The coming of age of legal technology

What will be needed to take Legal Tech to a new level?

Authors[1]

Abstract

Many useful applications have been developed over the years for legal research and documentation. Technical opportunities are more extensive than they were two or three decades ago. Legal sources are certainly more accessible now and more diverse. Documents no longer consist of fixed series of characters only, but may interact with their users. Computer applications for legal practice, part of the broader 'Legal Tech' concept, are gaining popularity amongst lawyers. It is, therefore, interesting to examine what the present possibilities of Legal Tech are now, and also what the future may hold.

Application types can be distinguished by the complexity ('intelligence') of the processing involved and by the degree of influence a user has on the output. Decision support systems and programmed decision systems can be quite intelligent, but differ in the degree of user input. For the fully intelligent programs that do not require much user input, there is the question of explainability. To 'feed' as well as assess these programs, jurimetrics research is necessary. Jurimetrics is the empirical, usually quantitative, study of law. By means of jurimetrics research, legal decisions can be analysed and predicted.

Given all this, can computers already take over decision-making in the field of law? Although building ('artificially intelligent'; 'robot-') applications containing self-learning algorithms is in itself possible these days, that does not mean these programs can match human decision-making or sufficiently explain and justify attained results. As it is a function of the law not only to build on existing legal dogmas but also to keep in step with developments in society, decisions may be needed that are essentially different from those taken in the past. Legal decision-making is a creative process that requires emotional skills. At present there are still technological limitations as well as numerous practical and theoretical problems to really replicate human decision-making. To overcome these, we argue that a new phase of technological development would be necessary, offering fundamentally new possibilities. For the foreseeable future, therefore, the conclusion must be that handing over legal decision-making to computers is not desirable.

[1] Authors affiliations

1. Introduction

The assertion that information technology, today and even more so in the future, induces significant changes to their work[2] will be stating the obvious for most lawyers. In the digital age, lawyers work differently. Computers are just as much an integral part of the office environment for lawyers, as for other professions. One aspect that stands out for lawyers in particular is the high degree to which legal documentation has been digitized. It all started with the digitalization of legislation and case law reports, but in the course of the years an increasing number of law guides as well as legal text books were added to the document collections available digitally. The digitization of the remainder of the 'paper' libraries that many law firms currently possess can be seen as a final step in these developments.

Digitizing legal sources and making these available by means of automatic retrieval systems already has a history of over 40 years. In many cases, it was the first substantive application of information technology lawyers encountered. And despite the – often temporary – popularity of other products of legal computer science in the past decades (such as automated document production and rule-based support systems), opening up legal information in digital format is probably still the most widely use legal IT application[3].

In this paper, we will first address developments that have taken place with respect to these systems for searching and retrieving legal information. As a result of the opportunities offered by information technology, the term 'document' these days is used in a new way. When embedded in software, documents can undergo a transformation, from fixed, straightforward texts to interactive, 'dynamic' or 'modelled' content.[4] The 'reader' can answer questions from the document and can often also ask questions him or herself. That is why we will also discuss some other types of applications, such as legal knowledge-based systems and advice systems that are based on results from 'big data' research, for instance on case law data. From that it will become clear that digitizing legal documentation is an essential prerequisite for the realization of other, future applications. It forms the basis for the development of *Legal Tech*.

2. Digitizing legal sources

Digital legal sources have already been around for a long time. A first information retrieval system for statutory law texts was constructed as early as 1960, by Prof. John Horty of the University of Pittsburgh.[5] It was, however, not very user friendly. The system was 'batch oriented': the user had to ask questions, which were not answered immediately but only much later (for instance: during the night), after which the results could be picked up and

[2] Szostek, Dariusz, 'The Concept of Legal Technology (LegalTech) and Legal Engineering', in: *Legal Tech*, Nomos Verlagsgesellschaft mbH & Co. KG 2021, p. 24.
[3] Bues, Micha-Manuel and Emilio Matthaei, 'LegalTech on the rise: technology changes legal work behaviours, but does not replace its profession', in: *Liquid Legal*, Springer, Cham 2017, p. 98.
[4] Lauritsen, Marc, 'Current frontiers in legal drafting systems', working paper for the *11th International Conference on AI and Law*, July 2007.Lauritsen 2007, p. 9.
[5] Bing, Jon (Ed.), *Handbook of Legal Information Retrieval*, New York: Elsevier Science Inc. 1984, p. 257-260.

reviewed. In the years that followed, more user-friendly commercial applications emerged, such as Lexis (by Mead Data Central) and Westlaw (by West Publishing Company) in the USA. Lexis became more extensive as it later included data from other jurisdictions, for instance from the United Kingdom, France, Australia and New Zealand. A drawback in those early years was that submitting queries to and retrieving documents from databases like these required the use of a telephone connection, using a 'modem', which in itself already made the process slow and expensive. With the increased price users had to pay for data retrieved from the system, it is perhaps not surprising that many lawyers did not embrace this novelty right away, although it did gain popularity amongst legal academics.[6]

The popularity of these systems grew quickly in the 1980s, when many legal databases became available on CD-ROM disk. During this period, governmental organizations, universities and larger law firms increasingly used digital legal sources. The number of actual users in the early days remained relatively stable, as the data collections were mainly accessed from computers in libraries, but then grew strongly as the internet age developed. Currently, practically all legal sources – legislation, case law reports, journals, law guides, textbooks as well as most other legal literature – are available digitally and can be consulted via the internet. Some of these sources are made available by the government and by the judiciary and are 'open access' – free to be used by everyone. For the majority of 'regular' sources, a subscription obtained from one of the legal publishers is necessary, just as this used to be the case previously with respect to sources on paper. Nonetheless, options to conduct legal research using nothing but open access sources definitely exist and are becoming increasingly popular, which in turn acts as a stimulus for authors to publish their work in such open sources.[7]

Digitizing legal documentation comes with important advantages. Not only does the digital carrier save space (because of the highly increased information density, compared to paper), but it also introduces powerful new possibilities for the retrieval of the information. In the early days, search and retrieval systems for digital data used to be limited to keywords that were manually added to the content. Nowadays, there are many ways to search and retrieve data intelligently and automatically, which has dramatically improved the accessibility of the material and of the information contained in it. A document can now be retrieved based on practically any conceivable characteristic, including the words or sentences that are part of it. This is also the case for searches *within* a document, which opens the possibility to highlight and jump to relevant passages straight away. Furthermore, the digital format means that material can be accessed from any work spot, without the need to visit a library.

---

[6] Jones, Chris, Maria Zenios & Jill Griffiths, 'Academic Use of Digital Resources: Disciplinary Differences and the Issue of Progression', *Networked Learning 2004 Conference*, Lancaster, United Kingdom 2004, p. 5; see also Novak, Matthew S., 'Legal Research in the Digital Age: Authentication and Preservation of Primary Material', in: *The Marvin and Virginia Schmid Law Library*, 2010, p. 19 and Combrink-Kuiters, C.J.M., *Kennis van Zaken*, 1998, p. 75.

[7] Van Dijck, Gijs, 'Legal Research when Relying on Open Access: a Primer', *Law & Method,* April 2016.

