Wirtschaftsinformatik 2022 Proceedings | Track 18: Design, Management & Impact of AI-based Systems

Jan 17th, 12:00 AM

Towards Personalized Explanations for AI Systems: Designing a Role Model for Explainable AI in Auditing

Jonas Rebstadt
German Research Center for Artificial Intelligence, Germany, jonas.rebstadt@dfki.de

Florian Remark
Strategion GmbH, florian.remark@strategion.de

Philipp Fukas
University of Osnabrück, philipp.fukas@strategion.de

Pascal Meier
German Research Center for Artificial Intelligence, Germany, pascal.meier@dfki.de

Oliver Thomas
German Research Center for Artificial Intelligence, Germany, oliver.thomas@dfki.de

Towards Personalized Explanations for AI Systems: Designing a Role Model for Explainable AI in Auditing

Jonas Rebstadt[1,2], Florian Remark[2], Philipp Fukas[1,2], Pascal Meier[1], and Oliver Thomas[1]

[1] German Research Center for Artificial Intelligence, Osnabrück, Germany
{jonas.rebstadt,philipp.fukas,pascal.meier,oliver.thomas}@dfki.de
[2] Strategion GmbH, Osnabrück, Germany
{jonas.rebstadt,florian.remark,philipp.fukas}@strategion.de

Abstract. Due to a continuously growing repertoire of available methods and applications, Artificial Intelligence (AI) is becoming an innovation driver for most industries. In the auditing domain, initial approaches of AI have already been discussed in scientific discourse, but practical application is still lagging behind. Caused by a highly regulated environment, the explainability of AI is of particular relevance. Using semi-structured expert interviews, we identified stakeholder specific requirements regarding explainable AI (XAI) in auditing. To address the needs of all involved stakeholders a theoretical role model for AI systems has been designed based on a systematic literature review. The role model has been instantiated and evaluated in the domain of financial statement auditing using focus groups of domain experts. The resulting model offers a foundation for the development of AI systems with personalized explanations and an optimized usage of existing XAI methods.

Keywords: Explainable AI, Auditing, Role Model, Personalization

1 Introduction

The use of Artificial Intelligence (AI) in the context of financial statement auditing, in the following shortened as auditing, has been discussed for many years in the scientific community and experienced its first peak in the 80s and 90s of the last century [1–3]. During that period authors such as Hansen and Messier [1], Bailey et. al. [2] or O'Leary et. al. [3] dealt mainly with expert systems for audit support systems. Using the term decision support systems (DSS) many systems have been accelerated in auditing and in different other domains [4] with the aim to close the gap between a business problem and an information delivery system [5]. Through the development of more and more powerful hardware as well as new procedures and AI methods, the use of AI in auditing has become increasingly important in recent years and is currently experiencing a strong renaissance [6]. AI-based auditing software promises great potential both in terms of performance [7–9] and in terms of potential savings of personnel resources in

17th International Conference on Wirtschaftsinformatik,
February 2022, Nürnberg, Germany

the context of annual audits [10]. However, although many characteristics of the auditor's tasks seem to be well suited to the use of AI, auditing is still lagging behind in the practical application of AI compared to other industries [9, 11, 12].

Moreover, software systems in the field of auditing are subject to strict regulations and laws, which do not explicitly regulate the use of AI yet [13]. Although these regulations and laws are basically country-specific, they are becoming increasingly standardized in compliance with the International Standards of Auditing (ISA) [14]. This area of conflict between technological possibilities and domain specific regulations combined with ethical implications that affect all stakeholders in the auditing environment [11] indicate special requirements for the development of AI systems in auditing. Even if, for example, artificial neural networks enable completely new audit activities, these must be able to be justified in a court of law in case of doubt. According to Munoko et. al. [11] this conflict is mainly based on three aspects, (1) a lack of transparency regarding AI-based decision making, (2) different implications for all actors involved in the audit process and (3) a lack of legal clarity. But in spite of its high relevance from both, a practical and academic research perspective, a consideration of these challenges in the literature is currently lacking.

To contribute to the solution of this problem we identify the requirements and implications regarding XAI in auditing. Due to the complexity of these requirements, we have encountered the need to differentiate explanations depending on the involved stakeholders of the considered AI systems. Tackling this problem, we developed an initial role model, which is instantiated and evaluated in the domain of auditing. The role model lays the foundation for the development of AI systems with personalized explanations for users, examiners and other stakeholders. Therefore, this paper addresses the following research question:

RQ: *How can a role model for Explainable AI in financial statement audit be designed?*

To answer this research question, we initially conducted seven expert interviews with experienced auditors. Using these interviews, we identified requirements, which can hardly be addressed solely using one explanation approach. Our solution for this problem is a role model which can support the introduction of individual explanations for the different stakeholders involved. To identify the most relevant stakeholders for AI systems we have conducted a systematic literature review to collect and aggregate all relevant roles as foundation for a stakeholder-specific personalization of explanations in AI systems. In a next step the identified roles have been structured and instantiated for the auditing domain based on the roles identified in the expert interviews. This domain specific model has been evaluated and stepwise improved by using the method of focus groups consisting of experienced auditors and IT experts from leading audit firms. In the following two sections the theoretical background and the methodical foundation is provided. In Section 4 the results of the expert interviews, the literature review, and focus groups are presented, laying the foundation for the development of personalized explanations in AI systems as discussed in Section 5. Finally, the paper is concluded in Section 6 and an outlook for further research is provided.

