European Review of Digital Administration & Law - Erdal
2026, Volume 7, Issue 1, pp. 61-79
ISSN 2724-5969 – ISBN 979-12-218-2846-7 – DOI 10.53136/9791221828467 08

*Context Matters: The Challenges of Aligning Legal Justification with Explainable AI**

Henrik Palmer Olsen
(Professor, Faculty of Law, University of Copenhagen)

Ragna Aarli
(Professor, Faculty of Law, University of Bergen)

Synne Sæther Mæhle
(Professor, Faculty of Law, University of Bergen)

Thomas Hildebrandt
(Professor, Department of Computer Science, University of Copenhagen)

Aysel Küçüksu
(Assistant Professor, Faculty of Law, University of Copenhagen)

ABSTRACT In view of the global interest in deploying artificial intelligence technology to enhance efficiency in public administration, one issue which seems to have received less attention is the importance of context sensitivity. A large part of public administration is regulated by law and concerns the legal rights and obligations of individual citizens. Legal decision-making in public administration takes place under the rule of law requirement, that such decisions must be accompanied by reasons that explain the decision. In this paper we show that justificatory explainability should not be confused with technological explainability (often referred to as XAI). Instead, justificatory explainability is a requirement that decisions in law are based on reasons that explain how the facts of a case relate to the legal rule that is applied in the case. The synthesis of fact and law in relation to the individual decision being made requires situational awareness of the context in which the rule operates. AI systems that are being introduced to interact with legal decision-making processes in public administration must be designed in a way which ensures that such context sensitivity is not erased from the decision-making process.

KEYWORDS: AI in public administration - Legal explainability - Justificatory reasoning in law - Context-sensitivity - Rule of law - Automated decision-making

TABLE OF CONTENTS: 1. Introduction: Justificatory Explainability and Situational Awareness. – 2. A Starting Point: The Need for Situational Awareness in Legal Decision-Making. – 3. The Black Box Problem and the Distinction between Causal Explanation and Justificatory Explanation. – 4. Legal Machines: Three Different Approaches. – 4.1. Fixed Rule Systems. – 4.2. Expert Systems. – 4.3. Large Language Models and Retrieval Augmented Generation. – 5. The Legal Requirement for Public Administration to Give Reasons. – 6. Elements in Justificatory Explanations for Legal Decisions. – 6.1. Identifying and Interpreting Legal Rules. – 6.2. The Role of Factual Information in Justifying Legal Decisions. – 7. (X)AI for Justificatory Explainability. – 7.1. Fixed Rule Systems. – 7.2. Expert Systems. – 7.2.1. Case-Based Reasoning Systems. – 7.2.2. Rule-Based Reasoning Systems. – 7.3. Large Language Models and Retrieval Augmented Generation. – 8. Conclusion: The Need for Further Research on How AI Can Support Justificatory Explainability.

1. *Introduction: Justificatory Explainability and Situational Awareness*

Legal reasoning is the process of providing justificatory explanations for legal decisions. Rather than simply declaring a decision, most, if not all legal systems, generally require (either via formal legislation or established practice) that decisions in law must be accompanied by reasons that support and explain these decisions.[1] To do so, decision makers need to show what law has been used as basis for the decision and how it has been

*Article submitted to double-blind peer review.

We would like to thank the participants of the LEXplain kick-off Conference in Paris in May 2025 for their highly valuable feedback on earlier versions of this article. We would also like to thank Rikke Hedegaard Rasmussen for her help in formatting the submission.
This work received support from the Research Council of Norway under Grant Agreement No. 352412, project title 'LEXplain: Explainability requirements for AI used in legal decision-making'.

[1] There may of course be exceptions, for example to shield state security interests and other sensitive information i.e. note 2.

applied to the case at hand. This involves both an explicit identification of the legal rule(s) used and an explanation of how the rule(s) in question is seen to apply to the concrete factual circumstances in the case at hand.[2] Often this involves an interpretive process in which the (relevant) facts of the case are subsumed under pre-existing fact categories set out in the law (i.e. the legal description of the facts that give rise to an application of the legal consequences contained in the rule), and subsequently inferring the decision as a corollary from this legal qualification of the facts.

Reason-giving is widely considered as essential for justificatory explainability in a rule of law context.[3] Avoiding and tempering arbitrariness in legal decisions, as a way of upholding the rule of law is a central goal for reason-giving, thereby tempering executive power.[4] Providing reasons for a legal decision is also generally thought to improve the decision's quality in terms of thoroughness and persuasiveness. When required to provide justification, decision-makers need to engage the case in more depth, to produce express lines of arguments in support of the decision. Such extended reflection on the case is also prompted by the fact that decision-makers may be asked to defend the decision through instruments of administrative and/or judicial appeal.

Explainability requirements also have positive effects on governance more generally, i.e., beyond the justification provided in relation to individual decisions. Requiring accessible reasons for a legal decision also supports overall normative coherence in the legal system, thereby enhancing equality before the law. Such systemic benefits are important to the continued support for the legitimacy of government. Research also highlights that providing reasons promotes respect and empathy for the individuals affected by a legal decision, as they are more clearly seen as fellow citizens by the decision-maker, when they must justify their decision.[5] Last, but not least, reason-giving, when decisions are made public, creates governance transparency because it allows for probing and review of the reasoning that underlay the decision.[6] This in turn enables democratic debate about legal decisions, the law they rely upon and this in turn enhances law's democratic legitimacy and facilitates the possibility of critique and legislative change.

The benefits of providing justificatory explanations for legal decisions are many and obvious. But producing such justifications is costly because they require time consuming skilled legal labor to produce. In many jurisdictions throughout the world, governments are therefore looking to artificial intelligence systems as a possible way to make decision making less costly. In academic research there has also been a rise of work that seeks to find ways of leveraging the power of AI to various legal tasks.[7]

Yet, contemporary developments in AI, such as deep neural networks, have led to increasingly complex AI models that are not transparent to human cognition, thereby giving rise to a problem of intransparency, and hence explainability when used.[8] Moreover, these models also often suffer from a lack of robustness because they rely on statistical modelling of past events and do not have 'here and now' real world understanding.[9] Finally it

[2] A. Bibal, M. Lognoul, A. de Streel and B. Frénay, *Legal requirements on explainability in machine learning,* in *Artificial Intelligence and Law*, vol. 29, 2021, 149–169, https://doi.org/10.1007/s10506-020-09270-4

[3] M. Cohen, *The rule of law as the rule of reasons,* in *Archiv für Rechts- und Sozialphilosophie*, vol. 96, no. 1, 2010, 1–16. DOI: 10.25162/arsp-2010-0001; J. Waldron, *Thoughtfulness and the Rule of Law*, Cambridge, MA, Harvard University Press, 2023.

[4] M. Krygier, *The Rule of Law: Pasts, Presents, and Two Possible Futures,* in *Annual Review of Law and Social Science*, vol. 12, no. 1, 2016, 199–229, http://dx.doi.org/10.1146/annurev-lawsocsci-102612-134103.

[5] J. Waldron, *Thoughtfulness and the Rule of Law*, cit.; S. Ranchordás, *Empathy in the Digital Administrative State,* in *Duke Law Journal*, vol. 71, no. 1, 2022, 1341–1389.

[6] It should be noted that protection of privacy often prevents public transparency. States differ in their efforts to balance transparency against privacy protection, and there is even variety between different public agencies within the same state. Publishing anonymized decisions is a way to overcome privacy related hindrance to public access to decision practice, but such large-scale anonymization is far from established practice everywhere.

[7] R. Susskind, *The Future of the Professions: How Technology Will Transform the Work of Human Experts*, Oxford, Oxford University Press, 2022.

[8] E. Sarid and O. Ben-Zvi, *Machine Learning and the Re-enchantment of the Administrative State,* in *Modern Law Review*, vol. 87, no. 2, 2024, 371–397, https://doi.org/10.1111/1468-2230.12843; F. Pasquale, *The Black Box Society*, Cambridge, MA, Harvard University Press, 2016.

[9] P.H. Padovan, C.M. Martins and C. Reed, *Black Is the New Orange Orange: How to Determine AI Liability,* in *Artificial Intelligence and Law*, vol. 31, no. 1, 2023,

is not clear how such AI models can deal with the value alignment needed for legal reasoning: i.e. how legal interpretation and application tasks can be performed in a way that aligns with the value-laden intent behind enacted legal rules and basic legal principles.

In response to this so-called black box problem, a new research field – explainable AI (XAI) has emerged. XAI approaches group into different categories like example-based XAI (explaining by presenting similar examples), contrasting-based XAI (explaining via counterfactual or adversarial examples), and saliency-based XAI (explaining by identifying important elements of information in the input), all of which have their pros and cons.[10] These are all generic technical solutions to mitigate the black box problem. XAI however is still only emerging as a research field and there is little to no research on how generic XAI mitigation can be adapted to existing explainability requirements in law.[11] Furthermore, the larger *problematique* is that even if XAI can make the black box sufficiently transparent, such technology may still not support the kind of justificatory explainability that is needed to support the rule of law.