The improved accessibility, by means of retrieval software, decreases the need for pre-selection by publishers. Even if the database contains lots of case law not matching the user's field of expertise, by refining the 'area of law' the unwanted 'hits' will quickly be removed from the list. The importance of traditional elements publishers used, such as a table of contents or a keyword index, has therefore diminished, which could be expected to lower the cost of publishing. However, adding documents to digital collections also places demand on the format, which can exceed that for paper publications. Usually, *metadata* need to be added, i.e., descriptive and/or technical data to enhance findability.[8] This provides users with the ability, for instance, to search for the newest additions (using the metadata field 'publication date') or to select data of a particular kind, for instance 'case law' or 'journal articles' (using the metadata field 'publication type'). Furthermore, added references to other digital documents, so-called *hyperlinks*, need to be kept up-to-date continuously in order to keep them functioning. Finally, retrieval, as well as filtering functionalities, need to be provided with the content in order to enable users to locate required documents quickly and precisely. Unfortunately, the direct costs of all this are usually not less than those for publishing on paper only. However, as digital content usually has a longer life span, because maintaining digital archives requires less effort and is therefore cheaper than maintaining archives full of paper documents, the total revenue per digital document could still be higher compared to that for publications, typically books and periodicals, on paper.

3. New options for digital content

Given the developments described in the previous section, in general the possibility to access digital documents through a retrieval system is seen as a major advantage, or even a necessity.[9] Still, there can be points of concern here as well. Not all retrieval systems have a user interface that is sufficiently understandable without specialist training, available options may differ and one publisher's retrieval system often only makes available their 'own' content. In the Netherlands, the market for legal information is relatively scattered. Lawyers usually not only have a need for sources from the large publishing houses, but also from smaller, 'niche' publishers as well as from public and governmental sources. Examples of the latter are unannotated case law, national and European legislation as well as official governmental publications.

In the past decade, this demand for diverse sources has been addressed by suppliers of so-called 'content integration systems'. In the Netherlands, two companies are active in that field[10], offering an online retrieval service that provides one-step access to the legal sources

[8] See for instance Kasdorf, William E. (Ed.), *The Columbia Guide to Digital Publishing*, New York: Columbia University Press 2003, p. 164.
[9] Margolis, Ellie & Kristen E. Murray, 'Say Goodbye to the Books: Information Literacy as the New Legal Research Paradigm', *U. Dayton Law Review* 38: 117, 2012.
[10] Legal Intelligence (www.legalintelligence.com) and Rechtsorde (www.rechtsorde.nl). Interestingly, these companies are owned by publishers Wolters Kluwer and Sdu/ELS, while their systems provide access to all sources, even those of their competitors.

of (practically) all publishers and to a large number of publicly available sources. The use of such additional services leads to extra costs, but to the user it also offers important advantages. The most noticeable of these is the uniform access to all sources, via one single search interface. Furthermore, the options for searching and for filtering results are more extensive than those provided by most publishers. This is in fact a necessity, because of the huge numbers of documents from the combined sources – often several millions. The addition of search support functions, such as a legal thesaurus containing legal terms, references and relationships between sources, introduces the option to retrieve, for instance, essential precedents by means of their popular name, even when that name is not present in the content itself. Most important, however, is that by combining all these sources it becomes possible to use them as one integrated – and interlinked – collection, a network of interconnected documents.

By using a retrieval system that offers access to more sources, a user in theory should be able to find a larger number of relevant results. However, this is certainly not always the case. The collection to be searched is larger, which puts higher demands on the methods to separate relevant from irrelevant information. Just like in search engines for the internet, such as Google, which can easily deliver thousands of hits following one single query, it is essential that in this list of hits the most relevant ones show up at the top. Yet even if that goal is achieved, and search engines pay a lot of attention to optimizing their 'ranking', it is practically impossible always to be able to view every document relevant to your query. A document can, for instance, be relevant because it is *related* to something you find, even though it does not (completely) comply with the original query. It might miss one or two of the search terms used, but nevertheless contains associated information important for the other results that were retrieved.

For that reason, it is essential that the *relationships* between documents are mapped, as much as is practically possible. By adding 'links' to related documents, in the form of metadata, whenever possible during pre-processing (i.e., every time a new document is added to the collection), the retrieval system will be able to use these data to improve search results, by adding related documents to the primary hits. Specifically, because such relationships often concern information from totally different sources, possibly from a completely different publisher, it is essential that as many of such sources are integrated, which is exactly what content integration systems aim to achieve. An example to illustrate this: imagine a lawyer is interested in recent developments in the field of corporate governance. In continental Europe, options to hold members of corporate boards, specifically of publicly traded companies,[11] accountable for managerial misbehaviour have been extended in the past two decades. A criterion for this 'directors' liability' that emerged from case law in the Netherland relates to the question if any 'serious reproach' can be made to the director. Our lawyer is aware of a key Dutch Supreme Court case in 1997, in which the Court specified valid grounds for accepting 'serious reproach'[12]. More recent

[11] 'Publicly traded' companies: companies that are listed and traded on a stock exchange. See Armour, John, B. Black, B. Cheffins & R. Nolan, 'Private enforcement of corporate law: an empirical comparison of the United Kingdom and the United States', *Journal of Empirical Legal Studies*, 6 (2009), p. 689.

[12] Dutch Supreme Court 10 January 1997, ECLI:NL:HR:1997:ZC2243 (*Staleman – Van de Ven)*.

cases will probably refer to this 1997 case. A legal retrieval system should be capable of showing all the more recent cases containing such references. One option to find these would be to perform a full text search using the case's known reference (for this Dutch example, that would be ECLI:NL:HR:1997:ZC2243). That would only work if that reference is present in a part of the document that is 'indexed' for full text search; for instance, in the 'body text'. If the reference is only present in a metadata field of a related document, the 'full text' search operation would probably fail to retrieve it. For that reason, it would be better if a dedicated process in the system would have mapped such relationships between documents in advance, taking into account all available metadata, making it possible to always show such *related documents* quickly and reliably.

This interconnecting of documents, based on mutual interlinks (references and citations) is one of the main advantages of the integration of document collections. When in a certain document B a reference to document A is found, this document A can be provided with a referral to document B as well. That way, 'networks' of connected documents of all kinds (case law, legislation, literature) can be formed automatically. By using these structures for making connected documents available together with primary search results (i.e. documents that contain all query terms the user entered), search results potentially become much more complete than would be the case when such information would not be used. Of course, an important requirement for this is that document connecting, based on references, is done in a reliable way.[13] Not every reference can be processed automatically. Notorious in this respect, for instance, are references to journal articles containing page numbers instead of article numbers. A page number, although relevant for 'paper' editions of magazines and books, has no real meaning for the digital editions that increasingly replace them, as these mainly consist of collections of separate digital documents. Given the importance of searching effectively, it can be expected that publishers will increasingly pay attention to such issues.[14] The increasing need to also refer to international sources, for instance within the European Union, could further complicate matters. The development of standards such as the European Law Identifier (ELI) and the European Case Law Identifier (ECLI) are an indication of growing awareness with respect to this, although the implementation of such standards in national information systems still provides challenges.[15]

[13] See for instance Van Opijnen, Marc, 'The European Legal Doctrine Identifier – a Missing Link?', in: Sebastiano Faro and Ginevra Peruginelli (Eds.), *La Dottrina Giuridica e la sua Diffusione*, Turin: G. Giappichelli Editore, 2017, p. 213-227.