2 AI and its Transparency in Financial Auditing

External audits are focusing on the increase of reliability in the audited company's financial statements in accordance to the legal regulations [15]. As in many industries, AI is relevant topic to support domain experts in their decision process. Starting already in the 1980s, expert systems, initial data analytics approaches and AI-based going-concern decision support systems have emerged [1, 16–19]. But as stated by Issa [9], the application of AI is still lagging behind other industries. One major reason for this are the regulatory and ethical implications of the use of AI in the auditing domain, as for example the need of transparency for the supported decisions [4, 11]. But due to the characteristic of some of the most present AI algorithms like artificial neural networks, decisions made by these systems are not always transparent. This forced developers and researchers to build models that are transparent and led to much research in the area of XAI. Similar to the ambiguity of AI definitions [20, 21], the research community tried to define relevant terms like explainability, interpretability, transparency, intelligibility or understandability in many different ways [22–27]. This research follows the differentiation of inherently interpretable and (post-hoc) explainable models using a second model trying to explain the first prediction model as proposed by Rudin [24]. Even though some authors define explainability as also dependent on the capabilities of the user to understand the reasoning of the system [28], in this paper these two terms are used mainly algorithm specific and independent of the user. The second main differentiation used in this paper is based on the focus of the explanation approach, either explaining a single decision of an algorithm (local) or trying to explain the whole model (global) [29]. Additionally to this, Guidotti et al. [29] mentioned in their categorization of existing explainable and interpretable methods the required consideration regarding the nature of users and existing time constraints if matching algorithms should be selected. To take care of these user and problem specifications the concept of understandability is used in this paper. The understandability of an algorithm should value the ability of an algorithm to explicate its reasoning process towards involved users. This has already been tackled by several studies. The overwhelming proportion of studies highlighted a fundamental added value of explanations over proposals without explanations [30–32]. In addition, several studies have compared different explanatory approaches, e.g., based on effectiveness, satisfaction, or ease of use [33–35]. For a more detailed view on studies regarding the evaluation of explanations Nunes and Jannach [36] offer an extensive review. Some of the presented studies already address the personalization of explanations for individual users, focusing on the adjustments of the specific items presented in explanations [33, 35]. But as shown by Tintarev et. al. [35], this type of personalization does not necessarily offer added value for users. Due to this, we focus on the more general stakeholder-specific personalization of complete explanations approaches for different stakeholders like end users or regulators. While end users might prefer local and time efficient explanations like SHAP-values [37], regulators might prefer global explanation approaches, for example by using surrogate models. In the following section the methodical foundation for the development of this type of personalization is described in more detail.

3 Research Approach

The research approach used for this paper is based on the Design Science Research (DSR) paradigm of Hevner et. al. [38], which is targeting at the development and evaluation of artifacts such as "constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices), and instantiations (implemented and prototype systems)" [38]. According to the guidelines they postulated, and the methodological approach suggested by Peffers et. al. [39] in total six steps have been performed to design a role model for XAI systems: *(1) Identify Problem & Motivate, (2) Define Objectives of a Solution, (3) Design & Development, (4) Demonstration, (5) Evaluation and (6) Communication*. Initially, we used expert interviews to *identify the problem (1)* and to *describe the solution (2)* based on the elicitation of requirements and a systematic literature review. In order to ascertain the requirements for XAI and to identify the stakeholders that are affected by it, a total of seven semi-structured interviews [40] were conducted with experienced experts from the field of financial auditing (auditors, audit managers, IT auditors and register accountants). The expert interviews were evaluated using inductive categorization within the framework of a qualitative content analysis according to Mayring [41].The interviews were first transcribed and then analyzed independently from each other by two researchers. Deviating results were discussed and defined by the research team. Setting up on the identified requirements, the foundation for the role model has been laid out by systematically analyzing current literature and deriving different roles, which have to be considered when designing XAI systems. To identify and collect all relevant research standardized steps of a systematic literature review were conducted [42]. The review scope was defined [43], central articles were identified and the topics were conceptualized to specify the key terms for the search string [44]. On the one hand, the search term is composed of the terms "Roles" and "Stakeholder" to ensure that research pointing to different roles in the AI development and application is identified. On the other hand, the terms "Explainable Artificial Intelligence", "Artificial Intelligence" and "XAI" are included in the search phrase since the technological scope of the review should only deal with AI. This results in the following search string *(Roles OR Stakeholder) AND ("Explainable Artificial Intelligence" OR ("Artificial Intelligence" AND "XAI")).* In total the following nine scientific databases were browsed: *ACM Digital Library, AISeL, EBSCO host, Emerad Insight, IEEE Explore, Science Direct, Springer Link, Web of Science and Wiley Online Library.* The initial search on the selected databases led to a number of 1086 search results. These results were scanned based on their titles and abstracts resulting in 113 possibly relevant results. Finally, duplicates were eliminated, and the articles were intensively analyzed towards our inclusion and exclusion criteria resulting in 63 relevant publications. A backward and a forward search yielded no further relevant articles [45]. Therefore, this set of research items was analyzed towards actors and roles involved in the development and usage of AI systems via a concept matrix approach [45, 46]. The identified roles have been used as foundation for the *design and development (3)* of the core artifact, the role model for XAI. The developed role model has been *demonstrated (4)* to a group of experienced auditors and IT experts in leading audit firms. The demonstration has been used to *evaluate (5)* the model and further improved in three iterations of interactive focus groups based on Morgan [47] and Sutton [48] to unfold the creative improvement in cooperative settings. Finally, the results have been used as foundation of this article, to *communicate (6)* the findings.

4 Designing a Role Model for Personalized Explanations

4.1 Elicitation of Requirements and Definition of Objectives

Munoko et. al. [11] and Fukas et. al. [49] have already identified explainability as critical success factor for the establishment of AI applications in the auditing domain. Due to this, the requirements elicitation in this paper does not target the general collection of requirements regarding the development of AI systems. It rather focuses on the identification of facets directly contributing to the need of explainability or interpretability mechanisms in AI systems for auditing. These requirements have been elicited as described in Section 3 and resulted in 18 requirements related to the explainability of AI systems in the auditing domain. Due to the explicit differentiation of several stakeholders, who are relevant for the development of AI systems, the section has been split in two parts. First, all requirements related to XAI in auditing are presented as shown in Figure 1 because of the relevance of the requirements itself for the design of XAI-Systems and therefore implicitly for all artefacts supporting these design processes. The requirements have been clustered and illustrated according to the percentage of the frequency of mentions. Subsequently, the involved roles are listed in more details as explicit foundation for the following steps in the research process.