Our paper proceeds in the following way. In section 2 we rely on Martha Finemans vulnerability theory to explain why legal decision making is a practice that requires situational awareness. In section 3 we distinguish between causal and justificatory (normative) explanations and show that the relevant kind of explanation for legal decision-making is justificatory explanation. Section 4 describes three different approaches to automation and shows how different kinds of automation allow for different kinds of justificatory reasoning. Section 5 elaborates on what the legal requirement for public agencies to explain their decisions entails. This section uses Danish Law as an example, but many legal systems have similar legal requirements for public administration. Section 6 focuses on the role of facts in legal decision making with a focus on its interplay with justificatory explanations Section 7 concludes by analyzing the various advantages and challenges of the three approaches to automated processes in legal decision-making.

2. *A Starting Point: The Need for Situational Awareness in Legal Decision-Making*

Explanations in law need to be meaningful and robust in a democratic sense: They must reflect a situational understanding of the human life that they are about and be susceptible to critical scrutiny and when relevant, appealable to a higher instance.[12] A central theme that emerges as a result is that of human vulnerability – a key concern in many areas of administrative decision-making – that manifests to varying degrees within concrete situations and contexts.

In a bureaucratic environment, legal decision-makers exercise judgment within a framework of rules, categories, routines, and even collegial expectations. The individual caseworker does not personally decide how to define vulnerability, which vulnerabilities it will be relevant to consider or how specific vulnerabilities should affect the right or obligations of citizens. The applicable law determines this. However, it is the case worker who interprets and evaluates the specific situation considering these legal categories. This interpretive activity is grounded in a type of situational understanding or awareness, shaped not only by professional experience and institutional routines, but also by an intuitive, foregrounded sense of the citizen's particular circumstances.[13] In contrast, AI systems cannot grasp a concrete situation in the way humans do, because they are not embedded in the world and thus they possess a more functional, rather than existential, understanding of human life.[14] Inspired by Martha Fineman's vulnerability theory,[15] but

133–167.

[10] A. Madsen, S. Reddy and S. Chandar, *Post-hoc Interpretability for Neural NLP,* in *ACM Computing Surveys*, vol. 55, no. 8, 2022, DOI: 10.1145/3546577.

[11] K.M. Richmond, S.M. Muddamsetty, T. Gammeltoft-Hansen, H.P. Olsen and T.B. Moeslund, *Explainable AI and Law: An Evidential Survey,* in *Digital Society: Ethics, Socio-Legal and Governance of Digital Technology*, vol. 3, no. 1, 2024, https://doi.org/10.1007/s44206-023-00081-z.

[12] H.P. Olsen, J.L. Slosser and T.T. Hildebrandt (eds.), *Artificial Intelligence, Humans and the Law*, London, Routledge, 2025.

[13] H.P. Olsen, J.L. Slosser and T.T. Hildebrandt (eds.), *Artificial Intelligence, Humans and the Law*, cit.

[14] H.P. Olsen, J.L. Slosser and T.T. Hildebrandt (eds.), *Artificial Intelligence, Humans and the Law*, cit.

[15] M.A. Fineman, *The Vulnerable Subject and the Responsive State,* in *Emory Law Journal*, vol. 60, no. 1, 2010, 251; M.A. Fineman, *Vulnerability and Social Justice,* in *Valparaiso University Law Review*, vol. 53,

applying it with an emphasis on situational awareness, we argue that this gap poses a core challenge for explainable AI in legal decision-making. This is especially true for cases requiring discretion that are often precisely those where context-sensitivity, situational awareness, and a nuanced understanding of vulnerability matter most. An overarching aim for law and AI, therefore, must be to identify what kind of AI systems are better suited to support (but not replace) legal decision-making that requires situational judgment.

As man-made constructs, neither the law, nor algorithmic decision-making is entirely objective or neutral. Thus, any combination of law and algorithmic processing in the context of individual decision-making that does not probe either's fundamental building blocks may be problematic. When it comes to the legal literature, a number of heterodox streams have highlighted how, in an effort to subsume the lived human experience under pre-determined legal categories, legal rules might (mis)read reality,[16] perpetuate unrealistic ideas of the human existence,[17] or even challenge the ontological conditions that enable the exercise of agency and free choice.[18] These observations align with Martha Fineman's proposition to view vulnerability as a universal human experience, whose manifestation varies both between individuals and within the same individual over time.[19] Her theory sees vulnerability as an enduring and essential element of the human condition and stands in contrast to the view that vulnerability can be fully captured and circumscribed by pre-enacted law. Vulnerability then is a holistic condition of human life that will influence how specific situations are perceived in the 'now', including in the context of legal judgment. This condition must also be considered in the context of AI systems designed to support legal decision making. Early examples of AI adaptation in the context of law have shown that AI systems are not easy to integrate into a rule of law culture.[20] Famously, in the so-called SyRI case from the Netherlands, the use of an algorithmic social welfare fraud detection system was found by the Hague District Court to violate the European Convention of Human Rights, because it unfairly profiled people based on their place of living and their socio-economic background.[21] A similar use of an algorithmic system in US State of Michigan was found to be fraught with error, unjustly depriving thousands of citizens of unemployment benefits.[22] And in another AI system scandal, concerning fraud detection in relation to childcare benefits, the Dutch government was even led to resign.[23]

Our point is that Fineman's vulnerability theory is useful to support continued awareness of the need for situational awareness when making judgments in law, including (and perhaps particularly) when legal decision making is supported by AI.[24]

no. 1, 2019, 341.

[16] M.A. Fineman, *The Vulnerable Subject and the Responsive State,* in *Emory Law Journal*, vol. 60, no. 1, 2010, 251; A. Küçüksu, *Fineman in Luxembourg: Empirical Lessons in Asylum Seeker Vulnerability from the CJEU,* in *Netherlands Quarterly of Human Rights*, vol. 40, no. 3, 2022, 290–310, DOI: 10.1177/09240519221113463; A. Küçüksu and G. Ünüvar, *Vulnerability Theory as a Paradigm Shift in International Investment Law: Reimagining the Role of the State,* in *Business and Human Rights Journal*, vol. 8, no. 3, 2023, 309–328, https://doi.org/10.1017/bhj.2023.38.

[17] F. ES. Falster, *Ableisme – Magt, menneskesyn og myten om den normale krop*, København, Samfundslitteratur, 2024; P. Ringø and E.S. Falster, *Beyond the Buzzword: Unpacking 'Special Needs' in Denmark*, in *Scandinavian Journal of Disability Research*, vol. 27, no. 1, 2025, 173–185, DOI: 10.16993/sjdr.1060.

[18] S.A. Teo, *The Unbearable Likeness of Being: How Artificial Intelligence Challenges the Social Ontology of International Human Rights Law*, in *Cross-disciplinary Research in Computational Law*, vol. 2, no. 1, 2025, 1–25.

[19] A. Küçüksu, *Fineman in Luxembourg: Empirical Lessons in Asylum Seeker Vulnerability from the CJEU*, in *Netherlands Quarterly of Human Rights*, vol. 40, no. 3, 2022, 290–310, DOI: 10.1177/09240519221113463; Z. Godzimirska, A. Küçüksu and S. Ravn, *From the Vantage Point of Vulnerability Theory: Algorithmic Decision-Making and Access to Regional Human Rights Courts*, in *Nordic Journal of Human Rights*, vol. 40, no. 1, 2022, 235–249, DOI: 10.1080/18918131.2022.2078028.

[20] M. Zalnieriute, L. Bennett Moses and G. Williams, *The Rule of Law and Automation of Government Decision-Making*, in *Modern Law Review*, vol. 82, no. 3, 2019, 425–455, http://dx.doi.org/10.2139/ssrn.3348831; S. Deakin and C. Markou (eds.), *Is Law Computable? Critical Perspectives on Law and Artificial Intelligence*, London, Hart Publishing, 2020.

[21] S. Bekker, *Fundamental Rights in Digital Welfare States: The Case of SyRI*, in O. Spijkers and Vv.Aa. (eds.), *Netherlands Yearbook of International Law*, The Hague, T.M.C. Asser Press, 2021.

[22] S. Ranchordás, *Empathy in the Digital Administrative State*, in *Duke Law Journal*, vol. 71, no. 1, 2022, 1341–1389.

[23] *See also* Dutch Childcare Benefits Scandal, Wikipedia, https://en.wikipedia.org/wiki/Dutch_childcare_benefits_scandal.

[24] S.A. Teo, *The Unbearable Likeness of Being: How*

For the purposes of this article, we instrumentalize Fineman's concept of vulnerability in the concrete decision-making context by translating it to the idea of 'situational understanding', which some of us have previously defined as 'the ability to meaningfully relate facts and judgment of evidence' to legal reasons and to produce arguments specifically related to that situation.[25] As our article will show, explainability requirements related to decision-making in public administration has evolved to incorporate situational understanding in various steps along the way. Our aim is to explore ways in which we can preserve this type of context-sensitivity in the translation of justificatory explainability from the manual to the automated domain.

In our study, we have a special focus on what may be called mass administration, i.e. the type of administration where a public body decides a large number of cases about citizens' rights or duties applying the same legal provisions in every case. Typical examples of this are decisions about the right to welfare service, tax duties, land use permissions, etc. The emphasis on mass administration reflects our theoretical assumption that in many cases, there will be a universality to the way human vulnerability manifests and is experienced, which would in turn permit the articulation of certain standards around the bureaucratic practice of situational understanding.

