

AI Checklist

Lextel

AI

Practical AI Ethics Checklist

Versione 1.0

Marzo 2025

Summary

- 1 Project description.....2
- 2 Technical Solution.....3
 - 2.1 AI Model for Legal Documents Retrieval 3
 - 2.2 Generative Model for Drafting Legal Opinions 4
 - 2.3 SYSTEM TRAINING 5
- 3 EU AI Act Compliance Check6
- 4 Pratical AI Ethics Checklist (from EU Experts Group).....7

1 Project description

Lextel.AI is a tool that aims at speeding up the research on Law topics leveraging on AI technologies, enabling attorneys to focus on high value tasks in dealing with customers' requests across all practice areas.

Lextel.AI is a web platform which both:

- empowers professionals' research thanks to an extensive database hosting top class documents;
- proactively creates summarizations and legal opinions elaborating on documents selected by the attorneys thanks to AI algorithms

To effectively support lawyers and law firms, Lextel.AI combines 3 different technologies:

- a complex database, storing documents collected from several tested sources, and a mining solution that provides near real-time search and analytics for all types of data (structured or unstructured text)
- large language models, supported by best of breed technologies, trained on complex legal tasks with the supervision of a professional team
- a knowledge graph tool that connects different contents, according to relevance and source parameters, and generates list of related or cited documents

Lextel.AI database contains contributions coming from different sources, archiving from:

- Norms: EU, Italian, Regional codes, laws
- Praxis: Ministerial, National and EU Agencies, Public institutions and bodies
- Jurisprudence/Cases: High Courts (Corte Costituzionale, Cassazione, Corte di Giustizia) Tribunals (Merito, T.A.R., Taxes)
- Sectorial Magazines and publications on different topics (Civil Law, Penal law, taxes etc)
- Authoral contributions: Commented codes, manuals, norms interpretations

2 Technical Solution

LextelAI solution leverages on different components, designed to:

1. improve lawyers' information retrieval processes
2. develop, through generative AI algorithms, solid legal opinions on the searched topics.

2.1 AI Model for Legal Documents Retrieval

The solution relies on the implementation and use of a Transformer model, specifically a BERT base model, trained and optimized for information retrieval within the Italian legal domain. The model has been adapted to handle complex and specific queries typical of legal language and is used within a hybrid search system that combines BM25 and vectorization to enhance result accuracy.

Model Architecture

BERT Base Fine-Tuned

The base model used is BERT (Bidirectional Encoder Representations from Transformers) in its "foundation" configuration, which includes 12 transformer layers, 768 hidden units, 12 attention heads, and approximately 110 million parameters. This model has been further fine-tuned on synthetic sentence pairs generated for the Italian legal domain to specialize the model's semantic similarity function. Fine-tuning was conducted using a contrastive loss function, optimizing the model's ability to distinguish between similar and dissimilar sentences in legal contexts.

Retrieval System (Hybrid Search)

The retrieval system is designed to combine the strengths of both BM25-based searches and vector embeddings.

BM25: This is a probabilistic information retrieval algorithm commonly used to measure the relevance of documents to a query. It provides a robust search function, based on the presence and frequency of keywords.

Vectors: The embeddings generated by the fine-tuned BERT model are used to capture semantic similarity between the user query and text segments (chunks) within the corpus. The combination of results between BM25 and vector similarity is dynamic, with the relative weight of the two approaches varying based on the number of words in the user's query. Longer queries tend to favor the BM25 approach, while shorter queries or those with specific terms favor vector similarity.

Complex Legal Expressions Handling

A key aspect of the retrieval system is managing legal expressions composed of multiple words. During the preprocessing phase, these expressions have been identified and treated as unique entities, maintaining semantic coherence during query and text processing. This approach prevents the fragmentation of expressions, improving the accuracy of search results.

Text Segmentation and Vectorization

The text within the legal corpus is segmented into chunks, meaning groups of paragraphs that maintain a coherent informational unit. This segmentation is crucial to avoid calculating semantic similarity on text segments that are too brief or unrepresentative of the broader legal context. Each chunk is then vectorized using the same fine-tuned BERT model, generating an embedding that represents the chunk's semantic content. During the search process, the user query is also transformed into an embedding, and the similarity between this and the chunk embeddings is calculated to determine the relevance of the results.