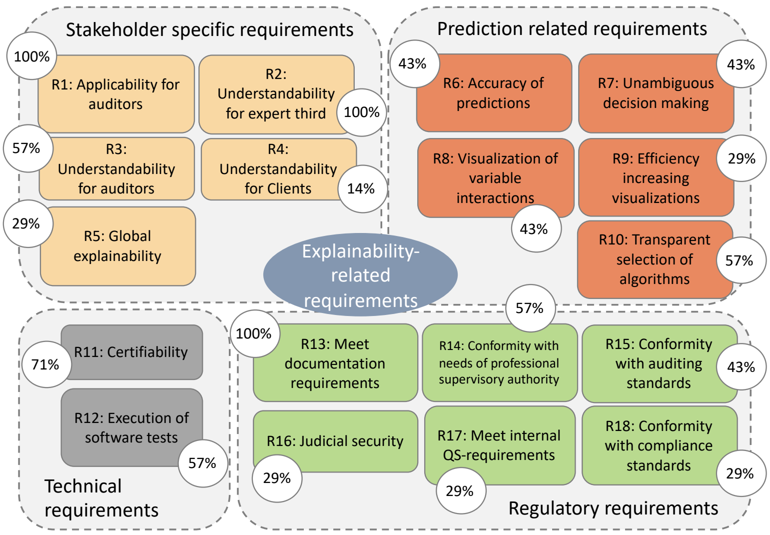

Figure 1. Explainability Related Requirements

Stakeholder specific requirements: The first cluster contains all the requirements that arise specifically from the needs of the stakeholders involved in an audit. The first two requirements, "Applicability for auditors" (R1) and "Understandability for expert third parties" (R2), stand out in particular as they were named independently of each other by all the experts surveyed. The first requirement targets the support of the AI systems for the specific use case of the auditor followed by (R2), which can also be derived from the basic requirements for the auditor's documentation duties (ISA 230 [14]). R3 and R5 demand the understandability of the proposed decisions once from the auditor's perspective (R3) and once from the client's perspective (R5). The last requirement missing from this category, R4, is the global explainability of AI systems. It demands the explanation of the whole model instead of single decisions for a founded validation of the system.

Prediction related requirements: This requirement cluster deals with the design and quality criteria of the XAI components in AI-based audit systems. Requirements R6, R7, and R8 were each mentioned by three of the seven respondents and relate to the availability of accurate information (R6), the making of unambiguous decisions by the system (R7) and the requirement for the visualization of variable interactions (R8). R10 relates to the disclosure of the algorithms used in the forecasting process. The last requirement in this cluster mentioned less frequently by the experts surveyed, is the requirement for a simplified presentation of results (R9). This means in particular that complex results may cause the auditors more work than they can save in terms of resources and is therefore an efficiency criterion.

Technical requirements: The two mainly technically oriented requirements derived from the expert interviews are R11 (Certifiability) and R12 (Execution of software tests). They are primarily focused on compliance with predefined quality standards. Although there are also technical components in other derived requirements, R11 and R12 are the only ones that have the technical aspects in the foreground of the consideration. Certifiability (R11) makes it easier for auditors to assess the quality of AI software even without technical background knowledge. The performance of software tests prior to the deployment of newly developed AI software is in line with known process models of classic software development and is a prerequisite for ensuring smooth and efficient productive use.

Regulatory requirements: The last group of requirements identified for the use of XAI in auditing primarily includes regulatory aspects. As expected, these are very present in the highly regulated environment of auditing although the average relative frequency of mentions seems to be very heterogeneously distributed. The experts agree on the fulfillment of existing documentation requirements (R13). Conformity with the requirements of the professional supervisory authority was also mentioned quite frequently (R14). Conformity with auditing standards plays an important role as well (R15). This is plausible, as international auditing standards (ISA) represent guidelines for the performance of an audit [14]. Furthermore, in the area of regulatory requirements the court security of XAI components as well as internal company requirements for quality assurance standards (R17) and compliance standards (R18) were named by the respondents.

In summary, there is a very differentiated picture of a total of 18 requirements in four categories, some of which differed widely in their format and frequency of mention. It is striking that all of the requirements mentioned do not include any explicit professional or legal regulations on the use of AI or even on the use of XAI. The experts surveyed often cited generic regulations that did not arise against the background of the use of AI but must be applied due to a lack of concrete specifications. Setting up on the identified requirements, the second part of the elicitation focused explicitly on the exploration and extraction of possible roles in the auditing ecosystem that are affected by the use of XAI and therefore need to be included in the role model, developed in the following section. In total 11 different stakeholders have been mentioned at least once. The roles most frequently identified in connection with XAI applications in auditing are those of the auditor (5 mentions), the client (5 mentions), the professional supervisory authority (5 mentions), the certifier (5 mentions) and, in particular, the professional association (7 mentions). This seems plausible, as these players have a direct influence on the audit process and the technology used in audit processes. The profession association as the setter of auditing standards, shapes the framework for the audit. The auditor carries out the audit using the software provided to him, which may be certified. This process is supervised by the professional supervisory authority. In the end the client is always the actor who has to deal with the result of the audit. It is therefore logical that he has an interest in the explainability of the decision-making process with regard to the audit report, too. The roles less frequently mentioned by the experts are also significantly involved in the performance of an audit but tend to exert an indirect influence on the audit as such. First of all, the IT departments (2 mentions) and policy departments (4 mentions) within audit firms should be mentioned here as well as the audit firms themselves (3 mentions). In addition, the legislator (4 mentions), who is responsible for the entire legal framework and for final decisions in cases of dispute, also plays a role, as does the software manufacturer (3 mentions), who must of course meet the requirements for the explainability of auditing software. One expert even suggested an Ethical Commission (1 mention) to derive new guidelines regarding XAI.

Using the results of the expert interviews presented in this section requirements regarding XAI in auditing have been identified and the need of a stakeholder specific differentiation has been elaborated as problem motivating the need of further research. Furthermore, the addressability of the identified XAI-related requirements and based on that the possibility to include all relevant stakeholder with their individual into the development and the evaluation of AI-System in the domain of auditing can be defined as objective for the role model.

4.2 Designing a Role Model for XAI-Systems

As already mentioned by Spraque [50], people with different backgrounds have different perspectives on the design and usage of DSS and especially AI-based Systems. The system itself and the problem it intends to solve can lead to an ambiguity of understandability, if the knowledge level of the involved stakeholders is neglected. Next to the identified need of differentiated explanations in the domain of auditing, Weller [51] stated that the requirements regarding explainable systems differ due to stakeholder specific goals and experiences independent of the domain under consideration. Therefore, explanations can vary in their degree of complexity regarding the needed domain or machine learning knowledge [52]. Based on this general need of personalized explanations, we derived our artifact by (a) developing a general role model based on the current state in the scientific literature and (b) instantiating it for the auditing domain. As presented in Section 3, we conducted a systematical literature review, collecting all roles and stakeholder interacting with XAI systems mentioned in the literature. Due to the varying naming of the same roles and stakeholders we have aggregated the single names resulting in a total of 12 roles. In Table 1, these roles are shown including a short description, the total number of appearances and a selection of appearances in the literature. The complete concept matrix can be found in Rebstadt et al. [53].