3. *The Black Box Problem and the Distinction between Causal Explanation and Justificatory Explanation*

AI models using deep neural networks and other sub-symbolic approaches give rise to the intransparency (or black box) problem of not being able to inspect the inner workings of the neural AI model because it is hyper complex. Hence, given an input such a model will provide an output, but it will not be possible to inspect how the input was processed by the model to arrive at the output provided. This poses a problem in a scenario where such models are used to make or support legal decisions that need to be explainable. If an AI model is used to produce a decision and if it is not possible to derive from the model how the decision was arrived at, then such a model will be incompatible with the rule of law because it cannot be seen if and how the decision was based on law.[26]

Accordingly, much of the literature on black box AI has focused on the perceived need to 'open up' the black box.[27] The question we raise here is whether this is the right approach to achieve legal explainability. Our reason for raising this question is that we think that opening the black box of complex sub-symbolic AI will not provide the right kind of explanation for law. To support this view, we introduce a distinction between causal and justificatory explanations.

Causal explanations cite prior events (A) as causes for subsequent events (B) and relies on an underlying assumption that the relationship between the prior event and the subsequent event is necessary. B would not have happened if it were not for A. Moreover, causal explanations are related to natural events. To use Hans Kelsen's classic distinction between *sein* (is) and *sollen* (ought)[28] as a reference point: causal explanations are explanations that belong to the world of *sein*. Causality can explain relationships between various factual states of the world. In the context of computational processing of information, it makes sense to try to identify the processes whereby an information input is acted upon by a computational model to produce an output. The output on this view has the status of an effect which is caused by the input and the model. If we assume in this example that the input is some information from a case file and

Artificial Intelligence Challenges the Social Ontology of International Human Rights Law, cit., 1–25.

[25] H.P. Olsen, J.L. Slosser and T.T. Hildebrandt (eds.), *Artificial Intelligence, Humans and the Law*, Routledge, 2025.

[26] A related but different transparency problem relates to the lack of public access to information from public agencies. For an overview of this problem, see H.P. Olsen, T.T. Hildebrandt, C. Wiesener, M.S. Larsen and A.W. Ammitzbøll Flügge, *The Right to Transparency in Public Governance: Freedom of Information and the Use of Artificial Intelligence by Public Agencies*, in *Digital Government: Research and Practice*, vol. 5, no. 1, 2024, DOI: 10.1145/3632753.

[27] See for example R. Guidotti, A. Monreale and Vv.Aa., *A Survey of Methods for Explaining Black Box Models*, in *ACM Computing Surveys*, vol. 51, no. 1, 2018, 1–42, https://doi.org/10.1145/3236009 who makes a distinction between 'how' and 'why' a decision was made, says: '… just as it is often not straightforward to explain how an ADM [Automated Decision Making] system reached a particular conclusion, so it is also not straightforward to determine why that system reached that conclusion'.

[28] This distinction is found already in the first edition of H. Kelsen, *The Pure Theory of Law*, 1934.

that the output is a legal decision, then providing a full causal explanation for the legal decision would consist in describing how the input was processed by the model. But an explanation of how the computational machinery works, is different from providing a *justifying* explanation. Legal decisions are explained (justified) by normative *reasons* that support the decision and therefore belong to the realm of *Sollen*. Justifying reasons are different from mechanistic explanations of machinery because they are not (natural) causes of the event but are instead normative grounds that serve to legitimize the decision. Such reasons are normative prescriptions. The relationship between a set of facts and the legal reasons that apply to them on the one hand and a decision outcome on the other are therefore based on (deontic) logic and normative persuasion. Interpretation and inference are the instruments used to produce and validate (and critique) legal decisions – not the establishment of a causally connected series of events.

If we now come back to the various issues connected to justificatory explainability that we flagged in the introduction we can see that it is not a feature of being able to look into the minutiae of the inner workings of an AI model (its computation of vectorized representations of text, for example) that is required to justify legal decisions under the rule of law. Only a normatively meaningful form of legal reasoning will provide the justification needed and such reasoning is not immediately available from the high dimensionality computations that operate in deep neural networks.[29]

This does not mean that the quest for transparent computing should be abandoned. On the contrary, developing knowledge of the causalities that operate inside AI machines is indeed desirable and may contribute to technological solutions that can advance the possibility of critically assessing the reliability of such models and possibly even their alignment with legal values that reign in the specific domain of law in which it is used. But our point is that ultimately XAI in and of itself is not enough: Decisions in law must be based on valid legal reasoning, which infer decisions by subsuming the factual situation under an applicable rule, a process that sometimes involves discretion. Systems used to produce or support such decisions need therefore to be able to produce output that is legally meaningful to case workers and furthermore to be robustly aligned with the values underpinning legal regulation, to ensure reliability.

It is important, however, to also recognize that legal rules differ and that some rules are simple to apply, which is why a certain level of automation is already used in public administration. In the following section we take a closer look at the different kinds of AI systems and how they may support legal decision-making in public administration through automation.

4. *Legal Machines: Three Different Approaches*

Computational decision support can be based on different kinds of computational technology. We distinguish between three different forms of computational machinery: fixed rule systems, expert systems, and machine learning based systems.

4.1. *Fixed Rule Systems*

Fixed rule systems can be used for very simple and clear legal rules. Such systems mirror traditional views of formal inferences in law: A legal rule sets out certain conditions and effects such that when the factual conditions are in place the legal effect kicks in. Fixed rule systems operate on a strict 'if this, then that' logic. Here is a simple example:

'Every Danish citizen above the age of 68 who has permanent residence in Denmark has the right to a state paid old age pension' (legal rule). A is a Danish citizen above the age of 68 and has permanent residence in Denmark (fact). A has a right to state paid old age pension (legal effect).'

Contrariwise: if A is not above the age of 68 or does not have permanent residence in Denmark, A does not have a right to state paid old age pension.

In this simple scenario the legal rule combined with the factual information about A justifies the decision that a state paid old

[29] By high dimensionality we refer to the computational complexity in large language models. GPT 4 for example is estimated to operate with 1,8 trillion parameters Wikipedia, *GPT-4*, https://en.wikipedia.org /wiki/GPT-4 which function as statistic operators to compute inputs and generate outputs word by word. This is very different from how humans use concepts and theories to generate text (legal decisions).

age pension is awarded to A. And similarly, someone who does not meet that condition, for example because they are 67 years, is a German citizen or has permanent residence in Norway will not have that right. As long as the validity of the legal rule and the truth of the facts are not disputed, it will be possible to automate the administration of the rule because the justification follows immediately from the clear criteria of the legal rule as a matter of logical inference.

This kind of automation is relatively easy to set up. It can be made fully transparent by documenting the code and making it open for inspection. Because the legal rule and the corresponding fact situations that the rule applies to are very clear, such systems can meet basic criteria of inference-based justifiability. Systems of this kind are already used in public administration for automated decision-making in some areas of taxation and social welfare.[30]

Fixed rule systems then can be used for decision-making in cases where the law is unambiguous. Precisely because such systems can only be implemented to execute unambiguous legal rules, they have obvious limitations. Most importantly they are not capable of processing, more complex legal cases, where facts are less clear and/or where facts need to be assessed in relation to more open-ended legal provisions, such as when a statute has conditions with vague terms, or in cases where prior practice plays a role in identifying what the law is. In these situations, because the rules need interpretation, it is not possible to immediately infer an explainable decision from the 'legal rule-factual situation' compound. Even in the clear example mentioned above, discretion may sometimes be necessary. It could be, for example, that it is not clear where the person has permanent residence. Logic based systems, therefore also need to allow for manual case handling intervention, so that unclear cases can be dealt with outside the system.

4.2. *Expert Systems*

Expert systems are designed to mimic more complex problem-solving than what can be achieved with the fixed rule systems used for fully automated decision-making. We distinguish between two subtypes of expert systems: rule-based reasoning systems and case-based reasoning systems.

One example that we will return to in more detail below, is the use of rule-based reasoning systems. This is an interactive expert system, where a human user provides the facts that cannot be automatically computed and may also select among a number of multiple rules, which rule to apply in an intermediate reasoning step. The rules selected may be connected to additional information such as most recent case law or legal guidelines, for example as set out in a circular. Such a system helps the caseworker comply with the law.

Rule based reasoning systems can be implemented in a decision structure, for example with the use of a Dynamic Condition Response (DCR) Graph. The design tools for mapping of law texts to DCR Graphs use a combination of Natural Language Processing (NLP) techniques for highlighting actions, rules and roles in the text and user interfaces allow domain experts and lawyers to validate and correct the mapping. The execution tool provides an open Application Programming Interface (API) that can be used to either prompt a case worker to e.g. input a fact or a sub-decision or to progress in the inference by executing a rule for which all conditions are fulfilled (where fulfilment can be validated manually or by an AI agent).

