

AI IN SOCIAL SECURITY: STATE OF THE ART, PROMISES AND CHALLENGES

The document is commissioned to support the event ‘Shaping Tomorrow: AI in Social Security and Work Capacity Assessment’ to be held at the European Economic and Social Committee (EESC), Brussels, March 06, 2024. The event is organised by the European Union of Medicine in Assurance and Social Security (EUMASS) in collaboration with the Belgian Ministry of Social Affairs (Federal Public Service FPS) and the National College of Insurance Medicine (Belgium), with the support of the Belgian Presidency of the Council of the European Union (1 January to 30 June 2024).

AUTHOR

Anastasiya Kiseleva

LINGUISTIC VERSIONS

Original: EN

DISCLAIMER AND COPYRIGHT

The opinions expressed in this document are the sole responsibility of the author and do not necessarily represent the official position of the EUMASS or FPS. Reproduction and translation for non-commercial purposes are authorised, provided the source is acknowledged and the author is given prior notice and sent a copy.

The manuscript was completed in December 2023.

Table of Contents

EXECUTIVE SUMMARY	3
1. INTRODUCTION.....	4
2. SURVEY OVERVIEW.....	6
2.1. OBJECTIVES.....	6
2.2. SCOPE.....	6
2.3. METHODOLOGY.....	7
2.4. RESPONSES RECEIVED	8
3. KEY FINDINGS	9
3.1. RESPONDENTS PERCEIVE AI IN A PRIMARILY POSITIVE WAY	9
3.2. REMAINING CONCERNS ABOUT AI, ITS NATURE, CAPACITIES AND RISKS	10
3.3. RESPONDERS BELIEVE THAT AI APPLICATIONS IN SOCIAL SECURITY SHOULD BE INTERNALLY AND EXTERNALLY VERIFIED BEFORE BEING USED TOWARD INDIVIDUALS	12
3.4. RESPONDERS PREFER BEING SUFFICIENTLY INFORMED ABOUT AI APPLICATIONS THEY USE.....	13
3.5. RESPONDERS DEMONSTRATE THEIR DEMAND AND WILLINGNESS TO BE SUPPORTED BY AND COOPERATE WITH AI DEVELOPERS	16
4. RECOMMENDATIONS.....	18
4.1. RAISING AWARENESS	18
4.2. RIGOROUS CONTROL OF SAFETY AND QUALITY IS STRONGLY REQUIRED	21
4.3. TRANSPARENCY IS THE KEY FACTOR FOR USERS' ACCEPTANCE	26
4.4. RELIABLE IMPLEMENTATION OF AI SHALL BE BASED ON COOPERATIVE EFFORTS	30
APPENDIX 1. INTRODUCTION TO THE SURVEY	33
APPENDIX 2. LIST OF QUESTIONS POSED TO THE POTENTIAL USERS OF AI	35
APPENDIX 3. LIST OF QUESTIONS POSED TO THE CURRENT USERS OF AI	39
APPENDIX 4. TABLE ON AI TRANSPARENCY EXPECTATIONS AND EXPERIENCES.....	44

Executive Summary

This report presents the current landscape of AI in social security and work capacity assessment. It is based on the survey conducted among stakeholders involved in the field to assess their perceptions, concerns, and experiences regarding AI applications. This report describes the results of the survey and translates them into recommendations for the reliable and effective implementation of the technology in the sector.

The survey revealed a strongly positive approach among respondents regarding the potential of AI in social security. However, despite the optimism, there are remaining concerns of participants regarding AI nature, capabilities, and associated risks. They emphasise the importance of rigorous verification processes—both internal and external—during the whole life cycle of AI applications in social security settings. In addition, participants indicated a clear preference for being sufficiently informed regarding employed AI applications. Finally, responders demonstrate their strong demand and willingness to be supported by and cooperate with AI developers.

This report underscores the need for a balanced approach, emphasising the potential benefits of AI while addressing concerns and advocating for transparent, ethical, and collaborative practices in the integration of technology within social security systems. Based on the described findings, key recommendations emerge:

- **Raising Awareness**: it is imperative to support and facilitate initiatives aimed at educating stakeholders, including users, policymakers, and professionals, about AI capabilities, limitations, and implications within social security. This awareness-building process is essential to mitigate misconceptions and enhance informed decision-making.
- **Rigorous Safety and Quality Control**: A robust framework for safety and quality control must be clearly defined and rigorously implemented. This framework should ensure thorough verification processes before deploying AI solutions in social security contexts, ensuring reliability and minimising risks to individuals.
- **Transparency**: Transparency stands as a key factor in gaining user acceptance and fostering trust. Therefore, efforts to promote transparency in AI applications—ensuring users are well-informed about the technology's functionalities and implications—are essential.
- **Cooperative Implementation**: Reliable implementation of AI in social security necessitates collaborative efforts between users, deployers, developers, regulators, political decision-makers and affected individuals.