Table 1. Overview of Roles and Their Characteristics

Role	Short Description	Hits	Reference Publications
End User	End Users directly use and interact with the AI systems on a regular basis but are not necessarily highly specialized domain experts.	53	[54–60]
AI Researcher	AI Researchers develop AI algorithms and drive the understanding of AI systems from a scientific point of view.	21	[54, 55, 57, 61–63]
Investor	Investors enable the development and use of AI systems by providing the necessary financial resources.	12	[55, 59–61, 64–66]
Data Provider	Data Providers make the data available, which is necessary for AI systems to operate.	4	[54, 55, 67, 68]
Data Subject	Data Subjects are the individuals, which are represented by the actual data underlying AI systems.	12	[54, 63, 65, 67–69]
Decision Subject	Decision Subjects are the individuals, AI systems are making decisions about.	9	[54, 62, 64, 65, 68, 70, 71]
Developer	Developers implement AI systems and integrate them into the corresponding IT infrastructure.	43	[54, 56, 57, 61, 62, 71, 72]
Ethicist	Ethicists explore ethical, social and philosophical implications of AI systems.	2	[54, 67]
Legislator	Legislators design legislation as well as standards and are responsible for the legal regulation of AI systems.	15	[54, 58, 61, 62, 69, 72]
Manager	Managers are domain experts and manage the overall application domain, in which AI systems operate. They do not necessarily directly interact with AI systems.	40	[54–57, 61, 62, 71]
Public	The Public is indirectly concerned with the use of AI systems in application domains.	3	[54, 59, 73]
Regulator	Regulators enforce the legal regulation of AI systems.	25	[56–60, 65, 69]

The first identified role is the *End User*, who directly interacts with the AI system on a regular basis. *End Users* are neither necessarily a technical nor a domain expert. If the usage of the AI system happens in the context of a company, they will have *Managers*, who coordinate the overall application domain and are responsible for decisions in the domain, in which the system operates. *Managers* do not necessarily directly interact with AI systems. They themselves may be directly or indirectly under the control of *Investors*, who enable the development and use of AI systems by providing the necessary financial resources. The development itself is done by *Developers*, who implement AI systems and integrate them into the corresponding IT infrastructure. The foundation for this is laid by *AI Researchers*, who develop algorithms and drive the understanding of AI systems from a scientific point of view. The core of all machine learning-based systems is the data it is trained on. *Data Providers* make this data available by extracting and integrating it from the *Data Subjects*. They represent the concrete data source, which is used by the systems. The data is the foundation for predictions about the *Decision Subject*, that is represented in the use cases of the application domain. *Legislators* design these policies and are responsible for the legal regulation of AI systems. These rules are enforced by the *Regulator*, who tries to ensure the interests of the *Public*, which itself is indirectly concerned with the use of AI systems in application domains. Finally, *Ethicists* explore ethical, social, and philosophical implications of AI systems. Even though these condensed roles are often mentioned without revealing their direct relations, they interact in the real world. All of them can have requirements for the system in general as well as for the explainability of AI-based systems in detail. Due to this, a differentiation in the design of explanations may be useful to achieve the required level of understandability for each person involved. While, in the case of auditing, for example the conformity with auditing standards as mentioned in Figure 1 is relevant for all roles, the existence of efficiency increasing visualizations is only relevant for roles like *End User* (*Auditors)*, directly interacting with the results of the AI-System. They need to interact efficiently with the system and for that they need to grasp the results and the preceding decision process under time pressure. On the other hand, *Data Provider* (*IT-Division*) do not need to understand the decision process in detail, but it is necessary to reveal the (most) important characteristics of the data for the AI-System to optimally address the data acquisition and the necessary preprocessing steps.

One central aspect for the design of explanations is the obligation and interaction between different roles and the AI system. In Figure 2 the obligation of an explanation between the roles and the AI system is visualized using arrows. Additionally, direct interactions with the AI system are marked using dashed circles. Finally, dashed lines describe indirect relations, which are not necessarily characterized by direct and directed communication. If the presented roles are instantiated for an existing AI system or an AI system that should be developed, a single person can occupy several roles especially if the user, who has to make a decision, interacts directly with the system. This general role assignment as shown in part (a) of Figure 2 has no claim to completeness but offers a foundation of roles relevant in the development of AI-based systems.

To address the specific constellations and requirements in the auditing domain the general model has been instantiated. The foundation for this artifact are the already identified roles in the expert interviews. This is shown in part (b) of Figure 2.

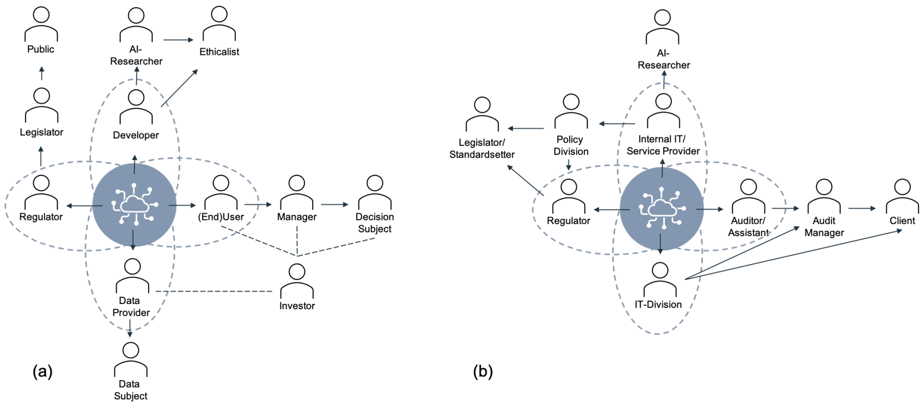

Figure 2. Roles of AI-based Systems in General (a) and in the Auditing Domain (b)

4.3 Demonstration and Evaluation the Developed Role Model

As described in Section 3 the general and the instantiated model have been demonstrated in three focus group meetings towards a selection of experienced auditors. During the first two of these digitally held meetings the relevance and the relation of the existing and potential new roles have been discussed. After each of these two sessions, a proposal of an adapted role model has been designed and presented in the next meeting, finishing with a successful assessment. In the third focus group, the possibly differentiation of the requirements described in Section 3 and possible stakeholder personalized explanation approaches have been successfully elaborated to evaluate the added value of the designed model.