Conclusion

The described retrieval system leverages an advanced and specialized approach for the Italian legal domain, combining traditional retrieval techniques with the power of semantic representation offered by Transformer models. The dynamic combination of BM25 and vector embeddings, along with the handling of complex legal expressions and informed text segmentation, enables the system to provide highly accurate results, effectively meeting the needs of legal professionals.

2.2 Generative Model for Drafting Legal Opinions

This paragraph outlines the architecture and functionality of a generative model designed for the automated drafting of legal opinions. The model, named Gemini-1.5-Flash, has been optimized to generate coherent and relevant texts based on excerpts from legal documents selected by the user, integrating these excerpts with extended context and related legal information. This system is designed to assist legal professionals in creating complex documents, ensuring precision and consistency at a low token cost.

Model Architecture

Gemini-1.5-Flash

The generative model used is Gemini-1.5-Flash, a state-of-the-art language model that offers an excellent balance between performance and cost. This model is particularly well-suited for legal domain applications due to its ability to handle extended contexts and generate high-quality textual outputs.

Model Input

The model uses several types of input to generate the legal opinion:

- a. **Selection of Text Excerpts:** The user manually selects text excerpts from relevant legal documents. These excerpts form the basis on which the model builds the legal opinion, ensuring that the final output is directly related to the selected sources.
- b. **References to Source Documents:** The model incorporates explicit references to the source documents, maintaining traceability and reliability of the generated opinion.
- c. **Context of Excerpts:** To improve understanding and processing of the excerpts, the model uses a context of 600 preceding words and 100 following words around the selected excerpts. This context helps the model better grasp the meaning and relevance of the excerpts as a whole, enriching the generated text with pertinent contextual information.
- d. **List of Legal Definitions:** The model can access a list of legal definitions uploaded to the system. These definitions are integrated into the legal opinion to ensure that legal concepts are used accurately and consistently.
- e. **Search for Similar/Related Case Law:** Using a Retrieval-Augmented Generation (RAG) approach, the model can search for case law similar or related to the selected excerpts. This further enriches the legal opinion with up-to-date and relevant jurisprudential references.

Advanced Prompt Engineering

An advanced prompt engineering approach is used to determine the model's behavior during text generation. This technique involves constructing specific prompts that guide the model in producing structured content that meets the user's expectations. The prompts may include directives on how to link the various excerpts, how to use legal definitions, and how to integrate related case law.

Few-Shot Learning

A key element in customizing the model's behavior is the use of few-shot learning. By exposing the model to a few examples of how text excerpts should be commented on, listed, and linked, the model quickly learns the desired writing structures and styles. This approach allows the model to adapt to specific stylistic and structural needs, improving the quality and relevance of the generated text.

Conclusion

The Gemini-1.5-Flash model represents an advanced and highly efficient solution for the automated generation of legal opinions. The combination of manual excerpt selection, extended contextualization, integration of legal definitions, and related case law searches offers unprecedented capability to produce complex and detailed legal documents. The use of advanced prompt engineering techniques and few-shot learning enables the model to generate highly personalized content, meeting the specific needs of legal professionals.

2.3 SYSTEM TRAINING

The solution leverages two artificial intelligence models, each with a specific purpose: one dedicated to information retrieval and the other to text generation.

The retrieval model was trained on **question-answer pairs** synthetically generated from a database of legal documents provided by a partner. Additionally, it was programmed to recognize **common legal expressions** as atomic units, thanks to a list developed by industry experts.

Subsequently, the model was further refined through **feedback collection** during its use by expert users, who, through a voting mechanism, provided their assessment on the relevance of the results.

The generative model, on the other hand, was trained using the **few-shot technique**. During the prompt engineering phase, examples consisting of input and expected output were included, developed in collaboration with legal specialists, to guide the generation of coherent and relevant responses.