1. Introduction

Artificial Intelligence (AI) is gradually transforming our lives. ‘This fast-evolving family of technologies can provide a wide array of economic and societal benefits across the entire spectrum of industries and social activities.’¹ Social security holds the paramount position among the sectors where the vast potential of AI must be harnessed.

Social security contributes to overall societal stability by reducing poverty, improving access to healthcare, and offering a sense of security and dignity to individuals and families. To continue promoting a more equitable and compassionate society where everyone has the opportunity to thrive, regardless of life's uncertainties, it is essential to support the further development of social security through cutting-edge innovations. Such support is guaranteed by the initiatives taken by the EU institutions, public bodies and key stakeholders in the sector.

In recent years, the EU has demonstrated its strong commitment to preserve the EU’s technological leadership and simultaneously ensure that new technologies are developed and function according to Union values, fundamental rights and principles. The European Commission highlights: ‘AI should be a tool for people and be a force for good in society with the ultimate aim of increasing human well-being.’²

In pursuing the described commitment, the European Commission issued the Proposal for the AI Act – the legislation for a harmonised EU approach to regulating AI, based on EU values and fundamental rights and aimed to give people and other users the confidence to embrace AI-based solutions, while encouraging businesses to develop them. Since April 2021, the legislative proposal has been heavily discussed. In December 2023, a political agreement on the AI Act was reached. In the words of Ursula von der Leyen, President of the European Commission, this historic moment transposes European values to a new era of AI-driven, but human-centric innovations.³

Starting in January 2024, the Belgian Presidency of the Council of the European Union will create the strongest platform to align the facilitation of AI innovations with the future-proof social agenda. ‘Building on the European Pillar of Social Rights, the Belgian Presidency aims to equip the EU with an ambitious social agenda to foster a European society that is

¹ The EC Proposal for the 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 as of April 21, 2021, COM/2021/206 final (EC Proposal for the AI Act), Recital 3.

² Explanatory Memorandum to the EC Proposal for the AI Act, 1.

³ European Commission, ‘Commission welcomes political agreement on Artificial Intelligence Act’ (Official Press Release, December 09, 2023), available at: < https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473> accessed December 15, 2023.

more inclusive, gender-equal and fair for all.⁴ It is aimed at delivering a stronger social safety net for European citizens and a labour market equipped for the future.⁵ ‘Simultaneously, the Presidency will uphold the EU’s ambition to advance the digital transition. It will prioritise a human-centered approach to digital transformation, especially relevant in the domain of AI. In pursuit of this goal, the Belgian Presidency will strive to finalise any remaining work on the AI Act.’⁶ The outlined directions of work during the Belgian Presidency of the Council of the European Union strongly correlate with the agenda of the key stakeholders in the social security sector.

The European Union of Medicine in Assurance and Social Security (EUMASS) is a European federation of national associations or organisations of doctors involved in insurance and social security medicine in their country. It aims to provide a platform for the exchange of experience and knowledge in the field of insurance medicine, mainly in the field of public social security.⁷ EUMASS with its deep-seated history of fostering collaboration and advancements in the field, has persistently worked towards enhancing the landscape of social security in the European context. As we venture further, it is imperative that we carry forward the mantle of innovation and inclusivity that has been the cornerstone of EUMASS’s endeavors.

Committed to facilitating responsible innovations, EUMASS makes significant efforts to realise the full potential of AI in insurance medicine in a trustworthy and reliable way. To explore the opportunities and potential risks associated with AI, EUMASS commissioned the survey to assess the perceptions, concerns, and experiences of insurance physicians and other relevant professionals concerning AI applications. This report represents the findings of the survey and based on them, recommendations for the effective and reliable implementation of AI in insurance medicine and work capacity assessment.

⁴ Official Program of the Belgian Presidency of the Council of the European Union (from the 1st of January to the 30th of June 2024), ‘Protect, Strengthen, Prepare’, available at: < https://belgian-presidency.consilium.europa.eu/media/3kajw1io/programme_en.pdf> accessed December 15, 2023, 5.

⁵ *ibid*, 26.

⁶ *ibid*, 37.

⁷ EUMASS, ‘What do we Stand for?’, a word of the EUMASS President Dr Jean-Pierre Baron Schenkelaars as of July 2022, available at: < <https://eumass.eu/wp-content/uploads/2023/10/DEF-1-E-WORD-PRESIDENT.pdf>> accessed December 15, 2023.