[14] In this respect it is interesting to note that in the Netherlands 'Juriconnect', a specific national standard, was developed around 2007. It is currently used for referring to several types of legal documents and has been adopted by the Dutch government (for legislation), the judiciary (for case law) as well as most publishers of Dutch legal sources; see http://www.juriconnect.nl.

[15] See Filtz, E., S. Kirrane, & A. Polleres, 'The linked legal data landscape: linking legal data across different countries', in: Artificial Intelligence Law 29 (2021), p. 485–539.

4. Developments in documentation

Developments with respect to legal documentation have been highly influential to the law itself, and therefore also to the functioning of governmental organizations. Of course, law already existed even before the use of documentation, but the invention of writing and the registration of rules and decisions made possible by that caused many changes. It provided new opportunities to rulers to impose their will on their subjects and civil servants. Written law is *the* instrument of the legislature. The invention of the printing press made the exertion of power on an even larger scale possible. Printed law became specifically important as a means to instruct civil servants. This meant that the executive power could grow to large organisations. The development of state bureaucracies, whether functioning in autocratic regimes or constitutional democracies, would not have been possible without printing press technology. After all, only if all or nearly all people that are subject to legal rules are able to know what these rules are, those states can function.

But what is the effect of digital information technology? Does that just lead to upscaling? Is the internet simply a faster version of printed documents? We would argue it is more than that, mainly because digital documents can be essentially different from fixed writing. Information technology not only introduced new communication options, but also automatic processing. Computer software can be used to perform calculations, to deliver information *depending on* a user's input. Software is (potentially) interactive. It is this interactive aspect that alters a document from being fixed and passive to being dynamic as well. The user poses a question, the document provides an answer. Such documents, for instance in the form of 'knowledge-based systems' or 'decision support systems', have developed rapidly, such as in public administration.[16] They are applied at all levels and in many areas to 'help' civil servants and citizens to apply as well as to explain legal rules and to foster compliance with these rules. It is not an overstatement to claim that these new developments are radically changing the nature of the law and the constitutional state.[17] When legislators are capable of formulating rules that can be applied without, or with very little, human intervention, the legal and political power of the legislative authorities will increase substantially. The same is true for agents of the administrative power, who can also create software capable of automatic decision-making that will be applied without further human intervention. The option to use such technologies, capable of applying the law interactively as well as on a large scale, has put pressure on the 'Trias Politica' and has already created a need for a fourth, monitoring power (Ombudsperson, Parliamentary Commissioner for Administration, Court of Audit, etc.). In the next section, the form software like this could take will be discussed, as well as its potential and limitations.

[16] See for instance the examples mentioned in Juan Gustavo Corvalán, Juan Gustavo, 'Digital and Intelligent Public Administration: transformations in the era of artificial intelligence', in: *A&C-Revista de Direito Administrativo & Constitucional* 18.71 (2018), p. 55-87.

[17] De Mulder, R.V. & A. Meijer, 'From Trias to Tetras Politica: The Need for Additional Checks and Balances. Illustrated by Immigration Law.', in: I. Snellen, M. Thaens & W. van den Donk (eds.), *Public Administration in the Information Age: Revisited*, Amsterdam: IOS press, p. 36-45.

5. Types of computer applications; 'Legal Tech 1.0, 2.0, 3.0'

Traditionally, computer science distinguishes Transaction Processing Systems, Decision Support Systems and Programmed Decision-Making Systems.[18] This distinction, on the one hand, is based on the level of processing complexity, or 'intelligence', a program has and, on the other hand, on the extent to which users can influence the program's output. Transaction Processing Systems merely record information (for instance on the ownership of certain real estate), are not intelligent and provide the user with only limited means to control the system's output and the impact it has in the real world. Decision Support Systems leave decision-making to the user, but may offer complex processing that requires intelligence. Programmed Decision-Making Systems take everything out of the user's hands. Once the input is available, the system will make the decisions even if the users have no idea of the computations necessary for this. These days there are already many examples of these systems. In government administration, there are systems that decide about tax endowments to citizens on the basis of all the available data such as address, education, income and employment. Finally, in this summary of computer applications, are there any computer applications that are not intelligent, but do leave most of the control to the user? Actually, there are: word processing applications are an example of that category. The user does all the work, the software just records it in a form also chosen by the user.

Similarly, within Legal Tech several – existing or future – options could be distinguished. We will group these in three distinct development phases for this field.[19] Legal Tech 1.0, the first development phase, only comprises those applications that support human actors in the current legal system.[20] The second phase, Legal Tech 2.0, provides 'disruptive' innovation because part of the work of human actors in the system is replaced by the technology. Automated contract drafting is one example. Software (expert system-like applications) in that case contains rules to make that possible. This type of technology can also be applied to help non-legally-trained consumers solve their legal problems.[21] In these first two phases, the existing legal infrastructure is left unaltered.

Phase 3.0[22] of Legal Tech disrupts the legal world even further. New computer applications can take over existing practice to a high degree. If tax regulations are programmed in computer code (which in fact represents a Programmed Decision-Making System), tax administration would be easier. Documents are provided with intelligence and interactivity and could, for instance, keep track of who creates, alters and consults them. Legal robots – or 'agents' – programmed as self-learning systems, might eventually be capable of taking

[18] Sharkey, Ultan & Thomas Acton, 'Innovations in information systems from transaction processing to expert systems', in: *Proceedings of the Global Conference of Innovation and Management, Wuhan* 2012.
[19] Szostek, Dariusz, 'The Concept of Legal Technology (LegalTech) and Legal Engineering', in: *Legal Tech*, Nomos Verlagsgesellschaft mbH & Co. KG 2021, p 19-28.
[20] Goodenough, Oliver R., *Legal Technology 3.0,* in Huffpost.com, April 6, 2015.
[21] In 'The Rise and the Fall of the Legal Expert System', (in: *European Journal of Law and Technology*, Vol. 1, No 1 (2010)) Philip Leith argues that the idea that the logic of for instance a decision tree could properly represent reality is still popular, but definitely incorrect.
[22] According to Goodenough (2015), in this phase the technology not only does the work, it substantially replaces the legal infrastructure. Contracts, compliance systems, and dispute resolution systems are created in computer code, rather than natural language, and are able to operate within their own encoded systems.

legal decisions autonomously, as some authors wish to make us believe.[23] Online Dispute Resolution would then be an even more attractive proposition providing creative solutions.[24]

However, for all three phases of Legal Tech, but specifically for phase 2 and 3, extended knowledge about the actual functioning of the legal system and of the (human) actors in it is of vital importance. Unfortunately, that type of knowledge is often incomplete, or even lacking altogether. That is why research aimed at increasing that knowledge is, in our view, closely connected to Legal Tech and in many cases even vital for its successful application in legal practice. This requires what is called 'jurimetrics' research. Some examples of that type of research will be given in the next section.