3 EU AI Act Compliance Check

Below is the verification of compliance with EU AI law

Presentation data: July 19, 2024. 11:04 (UTC, Coordinated Universal Time)

Entity type: Distributor

System changes: None of the above

Systems excluded: None of the above

Prohibited systems: None of the above

High risk: Attachment 1 Section B: None of the above

High risk: Attachment 1 Section A: None of the above

High risk: Attachment 3: None of the above

Name of the artificial intelligence system: Lextel.AI (APP-AI-01)

Verification Results

To verify compliance with EU AI law, we took into account the main provisions and requirements of the legislation, in particular paying attention to critical aspects such as transparency, data management, security and protection of fundamental rights. The documentation of the artificial intelligence system in question was carefully analysed, examining the privacy policies, the data collection and processing methodologies, and the measures adopted to guarantee fairness and non-discrimination.

A series of tests and cross-checks were carried out to verify that all system functions respected the guidelines established by law. The result of the analysis showed that the system under investigation is generally compliant with the requirements of the EU AI Law. However, some areas for improvement have emerged,

The result is the following:

General Obligations for AI Models

"You must adhere to these obligations for General Purpose AI models under Article 53. In summary, you must:

- Create and maintain technical documentation of the AI model and make it available to the AI Office upon request.
- Create and maintain documentation for providers integrating AI models, balancing transparency and protection of intellectual property.
- Implement a policy to comply with Union copyright law.
- Publish a public summary of the AI model's training data following a template provided by the AI Office.

Additionally, consider whether the AI is used as a system or as a component of one. If so, the obligations related to high-risk AI systems may apply directly or indirectly under Recital 85.

Transparency Obligations: Summary Content

You must comply with the transparency obligations under Article 50:

- Ensure that the outputs of the AI system are marked in a machine-readable format and detectable as artificially generated or manipulated.

This does not apply to content authorized by law."

EU AI Act Compliance Checker output:

[Verifica della conformità alla legge sull'intelligenza artificiale dell'UE | Legge sull'intelligenza artificiale dell'UE \(artificialintelligenceact.eu\)](#)

4 Practical AI Ethics Checklist (from EU Experts Group)

Trustworthy AI Assessment List

Human agency and oversight

Fundamental rights:

1. Did you carry out a fundamental rights impact assessment where there could be a negative impact on fundamental rights? Did you identify and document potential trade-offs made between the different principles and rights?

Answer: Yes, we assessed the impact on fundamental rights, without detecting any particular issues.

2. Does the AI system interact with decisions by human (end) users (e.g. recommended actions or decisions to take, presenting of options)?

Answer: The system empowers users' information search and proposes documents abstracts, related articles and extracts, helping the clients to develop a legal opinion upon the searched item: there is no direct intervention of the system on users' decision-making process. It does not recommend actions and leaves the decision to consider the paper to be left exclusively to the user

3. Could the AI system affect human autonomy by interfering with the (end) user's decision-making process in an unintended way?

Answer: No direct impact: the tool simply supports legal professional activities, during search processes on legal information, ensuring human autonomy

4. Did you consider whether the AI system should communicate to (end) users that a decision, content, advice or outcome is the result of an algorithmic decision?

Answer: Lextel.AI outputs clearly state that they are the results of algorithmic processing and require users' validation. Users are clearly informed since the moment of subscribing to the service, that the final documents are the product of the use of GAI technologies. Moreover, users are informed that GAI is engaged through click action over the button "Start AI Elaboration"

5. In case of a chat bot or other conversational system, are the human end users made aware that they are interacting with a non-human agent?

Answer: Not Applicable. No chatbot are available in Lextel.AI

Human agency:

6. Is the AI system implemented in work and labour process? If so, did you consider the task allocation between the AI system and humans for meaningful interactions and appropriate human oversight and control?

Answer: Yes, the AI system supports legal work processes. The AI is invoked exclusively downstream of the user's selection of legal content and it is necessary for analyses of legal challenges.

7. Does the AI system enhance or augment human capabilities?

Answer: Yes, tool powers up human capabilities to retrieve and process information from different sources and unstructured formats, increasing human synthesis skills in the face of reading numerous legal documents

8. Did you take safeguards to prevent overconfidence in or overreliance on the AI system for work processes?

Answer: Not applicable. The system does not trigger/manage any process. The AI outputs are merely used as support of legal activities.