Another example of an expert system is AI systems that perform so-called case-based reasoning system. Based on expert knowledge, it is possible to extract information from legal judgments and to use this information as relevant input to the process of decision-making in new cases. An early example of this approach is Kevin Ashley's system Hyper-Parameter Optimization (HYPO),[31] which was used to provide recommendations for litigators and

[30] See for example M. Zalnieriute, L. Bennett Moses and G. Williams, *The Rule of Law and Automation of Government Decision-Making*, in *Modern Law Review*, vol. 82, no 3, 2019, 425–455, http://dx.doi.org/10.2139/ssrn.3348831; M. Suksi, *Creating Legal Basis for Automated Decision-Making: A Review of Section 53 e of Finland's Administrative Procedure Act and Beyond*, in *European Review of Digital Administration & Law*, vol. 5, no. 2, 2024, 107–116, 2024 which emphasizes that only rule-based systems are allowed for automated decision-making in Finland.

[31] HYPO is famous for being the first computational legal expert system and has its own page on Wikipedia, *HYPO (CBR)*, https://en.wikipedia.org/wiki/HYPO_CBR. The system is described in more detail see K.D. Ashley, *Reasoning with Cases and Hypotheticals in HYPO*, in *International Journal of Man-Machine Studies*, vol. 34, no. 6, 1991, 753–796.

decision-makers in the field of trade-secret law. By annotating known judgments in the field (that is, by categorizing the operative elements in trade secret judgment text), a database is created that contains information about the relationship between the reasoning applied in each case and the outcome of the judgment. With the arrival of a new case, which is similarly annotated, the system can give recommendations about the most likely outcome based on the elements that count for or against a specific outcome. Since the model is rule-based and operates with recognizable legally relevant factors, a human decision-maker will be able to see how the recommendation was generated and would be able to use the elements from former decisions to construct a reasoned decision in the new case, thereby fulfilling the requirement to produce a justificatory explanation for the outcome. Case based reasoning systems are suited to provide a decision maker with a recommendation based on the weights of the various elements[32] and can serve as a basis for generating justificatory reasoning for the decision in a new case.

It should be noted that traditional Case-Based Reasoning (CBR) models rely on manual annotation. Such an approach is not very efficient for analysis of large datasets. Ashley's HYPO dataset consisted of 30 cases (judgments) and a total of 13 dimensions (elements in the reasoning of judgments that count for or against a specific outcome). Annotating 30 cases using a 13-dimension system is a manageable task, but when case law data reaches the 100's or 1000's and when the complexity of the argumentative process increases because the decisions have a more discretionary character or when decision criteria are not easy to identify in an unequivocal way, then manual annotation becomes less manageable. This is when machine learning technology becomes a more relevant approach.[33]

4.3. *Large Language Models and Retrieval Augmented Generation*

Machine learning is basically a method used to build computational models. Most recently it has been used to develop neural networks, which led to the large language model breakthrough in natural language processing, which is what we will focus on in this section. Neural networks are highly abstract representations of the data and data relationships that the model has been trained on. Neural networks are generally seen as intransparent because the units in the network do not correspond to anything that is recognizable to human cognition. These models are therefore not immediately explainable.

Large Language Models are developed by exposing an algorithm to a very large text corpora with the purpose of learning probabilities of relations between an input sentence and the most likely next word. This results in a model, which, given an input text (a prompt) – results in an output: the generation of a text, which is a response to the prompt. LLM based AI systems can thereby interact with humans in natural language. LLM's can also be adjusted (fine-tuned) to produce justificatory language, for example by being trained on a curated dataset of pre-existing sources of law (judgments and legislative text). LLM's are based on neural networks and therefore suffer from the same kind of inexplicability as other neural networks. Moreover LLM's are known to not always be reliable and trustworthy. Basic LLM's are therefore not well suited to automate legal decision-making.

Combining LLMs with a knowledge base (a digital law library), it is possible to construct Retrieval Augmented Generation (RAG) models, that can in a first step retrieve documents from the knowledge base using search engine techniques and then, in a subsequent step, generate an answer from the input obtained by combining a text encoding of the question with the retrieved documents. This goes some way towards remedying the problem that LLM's lack inherent meaningful reasons for the classification it provides, since the LLM can now include retrieved documents (for example legislative text, case law, administrative guidelines, etc) and process the case in the context of these documents to produce a text that will look like a justificatory reasoning for its classifications

[32] This can be compared to the internal rules of thumb, that are known to exist informally in many bureaucratic administrations.

[33] Ashley himself has recently gone in this direction, see J. Savelka, M.A. Gray, W. Oliver and K.D. Ashley, *Using LLMs to Discover Legal Factors*, in *Legal Knowledge and Information Systems*, vol. 395, 2024, 60–71, DOI: 10.48550/arXiv.2410.07504.

(proposed decisions). In other words, AI technology as it stands today can produce decision proposals in natural legal language.[34]

Still, however, RAG systems, even those sitting at the forefront of technological development, are not free from legal error or questionable reasoning. It is well documented that RAG systems may hallucinate[35] and thereby produce erratic errors in the proposed decision text. Moreover, RAG systems remain opaque in relation to how this text, which human agents can read as a justification, was produced and how it relates to the retrieved documents. We know through introspection how we ourselves (human agents) use reason to arrive at a justification for the decisions we make. We also assume that other human agents reason in the same way because we can have conversations with other human agents about those justifications and we can engage in critical back and forth discussion about the strengths and weaknesses of the justifications we provide. Such reasoning is also embedded in an institutional setting, which makes legal reasoning robust. It does not depend on one individual subject but is part of a communicative and interpretive community. When AI in the form of legal LLM/RAG is introduced into this community it raises questions about how such technology may integrate with the community in a way that preserves the trustworthiness of the institution as a whole.[36] This is a matter that remains to be explored in further depth, and which is becoming increasingly urgent as governments press for more uptake of AI in public administration.

5. *The Legal Requirement for Public Administration to Give Reasons*

Legal decision-making takes place in many different settings and the conditions that apply to these settings vary. Some decisions are made by the highest courts in a country on appeal from a lower court. The amount of time and effort put into these decisions from both litigating lawyers, judges and clerks is significant and may run into thousands of work hours for skilled lawyers working on the case over a long period of time. At the other end of the spectrum are routine decisions made in public administration, some of which may be decided on the basis of just one or two hours of work.

The variation in impact, institutional setting and time allocated to making legal decisions influences the way legal justifications are being produced in judicial and administrative practices. We focus on legal decision-making in administrative practice primarily in areas of mass administration. As mentioned before, we focus on the legal requirement that applies to administrative agencies in regard to their obligation to provide reasons to the recipients of their decisions. This obligation corresponds to a right for citizens to receive a justificatory explanation.

To establish a baseline for what kind of justification is required, we use Danish law as an example. In art. 24 of public administration act, it is stated that:

> '(1) The grounds for a decision must include a reference to the statutory rules relied upon in making the decision. To the extent that a decision based on such rules is discretionary, the grounds must also state the main considerations for the exercise of that discretion.
>
> (2) If required, the grounds must also include a brief account of the information on the facts of the case considered to be of considerable significance to the decision made.'

Similar requirement of the duty to give grounds/reasons for administrative decisions can be found, in one guise or another, in most

[34] Hence while we agree that LLM's, including legal RAG's (which is at root a machine learning technology) will necessarily 'create a number of black boxes for administrative decision making that result in decisions the reasons of which cannot be known to individuals and business enterprises' M. Suksi, *Creating Legal Basis for Automated Decision-Making: A Review of Section 53 e of Finland's Administrative Procedure Act and Beyond*, in *European Review of Digital Administration & Law*, vol. 5, no. 2, 2024, 107–116, we also wish to point out that if the generated text plausibly look like reasons that can be adopted by human decision-makers, such machines can still be a helpful support mechanism. As long as the machine proposed reasons are put in natural language, they are transparent in the sense that they can be tested for their plausibility by human decision-makers and they can be made subject to judicial review. Our concerns are more related to the robustness of this technology as it stands today.

[35] V. Magesh, F. Surani, M. Dahl, M. Suzgun, C.D. Manning and D.E. Ho, *Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools*, in *arXiv:2405.20362v1*, 2024.

[36] See for an analysis of this question: J.L. Slosser, B. Aasa and H.P. Olsen, *Trustworthy AI: a Cooperative Approach*, in *Technology and Regulation*, 2023, 58–68.

legal systems.[37]

The fact that administrative law itself, expressly and by an act of parliament, prescribes a duty for administrative agencies to produce meaningful explanations to the subjects of their decisions underscores the democratic importance of explainability in (administrative) law. Decisions that do not meet the explainability threshold set out in the Danish public administration act rarely (if ever) make it to court, but there is a rich Ombudsman practice, which elaborates on the requirements in terms of form, precision, extent etc., required of administrative agencies explanations for the decisions they make.[38] Issues such as the degree of involvement of the citizens in the decision-making process, the amount of discretion afforded to the administration and whether the decision involves weighing and balancing of competing interests will affect what is required to meet the explanation duty. The explanation requirement, then, really is a kind of legal standard, the application of which must be adapted to the kind of decision being made.

This explanation requirement set out in domestic administrative law is often overlooked in the AI and Law literature, where the concept of explanation is more frequently discussed in the context of the requirements under the General Data Protection Regulation (GDPR) in article 15 and/or the new AI Act article 86.[39] While these EU laws are relevant and important, domestic administrative law, like the Danish public administrative act is both broader in scope and involves more specific requirements that have been developed over a longer time through Ombudsman practice. We wish to therefore also highlight the importance of studying the requirements that flow from these domestic legal obligations and the need for AI Systems to build in a way that allows for compliance with these requirements.