6. The need for jurimetrics research

Legal computer applications can only come into being as a result of research and development. Research performed by lawyers (partly using new skills) is a prerequisite to gather the knowledge that is necessary for the new applications and to study and assess the effects.

Jurimetrics is the empirical study of the law. For traditionally trained lawyers, assessing statements empirically is not common practice. Instead, statements by legislators, judges and legal commentators are analysed 'hermeneutically', aimed at their interpretation. This type of interpretation is of a non-empirical kind: it expresses norms, 'what ought to be', rather than 'what is'. Results of that process become part of the system of demands and authorizations, issuing from state organizations, and embodied in legal rules. Jurimetrics, on the other hand, aims at making factual, empirically verifiable statements using mathematical models.[25] It studies all aspects of the law: form[26], meaning and pragmatics (i.e., establishment and functioning). Jurimetrics is 'about' the law, it is not part of it.

Jurimetrics research covers quite a broad area[27]. This type of research is vital to develop new applications. Without jurimetrics, predicting legal decisions, whether by judges or otherwise, is not possible. In order to develop software based on existing judgements that is capable of analysing which factors led to those decisions, and then in turn is able to predict future cases based on those factors, legal and methodological expertise is necessary. Jurimetrics provides the required competences.

---

[23] See for instance Marchant, Gary E., 'Artificial intelligence and the future of legal practice', in: *ABA SciTech Law* 14.1 (2017), p.22. A more balanced vision can be found in for instance Markovic, Milan, 'Rise of the Robot Lawyers', in: *Arizona Law Review,* 2019, 61: 325 and in Dobrev, D., 'The human lawyer in the age of artificial intelligence: Doomed for extinction or in need of survival manual', in: *Journal of International Business and Law*, 18(1) (2018), p. 39-68.

[24] The 'cognitive computer technology' developed by IBM as part of the Watson project, applied in the field of law by for instance the company Ross Intelligence Inc., is an example of this.

[25] See for instance Lave, C.A. & J.G. March, *An Introduction to Models in the Social Sciences,* New York 1975.

[26] See for instance Van Noortwijk, C. *Het Woordgebruik Meester (Legal Word Use)*, Lelystad: Vermande 1995.

[27] De Mulder, R., C. van Noortwijk & L. Combrink-Kuiters, 'Jurimetrics Please', in: *European Journal of Law and Technology*, Vol 1, Issue 1, 2010.

Knowledge of the substantive law is not the same as knowledge of the practical application of legal rules in society. To be successful in legal computer science, knowledge of legal reality is a necessity. Computer science as a tool for an acting human being – acting effectively, i.e. achieving the intended effect for concerned parties in the real world – is only possible if it is based on reliable knowledge about reality. An important aspect of jurimetrics research is the *analysis*, and based on the results of that analysis, the *prediction*, of court rulings. This research has tremendous potential for advisory applications, as well as for decision support and for programmed decision-making.

A 'basic recipe' for analysing and predicting court rulings could involve the following steps[28].

1. Choose a legal domain, for which case law will be analysed.
2. Choose a legal question (or 'legal item') that belongs to that domain. Formulate the question in such a way, that it can be answered by 'yes' or 'no'.
3. In case law reports, search for court rulings (of a particular type or level) in which a decision is made about this legal question.
4. In all cases that were found, search for factors (factual, legal or combined) that could possibly have played a role in the decision.
5. For each selected case, determine which of the identified factors played a role and which did not (this is the step of 'coding' the factors).
6. For each factor, calculate to what extent (i.e., how strongly) it has influenced the decision (in other words, what is the 'weight' of the factor).
7. Based on these factor weights, calculate the prediction for a new case.
8. As a validity check, for each of selected cases it can be assessed to what extent that case could have been predicted from the other cases in the set.

In this approach, the focus is on determining the relationship between (the presence of) certain factors and the final decision that is made regarding the 'legal item'. There will probably be factors that influence this decision considerably, while other factors might be of little or no influence. To establish the impact that a single factor *F* has, a sufficiently high number of cases needs to be included in the analysis. For each of these cases, the value – or 'weight' – of each single factor must be established, together with the decision (positive or negative) reached in the case. Based on these value pairs (factor – decision), a suitable measure of association between each factor and the final decision – a 'correlation coefficient', such as the Pearson PMC – can be calculated. A correlation coefficient is a number that indicates the strength of the connection between two factors or phenomena.[29] Factors that show a strong, or at least significant, correlation, either positive or negative,

[28] The method is explained in detail in De Mulder, R.V, C. van Noortwijk & L. Combrink-Kuiters, 'Jurimetrics Please', in: *European Journal of Law and Technology*, Vol 1, Issue 1, 2010.
A good example of such research can be found in the dissertation of Combrink-Kuiters (C.J.M. Combrink-Kuiters, *Kennis van Zaken; een jurimetrisch onderzoek naar rechterlijke besluitvorming inzake voogdij en omgang*, with a summary in English, Deventer: Gouda Quint 1998).

[29] Correlation Coefficients typically have a value in the range [-1..+1], with -1 representing perfect negative correlation (if the factor is absent, the decision will always be positive) and +1 perfect positive correlation (always a positive decision if the factor is present). The zero point, exactly between these two extremes, signifies the absence of any correlation between factor and decision.

could play a role in predicting the final decision for a new case. The reliability of this prediction can be expected to increase with the number of factors that show such significant correlations.[30]

7. Machine Learning

Research as described in the previous section is relatively uncomplicated, but the manual coding of factors, given the need for a sufficiently high number of cases in the research set, can be very time consuming.[31] A significant improvement of the process, at least regarding efficiency, would be achieved when factors could be identified and 'coded' automatically. In order to achieve the automation of this coding process, a computer algorithm would be necessary that is able to discern, for each case in the dataset, if a certain factor is present in it or not (and possibly even to what extent). As factors need to be identified from the running text of case reports, it would be virtually impossible to provide the algorithm with all possible word formulations that could represent the different factors.

The use of methods from the field of Artificial Intelligence, specifically 'machine learning', could be a solution to that problem. These algorithms can be developed along several different lines. The first option is to make use of what is called *supervised machine learning*[32], where the algorithm needs to be 'trained' before it can be applied: it has to learn how to carry out the task. Specifically, a method is needed to teach the algorithm when, in a decision, a certain factor is present or not. The Codas-algorithm[33] (which was originally developed as part of an application for automatic classification and/or ranking of legal documents) achieves this by supporting the interactive compilation of a set of example documents (in this case, containing a particular factor) as well as a set of counter examples (in which the factor does not play a role). This use of examples and counter-example builds a 'Bayesian Network', aimed at predicting the probability of a particular judgment in a new case.[34] . The use of examples and counter examples establishes a partial classification of the set of documents, which forms the basis for the training of the algorithm. Basically, when trained human assessors succeed in deciding upon the presence or absence of a factor in texts of court rulings, a suitably trained algorithm in principle should also be capable of achieving the same. Automating the coding process means that data needed for calculating

[30] In Combrink-Kuiters 1998, percentages of up to 96% of correctly predicted cases where found, an improvement of 20 to 40 percent points compared to the a-priori probability for the same cases (Combrink-Kuiters, C.J.M., *Kennis van Zaken; een jurimetrisch onderzoek naar rechterlijke besluitvorming inzake voogdij en omgang (On the Case; A jurimetrical study into judicial decision making in cases concerning custody and access)*, Deventer: Gouda Quint 1998, p. 281).