Human oversight:

9. Did you consider the appropriate level of human control for the particular AI system and use case?

Answer: Yes, the user is in control and takes last actions and strategy adopted for managing legal challenges

10. Can you describe the level of human control or involvement?

Answer: Users obtain retrieved documents, select relevant extracts and review tool elaboration. The papers produced by the AI are exclusively based on the texts selected by the user during the normal legal research process. The level of human control over AI is total, since the use of AI outputs is fully under users' responsibility

11. Who is the "human in control" and what are the moments or tools for human intervention?

Answer: The user is always in control. Lawyers enter search topics, evaluate system outputs and decide how to use them

12. Did you put in place mechanisms and measures to ensure human control or oversight?

Answer: Yes, user is the only responsible for how to use tool outputs. AI outputs are generated after clicking over the button "Start Legal Elaboration". Generative AI outputs are not mandatory for the user.

13. Did you take any measures to enable audit and to remedy issues related to governing AI autonomy?

Answer: We keep logs of all API calls to the AI systems used, that are periodically analysed to monitor tool proper functioning

14. Is there a self-learning or autonomous AI system or use case? If so, did you put in place more specific mechanisms of control and oversight?

Answer: No autonomous or self-learning AI system is employed. AI learning is bounded to the tool database documents upgrade

15. Which detection and response mechanisms did you establish to assess whether something could go wrong?

Answer: We keep logs of all API calls to the AI systems used. Users can provide feedback on AI outputs through "like/unlike" buttons

16. Did you ensure a stop button or procedure to safely abort an operation where needed? Does this procedure abort the process entirely, in part, or delegate control to a human?

Answer. Not applicable

Technical robustness and safety

Resilience to attack and security:

17. Did you assess potential forms of attacks to which the AI system could be vulnerable?

Answer. Regular VAPT sessions are planned (6 months)

18. Did you consider different types and natures of vulnerabilities, such as data pollution, physical infrastructure, cyber-attacks?

Answer. Yes, VAPT considers different vulnerabilities

19. Did you put measures or systems in place to ensure the integrity and resilience of the AI system against potential attacks?

Answer: Yes, specific components have been deployed such as for instance: firewalls, identity management tools, secure communication protocols, backup plan.

20. Did you verify how your system behaves in unexpected situations and environments?

Answer. Yes, test sessions are executed on different environments before deploy every tool upgrade

21. Did you consider to what degree your system could be dual-use? If so, did you take suitable preventative measures against this case (including for instance not publishing the research or deploying the system)?

Answer. Not applicable

Fallback plan and general safety:

22. Did you ensure that your system has a sufficient fallback plan if it encounters adversarial attacks or other unexpected situations (for example technical switching procedures or asking for a human operator before proceeding)?

Answer: Not applicable. No chatbot interaction is in place

23. Did you consider the level of risk raised by the AI system in this specific use case?

Answer. Yes, no specific needs have been identified

24. Did you put any process in place to measure and assess risks and safety?

Answer. Aligned with company standards (Policies & Procedures)

25. Did you provide the necessary information in case of a risk for human physical integrity?

Answer: Not applicable

26. Did you consider an insurance policy to deal with potential damage from the AI system?

Answer. In line with company standards, insurance policy covers both damage and Cybersec risks. If in the future, a dedicated company will be set up to market LextelAI solution, a similar policy will be subscribed by the newco.

27. Did you identify potential safety risks of (other) foreseeable uses of the technology, including accidental or malicious misuse? Is there a plan to mitigate or manage these risks?

Answer. Yes, risk analyses are always conducted on IT and Privacy risks for every application/service managed

28. Did you assess whether there is a probable chance that the AI system may cause damage or harm to users or third parties? Did you assess the likelihood, potential damage, impacted audience and severity?

Answer. Yes, IT and Privacy risks are conducted

29. Did you consider the liability and consumer protection rules, and take them into account?

Answer. Yes. Services are available only for lawyers, that are fully responsible for tool outputs use. The service is not intended for consumer market

30. Did you consider the potential impact or safety risk to the environment or to animals?

Answer. Not applicable

31. Did your risk analysis include whether security or network problems such as cybersecurity hazards could pose safety risks or damage due to unintentional behaviour of the AI system?