6. *Elements in Justificatory Explanations for Legal Decisions*

What does it take to justify a decision in law? In the most basic sense, every decision must as a minimum set out 1) the rule(s) which form the legal basis for the decision, 2) the relevant facts on which the decision is made and 3) explicit and implicitly explain how the legal rules apply to the facts of the case.

6.1. *Identifying and Interpreting Legal Rules*

Every legal rule consists of two elements. The first element is a description of the conditional facts contained in the rule. This involves a description of the fact that must be present for the legal consequence of the rule to take effect. The second element of the legal rule is a description of the legal consequences of the rule. The legal consequence is the legal effect that follows from applying the rule to the facts.

It may seem straightforward to identify a legal rule and its two elements, but as Joseph Raz has pointed out, legal rules (or norms as he prefers) are in fact composed of elements of various sources of law.[40] A statutory rule is composed not only of the text of the article in question but must be read and understood in the context of the entire statute. The statute must be read and understood (and interpreted) in light of its legislative history. Statutes also sometimes overlap, leading to a need for inter-statutory interpretation to establish the content of the norm in question. Moreover, administrative and judicial practice also influences how statutory acts are understood and applied. To set out 'the rule(s) that form the legal basis for the decision' may therefore not be as straightforward as simply citing the text of an article in a statute: the entirety of relevant legal material must be viewed and interpreted as a whole when formulating the legal rule.

In mass administration, which we, as mentioned above, focus on here, there is however a high level of repetition and similarity between cases, which may provide for a situation of rule stability and clarity in most cases. For mass administration, therefore, automation can be very helpful. It can ease workload for case workers and provide an instrument for more homogenous

[37] H.P. Olsen, J.L. Slosser and T.T. Hildebrandt (eds.), *Artificial Intelligence, Humans and the Law*, Routledge, 2025.

[38] For an overview, see N. Fenger, *Forvaltningsloven med kommentarer*, Copenhagen/København, Djøf Forlag, Jurist- og Økonomforbundet, 2020.

[39] L. Metikoš and J. Ausloos, *The Right to an Explanation in Practice: Insights from Case Law for the GDPR and the AI Act*, in *Law, Innovation and Technology*, vol. 17, no. 1, 2025, 205–240.

[40] J. Raz, *The Concept of a Legal System*, 2nd ed., Oxford, Oxford University Press, 2003, 70.

decision quality. Fixed rule systems can be an option for automating the most simple rules, but in areas where societal developments influence legal interpretation over time[41] or in areas marked by more factual complexity, fixed rule systems may not be useful. Public administration in areas such as social welfare law, planning law, environmental law, etc. often involves some amount of discretion in order to capture the relevance of some facts for the application of law.

Establishing facts can be at least as complicated as establishing applicable law[42] and the law can never be fully understood in isolation from facts. Facts, moreover, are often messy and difficult to capture. Decision-makers may have to rely not only on information provided by documents from, for example, doctors, schools, employers, the police, social workers, neighbours, tax authorities, etc. (depending on the case) in order to piece together 'the facts of the case'. Those subject to the decision will also often have to be consulted to verify or contest the facts of the case. This complex process of establishing the facts of a case shows that legal systems are embedded in broader social contexts and reflects an understanding that situational awareness is required to determine the factual situation in individual cases.[43] For an AI system to be compatible with the need to determine the specific factual situation in each individual case, it has to be built in such a way that it allows for situational awareness to be integrated into the overall architecture of the decision-making process.

Not only must a decision maker correctly identify the law and the facts, but it should also be noted how law and facts depend on each other and interact to produce the final stage of the decision, where it must be shown how the legal rule applies to the facts of the case. This dependency is in one sense a matter of logic: If the facts of the case do not fit under the conditional facts as set out in the legal rule, then the legal rule does not apply and if the legal rule does not apply, then the legal consequence contained in the rule cannot follow. In other words, the final decision has to be able to be presented as a syllogism. But as every lawyer knows, logic alone is not enough. Legal decision-making requires that the individual facts of the case belong to the category of facts that the rule applies to. This requires interpretation of the law and often also of the facts.

A simple example from Danish criminal law may illustrate the point made above. Section 244 of the Danish criminal code sets out that 'Anyone who uses violence against, or otherwise attacks the body of, another person is punished with a fine or imprisonment for up to 3 years.' In a case from 2006 the Supreme Court had to decide whether spitting at another person could be considered as an attack on another person. To make that decision it is necessary to both interpret the meaning of the word 'attack' and the impact on another person of the act of being spat on. To arrive at a decision in the case, the judges had to connect the interpretation of law to a broader social understanding of spit and spitting.[44] In the following we shall dig deeper into the broader question of how facts shape legal decision making.

6.2. *The Role of Factual Information in Justifying Legal Decisions*

Justificatory reason must relate to the facts of the case which is under consideration and those facts must be seen to be true. Factual information can be considered the most crucial variable in administrative legal decision-making. Factual information reflects the diversity of life, and the fact of each case is in one way or another, up close always unique. What makes cases similar is an effect of the way law circumscribes and delimits the *relevant* facts for the purpose of applying the legal rule in question. There is then an initial phase in every decision-making process where it must be decided what the unique facts are,

[41] C. Harlow and R. Rawlings, *Proceduralism and Automation: Challenges to the Values of Administrative Law*, in E. Fisher, J. King and A. Young (eds.), *The Foundations and Future of Public Law*, Oxford, Oxford University Press, 2019.

[42] See S. Slettan and T.M. Øie, *Forbrytelse og straff. Lærebok i strafferett*, Oslo, 1997, 28: 'In real life, determining the facts often cause more problems than determining the law and decisions are of course determined by what facts are used as a basis (minor premise) in the legal reasoning that is used to justify the decision' (author's translation of original: '[i] det levende livet volder nok faktum vanligvis større problemer enn jusen. Ofte står og faller saken med hvilket faktum som legges til grunn').

[43] See section 2 above.

[44] The Danish Supreme Court found that spitting someone in the face constitutes violence under section 244. Danish Supreme Court judgment, available at: https://pro.karnovgroup.dk/b/documents/7000355484.

but also what facts are relevant and not relevant for the legal decision. Determining what constitutes facts and which facts are relevant may sound like two separate operations. However, as we will return to below, there is an inseparable connection between fact determination and legal rule construction that is not always straightforward to operationalize neither in procedural law nor in algorithmic language.

Mapping out the facts of the case is nevertheless an indispensable starting point for legal decision-making. When determining what constitutes the unique facts, a central question is who is able and responsible for collecting or documenting the actual conditions required by the law. In contemporary digitized public administration, information can often at least partly be drawn from public registers, and some decisions can be automatically delivered without case workers involved and without significant efforts for documentation from the citizen's side. More often, citizens will have to provide documentation themselves to prove that the actual conditions required by the rule are met. Case workers will then subsequently assess the quality of the provided information and decide whether the case is sufficiently clarified and validated for a decision to be made.

In present legal frameworks for handling administrative cases, a duty to provide guidance and a duty to ensure that the case is properly clarified before a decision is made follows from the duty for public administrators to conduct an investigation of the relevant facts before making a decision.[45] This means that case workers have an obligation to clarify what information the law requires and to request any missing documentation before issuing a decision. Facts are therefore often established in collaboration: Some of the facts may originate from the citizen herself (for example from an application form), other facts may originate from other public authorities, doctors or schools or from information held in public databases (for example address, tax information, social welfare service records etc). Only when legal rules are clear and no discretion or interpretation is needed can the collection of factual information be fully automated. In other situations, and in line with the principle of vulnerability and situational context sensitivity that we have introduced earlier, facts must be seen in the context of the life situation that is being regulated and cannot be reduced to computationally isolated units of information.

The collaborative context for fact-finding implies that the abstract factual descriptions that must be met for the legal effect to be applicable must interact with a multitude of individual perceptions of what such facts may involve. Legal texts are designed to communicate with such a multitude of users and are written with the use of general expressions for the actual conditions that must be met for a legal effect to occur. It is in other words not necessarily obvious what fact situations falls within and what situations falls outside such generically formulated terms.[46] Therefore even the initial process of mapping and identifying relevant facts requires an assessment of the scope and limitations of the legal requirements.

The interplay between facts and law in legal decision-making has aptly been pointed out from a principle-based perspective by Stephen Waddams:

> '…The facts of a case are defined in relation to legal principles, but the principles themselves are formulated in relation to facts, real or hypothetical. Facts may be stated at countless levels of particularity, and legal issues and legal rules may be formulated at countless levels of generality. No map or scheme could possibly classify all imaginable facts, for there is no limit to the number of facts that may be postulated of a sequence of human events. The selection of legally relevant facts is a matter not of empirical investigation but of judgment, and not wholly separable from the formulation of the applicable legal rule'.[47]

An overall organizing theory of fact-finding in legal contexts has been requested by one leading legal theorist,[48] but such a

[45] Referred to as 'undersøkelsesprinsippet'/ 'officialprincippet' in Norway/Denmark.

[46] For an illustration see the 'spit' example, *supra* note 18.