[31] At [institution] and elsewhere, since around 1990 dozens of students have performed this type of research.

[32] Yao, Mariya, Marlene Jia and Adelyn Zhou, *Applied Artificial Intelligence*, New York (USA): Topbots 2017, p. 12-19.

[33] This algorithm is described in more detail in for instance L. Combrink-Kuiters, R.V. De Mulder & C. van Noortwijk, 'CODAS as a Tool for Jurimetrical Research', in: *Proceedings of the 16th Annual BILETA conference,* Edinburgh: British and Irish Legal Technology Association 2001.

[34] See Pearl, J., 'Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning' (UCLA Technical Report CSD-850017), in: *Proceedings of the 7th Conference of the Cognitive Science Society*, University of California, Irvine, 1985, CA., p 330.

the probability of a certain decision in a new case can be gathered and used more quickly, and with considerably less effort, compared to manual coding.

If preselecting example documents for some reason is not possible, it might even be possible to revert to a different type of machine learning: *unsupervised learning*. This method aims at discovering structures, or patterns, in data that are in themselves unstructured, i.e., data that have *not* been labelled – in whole or in part – as belonging to a certain class, for instance, the class of documents containing a specific factor. In other words, no initial training by a human assessor is required. A common application of unsupervised learning is *clustering*, where input data is classified in certain groups, depending on some form of similarity between the separate items (such as documents, or sets of measurements). It is different from the forms of supervised learning described above because the exact number of classes that will be formed is not known in advance and also the exact contents of each class are not predefined. Using different forms or measures of similarity might lead to different classes being formed, and also to documents being assigned to more than one class. By correlating the final decision in a case with the classes it has been assigned to by the clustering algorithm, a model for predicting new cases could be built.

Another example of the use of unsupervised machine learning, again with respect to sets of legal cases, would be to estimate the 'importance' of each separate case by classifying the cases in, for example, five groups (ranging from 'very important' to 'not important'). The classification could be derived from, among other things, the number of times the case is cited in other cases and maybe in legal literature, if that is part of the database as well.

Clustering algorithms are sometimes used iteratively. For instance, in a first-round sets of two documents that are sufficiently similar are combined into one document pair, at the same time removing the separate documents. In the next round, similarities between the round1-pairs and remaining single documents is used again to combine these into larger sets, and so on. In itself this can be a powerful method, but one risk is that if the number of iterations is too high, certain successful clusters might grow out of proportion and eventually gobble up all the remaining documents and eventually even all the other clusters. This result does not provide much information. The algorithm should apply rules to avoid that happening, for instance by raising similarity thresholds when clusters grow larger.[35]

Supervised as well as unsupervised machine learning have their drawbacks: respectively the time and effort to select suitable training data and the lack of control of the direction a process might head in. This has led to the development of certain hybrid forms such as 'semi-supervised learning' and 'active learning'. A common characteristic of these hybrid forms is that classifications already present in the data are used as a starting point, even if such classifications are incomplete or contain errors. The user is queried about inconsistencies that become apparent, or asked for other input, with the aim to diminish errors by improving the classifications present in the data. Although this often works quite well in practice – examples are online shops that recommend products you might also be

[35] See Van Noortwijk, C., *Het Woordgebruik Meester (Legal Word Use)*, Lelystad: Vermande 1995, p. 245-263.

interested in, as well as legal search systems that point out documents related to your query but lacking one or two of the query terms – systems for supervised learning probably still create the most business value.[36]

8. The importance of explaining decisions to users

Our own decision support systems JURICAS[37], created in the 1980s, was definitely interactive as it would produce output directly depending upon the input of the users. The resulting decisions were, however, all pre-programmed by the authors of the system. Since then, developments in Artificial Intelligence have made it possible to deal with more complicated, intelligent tasks. This means, among other things, that the authors of the software do not need to know in advance what precisely the decision-making tasks of the users will be. Such applications, specifically those using forms of machine learning, have enjoyed a spectacular growth in the past decade.[38] This has prompted the need to understand how the results in such systems are obtained. Users feel that a system should be able to *explain* how it reached a certain decision. In other words, results should be interpretable.[39] Burkart & Huber (2021) mention the following reasons for explainability[40]:

- Trust – needed for users to accept the prediction model;
- Causality – by pointing out underlying relationships between input and output, users get a sense of causality;
- Transferability – when the user knows that the model generalizes well, it can be used for future decision-making (using yet unseen data);
- Informativeness – does the system function well in real-world situations, not merely in the context of its training data?
- Fair and ethical decision-making – people need to know the reasons for decisions made about them, in order to conform to legal and ethical norms and standards.
- Accountability – it should be possible to hold algorithms in decision-making processes accountable for their actions, which requires explanation and justification;
- Possibility to make adjustments – when insight exists in parameters of the decision process, these parameters can be adjusted to improve results;

[36] Yao, Mariya, Marlene Jia and Adelyn Zhou, *Applied Artificial Intelligence*, New York (USA): Topbots 2017, p. 137.
[37] See for instance Van Noortwijk, C., P.A.W. Piepers and J.G.L. van der Wees, 'The JURICAS-system in practice: decisions in a social security environment', in: C. van Noortwijk, A.H.J. Schmidt & R.G.F. Winkels (Eds.), *Legal knowledge based systems, aims for research and development*. Lelystad: Koninklijke Vermande 1990, p. 79-86.
[38] Wirtz, Bernd W., Jan C. Weyerer & Carolin Geyer, 'Artificial Intelligence and the Public Sector—Applications and Challenges', in: *International Journal of Public Administration*, 42:7 (2019), p. 596.
[39] Biran, Or, and Courtenay Cotton, 'Explanation and justification in machine learning: A survey', *IJCAI-17 workshop on explainable AI (XAI)*, Vol. 8, No. 1, 2017, p. 1.
[40] Burkart, Nadia, and Marco F. Huber, 'A survey on the explainability of supervised machine learning', in: *Journal of Artificial Intelligence Research* 70 (2021), p. 249.

- Proxy functionality – explainability of a system or process can serve as a proxy for characteristics that relate to it, but are difficult to quantify, such as safety, fairness, privacy and robustness.

With respect to the forms of machine learning described in the previous sections, interpretability is not always easy to establish. Even with supervised machine learning, where the algorithm bases its conclusions on – often operator-provided – training data, the number of characteristics, or 'features', from these training data that are used in the calculations can be very high. This makes it hard to establish the role of individual features, while the model applied to these data can also be complicated.[41] Consequently, providing an interpretation of the methods that are used and the data that are processed could still fail to provide a sufficient explanation of the results of the algorithm.