There are five major differences between the general role model and the audit specific one that emerged during the focus groups. The first major subject is the aggregation of the *Data Subject* role and the *Decision Subject* role into the *Client* role. Due to the strict regulations specified in most contracts regarding the usage of client data for other decisions, *Data Subject* and *Decision Subject* need to be represented by the same *Client.* Second, the *investor* has been removed, because the dependencies regarding the decisions of the AI system are limited due to the direct liability of the auditors. Third, the *Policy Division* has been included in the role model based on the high impact regarding regulations and the resulting need of explainability for AI systems. Fourth, the role of the *Ethicist* has been removed because it was not considered as relevant in the focus groups. As last major difference the *Public* has been excluded in the evaluation phases because of the limited direct influence. But this last change has been part of discussion and may be reconsidered in further steps, owing to the general high relevance of audits for the public and especially investors of the clients being tested. In the last two sections the designed and evaluated core artifact of the described research process has been presented. All stakeholder that are mentioned in the current scientific literature have been defined and aggregated in the presented role model.

5 Discussion

Despite the high potential of AI for the auditing domain [9, 74–76] and the determination of XAI as critical success factor due to domain specific regulations [11], the development of AI-Systems in the auditing domain is still lagging behind [9]. This publication offers contributions to the theory and practice, trying to increase the understanding of the existing requirements and to reduce the barriers in the development of AI-based systems. From the theoretical perspective, one central finding derived from the expert interviews is tackling the relation of domain-related regulations and AI-specific regulations. Based on the regular recitation of domain specific regulation which did not arise in the context of AI use, from the view of the interviewed experts, audit and therefore domain specific regulations override AI-specific regulations in the corresponding domain. Next to this, as already mentioned by Oh [77], the development of AI systems requires the consideration of various stakeholders instead of simply taking a general user-centric perspective. This is especially relevant for the specification of explanation approaches. As shown by Tintarev et al. [35], the stakeholder-specific personalization of explanation on a low level of granularity that has been implemented and evaluated without necessarily adding value for the considered users. This may change if the stakeholder-specific personalization is considered on a more generalized level of complete explanation approaches for different stakeholders like end users or regulators. The presented research provides an extension of the existing design knowledge for the development of XAI-Systems by offering a general and an audit specific role model, [78]. Following Gregor and Hevner [79] the design knowledge in this publication can be argued as nascent design theory, due to its abstraction of specifically instantiated role model into operational principles, contributing to the prescriptive knowledge by describing elements of artifact design.

In addition to the theoretical contribution, the publication offers practical implications, showing a starting point for the development on AI-Systems with individually personalized explanations due to their differing requirements towards the design of explanations. If one considers an AI-based tool for auditors focusing on the detection of fraudulent activities, several stakeholders will possibly be involved in the development, use and evaluation of this tools as shown in Figure 2 (b). These requirements influence the choice of explanation algorithms as can be seen quite nicely on the differences between auditors and the policy division or regulators. The policy division and in the second step the regulators need to understand the functionality of the whole models good enough to evaluate the quality of the system and secure the transparency of all decisions. Regarding that, inherently interpretable models or surrogates with algorithms like decision trees or rule-based systems might be preferred. In comparison auditors have additional requirements in the direct use of the system. Based on the existing time pressure and their possibly lacking technical knowledge, the consideration of explanations like rulesets with several AND connections for each decision is not feasible. Therefore, the inclusion of an additional post-hoc explanation like SHAP values might be the best solution to help the auditors to understand the most influential factors for a possible fraudulent action and offer a starting point for further investigations. Additionally, the consideration of involved roles helps to include all stakeholders, who might be affected by decisions of the system and to offer personalized explanations. This increases the fairness and reduces the possible discrimination of such systems [73]. For stakeholder groups impacted by the system, requirement specific explanations can on one hand increase the trust in such systems and, on the other hand, simplify the filing of reasoned objections if necessary [73].

But even though the results of this publication seem to promisingly in the demonstration and evaluation rounds, there are some (methodical) limitation to mentions. First, mainly due to the quite small domain, only seven experts have been interviewed for the requirements elicitation. The authors tried to compensate this drawback by a balanced and purposeful selection of interviewees targeting the financial audit from different perspectives and backed with much professional experience. Next to this, the evaluation of the specified role model has only been done in the auditing domain, with a limited direct transferability and generalizability towards other audit types and domains, based on its specific regulations. But, due to its highly regulated character, the audit domain might offer early insights for many other domain as AI in general will get increasingly regulated as for example driven by the European Union [80].

6 Conclusion and Outlook

In this paper, we identified requirements regarding XAI for the efficient and compliant use of AI applications in auditing. The results show, that in the tension between technological possibilities and regulatory restrictions, the application of XAI is essential. We identified the need to differentiate explanations based on the ambiguity of all stakeholders involved. To address this issue, we developed a general role model for XAI systems based on a literature review. The model has been instantiated and iteratively evaluated for the auditing domain using the methods of expert interviews and focus groups. The developed role model opens up the opportunity to develop AI-based systems with stakeholder appropriate explanations. As a next step, a prototypical system is to be implemented and evaluated with different stakeholders to verify the added value of personalized explanations for each actor. The choice of explanations as well as the evaluation of these approaches can be based on the research of several publications [81–84] focusing on structure and drawbacks of explanations out of different perspectives. Since our research work was conducted in Germany and the interviewees are involved in auditing companies from the German economic area the generalizability of our findings may be limited to German audit companies. Due to the fact that international standards and requirements for an annual audit become increasingly harmonized, it can be assumed that findings regarding XAI will also converge in the future. Furthermore, we are convinced, that the instantiation of the role model as well as the idea of personalized explanation approaches are promising for other domains. Even though not all AI systems have the need to address this additional effort in the development process, all systems tackling high risk decisions should take these ideas into account.

