experts was obtained.

5.7. User interface

The ADF encapsulates knowledge of the domain to add support in handling the analysis of the cases in the legal domain. To fulfill this task, a forms-based interface was designed in conjunction with some of the case handlers who carry out the task and so are the target users of the implemented system.

- The interface is designed to take as input the base level factors which correspond to the questions in the checklist used by the case handlers.
- These questions are organised in an order which makes good sense in terms of the task. First the questions related to the claimant's actual knowledge of the hearing loss are displayed. The answers to questions are used to limit the options provided in later questions and, where possible, to provide automatic answers to other questions.
- Three to four questions are used per screen to maintain simplicity.
- The input to the questions uses a drop down list with the given options (facts), or radio buttons when one option needs to be selected, or checklists for multiple options. Text boxes are also provided to input information particular to individual cases, such as names and dates, or to allow further information for some questions.
- All the questions must be answered, but to minimise effort default answers can be provided for some questions (which must be explicitly accepted and can, of course, be overwritten by the users).

The designed interface enables the ADF to act as a decision augmentation tool for the particular task. Where the information is already available, for example, from an on-line claim form, some parts of the form can be pre-completed. A sample screen is shown in Fig. 1: the gender is pre-completed, but can be changed if necessary from a drop down menu.

6. Discussion

Developing the application was intended to realise a number of goals, each offering a perspective for assessment. Note that the system has not been fielded: it was intended as a feasibility study and the programs are prototypes. Validation of its actual (as opposed to potential) practical utility must await the fielding of a robust system engineered for operational use.

- The ANGELIC methodology had previously been applied only to academic examples. The desire here was to see whether it would also be effective when applied to a reasonably sized, independently specified, domain, intended to produce a system for practical use.
- Weightmans wished to come to a better understanding of the technology and what it could do for them and their clients.

The screenshot shows a web form for 'Weightmans NIHL Limitation'. The form is titled 'Part 1: Claimant'. It contains the following fields and controls:

- Claim Number:** A text input field.
- Claim Date:** A date picker with a calendar icon and the format 'DD/MM/YYYY'.
- First Name:** A text input field.
- Surname:** A text input field.
- Date of Birth:** A date picker with a calendar icon.
- Gender:** A dropdown menu with 'Female' selected.
- Working History:** A section with four sub-fields: 'Start' (date picker), 'End' (date picker), 'Employer' (text input), and 'Occupation' (text input).
- Add another employer:** A button below the Working History section.
- NEXT:** A button at the bottom of the form.

Fig. 1. Screen for User Interface.

- The methodology was designed to encapsulate knowledge of the domain using techniques representing the state of the art in computational argumentation. It was desired to see whether a domain encapsulated in this way could be the basis for a particular, practically useful, task in that domain.

Encouraging results were produced for each of the above objectives, sufficient to motivate further exploration and progress towards a fielded system. The methodology proved to be applicable to the new domain without significant change, and could be used with the sources provided. The time required with the domain experts was not prohibitive, and after an initial session introducing the academics and the lawyers to each other's worlds no real difficulties were encountered with this aspect of the project. Some desirable additional information that should be recorded about the nodes (such as provenance) was identified. The result was the specification and development of a set of tools to record and support the use of the methodology – the ANGELIC Environment [1].

Weightmans were encouraged that these techniques could indeed prove useful to their business. They are currently considering, with the University of Liverpool, other software providers and other commercial and industrial organisations, how best to take their investigations further, and to produce fielded systems. Since the original submission, several other legal areas have already been successfully represented using the ANGELIC methodology.

The key to most areas of law is to triage the issues and then make the decision right first time every time. The knowledge encapsulated in the ADF augments human lawyers well. In commercial and practical terms it means Weightmans can deliver an improved service through being better, faster and cheaper.

Implementation is being investigated with several tools and interfaces in order to find a solution that best integrates with a lawyer's working practice (an important element in success, as shown by Softlaw) and that also integrates with Weightmans case management system to prevent multiple systems being required, and to ensure that support is available at the most opportune time. Rather than simply developing a series of different functions for varied outcomes, Weightmans is working on connectivity for clients and lawyers, blending new and creative solutions to work seamlessly alongside each other, offering a smoother, more efficient and consistent approach to claims handling. A key feature of the use of ADFs is that they promote transparency in decision making since the specific AI techniques being used lend themselves particularly well to providing justifications for the decisions recommended by the decision engine and indeed at the testing stage, this was particularly appreciated by the lawyers.