[47] Although S. Waddams, *Dimensions of Private Law: Categories and Concepts in Anglo-American Legal Reasoning*, Cambridge, Cambridge University Press, 2003, remarks is provided in the context of private law in an Anglo-American setting, we find that the observation is true for law in general.

[48] See e.g., W. Twining, *Taking Facts Seriously*, in *Law in Context*, Oxford, Clarendon Press, 1997, 89–109, requesting a coherent theory of evidence, proof and

theory does not exist as of today. The intricate relationship between facts and law remains aloof and difficult to handle in a legally unambiguous manner, be that with or without the help of artificial intelligence. However, when the interaction between the parties has already moved to a digital platform, there is little doubt that machines have the potential to be trained or programmed to explain what information is missing and can assist case workers in the decision-making process.[49] Requests for and selection of legally relevant facts are, from our point of view, tasks machines *can* perform, potentially in some situations also in a more consistent manner than a human being. There is nevertheless no doubt that there are several pitfalls associated with transferring even these kinds of initial judgments in the decision-making process to machines.

One issue would be whether the relevant facts being requested by the law are feasible for machine processing. E.g., when determining whether a certain disease may entitle a citizen to receive disability benefits under the Norwegian National Insurance Act, the factual concept of 'disease' in Section 12-6, second paragraph, is based on a definition that is 'scientifically based and generally recognized in medical practice'. This is not a definition that easily lends itself to programming. Laws are written in natural language that is often open ended because it is meant to interact with human understanding and social practice developing over time, but this also means that machines will have a limited capacity to process those laws, as they do not have (and cannot be encoded with) the background knowledge and experience needed to perform judgment in individual instances.[50] To avoid that decisions are based on outdated information, decision support systems should be build in ways that makes them updatable, for example via interaction modules that asks for human input or connect to external databases, allowing access to updateable information about e.g. scientifically based and generally recognized medical practice (to use the same example), to ensure that the fundamental requirement that legal decisions must be based on *true* information is met.

The information used by public agencies to justify their decision must not only be relevant but also *true*. The truth requirement means that the scope for using computerized 'predictions' based on proxy information is limited. It may e.g. be relevant for a legal decision whether a person is or is not cohabiting with another person. It would be tempting to draw information from the central housing register to determine whether the person is living alone or cohabiting with others at his or her address because this is usually a good representation of whether people do in fact cohabit. The problem is that the information in the central housing register does not always reflect the realities of everyday life. Even if it's an obligation to register information in the central housing register that reflects the truth about residence, the information may still turn out to not be true. Moreover, the concept of 'residence' may itself be unclear to a specific life situation. When there is a crisis in a marital relationship or when a spouse is stationed abroad for a period, the time spent 'at home' begins to decrease, but when exactly one no longer 'resides' at a specific address may be uncertain. In cases like these the proxy may not correspond to the true facts, and the true facts may not clearly fall within or without the legal category and hence a decision based on information from the housing register alone would result in an *unjustified* decision because it would not reflect true information about co-habitation. AI legal decision support, which use proxies to build 'predictions' should therefore be designed in a way that allows for unusual situations to be dealt with outside the system.[51]

If AI-based legal decision support systems make decisions that are interpretable and

fact-finding that potentially could be developed from fact-oriented sub-theories such as Bentham's theory of evidence and Wigmore's theory of proof.

[49] See for example a recent Danish legislative proposal (Danish Parliament, *Lovforslag L 152 (2024–25)*, available at: www.folketingstidende.dk/samling/20241/lovforslag/L152/20241_L152_som_fremsat.pdf.) which seeks to provide a legal basis for use of AI to assist caseworkers in assessing the need for additional documentation in cases regarding application for a special heightened educational support for handicapped students.

[50] For a more comprehensive philosophical analysis of the same point see J. Wang, *Is Artificial Intelligence Capable of Understanding? An Analysis Based on Philosophical Hermeneutics*, in *Cultures of Science*, vol. 4, no. 3, 2021, 135–146.

[51] This could be done by using default reasoning, see e.g., R. Reiter, *A Logic for Default Reasoning*, in M.L. Ginsberg (ed.), *Readings in Nonmonotonic Reasoning*, San Mateo, CA, Morgan Kaufmann Publishers, 1987, 68–93.

explainable to most people, traditional methods of handling expert knowledge can be used to justify the use of such models as part of the decision-making process. This has made explainable AI very attractive to the legal field.[52] Explainable or not, use of technology to streamline, simplify and improve case processing introduces potential new sources of errors also in the fact-finding process, e.g. based on bias in training data or faulty construction in the algorithms or errors in the input and processing of new data. Such errors are, if materialized, likely to be repeated multiple times and can be difficult to detect, thus representing new challenges for the rule of law which must be addressed before such systems are implemented.

Uncertainty regarding the relevance, sufficiency and truth of the factual circumstances that form the basis for a legal decision is not uncommon. Classic administrative law mechanisms for correcting errors, such as rectification, reversal, or withdrawal of decisions, illustrates both that it is a known problem and that legal mechanisms have been developed to address the issue. As these mechanisms are important elements of the rule of law, AI systems used to support legal decision- making should be built in a way that allows for these repair mechanisms to function

7. *(X)AI for Justificatory Explainability*

Having analyzed the function of and requirements related to justificatory explanations in law, we can now return to the question of how AI may support legal decision-making in a trustworthy way. In section 4 above we distinguished between three basic AI approaches to legal decision support.[53] We shall now take a closer look at how each of these three approaches may or may not support justificatory explainability for legal decision making.

7.1. *Fixed Rule Systems*

Fixed Rule Systems are based on logic rules that have been manually coded. Such systems can be seen as the most basic example of Symbolic AI, which was the dominant approach in early AI (i.e. during the 1950s-1980s). Manually coded rule systems are in use many places and work well for simple or very clearly defined decision-tasks. They can handle many tasks automatically and can also be set up to send cases that cannot be processed automatically (i.e. which do not fulfill the criteria for being subject to automated decision-making) to a human case worker. If the code is open for inspection and well-documented, such systems can in principle be made fully transparent and explainable and do not then create any challenge with regard to the need to provide justificatory explanations. The problem with this kind of system for legal decision making is however at least two-fold. First, the applicability of a legal rule often depends on information in the context of a concrete case that is not available to the automated system. Moreover as relevant contextual factors may evolve over times, fixed rule systems can become outdated. Secondly, some legal rules may have exceptions which are difficult to code correctly in traditional fixed rule programming languages.

For these reasons, the responsible use of Fixed rule based systems is restricted to simple and unambiguous rules, which is why it cannot support decision-making in areas where there is a need to conduct legal interpretation or exercise discretion, or where the law is very complex or has frequent changes.[54] It is against this background that (interactive) expert systems have evolved.

7.2. *Expert Systems*

Expert systems can take a number of forms, here we shall focus on two different approaches: Case-Based Reasoning models and Knowledge Based Reasoning.[55]

[52] K.M. Richmond, S.M. Muddamsetty, T. Gammeltoft-Hansen, H.P. Olsen and T.B. Moeslund, *Explainable AI and Law: An Evidential Survey*, in *DISO*, vol. 3, no. 1, 2024, DOI: 10.1007/s44206-023-00081-z.

[53] As mentioned in section 1, we do not consider outcome prediction/recommendation with Machine Learning to be a viable approach to legal decision support. Neither do we consider LLM's without retrieval augmentation via an up-to-date database covering at least all valid legislation from the area of law relevant to the decisions that are being supported by the system.

[54] There have unfortunately been several examples of implementations of rule-based systems in Denmark implementing complex rules, e.g. in social services (KY) and tax (EFI), which have turned out to be errorneous and lead to harm to citizens and huge development expenses.

[55] Also referred to as Knowledge Representation and Reasoning (KRR) and Knowledge Based Systems (KBS).

7.2.1. *Case-Based Reasoning Systems*

Case-Based Reasoning (CBR) systems are Artificial Intelligence (AI) tools that assist in resolving problems (here: making a legal decision) by retrieving, adapting, and reusing solutions from similar past problems (here: past legal decisions of the same kind). CBR has been known since the 1990s and are inspired by the human practice of solving problems by learning from earlier problem-solving cases and applying the knowledge from that previous case to solve a new problem. CBR lends itself intuitively to legal decision-making practice because of its clear resemblance to analogical and precedent based approaches to decision-making.[56]

CBR differs from rule-based problem-solving used for automated administration of clear rules, in which knowledge of how to solve a problem is encoded to enable automated decision-making. CBR cannot perform automated decision-making but can suggest nearest relevant prior case(s) from which a decision maker may draw inspiration and possibly analogize to decide the new case. The following figure inspired by Craw[57] illustrates how CBR can be implemented in a workflow:

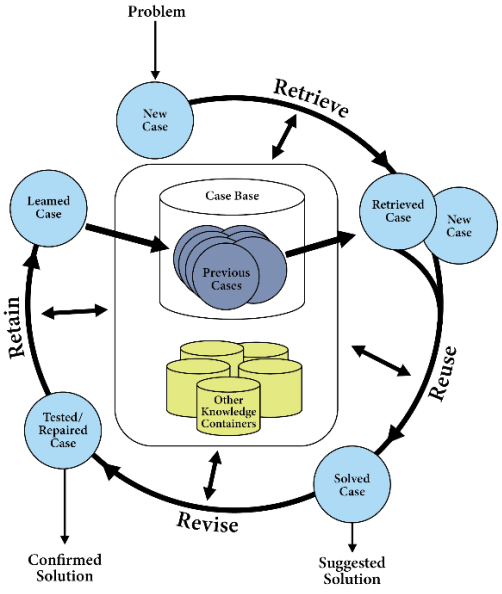

Figure 1: Case-Based Reasoning System

When a new case enters the office, the most similar previous cases(s) are retrieved from the case base. Retrieval can be done in many different ways. Hand coded expert systems were used in the early days of CBR,[58] but today, using natural language processing, may yield better results, especially if implemented with reinforcement learning to adapt to the specific use domain.