References

1. Hansen, J. V., Messier, W.F.: Expert systems for decision support in EDP auditing. Int. J. Comput. Inf. Sci. 11, 357–379 (1982). https://doi.org/10.1007/BF01001957.
2. Bailey, A.D., Meservy, R.D., Duke, G.L., Johnson, P.E., Thompson, W.: Auditing, Artificial Intelligence and Expert Systems. In: Decision Support Systems: Theory and Application. pp. 265–319. Springer Berlin Heidelberg, Berlin, Heidelberg (1987). https://doi.org/10.1007/978-3-642-83088-4_10.
3. O'Leary, D.E., O'Keefe, R.M.: The impact of artificial intelligence in accounting work: Expert systems use in auditing and tax. Ai Soc. 11, 36–47 (1997).
4. Richard Ye, L.: The value of explanation in expert systems for auditing: An experimental investigation. Expert Syst. Appl. 9, 543–556 (1995). https://doi.org/10.1016/0957-4174(95)00023-2.
5. Vierck, R.K.: Decision Support Systems: An MIS Manager's Perspective. MIS Q. 5, 35 (1981). https://doi.org/10.2307/249326.
6. Issa, H., Sun, T., Vasarhelyi, M.A.: Research Ideas for Artificial Intelligence in Auditing: The Formalization of Audit and Workforce Supplementation. J. Emerg. Technol. Account. 13, 1–20 (2016). https://doi.org/10.2308/jeta-10511.
7. Kokina, J., Davenport, T.H.: The Emergence of Artificial Intelligence: How Automation is Changing Auditing. J. Emerg. Technol. Account. 14, 115–122 (2017). https://doi.org/10.2308/jeta-51730.
8. Downar, B., Fischer, D.: Wirtschaftsprüfung im Zeitalter der Digitalisierung. In: Handbuch Industrie 4.0 und Digitale Transformation. pp. 753–779. Springer Fachmedien Wiesbaden, Wiesbaden (2019). https://doi.org/10.1007/978-3-658-24576-4_32.
9. Issa, H., Sun, T., Vasarhelyi, M.A.: Research Ideas for Artificial Intelligence in Auditing: The Formalization of Audit and Workforce Supplementation. J. Emerg. Technol. Account. 13, 1–20 (2016). https://doi.org/10.2308/jeta-10511.
10. Frey, C.B., Osborne, M.A.: The future of employment: How susceptible are jobs to computerisation? Technol. Forecast. Soc. Change. 114, 254–280 (2017). https://doi.org/10.1016/j.techfore.2016.08.019.
11. Munoko, I., Brown-Liburd, H.L., Vasarhelyi, M.: The Ethical Implications of Using Artificial Intelligence in Auditing. J. Bus. Ethics. (2020). https://doi.org/10.1007/s10551-019-04407-1.
12. Gierbl, A.S., Schreyer, M., Leibfried, P., Borth, D.: Künstliche Intelligenz in der Prüfungspraxis - Eine Bestandsaufnahme aktueller Einsatzmöglichkeiten und Herausforderungen. Expert Focus. 2020, 612–617 (2020).
13. Munoko, I., Brown-Liburd, H.L., Vasarhelyi, M.: The Ethical Implications of Using Artificial Intelligence in Auditing. J. Bus. Ethics. (2020). https://doi.org/10.1007/s10551-019-04407-1.
14. International Auditing and Assurance Standards Board, (IAASB): Handbook of International Quality Control, Auditing Review, Other Assurance, and Related Services Pronouncements. (2016).
15. ISA 200: ISA 200: Overall Objectives of the Independent Auditorand the Conduct of an Audit in Accordance with International Standards on Auditing. In: International

Federation of Accountants (IFAC) (ed.) International Standards on Auditing (ISAs). Institut der Wirtschaftsprüfer (IDW) Verlag GmbH, Düsseldorf (2009).
16. Bailey, A.D., Meservy, R.D., Duke, G.L., Johnson, P.E., Thompson, W.: Auditing, Artificial Intelligence and Expert Systems. In: Holsapple, C. and Whinston, A.B. (eds.) Decision Support Systems: Theory and Application. pp. 265–319. Springer, Berlin (1987). https://doi.org/10.1007/978-3-642-83088-4_10.
17. IAASB: Exploring the Growing Use of Technology in the Audit, with a Focus on Data Analytics FEEDBACK STATEMENT. Indep. Regul. Board Audit. 8873 (2017).
18. Yeh, C.-C., Chi, D.-J., Lin, Y.-R.: Going-concern prediction using hybrid random forests and rough set approach. Inf. Sci. (Ny). 254, 98–110 (2014). https://doi.org/10.1016/j.ins.2013.07.011.
19. Chen, S.: An effective going concern prediction model for the sustainability of enterprises and capital market development. Appl. Econ. 51, 3376–3388 (2019). https://doi.org/10.1080/00036846.2019.1578855.
20. Russell, S., Norvig, P.: Artificial intelligence: a modern approach. (2002).
21. McCorduck, P., Cfe, C.: Machines who think: A personal inquiry into the history and prospects of artificial intelligence. CRC Press (2004).
22. Lipton, Z.C.: The mythos of model interpretability. Commun. ACM. 61, 35–43 (2018). https://doi.org/10.1145/3233231.
23. Tomsett, R., Braines, D., Harborne, D., Preece, A., Chakraborty, S.: Interpretable to Whom? A Role-based Model for Analyzing Interpretable Machine Learning Systems. (2018).
24. Rudin, C.: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat. Mach. Intell. 1, 206–215 (2019). https://doi.org/10.1038/s42256-019-0048-x.
25. Dhurandhar, A., Iyengar, V., Luss, R., Shanmugam, K.: TIP: Typifying the Interpretability of Procedures. 1–39 (2017).
26. Clinciu, M.A., Hastie, H.F.: A Survey of Explainable AI Terminology. Proc. 1st Work. Interact. Nat. Lang. Technol. Explain. Artif. Intell. 8–13 (2019).
27. Moore, J.D., Swartout, W.R.: Explanation in Expert Systems: A Survey. ISI Res. Rep. (1988).
28. Van Lent, M., Fisher, W., Mancuso, M.: An explainable artificial intelligence system for small-unit tactical behavior. Proc. Natl. Conf. Artif. Intell. 900–907 (2004).
29. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D.: A survey of methods for explaining black box models. ACM Comput. Surv. 51, 1–45 (2018). https://doi.org/10.1145/3236009.
30. Muhammad, K., Lawlor, A., Smyth, B.: A live-user study of opinionated explanations for recommender systems. Int. Conf. Intell. User Interfaces, Proc. IUI. 07-10-Marc, 256–260 (2016). https://doi.org/10.1145/2856767.2856813.
31. Suermondt, H.J., Cooper, G.F.: An evaluation of explanations of probabilistic inference. Comput. Biomed. Res. 26, 242–254 (1993). https://doi.org/10.1006/cbmr.1993.1017.
32. Wang, N., Pynadath, D. V., Hill, S.G.: Trust calibration within a human-robot team: Comparing automatically generated explanations. ACM/IEEE Int. Conf. Human-Robot Interact. 2016-April, 109–116 (2016). https://doi.org/10.1109/HRI.2016.7451741.
33. Gedikli, F., Ge, M., Jannach, D.: Understanding Recommendations by Reading the