User testing has shown that it is important to spend time refining the questions that become the visible part of the software realisation of the ADF to reduce ambiguity as much as possible and to provide guidance or an explanation on some of the more nuanced questions. Weightmans has partnered with software company Kira Systems, the leading provider of contract review and analysis software, to provide the data extraction capabilities required for the decision engine. Kira automatically identifies and extracts material contract metadata into an exportable summary, which then is fed into the decision engine to assist with decision making and argument identification. The technology has been proven to reduce the amount of time spent on individual cases, while maintaining a high standard of legal advice and a clear audit trail – all passing on value and delivering consistency for clients.

From the academic standpoint, as well as confirming the usefulness and applicability of the ANGELIC methodology, the customisation for a particular task showed that the general knowledge encapsulated in the ADF can be deployed for a specific task by the addition of a suitable interface.

7. Concluding remarks

The application of the ANGELIC methodology to a practical task enabled the academic partners in this project to demonstrate the utility of the methodology and identify possible extensions and improvements. The legal partners in the project were able to improve their understanding of the technology, and what it could do for their business, as well as what the development of an application capable of actual business use would require of them. For the kind of application described here, the argumentation is all important: the system is not meant to make a decision as to, or a prediction of, entitlement. Rather the case handlers are interested in whether there are any plausible arguments that could be advanced to challenge or mitigate the claim, or whether the arguments suggest that the claim should be accepted.

We believe that the success experienced for this task and domain is reproducible and look forward to using the methodology and supporting tools to build further applications, and to evaluating their practical utility when fielded. It should however, be recognised that the application developed here addresses only part, albeit a central part, of the pipeline. There is still a gap between the unstructured information which appears in a case file and the structured input necessary to drive the program. In the above application this step relies on the skills of the case handlers, but there are other developments which could potentially provide support for this task, such as the tools developed by companies such as Kira Systems for contract analysis and lease abstraction.³¹ It is to be hoped that this kind of machine learning tool might provide support for this aspect of the task in future. Similarly, the interface is currently hand crafted and one off. It is likely that the process of developing a robust implementation from the animated specification provided by the Prolog program could benefit from tool support, such as that available from companies such as Neota Logic.³² Weightmans' initial approach was to use NIHL as a proof of concept to see if the academic theory could be applied to their real world problems. If it was successful, it was always Weightmans' intention firstly to extend the principles to other domains which they are now doing with consultancy from the Department of Computer Science at the University of Liverpool and, more importantly, to turn the ADF into a tool that could be used by case handlers on a day to day basis. This second aspect, on which they are also working, has involved collaboration with other software providers.

³¹See <https://kirasystems.com/>.

³²<https://www.neotalogic.com/>.

What has been described is essentially an exploratory study, but one which provides much encouragement and suggests directions for further exploration, and the promise that eventually robust decision support tools based on academic research will be used in practice.

Acknowledgements

The authors would like to thank Jim Byard, Partner at Weightmans, who acted as the lead expert on the project.