Applying the reasoning of the retrieved case to the new case results in a suggested solution (recommendation) that may be accepted, revised or rejected by a human case worker depending on whether he/she assert that the recommended decision is right and sufficiently well explained or whether it is in need of revision.[59] This step again involves automation, whereby justificatory text from one or more previous decisions are lifted out and inserted into a decision template for the new case.

In deciding whether to use or revise such a text, the case worker may rely on other knowledge (illustrated in figure 1 as 'other knowledge containers'). This additional feature is crucially important because it allows for human assessment – an assessment that contain two important elements: 1) Review of the factual situation in the case. This should ensure that all relevant facts are properly accounted for in the decision (i.e. that no factual information is missing) and that those facts are sufficiently well understood (i.e. that individual facts are understood contextually in relation to the overall situation in the case). 2) Review of the final decision to ensure that the case worker is satisfied that the decision can be signed off (confirmed solution in figure 1 above).

Once the decision is released to the citizen, a copy is kept and added to the previous case collection in the case base. Cases that are filed should be automatically marked so that they can be removed if at some later stage they are overridden on appeal.

[56] As mentioned above CBR for law was first developed by Kevin Ashley. See K.D. Ashley, *Reasoning with Cases and Hypotheticals in HYPO*, in *International Journal of Man-Machine Studies*, vol. 34, no 6, 1991, 753–796.

[57] S. Craw, *Case-Based Reasoning*, in C. Sammut and G.I. Webb (eds.), *Encyclopedia of Machine Learning*, Springer, 2011.

[58] HYPO mentioned above is an example of that.

[59] It should be noted that the point of developing a CBR system is that the workload of checking and revising a recommended decision is much lighter than writing a decision from scratch, so even if decision making is not fully automated it still comes with advantages in terms of increased effectiveness. It could also be argued that automated information (previous cases) retrieval advances justice, as it may counteract human biases in choosing what prior cases to compare with.

7.2.2. *Rule-Based Reasoning Systems*

Rule-Based Reasoning Systems, also referred to as Knowledge-Base Systems or Knowledge Representation and Reasoning dates back to the 1960s.[60] Rule-Based Reasoning Systems is a development within Symbolic AI that moves beyond the early fixed, rule-based systems (which are the kind used in the Fixed rule systems described above) to more dynamic systems, where the rules are stored in a separate knowledge base, that can be dynamically updated. Rule-based reasoning systems then use a knowledge base (here a set of rules) and an inference engine to solve problems by mimicking human expert decision-making (here applying rules from the knowledge base to the facts of case). The development of Rule-based reasoning systems led to the first wave of optimism of AI in the 1980s, where many companies began to use expert systems for decision support. However, in the 1990s it became clear that the systems developed at the time did not scale well. A key reason for this was difficulties for experts to learn how to update the representation of rules as the knowledge base grew larger and more complex. While still being used for domain specific decision support, e.g. within healthcare, expert systems stepped into the background and instead focus turned to the development of sub-symbolic AI based on neural networks and machine learning.

However, during the last decade, it has become clear that sub-symbolic AI comes with severe challenges when it comes to explainability, as we have also pointed out above. This has prompted a renewed interest in the development of scalable symbolic AI that can be maintained by domain experts. An example of this is the Dynamic Condition Response (DCR) Graphs notation system, which is a rule graph notation that can be interpreted both by humans and computers. Originally build to describe the rules of distributed knowledge workflows in organizations[61] to support digitization that could be maintained and adapted by the knowledge workers themselves, DCR graphs were later extended by design tools for mapping rules of e.g. law texts and organizational policies and by execution tools supporting both validation of the rules during design and execution of the rules by a case management system.[62] The design tools for mapping of law texts to DCR Graphs use a combination of NLP techniques for highlighting actions, rules and roles in the text and user interfaces allowing domain experts and lawyers to validate and correct the mapping. The execution tool provides an open Application Programming Interface (API) that can be used to either prompt a case worker to e.g. input a fact or to progress in the inference by executing a rule for which all conditions are fulfilled.

The nodes of a DCR graph represent events or actions, and the edges of the graph represent different kinds of logical constraints. The events or actions can either be the occurrence of a fact, a computation based on data provided earlier, or a conclusion/action to be carried out during the inference process. The occurrence of a fact can be an input from an external database e.g. the determination of the age of a citizen, or from a case worker, for example after assessing that the illness of a child is critical in a case where a citizen has applied for additional childcare support. A computation could e.g. be the sum of additional expenses provided to a citizen for whom a right to additional childcare support has been established in a previous case management step. A conclusion could e.g. be that the citizen is eligible for a specific monthly amount for a given period or the action of transferring that amount to the citizen's bank account.

The edges of a DCR graph represent logical constraints and can be divided into two categories: The first category consists of constraints that describe pre-conditions (the 'if' side of the rule). The second category is then the legal effect (the 'then' side of the rule) that follows when the individual facts of the case match the description of the pre-conditions for applying the rule.

The pre-condition constraints can be used

[60] Well-known early example is the MYCIN diagnosis system from the 1970s and the Deep Blue chess system from IBM in the 1990s that won against the chess world champion.

[61] T.T. Hildebrandt and R.R. Mukkamala, *Distributed Dynamic Condition Response Structures*, IT University of Copenhagen, Programming, Logic and Semantics Group, 2011, https://pure.itu.dk/en/publications/distributed-dynamic-condition-response-structures.

[62] T.T. Hildebrandt, *EcoKnow: Effective, Co-Created & Compliant Adaptive Case Management for Knowledge Workers*, in *IEEE 22nd International Enterprise Distributed Object Computing Workshop (EDOCW)*, Stockholm, Sweden, 2018, 9–11, DOI: 10.1109/EDOCW.2018.00012.

to describe which fact/action/conclusion must be established/executed/inferred for another fact/action/conclusion to be at all possible to establish/perform. This could for instance be the validation of information about a bank account number being required before money can be transferred. It could also be the establishment of the fact that a child or young person has an age below 18. Note that such pre-conditioning constraints are not the same as a conclusion or effect: Money is permitted/possible to be transferred but not obliged to be transferred just because there is validated information about a bank account. Such distinction between permission and obligation is also referred to as deontic rules.

The effects are described by the other category of constraints. The most basic effect is the response effect, which is used to model that if one fact/action/conclusion is established/executed, then another fact/action/conclusion must eventually be established/executed/inferred. Such response effects can be given a deadline, which can then be used to model that the effect must happen within the deadline. A more complex effect could be that the establishment of a fact dynamically removes the need/possibility of establishing another fact or making a conclusion. An example of this could be that the establishment of the fact that additional childcare expenses are not needed to support the child. This would then remove the possibility of concluding that additional support should be provided. Such kind of rules are also referred to as defeasible rules.[63] An example of a DCR graph describing the constraints of an article from an act of Danish welfare law (article 86 in the 'Children's Act') could look like this:[64]

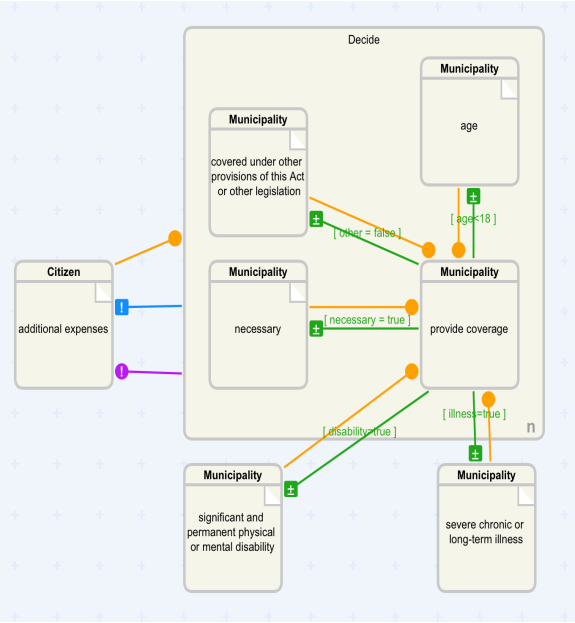

Figure 2: DCR-graph of Article 86 in the Danish Children's Act

The squared boxes are the nodes of the graph, representing either conclusions or facts. The participant responsible for establishing the fact or conclusion is written at the top. In the case, the participant can either be the citizen (who can only make additional expenses) and the Municipality.