A concept related to interpretability is that of justification. It might be useful if a system could compute and thus check if the conclusion or the result is good, in the sense of 'correct', or 'conforming to certain quality standards', even if it cannot explain exactly how that conclusion or result was reached. Justification could be possible even for non-interpretable systems.

Early examples of AI applications, such as the rule-based expert systems that were even used in the field of law[42], represented popular first attempts to craft 'intelligent' software applications. These already applied some form of explanation. At first, this was mainly achieved by incorporating a 'tracing' facility– showing which rules had been applied, in what order, on the way to the final result. Although useful for debugging purposes, tracing often does not clarify the ratio behind the rules themselves and behind the inference process that was responsible for selecting and applying them. Therefore, its use in properly explaining the outcome of a system to laymen is often limited. When this was recognized, attempts to incorporate forms of justification quite soon emerged.[43] This justification could be based on underlying domain models, but also on separate communication layers present in the system. In general, in (rule-based) systems like these, justification was often based on processed information.

Justifying the outcome of systems that base their results on a form of machine learning, however, is a different matter. A machine learning model is often capable of calculating and reporting numbers that illustrate 'model confidence', for instance in the form of the probability that the model outcome is correct given the training data. It is also possible that graphs can be produced that show the relationships between the 'features' involved, or their relative importance. In practice, however, these technical options to justify model performance are often perceived as less convincing than simple, easy to establish facts, such

[41] Adadi, Amina, and Mohammed Berrada, 'Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)', *IEEE access* 6 (2018), p. 52145.
[42] Examples of such systems can for instance be found in Susskind, Richard E., *Expert Systems in Law*, Oxford: Clarendon Press 1987.
[43] See for instance Swartout, William R., 'XPLAIN: A system for creating and explaining expert consulting programs', in: *Artificial intelligence* 21.3 (1983): 285-325.

as the strongest feature playing a role in the model.[44] Information about such a feature might be combined with information about the user's past choices, such as '77% of the outcomes you selected previously contained this feature' and about those of similar users.[45]

Take as an example the application of supervised machine learning to locate case law reports containing a particular factor[46]. What would be needed to explain the output of such a process? That output would probably consist of a set of documents, each of which is supposed to contain the case factor the researcher is interested in. Interpreting the output by reporting the full list of features – individual words, possibly also combinations of words – as well as the algorithm used by the process might not provide the information the researcher is looking for, specifically if the number of features used in the model is high and the algorithm is complicated. An approach aimed at *justifying* the output by reporting only a limited selection of features, for instance those with particularly high weights or with relatively low frequencies within each document, might be more useful. Another possibility would be to inspect individual documents with highest and possibly also lowest overall probabilities, verifying if the case factor is present in them or not, with the option to add such documents to the sets of positive or negative training documents and subsequently recalculating scores. The latter, iterative, method in fact is an essential element in the Codas software mentioned previously.[47]

The use of AI methods has become increasingly popular. The number of developers trained to apply them continues to grow and the number of AI 'toolkits', available commercially or 'open source', is huge. Given this growth, an increasing demand for explainability should not come as a surprise. The way in which that demand can best be met can vary. For legal applications, solutions aimed at justifying results, specifically if these are not in the first place technical but adapted to the legal issue at hand, currently seem to be the most promising in this respect.

9. To what extent can AI replace legal decisions made by humans?

Arguably, every legal decision is a creative action, based not only on written knowledge about the legal subject in question and its societal implications, but on a legal 'feeling', a being in tune with society[48]. This is obvious when cases concern issues that have never

[44] Herlocker, J. L., Konstan, J. A., & Riedl, J., 'Explaining collaborative filtering recommendations', in: *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, ACM 2000, p. 245.
[45] Papadimitriou A, Symeonidis P, Manolopoulos Y., 'A generalized taxonomy of explanations styles for traditional and social recommender systems', in: *Data Mining and Knowledge Discovery* 2012, 24(3), p. 555-583.
[46] This example corresponds to what was described in section 7: to find all documents in a set containing a particular factor, by providing examples and counter examples of such documents.
[47] See Combrink-Kuiters, C.J.M., R.V. De Mulder & C. van Noortwijk, 'CODAS as a Tool for Jurimetrical Research', in: *Proceedings of the 16th Annual BILETA conference*, Edinburgh: British and Irish Legal Technology Association 2001.
[48] Similar to Oliver Wendell Holmes Jr. in the American Common Law, for Dutch lawyers it was the scholar Paul Scholten who put forward that legal decisions are not just logical (Huppes-Cluysenaer *et al.*, *General Method of Private Law, English Translation of the First Chapter of the General Volume of the Asser-series on Dutch Civil*

come up before. This could be because of technological changes in society (e.g. the big data companies that have changed the way people interact with each other and purchase goods and information) or because of other radical changes (such as dominant ideas about same gender marriage or ideas about using the legal system in a manipulative or disruptive way by powerful actors). But even in cases that do not seem to differ a lot from earlier ones, legal decision-making requires more than just logical reasoning. In Dutch criminal law for example, a judge is even explicitly required not only to examine the evidence and the rational implication thereof for the decision to declare the suspect guilty, but also to be 'internally convinced' of the guilt of the perpetrator[49]. This requirement of emotional involvement and creativity is particularly clear when a legal decision has to be made by a court, that would imply to reverse, or overrule, or break with, or radically deviate from previous decisions made by other courts or itself.

Algorithms, as described in the previous section of this article, are typically incapable of reasoning 'outside the box'. To teach an algorithm what it is that causes a new case to be essentially different from all the previous ones, a different type of learning process is necessary. Theoretically speaking, a whole data set of judgements would be required for that, judgements in each of which eventually a new, overruling decision was taken. This set would then have to be compared to rulings that were never reversed by decisions in subsequent, similar cases. The algorithm would then need to learn what characteristics of the texts of decisions are an indication of the probability (possibly increasing in time) that a fundamental change in case law is on the way. Constructing suitable data sets for this would not be easy. It is further complicated by the large number of different jurisdictions (sovereign states). The number of cases referring to developments leading to fundamental changes in decisions is usually rather too low for statistical use. In other words: it would probably be difficult to find sufficient training data to effectively teach algorithms to recognize the need for a fundamentally new path. For that reason alone, the moment an algorithm can detect the need of a significant new development has not yet arrived[50].

An interesting question is whether clues for developments like these could even be found in legal texts at all. We would argue that this possibility does exist. One example would be when a Supreme Court explicitly questions whether a previous interpretation should be overruled, after which this change is implemented in one of the next rulings. This does not imply, however, that information from case law sources will ever suffice to reliably predict a radical change in case law. Furthermore, although we have argued that legal knowledge

---

Law, Written by Paul Scholten, DPSP Annual, II: New Translations, Volume 1 (2020), p. 306-434 comprises a recent translation in English of his seminal work; see for instance par. 26: 467-470).

[49] For this rule the French terms: "conviction intime" versus "conviction raisonnée" are the local parlance.