2. Survey Overview

2.1. Objectives

The survey pursued several objectives:

- Identify the current positioning of AI in social security and work capacity assessment;
- Assess awareness and perceptions of the stakeholders regarding the role of AI in social security and work capacity assessment;
- Recognise the promises and positive contributions of AI in social security and work capacity assessment;
- Determine the most pressing concerns related to the development and implementation of AI in social security and work capacity assessment;
- Based on the survey's findings, develop the preliminary recommendations for the trustworthy integration of AI in social security.

2.2. Scope

The survey was designed to gather both the experiences and perceptions of the EUMASS community regarding AI and its implementation. To consider the differences in perception of the technology between those who have already applied AI in their professional activities and those who have never done that, the survey (preceded by the introductory part, see [Annex I](#)) started with the question below:

Do you use an AI application at work within your role and your specialty?

If you do not use an AI application at this moment but have done it in the past, please reply 'yes' to share your experience. We use the term 'AI' in the broadest sense, when a computer can mimic human intelligence, learn (adapt the performance based on new data) and make decisions to perform a given task.

Depending on the answer to the first question, the survey then broke down into two types of questionnaires: addressed to ‘non-users’⁸ (see the full list of questions in [Annex II](#)) and to ‘users of AI’⁹ (see in [Annex III](#)).

Although the survey considered the differences between users and non-users of AI, to generate a common picture for both groups, the posed questions were grouped to cover similar topics:

- *information about a participant;*
- *assessment of the intent to start using AI or of the experience with it;*
- *views of the participants about the benefits and challenges of using AI in their work;*
- *assessment of responders’ expectations regarding transparency: about the technology, its features, benefits and warnings, how an AI application is developed, tested and validated;*
- *evaluation of the received or expected support from AI developers;*
- *exploring the demand and willingness to cooperate with AI developers;*
- *final recommendations and thoughts.*

2.3. Methodology

The survey design employed a mixed-method approach, utilising both quantitative and qualitative research techniques. The data was collected through several types of questions: closed-ended questions (yes/no, multiple choice, Likert scale) and open-ended questions.

Likert scale questions (1-10) were used to assess the attitude of the responders toward AI, starting from the negative (from 0 to 4), neutral (5) to positive (from 6 to 10).

To facilitate quantification and decrease the time expenditures for responding to the survey, the majority of questions required ‘yes/no’ answers. To enable responders to elaborate on the specific questions, most of the closed-ended questions included the margins for comments. The questions on final thoughts and recommendations were open-ended.

As AI is a rather complex and novel technology, to facilitate the comprehension of the posed questions, the author of the survey offered the webinar to clarify any questions that participants could have while responding to the survey. The webinar was scheduled to take place on October 30, 2023, with 2 hours reserved. As the attendees (2) did not have any uncertainties regarding the content of the survey, the webinar focused mostly on organisational matters and lasted approximately 20 minutes.

⁸ Responders who did not have the experience of using AI applications in their professional activities.

⁹ Responders who had the experience of using AI applications in their professional activities.

2.4. Responses received

The survey received a robust reaction from the EUMASS community, with a total of 45 submitted responses. The responses were collected over a month (October 11, 2023 – November 12, 2023), ensuring sufficient term for participants to familiarise themselves with the questions and contemplate their responses.

Among the survey participants, the *group of 'AI non-users'* prevails, comprising 34 non-users (75,56%), whereas there are 11 users (24,44%). The responders represent a broad **geography** of the EUMASS community, with Belgium (16), the Netherlands (5), France (5) and Switzerland (4) taking the majority of the participants. Most of the responders are *experienced*: 20 (44%) have experience of 16-30 years and 16 (36%) have experience of more than 30 years. The *roles of represented groups* include, inter alia, medical advisors (15), insurance physicians (6), and researchers (4).

3. Key Findings

3.1. Respondents perceive AI in a primarily positive way

The survey included the Likert scale (1-10) to assess the attitude of responders toward AI. The group of users was asked to assess their experience with applying AI in their work (negative (1-4), neutral (5) and positive (6-10)). In the majority, responders are either positive (45.45% in total),¹⁰ or neutral (45.45%)¹¹ about their experience with the use of AI. The group of non-users was asked to assess their intent to start using AI (no intent (1-4), neutral (5) and strong intent)). In this case, the positive attitude strongly prevails: 61.67% of responders have the intent to use AI in their professional practice.¹² A neutral position was demonstrated by 23.53% of participants.