Clouds. In: Aalst, W. van der, Mylopoulos, J., Rosemann, M., Shaw, M.J., and Szyperski, C. (eds.) E-Commerce and Web Technologies. pp. 196–208. , Toluose (2011). https://doi.org/10.1007/978-3-642-23014-1.
34. Bussone, A., Stumpf, S., O'Sullivan, D.: The role of explanations on trust and reliance in clinical decision support systems. Proc. - 2015 IEEE Int. Conf. Healthc. Informatics, ICHI 2015. 160–169 (2015). https://doi.org/10.1109/ICHI.2015.26.
35. Tintarev, N., O'Donovan, J., Felfernig, A.: Introduction to the special issue on human interaction with artificial advice givers. ACM Trans. Interact. Intell. Syst. 6, 1–12 (2016). https://doi.org/10.1145/3014432.
36. Nunes, I., Jannach, D.: A systematic review and taxonomy of explanations in decision support and recommender systems. User Model. User-adapt. Interact. 27, 393–444 (2017). https://doi.org/10.1007/s11257-017-9195-0.
37. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 2017-Decem, 4766–4775 (2017).
38. Hevner, A.R., March, S.T., Park, J., Ram, S.: Design science in information systems research. MIS Q. Manag. Inf. Syst. (2004). https://doi.org/10.2307/25148625.
39. Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S.: A design science research methodology for information systems research. J. Manag. Inf. Syst. 24, 45–77 (2007). https://doi.org/10.2753/MIS0742-1222240302.
40. Bogner, A., Littig, B., Menz, W.: Interviews mit Experten. Springer Fachmedien Wiesbaden, Wiesbaden (2014). https://doi.org/10.1007/978-3-531-19416-5.
41. Mayring, P.: Qualitative inhaltsanalyse. In: Handbuch qualitative Forschung in der Psychologie. pp. 601–613. Springer (2010).
42. vom Brocke, J., Niehaves, B., Simons, A., Riemer, K.: Reconstructing the Giant : On the Importance of Rigour in Documenting the Literature Search Process. In: Proceedings of the European Conference on Information Systems 2009. , Verona (2009).
43. Cooper, H.M.: Organizing knowledge syntheses: A taxonomy of literature reviews. Knowl. Soc. 1, 104–126 (1988). https://doi.org/10.1007/BF03177550.
44. Rowley, J., Slack, F.: Conducting a literature review. Manag. Res. News. 27, 31–39 (2004). https://doi.org/10.1108/01409170410784185.
45. Webster, J., Watson, R.T.: Analyzing the Past to Prepare for the Future: Writing a Literature Review. Manag. Inf. Syst. Q. 26, xiii–xxiii (2002). https://doi.org/10.1.1.104.6570.
46. Salipante, P., Notz, W., Bigelow, J.: A Matrix Approach to Literature Reviews. Res. Organ. Behav. 4, 321–348 (1982).
47. Morgan, D.L.: Focus groups as qualitative research. (1996).
48. Sutton, S.G., Arnold, V.: Focus group methods: Using interactive and nominal groups to explore emerging technology-driven phenomena in accounting and information systems. Int. J. Account. Inf. Syst. 14, 81–88 (2013). https://doi.org/10.1016/j.accinf.2011.10.001.
49. Fukas, P., Rebstadt, J., Remark, F., Thomas, O.: Developing an Artificial Intelligence Maturity Model for Auditing. In: 29th European Conference on Information Systems Research Papers (2021).
50. Sprague, R.H.: A framework for the development of decision support systems. MIS Q. Manag. Inf. Syst. 4, 1–26 (1980). https://doi.org/10.2307/248957.

51. Weller, A.: Transparency: Motivations and Challenges. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 11700 LNCS, 23–40 (2019). https://doi.org/10.1007/978-3-030-28954-6_2.
52. Burkart, N., Huber, M.F.: A survey on the explainability of supervised machine learning. J. Artif. Intell. Res. 70, 245–317 (2021). https://doi.org/10.1613/JAIR.1.12228.
53. Rebstadt, J., Remark, F., Fukas, P., Meier, P., Thomas, O.: Supplementary Material for the Publication " Towards Personalized Explanations for AI Systems : Designing a Role Model for Explainable AI in Auditing ," https://doi.org/10.5281/zenodo.5682459.
54. Kasirzadeh, A.: Reasons, Values, Stakeholders: A Philosophical Framework for Explainable Artificial Intelligence. (2021).
55. Brennen, A.: What do people really want when they say they want "explainable AI?" we asked 60 stakeholders. Conf. Hum. Factors Comput. Syst. - Proc. 1–7 (2020). https://doi.org/10.1145/3334480.3383047.
56. Baldauf, M., Fröehlich, P., Endl, R.: Trust Me, I'm a Doctor-User Perceptions of AI-Driven Apps for Mobile Health Diagnosis. ACM Int. Conf. Proceeding Ser. 167–178 (2020). https://doi.org/10.1145/3428361.3428362.
57. Ehsan, U., Liao, Q.V., Muller, M., Riedl, M.O., Weisz, J.D.: Expanding Explainability: Towards Social Transparency in AI systems. 1–19 (2021). https://doi.org/10.1145/3411764.3445188.
58. Bedué, P., Fritzsche, A.: Can we trust AI? An empirical investigation of trust requirements and guide to successful AI adoption. J. Enterp. Inf. Manag. ahead-of-p, (2021). https://doi.org/10.1108/jeim-06-2020-0233.
59. Meske, C., Bunde, E., Schneider, J., Gersch, M.: Explainable Artificial Intelligence: Objectives, Stakeholders, and Future Research Opportunities. Inf. Syst. Manag. 00, 1–11 (2020). https://doi.org/10.1080/10580530.2020.1849465.
60. Mehdiyev, N., Houy German, C., Gutermuth German, O., Mayer German, L., Fettke German, P.: Association for Information Systems Association for Information Systems AIS Electronic Library (AISeL) AIS Electronic Library (AISeL) Wirtschaftsinformatik 2021 Proceedings Track 8: Smart City & E-Government Explainable Artificial Intelligence (XAI) Suppo. (2021).
61. Beaton, B.: Crucial Answers about Humanoid Capital. ACM/IEEE Int. Conf. Human-Robot Interact. 5–12 (2018). https://doi.org/10.1145/3173386.3173391.
62. Ehsan, U., Wintersberger, P., Liao, Q.V., Mara, M., Streit, M., Wachter, S., Riener, A., Riedl, M.O.: Operationalizing Human-Centered Perspectives in Explainable AI. Conf. Hum. Factors Comput. Syst. - Proc. (2021). https://doi.org/10.1145/3411763.3441342.
63. Cheng, H.F., Wang, R., Zhang, Z., O'Connell, F., Gray, T., Harper, F.M., Zhu, H.: Explaining decision-making algorithms through UI: Strategies to help non-expert stakeholders. Conf. Hum. Factors Comput. Syst. - Proc. 1–12 (2019). https://doi.org/10.1145/3290605.3300789.
64. Chromik, M., Eiband, M., Buchner, F., Krüger, A., Butz, A.: I Think i Get Your Point, AI! The Illusion of Explanatory Depth in Explainable AI. Int. Conf. Intell. User Interfaces, Proc. IUI. 307–317 (2021). https://doi.org/10.1145/3397481.3450644.
65. Huynh, T.D., Tsakalakis, N., Helal, A., Stalla-Bourdillon, S., Moreau, L.: Addressing Regulatory Requirements on Explanations for Automated Decisions with Provenance—