Answer. Yes, we address and periodically measure the risk of security or network problems

32. Did you estimate the likely impact of a failure of your AI system when it provides wrong results, becomes unavailable, or provides societally unacceptable results (for example discrimination)?

Answer. Yes. Tool supports lawyers' activity. The lawyer can carry out his/her own activity also in case of tool unavailability or provides unacceptable results.

33. Did you define thresholds and did you put governance procedures in place to trigger alternative/fallback plans?

Answer. Yes, service major features are legal content search and AI content generation. In case of down AI components, service is limited to search functions

34. Did you define and test fallback plans?

Answer.

Accuracy

35. Did you assess what level and definition of accuracy would be required in the context of the AI system and use case?

Answer. Yes, we assessed the accuracy level. WE have agreements with our info providers to ensure formats and contents accuracy

36. Did you assess how accuracy is measured and assured?

Answer. Yes, during system training, a professional team compared the AI's answers with the expected answers that a legal would give on the specific topic

37. Did you put in place measures to ensure that the data used is comprehensive and up to date?

Answer. Yes, our agreements with providers include a daily update process

38. Did you put in place measures to assess whether there is a need for additional data, for example to improve accuracy or to eliminate bias?

Answer. We collect feedback from users to monitor AI outputs accuracy

39. Did you verify what harm would be caused if the AI system makes inaccurate predictions?

Answer. No harm is done, AI-generated answers are controlled by legal users.

40. Did you put in place ways to measure whether your system is making an unacceptable number of inaccurate predictions?

Answer. We collect feedback from users to monitor AI outputs accuracy. Periodically, we will cross-check feedback to identify viable solutions (if needed)

41. Did you put in place a series of steps to increase the system's accuracy?

Answer. Users feedback will help increase system's accuracy

Reliability and reproducibility

42. Did you put in place a strategy to monitor and test if the AI system is meeting the goals, purposes and intended applications?

Answer. Yes, periodically we are gathering feedback from users who are using our tool

43. Did you test whether specific contexts or conditions need to be taken into account to ensure reproducibility?

Answer. AI outputs depend on document set that the legal submitted to the system.

44. Did you put in place verification methods to measure and ensure different aspects of the system's reliability and reproducibility?

Answer. No

45. Did you put in place processes to describe when an AI system fails in certain types of settings?

Answer. Problem determination is supported by system log analysis

46. Did you clearly document and operationalize these processes for the testing and verification of the reliability of AI systems?

Answer. In our training and testing sessions we involve a professional team specifically competent for legal domain

47. Did you establish mechanisms of communication to assure (end-)users of the system's reliability?

Answer. No, there are no messages to communicate the system's reliability.

Privacy and data governance

Respect for privacy and data Protection:

48. Depending on the use case, did you establish a mechanism allowing others to flag issues related to privacy or data protection in the AI system's processes of data collection (for training and operation) and data processing?

Answer. Yes. Portal provides a contact point for reporting questions relating to privacy

49. Did you assess the type and scope of data in your data sets (for example whether they contain personal data)?

Answer. Yes. Assessments relating to the processing of personal data have been conducted and relationships have been regularized in accordance with the provisions of EU 679/2016

50. Did you consider ways to develop the AI system or train the model without or with minimal use of potentially sensitive or personal data?

Answer. Yes, data used to train the AI system is acquired in anonymized form

51. Did you build in mechanisms for notice and control over personal data depending on the use case (such as valid consent and possibility to revoke, when applicable)?

Answer. Yes, Service portal informs the users about data management purposes and process on how to manage consent and possibility to revoke.

52. Did you take measures to enhance privacy, such as via encryption, anonymization and aggregation?

Answer. Yes, service adopted measures to guarantee privacy, through the use of encryption systems both in communication with the end user and in archiving

53. Where a Data Privacy Officer (DPO) exists, did you involve this person at an early stage in the process?

Answer. Yes, DPO was involved in early stage project and during initial risk analysis

Quality and integrity of data:

54. Did you align your system with relevant standards (for example ISO, IEEE) or widely adopted protocols for daily data management and governance?