[50] In science, this subject of "newness" has been very interestingly dealt with by Thomas Samuel Kuhn (*The Structure of Scientific Revolutions*, 3rd edition, Chicago, IL: University of Chicago Press 1996). Legal historian Helen Gubby has investigated the legal and societal factors that were involved in the legal rules for patents in the early phase of the Industrial Revolution in England (Gubby, H., *Developing a Legal Paradigm for Patents: the attitude of judges to patents during the early phase of the Industrial Revolution in England (1750s–1830s)*, Rotterdam: Erasmus University 2011). Hartendorp shows how legal decisions are made by judges in everyday practice (Hartendorp, R.C., *Praktisch gesproken, alledaagse civiele rechtspleging als praktische oordeelsvorming*, Rotterdam: Erasmus University 2008).

should have an empirical, jurimetrics, base, it is not evident that such knowledge could be laid down in a form that computers can process in an appropriate way.

Finally, we would like to argue that handing over legal decision-making to computer programs, however intelligent, would be ill-advised. There are so many factors that do and should influence legal ruling that will not be dealt with by computers. There are creativity, intuition, empathy and other emotions, general knowledge, wisdom, common sense, kindness of heart and most likely a lot of other skills and faculties that are involved and should be involved in legal decision-making. Automatic application of most of these factors would require a phase of technology we have not yet reached.[51] Human decision makers are still invaluable in legal decision-making, although the availability of analytical results to support their decisions will be of benefit in many cases.

Conclusion

Legal computer science has become excellent at documenting primary legal sources and literature. Current *Legal Tech* has also proved to be valuable as a tool in decision support, whether with respect to decisions concerning advice, policy making or even in court. At all these levels, there are options to apply learning algorithms. Specifically forms of supervised machine learning, for instance capable of finding patterns in sets of example data and of predicting results based on those patterns, have shown promising results. A prerequisite would be the incorporation of functionality to explain and justify the results of such applications.

However, what still has to be done more in creating legal software and training effective learning algorithms is jurimetrics research. Empirical knowledge for *Legal Tech* applications is in general desirable and necessary.

The addition of AI methods and tools to Legal Tech applications as well as the use of jurimetrics research will gradually increase possibilities. However, the wider functionality of decision-making encompasses elements such as availability, speed, communication, transparency, human understanding and empathy, understanding of society and the ability to anticipate the future effects of a decision, which places restrictions on the application of *Legal Tech*. While there are, within the field of law, certainly functions for which the use of technology will be increasingly necessary or at least desirable, there are functions that can still be performed better, or only, by people.

Computer software, with or without learning algorithms, that can reliably point out that a decision should ‘definitely be taken differently’ is still no more than science fiction. Such software, we would like to conclude, would also not be desirable. It lacks too many aspects that are necessary or at least desirable for legal decision-making. Two of those would be

[51] For an explanation of the phases of technology referred to here, see R. De Mulder, *Een model voor juridische informatica* (A Model for the Application of Computer Science to Law), Lelystad: Vermande 1984, p. 95.

creativity and wisdom. Therefore, human decision makers are still indispensable. Nevertheless, the proper use of *Legal Tech* will help them to do their job in the challenging environment of a rapidly changing society.

References

Adadi & Berrada 2018

Adadi, Amina, and Mohammed Berrada, 'Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)', *IEEE access* 6 (2018): 52138-52160, https://ieeexplore.ieee.org/abstract/document/8466590.

Armour *et al.* 2009

Armour, J., B. Black, B. Cheffins & R. Nolan, 'Private enforcement of corporate law: an empirical comparison of the United Kingdom and the United States', *Journal of Empirical Legal Studies*, 6 (2009), pp. 687-722, https://doi.org/10.1111/j.1740-1461.2009.01157.x.

Bing 1984

Bing, J. (Ed.), *Handbook of Legal Information Retrieval*, New York: Elsevier Science Inc. 1984.

Biran & Cotton 2017

Biran, Or, and Courtenay Cotton, 'Explanation and justification in machine learning: A survey', *IJCAI-17 workshop on explainable AI (XAI)*, Vol. 8, No. 1, 2017, http://www.cs.columbia.edu/~orb/papers/xai_survey_paper_2017.pdf.

Bues & Matthaei 2017

Bues, Micha-Manuel, and Emilio Matthaei, 'LegalTech on the rise: technology changes legal work behaviours, but does not replace its profession', in: *Liquid Legal*, Springer, Cham 2017, p. 89-109 https://link.springer.com/chapter/10.1007/978-3-319-45868-7_7.

Burkart & Huber 2021

Burkart, Nadia, and Marco F. Huber, 'A survey on the explainability of supervised machine learning', in: *Journal of Artificial Intelligence Research* 70 (2021): 245-317, https://www.jair.org/index.php/jair/article/view/12228.

Combrink-Kuiters 1998

Combrink-Kuiters, C.J.M., *Kennis van Zaken; een jurimetrisch onderzoek naar rechterlijke besluitvorming inzake voogdij en omgang (On the Case; A jurimetrical study into judicial decision making in cases concerning custody and access)*, Deventer: Gouda Quint 1998.

Combrink-Kuiters *et al.* 2001

Combrink-Kuiters, C.J.M., R.V. De Mulder & C. van Noortwijk, 'CODAS as a Tool for Jurimetrical Research', in: *Proceedings of the 16th Annual BILETA conference*, Edinburgh: British and Irish Legal Technology Association 2001,

https://www.bileta.org.uk/wp-content/uploads/CODAS-as-a-Tool-for-Jurimetrical-Research.pdf.

Corvalán 2018

Corvalán, Juan Gustavo. 'Digital and Intelligent Public Administration: transformations in the era of artificial intelligence', in: *A&C-Revista de Direito Administrativo & Constitucional* 18.71 (2018), p. 55-87, http://www.revistaaec.com/index.php/revistaaec/article/view/857.

Dobrev 2018

Dobrev, D., 'The human lawyer in the age of artificial intelligence: Doomed for extinction or in need of survival manual', in: *Journal of International Business and Law*, 18(1) (2018), p. 39-68, https://heinonline.org/HOL/P?h=hein.journals/jibla18&i=47.

Van Dijck 2016

Van Dijck, Gijs, 'Legal Research when relying on open access: a primer', *Law & Method,* April 2016, DOI: 10.5553/REM/.000017, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2654144.

Filtz *et al.* 2021

Filtz, E., Kirrane, S. & Polleres, A., 'The linked legal data landscape: linking legal data across different countries', in: *Artificial Intelligence Law* 29 (2021), p. 485–539, https://link.springer.com/article/10.1007/s10506-021-09282-8.

Goodenough 2015

Goodenough, Oliver R, 'Legal Technology 3.0', in: *Huffpost.com*, April 6, 2015, https://www.huffpost.com/entry/legal-technology-30_b_6603658.

Gubby 2011

Gubby, H., *Developing a Legal Paradigm for Patents: the attitude of judges to patents during the early phase of the Industrial Revolution in England* (1750s–1830s), Rotterdam: Erasmus University 2011, http://hdl.handle.net/1765/22239.

Hartendorp 2008

Hartendorp, R.C., *Praktisch gesproken, alledaagse civiele rechtspleging als praktische oordeelsvorming*, Rotterdam: Erasmus University 2008, http://hdl.handle.net/1765/12202.