References

- [1] L. Al-Abdulkarim, K. Atkinson, S. Atkinson and T. Bench-Capon, Angelic environment: Demonstration, in: *Proceedings of the 16th International Conference on Artificial Intelligence and Law*, ACM, 2017, pp. 267–268.
- [2] L. Al-Abdulkarim, K. Atkinson and T. Bench-Capon, Accommodating change, *Artificial Intelligence and Law* **24**(4) (2016), 1–19. doi:[10.1007/s10506-016-9178-1](https://doi.org/10.1007/s10506-016-9178-1).
- [3] L. Al-Abdulkarim, K. Atkinson and T. Bench-Capon, A methodology for designing systems to reason with legal cases using Abstract Dialectical Frameworks, *Artificial Intelligence and Law* **24**(1) (2016), 1–49. doi:[10.1007/s10506-016-9178-1](https://doi.org/10.1007/s10506-016-9178-1).
- [4] L. Al-Abdulkarim, K. Atkinson and T. Bench-Capon, Statement types in legal argument, in: *Legal Knowledge and Information Systems – JURIX 2016: The Twenty-Ninth Annual Conference*, 2016, pp. 3–12.
- [5] L. Al-Abdulkarim, K. Atkinson and T. Bench-Capon, Noise Induced Hearing Loss: An Application of the Angelic Methodology, in: *Legal Knowledge and Information Systems – JURIX 2017: The Thirtieth Annual Conference*, 2017, pp. 79–88.
- [6] N. Aletras, D. Tsarapatsanis, D. Preoțiuc-Pietro and V. Lampos, Predicting judicial decisions of the European Court of Human Rights: A natural language processing perspective, *PeerJ Computer Science* **2** (2016). doi:[10.7717/peerj-cs.93](https://doi.org/10.7717/peerj-cs.93).
- [7] V. Alevén, Teaching case-based argumentation through a model and examples, PhD thesis, University of Pittsburgh, 1997.
- [8] K.D. Ashley, *Modeling Legal Arguments: Reasoning with Cases and Hypotheticals*, MIT Press, Cambridge, Mass, 1990.
- [9] K.D. Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*, Cambridge University Press, 2017.
- [10] K.D. Ashley and S. Brüninghaus, Automatically classifying case texts and predicting outcomes, *Artificial Intelligence and Law* **17**(2) (2009), 125–165. doi:[10.1007/s10506-009-9077-9](https://doi.org/10.1007/s10506-009-9077-9).
- [11] K. Atkinson, Introduction to special issue on modelling Popov v. Hayashi, *Artificial Intelligence and Law* **20**(1) (2012), 1–14. doi:[10.1007/s10506-012-9122-y](https://doi.org/10.1007/s10506-012-9122-y).
- [12] K. Atkinson and T. Bench-Capon, Relating the ANGELIC methodology and ASPIC+, in: *Proceedings of COMMA 2018*, IOS Press, 2018, pp. 109–116.
- [13] T. Bench-Capon, Neural networks and open texture, in: *Proceedings of the 4th International Conference on Artificial Intelligence and Law*, ACM, 1993, pp. 292–297.
- [14] T. Bench-Capon, Relating values in a series of Supreme Court decisions, in: *Legal Knowledge and Information Systems – JURIX 2017: The Twenty-Fourth Annual Conference*, 2011, pp. 13–22.
- [15] T. Bench-Capon, HYPO’s legacy: Introduction to the virtual special issue, *Artificial Intelligence and Law* **25**(2) (2017), 1–46. doi:[10.1007/s10506-017-9201-1](https://doi.org/10.1007/s10506-017-9201-1).
- [16] T. Bench-Capon, M. Araszkiwicz, K. Ashley, K. Atkinson, F. Bex, F. Borges, D. Bourcier, P. Bourguine, J. Conrad, E. Francesconi et al., A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law, *Artificial Intelligence and Law* **20**(3) (2012), 215–319. doi:[10.1007/s10506-012-9131-x](https://doi.org/10.1007/s10506-012-9131-x).
- [17] T. Bench-Capon and F. Coenen, Isomorphism and legal knowledge based systems, *Artificial Intelligence and Law* **1**(1) (1992), 65–86. doi:[10.1007/BF00118479](https://doi.org/10.1007/BF00118479).
- [18] T. Bench-Capon, G. Robinson, T. Routen and M. Sergot, Logic programming for large scale applications in law: A formalisation of Supplementary Benefit legislation, in: *Proceedings of the 1st International Conference on Artificial Intelligence and Law*, ACM, 1987, pp. 190–198.
- [19] D.H. Berman and C.D. Hafner, Understanding precedents in a temporal context of evolving legal doctrine, in: *Proceedings of the 5th International Conference on Artificial Intelligence and Law*, ACM, 1995, pp. 42–51.
- [20] L.K. Branting, Data-centric and logic-based models for automated legal problem solving, *Artificial Intelligence and Law* **25**(1) (2017), 5–27. doi:[10.1007/s10506-017-9193-x](https://doi.org/10.1007/s10506-017-9193-x).