Responders among non-users declared several benefits they see in applying AI with the time savings¹³ and tackling the shortage of medical advisors¹⁴ prevailing in the answers. Participants also specified as positive effects of AI: improving knowledge,¹⁵ increasing accuracy of decision-making,¹⁶ aiding with automated and administrative tasks,¹⁷ optimising the processes,¹⁸ and increasing objectivity of the assessment.¹⁹ One of the responders highlighted: *'We urgently need to embrace the huge possibilities in AI as a help in our daily practice.'*²⁰ Another potential user shared: *'I expect it [AI] to change our work field in the next few years.'*²¹

Current users of AI applications provided positive use-cases. For example, one of the responders from Belgium (staff insurance physician) specified that they apply an algorithm that determines how and when claimants should be contacted following their sick report. The satisfaction of the usage was rated as 8 out of 10 and the responder commented that the AI application *'appears to be a practical first-shift that reduces scarce physician capacity unnecessarily deployed on short-term sick leave.'*²²

¹⁰ The relevant numbers are: 18,18 % assessed as 6, 18,18 % assessed as 8, and 9,09 % assessed as 10.

¹¹ Several responses covered the use of Chat-GPT, which is rather promising, but is not designed for the specific purposes to use in social security, which might explain the neutral results.

¹² The relevant numbers are: 5,88% assessed as 5; 26,47% assessed as 7; 17,65% evaluated as 8; 11,76% assessed as 10.

¹³ ID 117; ID 17; ID 44; ID 56; ID 80; ID 94; ID 107; ID 117.

¹⁴ ID 36; ID 37; ID 80.

¹⁵ ID 94; ID 44; ID 56.

¹⁶ ID 97.

¹⁷ ID 117; ID 124.

¹⁸ ID 77; ID 84; ID 111, ID 62.

¹⁹ ID 117; ID 124.

²⁰ ID 64.

²¹ ID 107.

²² ID 72.

3.2. Remaining concerns about AI, its nature, capacities and risks

Despite the prevailing positive attitude toward the use of AI in social security and work capacity assessment, participants expressed their concerns regarding the technology. The outlined considerations relate to varied matters, including the nature and capacities of the technology, its impact on the profession and provision of services, reliability and safe implementation, ethical and legal issues, and availability of technical and organisation sources.

As for the nature and capacities of the technology, some of the responders described it as ‘SI, *intelligence superficiel, no AI*,²³ ‘*big brother is watching us*’²⁴ and one non-user participant suggested forbidding the use of AI applications because it is unnecessary, complicated and might result in the loss of competence.²⁵

Several responders from non-users expressed their worries about the impact of AI on the profession and provision of services. They mentioned possible consequences: ‘*loss of personal competence, no more individual solution for each case*,’²⁶ ‘*simple work could be replaced by AI which means job loss for some colleagues*,’²⁷ ‘*dehumanisation of social security*,’²⁸ ‘*resistance to new ways of working*.’²⁹

Reliability and safe implementation of AI is seen by participants as the crucial factor for adopting the technology by professionals in the social security sector. The majority of non-users (73,53%) expressed that they see challenges to the adoption and implementation of AI in their field of work, including the reliability of AI,³⁰ the quality³¹ and accuracy³² of algorithmic recommendations. Participants among users shared their experiences that prevented them from completing the implementation of AI in their work routines: ‘*the challenge was the lack of 100% reliability and getting user acceptance (interpretation of algorithmic recommendations required extra work for the medical advisors)*.’³³

Responders outlined the availability of technical and organisation sources as an important factor for the adoption of AI in their field of work. Among the technical challenges, participants mentioned ‘*the need for a technical team for coding and problem-solving*,’³⁴

²³ French, ‘superficial intelligence’, ID 23.

²⁴ ID 117.

²⁵ ID 41.

²⁶ ID 41.

²⁷ ID 84.

²⁸ ID 90.

²⁹ ID 94.

³⁰ ID 88.

³¹ ID 45.

³² ID 107.

³³ ID 68.

³⁴ ID 17.

'insufficient computerization,'³⁵ data availability,³⁶ security and integrity.³⁷ Highlighted organisational challenges include 'the need for a shift in organisational structure and routines, for expanding the number of personnel with computer competency,'³⁸ scarcity of financial and human resources for the technological implementation,³⁹ 'adaptation of IT tools used in institutions,'⁴⁰ indefinite time investments for training professionals to use AI.⁴¹

Among ethical and legal challenges, responders highlighted the privacy and data protection issues,⁴² including data breaches due to cyber-attacks⁴³ or loss of confidentiality,⁴⁴ risk of discrimination,⁴⁵ and allocation of responsibility in case of implication of errored algorithmic outcomes.⁴⁶

³⁵ ID 111.

³⁶ ID 56.

³⁷ ID 44.

³⁸ ID 62.

³⁹ ID 94.

⁴⁰ ID 107.

⁴¹ ID 126.

⁴² ID 12, ID 59, ID 64, ID 88, ID 107.

⁴³ ID 64.

⁴⁴ ID 117.