Answer. Yes, system follows ISO 27000 standards

55. Did you establish oversight mechanisms for data collection, storage, processing and use?

Answer. Yes, mechanisms have been established for periodic verify collection, storage, processing and use of data including through the analysis of system logs

56. Did you assess the extent to which you are in control of the quality of the external data sources used?

Answer. Yes, quality of the external data sources used is guaranteed by specific agreements with suppliers and through random checks

57. Did you put in place processes to ensure the quality and integrity of your data? Did you consider other processes? How are you verifying that your data sets have not been compromised or hacked?

Answer. Yes, quality and integrity of the data is guaranteed by appropriate security measures which also include the use of Security Operation Center and back-up plan

Access to data:

58. What protocols, processes and procedures did you follow to manage and ensure proper data governance?

Answer. Yes, data governance is guaranteed through the use of cloud systems, the release of specific user profile for system administrators and through specific organizational procedures

59. Did you assess who can access users' data, and under what circumstances?

Answer. Yes, user data can only be accessed by users and, in specific maintenance circumstances, by system administrators

60. Did you ensure that these persons are qualified and required to access the data, and that they have the necessary competences to understand the details of data protection policy?

Answer. Yes, System Administrators are qualified and have the necessary skills to understand the details of the data protection policy. Any access to data is logged. Administrators is verified annually for the renewal of their positions

61. Did you ensure an oversight mechanism to log when, where, how, by whom and for what purpose data was accessed?

Answer. Yes, As specified above, any access to the data is logged. There are log protection mechanisms that allow them to not be modified

Transparency

Traceability:

62. Did you establish measures that can ensure traceability? This could entail documenting the following methods:

- Methods used for designing and developing the algorithmic system.
- Rule-based AI systems: the method of programming or how the model was built.
- Learning-based AI systems; the method of training the algorithm, including which input data was gathered and selected, and how this occurred.
- Methods used to test and validate the algorithmic system:
- Rule-based AI systems; the scenarios or cases used in order to test and validate;
- Learning-based model: information about the data used to test and validate.

Answer. Yes, we can produce (on demand) evidence on methods, rules and test/training procedures used to develop our system

63. Outcomes of the algorithmic system: the outcomes or decisions taken by the algorithm, as well as potential other decisions that would result from different cases (for example, for other subgroups of users).

Answer. Yes, outcomes result from different legal cases

Explainability:

64. Did you assess:

- to what extent the decisions and hence the outcome made by the AI system can be understood?
- to what degree the system's decision influences the organisation's decision-making processes?
- why the particular system was deployed in this specific area?
- what the system's business model is (for example, how does it create value for the organisation)?

Answer. Tool has been designed and developed to meet legal domain needs and business model was built accordingly. AI outcomes language is implemented leveraging on specific legal jargon and legal dictionary that has been purposely loaded onto the tool

1. Did you ensure an explanation as to why the system took a certain choice resulting in a certain outcome that all users can understand?
- Did you design the AI system with interpretability in mind from the start?
 - Did you research and try to use the simplest and most interpretable model possible for the application in question?
 - Did you assess whether you can analyse your training and testing data? Can you change and update this over time?
 - Did you assess whether you can examine interpretability after the model's training and development, or whether you have access to the internal workflow of the model?

Answer. AI output template has been validated by our legal domain experts, the tool uses legal jargon and outputs are validated by users

Communication:

65. Did you communicate to (end-)users – through a disclaimer or any other means – that they are interacting with an AI system and not with another human? Did you label your AI system as such?

Answer. Yes, users are aware when they interact with AI-systems.

66. Did you establish mechanisms to inform (end-)users on the reasons and criteria behind the AI system's outcomes?

Answer. AI outcomes depend on legal users' documents selection

67. Did you communicate this clearly and intelligibly to the intended audience?

Answer. Yes, it is reported on the Portal and Customer care may provide information about this

68. Did you establish processes that consider users' feedback and use this to adapt the system?

Answer. Yes, we regularly collect users' feedback on each outcome

69. Did you communicate around potential or perceived risks, such as bias?

Answer. We have dedicated disclaimers

70. Depending on the use case, did you consider communication and transparency towards other audiences, third parties or the general public?