- [21] P. Bratley, J. Frémont, E. Mackaay and D. Poulin, Coping with change, in: *Proceedings of the 3rd International Conference on Artificial Intelligence and Law*, ACM, 1991, pp. 69–76.
- [22] G. Brewka, Weighted Abstract Dialectical Frameworks, in: *Workshop on Argument Strength*, 2016, p. 9.
- [23] G. Brewka, S. Ellmauthaler, H. Strass, J. Wallner and S. Woltran, Abstract Dialectical Frameworks revisited, in: *Proceedings of the Twenty-Third IJCAI*, AAAI Press, 2013, pp. 803–809.
- [24] G. Brewka and S. Woltran, Abstract Dialectical Frameworks, in: *Twelfth International Conference on the Principles of Knowledge Representation and Reasoning*, 2010.
- [25] A. Caliskan, J.J. Bryson and A. Narayanan, Semantics derived automatically from language corpora contain human-like biases, *Science* **356**(6334) (2017), 183–186. doi:[10.1126/science.aal4230](https://doi.org/10.1126/science.aal4230).
- [26] K. Clark, Negation as failure, in: *Logic and Data Bases*, Springer, 1978, pp. 293–322. doi:[10.1007/978-1-4684-3384-5_11](https://doi.org/10.1007/978-1-4684-3384-5_11).
- [27] F. Coenen and T. Bench Capon, *Maintenance of Knowledge-Based Systems: Theory, Techniques & Tools. Number 40*, Academic Press, 1993.
- [28] S. Danziger, J. Levav and L. Avnaim-Pesso, Extraneous factors in judicial decisions, *Proceedings of the National Academy of Sciences* **108**(17) (2011), 6889–6892. doi:[10.1073/pnas.1018033108](https://doi.org/10.1073/pnas.1018033108).
- [29] J. Eigel and D. Chen, Can machine learning help predict the outcome of asylum adjudications? in: *Proceedings of the 16th International Conference on Artificial Intelligence and Law*, ACM, 2017, pp. 237–240.
- [30] J. Goodman, *Robots in Law: How Artificial Intelligence Is Transforming Legal Services*, Ark Group, 2016.
- [31] P. Johnson and D. Mead, Legislative knowledge base systems for public administration: Some practical issues, in: *Proceedings of the 3rd International Conference on Artificial Intelligence and Law*, ACM, 1991, pp. 108–117.
- [32] D.M. Katz, M.J. Bommarito II and J. Blackman, A general approach for predicting the behavior of the Supreme Court of the United States, *PLoS one* **12**(4) (2017), e0174698. doi:[10.1371/journal.pone.0174698](https://doi.org/10.1371/journal.pone.0174698).
- [33] M. Lippi and P. Torrioni, Argumentation mining: State of the art and emerging trends, *ACM Transactions on Internet Technology (TOIT)* **16**(2) (2016), 10. doi:[10.1145/2850417](https://doi.org/10.1145/2850417).
- [34] G. Marcus, Deep learning: A critical appraisal, 2018, preprint, [arXiv:1801.00631](https://arxiv.org/abs/1801.00631).
- [35] S. Modgil and H. Prakken, The ASPIC+ framework for structured argumentation: A tutorial, *Argument & Computation* **5**(1) (2014), 31–62. doi:[10.1080/19462166.2013.869766](https://doi.org/10.1080/19462166.2013.869766).
- [36] H. Prakken, An abstract framework for argumentation with structured arguments, *Argument and Computation* **1**(2) (2010), 93–124. doi:[10.1080/19462160903564592](https://doi.org/10.1080/19462160903564592).
- [37] R. Pressman, *Software Engineering: A Practitioner's Approach*, Palgrave Macmillan, 2005.
- [38] E.L. Rissland and M.T. Friedman, Detecting change in legal concepts, in: *Proceedings of the 5th International Conference on Artificial Intelligence and Law*, ACM, 1995, pp. 127–136.
- [39] T. Routen and T. Bench-Capon, Hierarchical formalizations, *International Journal of Man-Machine Studies* **35**(1) (1991), 69–93. doi:[10.1016/S0020-7373\(07\)80008-3](https://doi.org/10.1016/S0020-7373(07)80008-3).
- [40] M. Sergot, F. Sadri, R. Kowalski, F. Kriwaczek, P. Hammond and T. Cory, The British Nationality Act as a logic program, *Communications of the ACM* **29**(5) (1986), 370–386. doi:[10.1145/5689.5920](https://doi.org/10.1145/5689.5920).
- [41] R. Van Kralingen, P. Visser, T. Bench-Capon and J. Van Den Herik, A principled approach to developing legal knowledge systems, *International Journal of Human-Computer Studies* **51**(6) (1999), 1127–1154. doi:[10.1006/ijhc.1999.0300](https://doi.org/10.1006/ijhc.1999.0300).