Herlocker *et al.* 2000

Herlocker, J. L., Konstan, J. A., & Riedl, J., 'Explaining collaborative filtering recommendations', in: *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, ACM 2000, pp. 241-250, https://dl.acm.org/doi/abs/10.1145/358916.358995.

Huppes-Cluysenaer *et al.* 2020

Huppes-Cluysenaer, Liesbeth, Termorshuizen-Arts, Marjanne, Steer, Cassandra, Scholten, Paul, *General Method of Private Law, English Translation of the First Chapter of the General Volume of the Asser-series on Dutch Civil Law, Written by Paul Scholten,*

DPSP Annual, II: New Translations, Volume 1 (2020), p. 306-434, https://paulscholten.eu/research/article/general-method-of-private-law/.

Jones, Zenios & Griffiths 2004
Chris Jones, Maria Zenios & Jill Griffiths, 'Academic Use of Digital Resources: Disciplinary Differences and the Issue of Progression', *Networked Learning 2004 Conference*, Lancaster, United Kingdom 2004, 8 pp., available via https://telearn.archives-ouvertes.fr/hal-00190271.

Kasdorf 2003
Kasdorf, William E. (Ed.), *The Columbia Guide to Digital Publishing*, New York: Columbia University Press 2003.

Kuhn 1996
Kuhn, Thomas S., *The Structure of Scientific Revolutions*, 3rd edition, Chicago, IL: University of Chicago Press 1996.

Lauritsen 2007
Lauritsen, Marc, 'Current frontiers in legal drafting systems', working paper for the *11th International Conference on AI and Law*, July 2007, http://www.capstonepractice.com/s/CurrentFrontiers.pdf.

Lave & March 1975
Lave, C.A. & J.G. March, *An Introduction to Models in the Social Sciences*, New York 1975.

Leith 2010
Leith, Ph., 'The Rise and the Fall of the Legal Expert System', in: *European Journal of Law and Technology*, Vol. 1, No 1 (2010), https://ejlt.org/index.php/ejlt/article/view/14.

Marchant 2017
Marchant, Gary E., 'Artificial intelligence and the future of legal practice', in: *ABA SciTech Law* 14.1 (2017), https://www.americanbar.org/groups/science_technology/publications/scitech_lawyer/2017/fall/artificial-intelligence-and-future-legal-practice/.

Markovic 2019
Markovic, Milan, 'Rise of the Robot Lawyers', in: *Arizona Law Review*, 2019, 61: 325, https://heinonline.org/HOL/P?h=hein.journals/arz61&i=340.

Margolis & Murray 2012
Margolis, Ellie. & Kristen E. Murray, 'Say Goodbye to the Books: Information Literacy as the New Legal Research Paradigm', *U. Dayton Law Review* 38: 117, 2012, https://heinonline.org/HOL/P?h=hein.journals/udlr38&i=123.

De Mulder 1984
De Mulder, R.V., *Een model voor juridische informatica (A model for the application of computer science to law)*, Lelystad: Vermande 1984, http://hdl.handle.net/1765/1218.

De Mulder & Meijer 2012

De Mulder, R.V. & A. Meijer, 'From Trias to Tetras Politica: The Need for Additional Checks and Balances. Illustrated by Immigration Law.', in: I. Snellen, M. Thaens & W. van den Donk (eds.), *Public Administration in the Information Age: Revisited*, Amsterdam: IOS press 2012, p. 36-45, doi: https://doi.org/10.3233/978-1-61499-137-3-36.

De Mulder, Van Noortwijk & Combrink-Kuiters 2010

De Mulder, R.V, C. van Noortwijk & L. Combrink-Kuiters, 'Jurimetrics Please', in: *European Journal of Law and Technology*, Vol 1, Issue 1, 2010, https://ejlt.org/index.php/ejlt/article/view/13.

Van Noortwijk *et al.* 1990

Van Noortwijk, C., P.A.W. Piepers and J.G.L. van der Wees, 'The JURICAS-system in practice: decisions in a social security environment', in: C. van Noortwijk, A.H.J. Schmidt & R.G.F. Winkels (Eds.), *Legal knowledge based systems, aims for research and development*. Lelystad: Koninklijke Vermande 1990, p. 79-86.

Van Noortwijk 1995

Van Noortwijk, C., *Het Woordgebruik Meester (Legal Word Use)*, Lelystad: Vermande 1995, with a summary in English, http://hdl.handle.net/1765/1433.

Novak 2010

Novak, Matthew S., 'Legal Research in the Digital Age: Authentication and Preservation of Primary Material', in: *The Marvin and Virginia Schmid Law Library*, 2010, 4, p. 19-25, https://digitalcommons.unl.edu/lawlibrary/4.

Van Opijnen 2014

Van Opijnen, Marc, 'The European Legal Doctrine Identifier – a Missing Link?', in: Sebastiano Faro and Ginevra Peruginelli (Eds.), *La Dottrina Giuridica e la sua Diffusione*, Turin: G. Giappichelli Editore, 2017, p. 213-227, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3089176.

Papadimitriou *et al.* 2012

Papadimitriou A, Symeonidis P, Manolopoulos Y., 'A generalized taxonomy of explanations styles for traditional and social recommender systems', in: *Data Mining and Knowledge Discovery* 2012, 24(3), p. 555-583, https://link.springer.com/article/10.1007/s10618-011-0215-0.

Pearl 1985

Pearl, J., 'Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning' (UCLA Technical Report CSD-850017), in: *Proceedings of the 7th Conference of the Cognitive Science Society*, University of California, Irvine, 1985, CA. pp. 329–334, http://ftp.cs.ucla.edu/pub/stat_ser/r43-1985.pdf.

Sharkey & Acton 2012

Sharkey, Ultan, and Thomas Acton, 'Innovations in information systems from transaction processing to expert systems', in: *Proceedings of the Global Conference on Innovations and Management*, Wuhan 2012, http://hdl.handle.net/10344/3172.

Susskind 1987

Susskind, Richard E., Expert Systems in Law, Oxford: Clarendon Press 1987.

Swartout 1983

Swartout, William R., 'XPLAIN: A system for creating and explaining expert consulting programs', in: *Artificial intelligence* 21.3 (1983): 285-325, https://doi.org/10.1016/S0004-3702(83)80014-9.

Szostek 2021

Szostek, Dariusz, 'The Concept of Legal Technology (LegalTech) and Legal Engineering', in: *Legal Tech,* Nomos Verlagsgesellschaft mbH & Co. KG 2021, p 19-28, https://doi.org/10.5771/9783748922834-19.

Wirtz, Weyerer & Geyer 2019

Wirtz, Bernd W., Jan C. Weyerer & Carolin Geyer, 'Artificial Intelligence and the Public Sector—Applications and Challenges', *International Journal of Public Administration*, 42:7 (2019), p. 596-615, https://doi.org/10.1080/01900692.2018.1498103.

Yao, Jia & Zhou 2017

Yao, Mariya, Marlene Jia and Adelyn Zhou, *Applied Artificial Intelligence*, New York (USA): Topbots 2017.