Answer. In our webinar channel we reach different audiences and clarify how our tool functions

71. Did you clarify the purpose of the AI system and who or what may benefit from the product/service?

Answer. Yes, it is in our main value proposition and it is the core of our communication campaign

72. Did you specify usage scenarios for the product and clearly communicate these to ensure that it is understandable and appropriate for the intended audience?

Answer. it is explained in our tutorial, live events, webinar

73. Depending on the use case, did you think about human psychology and potential limitations, such as risk of confusion, confirmation bias or cognitive fatigue?

Answer. Not applicable

74. Did you clearly communicate characteristics, limitations and potential shortcomings of the AI system?

Answer. Yes

75. In case of the system's development: to whoever is deploying it into a product or service?

Answer. Yes

76. In case of the system's deployment: to the (end-)user or consumer?

Answer. Yes

Diversity, non-discrimination and fairness

Unfair bias avoidance:

77. Did you establish a strategy or a set of procedures to avoid creating or reinforcing unfair bias in the AI system, both regarding the use of input data as well as for the algorithm design?

Answer: *We use external AI services and we do not have control over the algorithm design. Our applications do not feed AI system with sensitive data or other types of data that can lead to unfair bias.*

78. Did you assess and acknowledge the possible limitations stemming from the composition of the used data sets?

Answer: *Yes, for this reason the outcome of the AI processing is always controlled by human users.*

79. Did you consider diversity and representativeness of users in the data? Did you test for specific populations or problematic use cases?

Answer: *No, we did not consider this aspect because we do not use personal data*

80. Did you research and use available technical tools to improve your understanding of the data, model and performance?

Answer. Not applicable

81. Did you put in place processes to test and monitor for potential biases during the development, deployment and use phase of the system?

Answer: *No, we did not put in place processes to test and monitor for potential biases.*

82. Depending on the use case, did you ensure a mechanism that allows others to flag issues related to bias, discrimination or poor performance of the AI system?

Answer. Customer care can collect any issues from users

83. Did you establish clear steps and ways of communicating on how and to whom such issues can be raised?

Answer. Yes, there is an escalation process. Customer care has a clear contact point

84. Did you consider others, potentially indirectly affected by the AI system, in addition to the (end) users?

Answer: *No, for these restricted use cases we did not consider others that are not end users of the applications.*

85. Did you assess whether there is any possible decision variability that can occur under the same conditions?

Answer: *Yes, for this reason the user is asked to control the outcomes.*

86. If so, did you consider what the possible causes of this could be?

Answer: *Yes, it is the way Generative AI tools works. They provide stochastic answers, they are not deterministic.*

87. In case of variability, did you establish a measurement or assessment mechanism of the potential impact of such variability on fundamental rights?

Answer. *Not applicable*

88. Did you ensure an adequate working definition of “fairness” that you apply in designing AI systems?

Answer. Tool has been developed according to Visura standards

89. Is your definition commonly used? Did you consider other definitions before choosing this one?

Answer. It is a standard

90. Did you ensure a quantitative analysis or metrics to measure and test the applied definition of fairness?

Answer. No

91. Did you establish mechanisms to ensure fairness in your AI systems? Did you consider other potential mechanisms?

Answer. No

Accessibility and universal design:

92. Did you ensure that the AI system accommodates a wide range of individual preferences and abilities?

Answer. Tools that make use of AI systems are very tailored to specific use cases and applications.

93. Did you assess whether the AI system usable by those with special needs or disabilities or those at risk of exclusion? How was this designed into the system and how is it verified?

Answer. No, we did not assess it.

94. Did you ensure that information about the AI system is accessible also to users of assistive technologies?

Answer. No, we did not ensure that.

95. Did you involve or consult this community during the development phase of the AI system?

Answer. No

96. Did you take the impact of your AI system on the potential user audience into account?

Answer. No, we did not consider the impact on potential users with special needs.

97. Did you assess whether the team involved in building the AI system is representative of your target user audience? Is it representative of the wider population, considering also of other groups who might tangentially be impacted?