The box labelled Decide is a shorthand for all the boxes inside it, allowing us to define constraints related to all of them with one edge. The coloured lines are edges of the graph representing the rules. The orange edge with a circle at the end is a pre-condition, meaning that the fact or conclusion at the end without circle needs to be established before the fact/conclusion at the end with the circle can be established. The blue edge with a box and an exclamation mark is the response effect. That is, if an additional expense is reported by the citizen, then the municipality must establish the three facts inside the box labelled Decide (is the expense necessary?, what is the age? Can the expense be covered by other means?) and also establish the conclusion to provide coverage. The green edges are include/exclude effects that excludes the conclusion (so it cannot and must not be made) if the expression at the edge evaluates to false and includes the conclusion if the condition is true. If included, the conclusion can be made, when all facts have been established. The two facts outside the box Decide needs to be established at least once before the conclusion can be made, but if the municipality is not obliged to establish them each time an additional expense is

[63] D. Nute, *Defeasible Logic*, in *Web Knowledge Management and Decision Support*, in O. Bartenstein, U. Geske, M. Hannebauer and O. Yoshie (eds.), *Lecture Notes in Computer Science*, vol. 2543, Heidelberg, Springer, 2003, 151–169, https://doi.org/10.1007/3-540-36524-9_13

[64] Article 86, unofficial translation made by google translate: 'The municipal council shall provide coverage for necessary additional expenses for the care at home of a child or young person under the age of 18 with a significant and permanent physical or mental disability or a severe chronic or long-term illness. It is a condition that the additional expenses are a consequence of the disability or the severe chronic or long-term illness and cannot be covered under other provisions of this Act or other legislation'.

registered. Finally, the purple edge with the circle containing an exclamation mark is a special pre-condition, stating that if any of the facts/conclusions inside the Decide box are pending obligations, then a new expense cannot be reported by the citizen. This ensures that all the effects of a reported expense have been handled before a new one can be reported.

Expert systems, like rule systems, can as in the case with the DCR graphs be made fully transparent, by making the graphs openly available for inspection. It is even possible to interact with the graph and simulate cases to help understand it. They also allow for automatic execution of tasks, when certain rules are fulfilled, and for automating (and documenting) that case workers are asked to determine a fact. Expert systems may also provide decision-support for interpretation and discretion through automated weighting mechanisms. When a case cannot be decided automatically because it does not fulfil the conditions for automated decision-making, the case can be processed through a CBR model or some other weighting model. This can produce a transparent weighing of relevant case factors, which can be showed to a caseworker, who can then write up a decision after deciding whether all relevant information in the case has been fully considered and whether the proposed weighting of factors will produce a legally defensible (or a legally correct) decision.

Expert systems are in many ways a step forward from rule systems, but they still have their limitations. They require a labor-intensive effort to set up (as the rules and weights need to be exhaustively mapped to the specific use case (field of decision-making) and the number of factors and weights must be updated regularly as legal rules and the world to which legal rules are applied change over time. Its advantage is that it can be made fully transparent and, if the meaning of the rules and weights are well defined, explainable. Its disadvantage is that it relies on a 'mechanical' understanding of legal reasoning and processes documents by focusing on specific word usage. It may not be sufficiently context-sensitive and situationally aware, which, as we have argued above, is one of the fundamental characteristics of the current system we would like to preserve, as it acknowledges the inherent vulnerability of human beings. This disadvantage of expert systems is also one of the chief reasons why the recent breakthroughs in large language models and the ensuing development of RAG systems have become the preferable go-to tools for building automated legal decision support systems.

7.3. *Large Language Models and Retrieval Augmented Generation*

RAG systems rely on a combination of a knowledge base (whose content can to a large extent be automatically updated), automated information retrieval and text generation with LLMs. The use of meta-prompts to set up and configure legal AI systems and introducing AI agents to check generated output is gradually gaining more traction in various businesses including in legal information services.[65] AI driven legal advice robots are no longer uncommon and such technology may soon find its way into legal decision-making processes in public administration. While the technology has many advantages in terms of flexibility, adaptability and efficiency, it has also led researchers to raise questions about robustness and trustworthiness[66] different from the ones encountered in the context of the classic bureaucratic institution inhabited solely by humans.[67]

While RAG systems enhance reliability compared to the use of 'naked' off-the-shelve LLM's recent research has found that RAG systems, despite their improved reliability still exhibit flaws in the form of so-called hallucinations,[68] which undermine the trustworthiness of this technology. Magesh et al, who have inquired into the performance of professional legal RAG systems, sold by leading American legal information service providers, observe that:[69]

> 'These systems can be quite helpful when they work. But as we now illustrate in detail, their answers are often significantly flawed. We find that these systems continue to struggle

[65] See Boston Consulting Group, *AI Agents*, www.bcg.com/capabilities/artificial-intelligence/ai-agents.

[66] E. Sarid and O. Ben-Zvi, *Machine Learning and the Re-enchantment of the Administrative State*, in *Modern Law Review*, vol. 87, no. 2, 2024, 371–397, https://doi.org/10.1111/1468-2230.12843.

[67] J.L. Pearce, *Bureaucracy and Trust: A Review of Recent Volumes in the Russell Sage Foundation Series on Trust*, in *International Public Management Journal*, vol. 11, no. 4, 2008, 481–485.

[68] A.T. Kalai, O. Nachum, S.S. Vempala and E. Zhang, *Why Language Models Hallucinate*, in *arXiv*, 2025, https://arxiv.org/abs/2509.04664

[69] L. Mathiassen, *Reflective Systems Development*, in *Scandinavian Journal of Information Systems*, vol. 10, no. 1, 1998, 12, DOI: 10.17705/3sjis.01012

> with elementary legal comprehension: describing the holding of a case[70]:, distinguishing between legal actors (e.g., between the arguments of a litigant and the holding of the court), and respecting the hierarchy of legal authority. Identifying these misunderstandings often requires close analysis of cited sources. These vulnerabilities remain problematic for AI adoption in a profession that requires precision, clarity, and fidelity.'

The kind of errors made by RAG systems include retrieval errors (the system will find a wrong precedent or a wrong statutory rule and then generate an answer from that wrong source). Such errors occur because correct retrieval itself requires legal analysis; i.e. to retrieve a relevant legal source you need to perform a legal interpretation of the question. To do so, a human lawyer will rely on his/her background training and experience. Such training and experience are difficult to convey to a machine but has important consequences for text generation since generated answers based on retrieval errors will almost always lead to wrong answers.

Another error is the lack of understanding document composition. For example, a Legal RAG system may retrieve a relevant precedent but will fail to distinguish between arguments made by one of the parties to a case and the reasoning of the court. This may lead to the generation of wrong answers, when the RAG generates from the arguments of a losing party. Even when generating answers from the court's reasoning a RAG system may generate wrong answers because it may generate from lines of reasoning that are by the court in its reasoning.

Yet other problems can emerge because a decision does not always include the full information of a case. Hence retrieving a prior decision may not always be helpful, because the decision does not contain information about prior administrative steps that shaped the decision (for example by rejecting some information as irrelevant or requiring a citizen to hand in additional documentation for some factual information).

These issues are serious enough in themselves, but perhaps not so serious, that RAG systems should be dismissed altogether. Magesh et al, reports that the best performing RAG system (offered by Lexis) has an accurate response rate of 65% with 18% incomplete answers and hallucinations in 17% of their generated answers based on a test which used 200 different queries. How well human legal assistants perform on the same kind of tasks is not reported, but we know that the answer to that question will vary significantly depending among other things on the amount of time allocated to the task. Overall, therefore, the question lurking behind these data is whether Legal RAG's, despite their flaws, provide an overall advantage when used *in collaboration* with human experts rather than as stand-alone systems. No research exist that have tried to document if, and how, legal work performed by humans can be improved in quality and efficiency by *collaborating* with Legal RAG's.

When approaching this question, it is not sufficient however to simply focus on legal RAG's ability to generate answers to short model questions. The performance rate reported by Magesh et al. is based on short legal questions that are formulated in an abstract way.[71] Such questions may give some guidance to the formulation of legal principles, but this may be only of little help to a legal decision maker who has to decide on how to apply the principle to a concrete fact situation. To do so requires, as we have explained above situational awareness and the performance of *judgment*. In this context, the hallucination risk is even more problematic, as it can lead to misrepresentation of the case and thereby to wrongful decisions even if the system gets the law right.

8. *Conclusion: The Need for Further Research on How AI Can Support Justificatory Explainability*

There is a need therefore to further investigate what kind of AI systems can consistently and reliably support justificatory explainability for legal decision-making in public administration. Today, in 2026, traditional fixed rule systems are accepted and used in practice, but neither expert system nor LLM/RAG systems have so far been widely implemented.

[70] L. Zheng, *When Does Pretraining Help? Assessing Self-Supervised Learning for Law and the CaseHOLD Dataset*, 2021, https://doi.org/10.48550/arXiv.2104.08671.

[71] For example V. Magesh, F. Surani, M. Dahl, M. Suzgun, C.D. Manning and D.E. Ho, *Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools*, arXiv:2405.20362v1, 2024, discussing prompts such as 'Are the deadlines established by the bankruptcy rules for objecting to discharge jurisdictional?; 'What standard of review applies to abortion regulations under the US Constitution?'; and 'Do any special requirements apply to online dating services in Connecticut?'