⁴⁵ ID 37; ID 64.

⁴⁶ ID 94.

3.3. Responders believe that AI applications in social security should be internally and externally verified before being used toward individuals

Potential users of AI were asked whether they believe that AI applications in social security should be internally and externally verified before being used toward individuals. The majority of the responders (85,29%) positively replied to the question. They explained that such verification is essential⁴⁷ and *'will allow to trust the system.'*⁴⁸ A participant outlined: *'the impact of using AI in our decisions is very profound, therefore verification before use is indispensable.'*⁴⁹ Several responders highlighted that *'AI is a tool, all tools should have a quality check before being implemented.'*⁵⁰ Finally, one of the responders suggested that AI applications in social security should follow the same strict regulation as for the approval of medicines and authorisation of medical devices.⁵¹

⁴⁷ ID 80.

⁴⁸ ID 36; ID 90.

⁴⁹ ID 94.

⁵⁰ ID 64; ID 85.

⁵¹ ID 62; another responder also referred to the German proceedings of medical devices (ID 44).

3.4. Responders prefer being sufficiently informed about AI applications they use

Surveys for both groups of participants contained detailed sections with questions to assess the level of transparency experienced (by users) or expected (by non-users) during the use of AI applications. The submitted responses clearly demonstrate that transparency is extremely highly valued by both actual and prospective users of AI applications.

First, responders were asked about their views on the additional educating and training for using an AI application. The majority of prospective users (88,24%) positively responded to the question. As highlighted by one of the responders, *'it is obligatory, especially to ensure the feeling of being empowered by the AI and not overruled by.'*⁵²

Second, responders were queried about specific information grouped into several categories: general information about an AI application; usage instructions; benefits and risks of an AI application. Among the group of users, the majority focused on their experiences with generalised AI models (such as Chat-GPT). These models are not explicitly tailored for specialised fields like social security, lack external validation in terms of safety and quality, and do not exhibit a high level of transparency towards users.

To ensure more accurate statistics specifically relevant to the social security sector, user responses were divided into two groups: 1) encompassing all users; and 2) related to users of AI applications designed for specific social security purposes. The details of the responses regarding the transparency of participants from the latter group are elaborated in Annex 4.

Table 1 clearly illustrates that the vast majority of potential users prefer comprehensive information about an AI application. This includes details about its capacities, performance, target groups and contra-indications, the relevant benefits and risks, as well as sufficient instructions. One of the responders outlined: *'People need to know and must be aware what the abilities but also the shortcomings of AI are, in order to use it in the right (and ethically correct) way in their work.'*⁵³

Transparency expectations of users-to-be align with the positive experiences reported by the current users of AI applications designed for specific purposes. In most cases, this group of users was adequately informed about informational aspects regarding the AI application they used.

⁵² ID 94.

⁵³ ID 98.

Table 1. Transparency expectations and experiences of responders regarding AI applications

QUESTION	PERCENTAGE OF POSITIVE REPLIES		
	Non-users	All users	Users of AI applications designed for specific purposes
GENERAL INFORMATION ABOUT AN AI APPLICATION			
Nature, functions and intended purpose	94,12%	45,45%	100%
Performance characteristics, accuracy level and stability	91,18%	36,36%	100%
Target groups and contra-indications	94,12%	36,36%	100%
INSTRUCTIONS			
How to handle/filter the input data to be added to the AI application	85,29%	9,09%	33,33%
How to assess the correctness of the outcomes/suggestions provided by the AI application	82,35%	27,27%	66,67%
How the quality and safety of an AI application were tested and verified	85,29%	27,27%	100%
How the training, testing and validation data was selected and handled to ensure data quality and sufficient representativeness	82,35%	36,36%	100%
Embedded explanations of the algorithmic recommendations	88,24%	60,00%	66,67%
BENEFITS OF AN AI APPLICATION			
How an AI application performs in comparison to non-AI systems	82,35%	14,29%	33,33%
How an AI application performs in comparison to human specialists	82,35%	57,14%	100%
Claimed accuracy level of an AI application	91,18%	42,86%	100%
RISKS OF AN AI APPLICATION			
Risk of algorithmic opacity	82,35%	66,67%	33,33%
Risk of shifts in the accuracy level/data shifts	94,12%	66,67%	33,33%
Limited ability to verify the correctness of the specific AI outcome	94,12%	66,67%	33,33%
Risk of latent biases	94,12%	66,67%	33,33%
Information about how an AI provider handles the limitations, minimises and manages the risks	85,29%	33,33%	33,33%

In addition to the 'yes/no' questions described in Table 1, responders among non-users were given an open question regarding their transparency expectations. Overall, participants outlined that they wish to know about: the limitations and bias of the device,⁵⁴