A Case Study. Digit. Gov. Res. Pract. 2, 1–14 (2021). https://doi.org/10.1145/3436897.
66. Higgins, D., Madai, V.I.: From Bit to Bedside: A Practical Framework for Artificial Intelligence Product Development in Healthcare. Adv. Intell. Syst. 2, 2000052 (2020). https://doi.org/10.1002/aisy.202000052.
67. Nassar, M., Salah, K., ur Rehman, M.H., Svetinovic, D.: Blockchain for explainable and trustworthy artificial intelligence. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 10, 1–13 (2020). https://doi.org/10.1002/widm.1340.
68. Zednik, C.: Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. Philos. Technol. (2019). https://doi.org/10.1007/s13347-019-00382-7.
69. Felzmann, H., Fosch-Villaronga, E., Lutz, C., Tamò-Larrieux, A.: Towards Transparency by Design for Artificial Intelligence. Sci. Eng. Ethics. 26, 3333–3361 (2020). https://doi.org/10.1007/s11948-020-00276-4.
70. Arya, V., Bellamy, R.K.E., Chen, P.Y., Dhurandhar, A., Hind, M., Hoffman, S.C., Houde, S., Liao, Q.V., Luss, R., Mojsilović, A., Mourad, S., Pedemonte, P., Raghavendra, R., Richards, J., Sattigeri, P., Shanmugam, K., Singh, M., Varshney, K.R., Wei, D., Zhang, Y.: AI Explainability 360 Toolkit. ACM Int. Conf. Proceeding Ser. 376–379 (2020). https://doi.org/10.1145/3430984.3430987.
71. Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., Baum, K.: What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. Artif. Intell. 296, 0–57 (2021). https://doi.org/10.1016/j.artint.2021.103473.
72. Yu, B., Yuan, Y., Terveen, L., Wu, Z.S., Forlizzi, J., Zhu, H.: Keeping designers in the loop: Communicating inherent algorithmic trade-offs across multiple objectives. DIS 2020 - Proc. 2020 ACM Des. Interact. Syst. Conf. 1245–1257 (2020). https://doi.org/10.1145/3357236.3395528.
73. Berscheid, J., Roewer-Despres, F.: Beyond transparency. AI Matters. 5, 13–22 (2019). https://doi.org/10.1145/3340470.3340476.
74. Kokina, J., Davenport, T.H.: The Emergence of Artificial Intelligence: How Automation is Changing Auditing. J. Emerg. Technol. Account. 14, 115–122 (2017). https://doi.org/10.2308/jeta-51730.
75. Downar, B., Fischer, D.: Wirtschaftsprüfung im Zeitalter der Digitalisierung. In: Handbuch Industrie 4.0 und Digitale Transformation. pp. 753–779. Springer Fachmedien Wiesbaden, Wiesbaden (2019). https://doi.org/10.1007/978-3-658-24576-4_32.
76. Frey, C.B., Osborne, M.A.: The future of employment: How susceptible are jobs to computerisation? Technol. Forecast. Soc. Change. 114, 254–280 (2017). https://doi.org/10.1016/j.techfore.2016.08.019.
77. Oh, C., Kim, S., Choi, J., Eun, J., Kim, S., Kim, J., Lee, J., Suh, B.: Understanding how people reason about aesthetic evaluations of artificial intelligence. DIS 2020 - Proc. 2020 ACM Des. Interact. Syst. Conf. 1169–1181 (2020). https://doi.org/10.1145/3357236.3395430.
78. Gregor, S., Jones, D.: The anatomy of a design theory. J. Assoc. Inf. Syst. 8, 312–335 (2007). https://doi.org/10.17705/1jais.00129.
79. Gregor, S., Hevner, A.R.: Positioning and Presenting Design Science Research for

Maximum Impact. MIS Q. 37, 337–355 (2013).
80. COMMISSION, E.: Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS. (2021).
81. Cunningham, P., Doyle, D., Loughrey, J.: An evaluation of the usefulness of Case-Based explanation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 2689, 122–130 (2003). https://doi.org/10.1007/3-540-45006-8_12.
82. Sklar, E.I., Azhar, M.Q.: Explanation through argumentation. HAI 2018 - Proc. 6th Int. Conf. Human-Agent Interact. 277–285 (2018). https://doi.org/10.1145/3284432.3284473.
83. Lage, I., Chen, E., He, J., Narayanan, M., Kim, B., Gershman, S., Doshi-Velez, F.: An Evaluation of the Human-Interpretability of Explanation. 1–24 (2019).
84. Liao, B., Anderson, M., Anderson, S.L.: Representation, Justification and Explanation in a Value Driven Agent: An Argumentation-Based Approach. 1–24 (2018).