Answer. Yes, we involved a team of external legal domain experts since design phase and all along the development phase

98. Did you assess whether there could be persons or groups who might be disproportionately affected by negative implications?

Answer. Not applicable

99. Did you get feedback from other teams or groups that represent different backgrounds and experiences?

Answer. Yes, initial focus group involved people with different background

Stakeholder participation:

100. Did you consider a mechanism to include the participation of different stakeholders in the AI system's development and use?

Answer: Yes, different stakeholders were involved in all the phases of the projects.

101. Did you pave the way for the introduction of the AI system in your organization by informing and involving impacted workers and their representatives in advance?

Societal and environmental well-being

Sustainable and environmentally friendly AI

Answer. Yes, we made introductory meetings (both with internal and external audiences) on the use of AI, especially for specific needs where we are adopting it.

102. Did you establish mechanisms to measure the environmental impact of the AI system's development, deployment and use (for example the type of energy used by the data centers)?

Answer: No, we don't

103. Did you ensure measures to reduce the environmental impact of your AI system's life cycle?

Answer: No, we did not ensure these measures

Social impact:

104. In case the AI system interacts directly with humans:

- Did you assess whether the AI system encourages humans to develop attachment and empathy towards the system?

Answer. No, we did not assess this particular potential effect.

105. Did you ensure that the AI system clearly signals that its social interaction is simulated and that it has no capacities of "understanding" and "feeling"?

Answer: Not applicable

106. Did you ensure that the social impacts of the AI system are well understood? For example, did you assess whether there is a risk of job loss or de-skilling of the workforce? What steps have been taken to counteract such risks?

Answer. We are aware of these social impacts, we have implemented, and we are implementing systems where AI make the work easier for the users, without removing their need and reducing their centrality.

Society and democracy:

107. Did you assess the broader societal impact of the AI system's use beyond the individual (end-)user, such as potentially indirectly affected stakeholders?

Answer. No, we did not assess the broader societal impact of the specific AI-enhanced systems that we currently have in Production.

Accountability

Auditability:

108. Did you establish mechanisms that facilitate the system's auditability, such as ensuring traceability and logging of the AI system's processes and outcomes?

Answer. Yes, there are logs on the transactions and requests that are made.

109. Did you ensure, in applications affecting fundamental rights (including safety-critical applications) that the AI system can be audited independently?

Answer: Not applicable

Minimising and reporting negative Impact:

110. Did you carry out a risk or impact assessment of the AI system, which takes into account different stakeholders that are (in)directly affected?

Answer. We conducted an overall risk impact assessment

111. Did you provide training and education to help developing accountability practices?

Answer. We deployed and have plan to deploy several introductory sessions to educate on AI domain and our approach to it

112. Which workers or branches of the team are involved? Does it go beyond the development phase?

Answer. We involved: designers, developers, testers and customer care operators and company strategic committee

113. Do these trainings also teach the potential legal framework applicable to the AI system?

Answer. Yes

114. Did you consider establishing an 'ethical AI review board' or a similar mechanism to discuss overall accountability and ethics practices, including potentially unclear grey areas?

Answer. Company boards

115. Did you foresee any kind of external guidance or put in place auditing processes to oversee ethics and accountability, in addition to internal initiatives?

Answer. No

116. Did you establish processes for third parties (e.g. suppliers, consumers, distributors/vendors) or workers to report potential vulnerabilities, risks or biases in the AI system?

Answer. Yes, users know how to report vulnerabilities that they recognize with the use of the systems.

Documenting trade-offs:

117. Did you establish a mechanism to identify relevant interests and values implicated by the AI system and potential trade-offs between them?

Answer. No, we did not establish this kind of mechanism.

118. How do you decide on such trade-offs? Did you ensure that the trade-off decision was documented?

Answer. Not applicable

Ability to redress:

119. Did you establish an adequate set of mechanisms that allows for redress in case of the occurrence of any harm or adverse impact?

Answer. Yes, defined in service T&C

120. Did you put mechanisms in place both to provide information to (end-) users/third parties about opportunities for redress?

Answer. Not applicable