⁵⁴ ID 36; ID 90.

whether the model is accurate and trained on the accurate data;⁵⁵ whether the AI was well developed and the risk of error is insignificant,⁵⁶ how the data is managed,⁵⁷ how algorithms function and make decisions,⁵⁸ how the specific algorithmic outcome was generated and whether it is reliable,⁵⁹ about the algorithmic “black-box” and its supervision,⁶⁰ whether the use of the AI application is compliant with the GDPR,⁶¹ and whether the AI application went through the certification by the relevant authorities.⁶²

Finally, responders were asked regarding informing their clients or patients about the use of the AI application in certain cases. The majority of ‘non-users’ participants (82,35%) outlined that such informing is necessary. They elaborated on the responses mentioning that patients should be aware⁶³ and know that *‘their case is solved by software’*⁶⁴ and it should be *‘just like when you propose any kind of test.’*⁶⁵ However, several responders expressed that informing patients is not necessary *as soon as the responsibility for the end result stays with a clinician.*⁶⁶ One responder suggested a balanced approach: *‘patients should be warned that the assessment will be carried out with the help of the DI [digital instruments/information], but will be validated by a health professional.’*⁶⁷ ‘Users’ of AI were asked if they currently inform patients about the use of the AI application and none of them positively replied to this question.

⁵⁵ ID 107.

⁵⁶ ID 111.

⁵⁷ ID 39; ID 37. ID 101.

⁵⁸ ID 44; ID 80; ID 90; ID 94; ID 101.

⁵⁹ ID 94.

⁶⁰ ID 124.

⁶¹ ID 59; ID 84.

⁶² ID 12; ID 117.

⁶³ ID 88.

⁶⁴ ID 41.

⁶⁵ ID 36.

⁶⁶ ID 107; ID 94; ID 64.

⁶⁷ ID 117.

3.5. Responders demonstrate their demand and willingness to be supported by and cooperate with AI developers

Responders were queried about the mechanism for contacting AI providers/seeking their support when they have any questions or doubts about the use of AI and the implication of AI outcomes. Users were asked if they had such a mechanism, and only 2(18,18%) of them positively replied to this question. One of the users described: *'The AI application is developed within our organisation so that discussion and, if necessary, adjustments are always possible.'*⁶⁸ Importantly, this user expressed a strongly positive experience with AI, thus confirming the importance of the providers' support in smooth technological implementation. Similarly, the majority of non-users (76.47%) confirmed the demand to have a mechanism for contacting AI providers. Regarding the preferred frequency, participants shared different views: on demand;⁶⁹ constantly on-going⁷⁰ or during working hours;⁷¹ more frequent support during the initial phases of the technological implementation⁷² (*'it is similar to when you start working with a new device, for example, echograph'*⁷³).

Participants were asked about the involvement of the representatives of healthcare and social security professionals in the pre-market assessment and verification of AI applications. Both groups (users with 63.64% and non-users with 91.18%) strongly agreed that such involvement is necessary. One of the users shared the relevant experience: *'In developing the AI application, it proved very useful to involve professionals; outline the context of use and test and validate the use.'*⁷⁴ Potential users expressed that it is crucial to involve healthcare professionals, as they assess the recommendations provided by AI applications and take responsibility for the final decision.⁷⁵ One of the responders highlighted: *'I cannot accept a protocol without testing it myself.'*⁷⁶ Additionally, they expressed concerns about a potential superior position of developers toward users⁷⁷ and a conflict of interests.⁷⁸

Responders also expressed their willingness to participate in the pre-market assessment and verification of AI applications: 71.43% of users and 80.65% of non-users positively replied to this question. Several participants mentioned that they are ready to participate

⁶⁸ ID 72.

⁶⁹ ID 17; ID 41; ID 59; ID 126.

⁷⁰ ID 44; ID 77; ID 84; ID 124.

⁷¹ ID 90.

⁷² ID 37; ID 111.

⁷³ ID 111.

⁷⁴ ID 72.

⁷⁵ ID 124.

⁷⁶ ID 117.

⁷⁷ ID 64.

⁷⁸ ID 43.

in the assessment if they have the relevant resources (time capacity⁷⁹ or the relevant knowledge⁸⁰). One of the participants specified that the involvement of all the relevant stakeholders is necessary: lawyers, specialists in information technologies, experts in the AI field and developers of the technology.⁸¹

Besides the involvement of users in the pre-market assessment of AI applications, participants were asked about other forms of cooperation between providers and users of AI. In this regard, many responders specified the opportunity to give feedback to providers in order to improve the tools.⁸² Meetings between AI providers and users were also mentioned as a form of cooperation.⁸³ One of the responders also highlighted the necessity to have the constant support of humans (not AI) in the background.⁸⁴

⁷⁹ ID 37.

⁸⁰ ID 85.

⁸¹ ID 85.

⁸² ID 17; ID 36; ID 95; ID 117; ID 126.

⁸³ ID 37; ID 41.

⁸⁴ ID 84.

4. Recommendations

4.1. Raising awareness

The success of technological innovations is essentially reliant on how the primary stakeholders perceive the technology. Raising awareness among professionals in social security, their clients, policymakers, regulators, and society is the absolute must for fostering trust in the AI itself and in the human approaches to develop, validate and govern it. Trust is the precondition to the adoption of the technology by society, which can further spur innovations and lead to the development of AI tools that better serve the needs of social security beneficiaries.

The approach to raising awareness shall be multi-faceted and combine the measures involving training and outreach programs, public participation and feedback mechanisms, information dissemination, carrying out and demonstration of the case studies. EUMASS consistently applies the specified measures, and the findings of the survey demonstrate the importance of facilitating efforts in this direction.

- *Training and outreach programs* should serve to demystify AI. As the finding 1 of the survey demonstrates, misconceptions about the nature, possibilities and risks of AI remain existing. As the field of AI keeps advancing with enormous velocity, existing concerns may escalate if not handled promptly. It is crucial to ensure that the key stakeholders, mainly professionals in the social security sector and their clients, clearly understand what the technology is and how it may be implemented in a secure and reliable way. This knowledge should empower them to adapt and benefit from new developments rather than feel intimidated or left behind. Cross-disciplinary educating and training programs should be developed and delivered in an accessible and compressible way.
- *Public participation and feedback mechanisms* are necessary to ensure that the measures to raise awareness are tailored according to the specific needs of the intended audience. Surveys, including the current study, prove the utility of creating feedback mechanisms where the relevant stakeholders can share their experiences, concerns and suggestions regarding AI in social security. Events and thematic conferences also provide an efficient platform for exchanging knowledge, expertise and expectations. The EUMASS events such as '24th EUMASS Congress Strasbourg Medicine 2.0 in a Changing World' (September 28 to 30, 2023) and the upcoming symposium 'AI's Role in Defining the Future of Social Security and Work Capacity Assessment' (Brussels, March 06, 2023) are the examples of efficient participation and collaboration. Further surveys, events and conferences, as well as interactive methods like workshops, focus groups and public consultation may be

used to address the specific topic depending on the actual needs of the relevant stakeholders.

- *Information dissemination* addressed to the broadest audience is the key component for raising awareness regarding AI in the social security sector. Well-informed professionals tend to make better decisions, while their informed clients are more likely to adhere to algorithmic suggestions provided in certain situations. To engage with the public and disseminate information about AI in social security, various media channels of the broad spectrum should be utilised, including social media, workshops, webinars, podcasts, and interactive platforms.
- *Demonstration of the case studies* should be an integral component of the strategy aimed at raising awareness about AI in social security. As indicated by the survey findings, successful use-cases of applying AI have already appeared within the EUMASS community. These use-cases should be actively promoted to encourage other stakeholders to adopt similar best practices that have already proven effective. Showcasing these positive outcomes and improvements brought about by AI can significantly contribute to building trust and understanding among the public. Stakeholders involved in the successful use-cases should not only highlight the benefits of AI, but also share their strategies to develop and implement the technology. Emphasising the key factors that contributed to their positive outcomes will provide valuable insights for others aiming to replicate such successes.

For example, the participant of the survey with a highly positive experience of using AI application flagged that the technology was developed by mathematicians within the organisation in collaboration with insurance physicians, users support and necessary adjustments are available on demand, and the physicians are highly informed about the AI application, its capacities and risks. Starting from the development phase and through the whole life cycle, any concerns and demands of users are taken into careful consideration, which might explain the high level of satisfaction with the technology.

Neutral and negative use-cases should also be rigorously analysed for the identification of the factors that could hinder the successful implementation of the technology. For example, one of the users explained the neutral experience by the lack of full reliability of AI and the need for extra work for medical advisors to interpret algorithmic recommendations.⁸⁵ Definitely, the identified concerns should be further disseminated to guide developers in enhancing their algorithms and tailoring them to the expectations of users right from the start of creating the technology.

⁸⁵ ID 68.

Use-cases can be disseminated through various forms described previously: training and outreach programs, public participation and feedback mechanisms, and information dissemination channels.

4.2. Rigorous control of safety and quality is strongly required

4.2.1. *Safety and quality control: justification*

To be accepted by society, AI systems have to be safe and reliable. To guarantee these parameters, AI applications and their developers should be strictly controlled, especially in the areas where people's health, well-being, dignity and fundamental rights are affected. In social security and work capacity assessment, the cost of implicating errored or biased algorithmic recommendations is extremely high. It may cause material or physical harm to the individuals affected by the recommendations, their discrimination and stigmatization. On the contrary, when the relevant stakeholders are aware that AI systems have undergone rigorous testing and validation, their acceptance of the technology may exponentially increase. Finding 3 of the survey confirmed that professionals in the social security sector strongly demand AI applications to be internally and externally verified before being used towards individuals.

Work capacity assessment encompasses a broad range of procedures falling within different domains and requiring various types of expertise. 'This might include medical assessments to determine a person's work capacity or entitlement to a particular health treatment, assessments to determine whether someone is entitled to the allocation of social benefits, or needs-based assessments that determine whether they are eligible for practical assistance at home, at work or in education.'⁸⁶ Considering the differences, validation procedures applying to the specific AI solution would highly depend on its intended purpose and functionality.

4.2.2. *Safety and quality control of AI applications in social security and work capacity assessment: applicable procedures*

The current regulatory framework does not provide the validation procedure that is specifically designed for AI applications in social security and work capacity assessment. Depending on the intended purpose and functionality of an AI application, it might be regulated either by the existing rules for controlling medical devices, or by the future framework suggested by the EC Proposal for the AI Act.⁸⁷ The two scenarios are outlined below.

⁸⁶ Mark Priestley, 'Disability assessment, mutual recognition and the EU Disability Card, Progress and Opportunities,' the European Parliament Study, requested by the PETI committee, November 2022, available at: <[https://www.europarl.europa.eu/RegData/etudes/STUD/2022/739397/IPOL_STU\(2022\)739397_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/739397/IPOL_STU(2022)739397_EN.pdf)> accessed December 07, 2023.

⁸⁷ As the political agreement on the AI Act has been reached on December 09, 2023, the final text of the adopted legislation is expected to be published at the beginning of 2024. This report refers to the initial text of the proposal suggested by the European Commission in 2021, but the references may be further updated considering the changes made by the European Parliament and the Council that were implemented to the adopted law.

In scenario 1, when an AI application is used for a medical assessment of an individual, it might be qualified as a medical device.⁸⁸ In the EU, medical devices are regulated by the Medical Devices Regulation (the ‘MDR’).⁸⁹ Compliance with the MDR is mandatory for anyone developing and marketing a medical device. Every action that should be taken by the providers of AI medical devices is to be found in the said regulatory framework.

The MDR establishes a rigorous system of quality control – conformity assessment procedure and post-market surveillance. The existence of such a system implies the obligation of AI providers to demonstrate and maintain the safety and performance of medical devices throughout their lifecycle.⁹⁰ Performance means the ability of a device to achieve its intended purpose as stated by the manufacturer.⁹¹ Safety requirement can be interpreted as freedom from unacceptable risks⁹² and is associated with taking measures to minimise risks and keep the risk-benefit ratio positive.⁹³ Safety and performance are interconnected and are usually considered together as the necessary conditions for a medical device to be used in healthcare settings toward individuals.

Under the MDR, AI providers have numerous obligations: to develop technical documentation⁹⁴ and post-market surveillance plan,⁹⁵ carry out clinical evaluations of a device,⁹⁶ establish, document and implement a quality management system.⁹⁷ Assessment

⁸⁸ According to art. 2(1) of the MDR, ‘a medical device can be defined as any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the specific medical purposes. The medical purposes include, inter alia, diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease, injury or disability.’ AI applications are qualified as software and thus are covered by the definition when used for medical purposes.

⁸⁹ The framework consists of the two legislative acts: the Regulation 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices (the ‘MDR’) and the Regulation 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices (the ‘IVDR’). While most of the AI applications, particularly for the medical assessment in the social security sector, are covered by the MDR, this report refers only to the MDR.

⁹⁰ The MDR, art. 2(22).

⁹¹ The MDR, art. 5(2).

⁹² This definition is accepted in the ISO 14971:2019 “Medical devices Application of risk management to medical devices” approved on 2 May 2019 by the Association for the Advancement of Medical Instrumentation (AAMI) and on 10 May 2019 by the American National Standards Institute (ANSI).

⁹³ The MDR, Annex I, Chapter I, art. 1.

⁹⁴ The documentation shall be presented in a clear, organised, readily searchable and unambiguous manner and contain: general device description; information for users; design and manufacturing information; general safety and performance requirements; benefit-risk analysis and risk management; AI software verification and validation (MDR, Annex II.1).

⁹⁵ Post-market surveillance plan (‘PMSP’) shall address the collection and utilization of information about device’s real-life use, in particular: information concerning serious incidents, records referring to non-serious incidents and data on any undesirable side-effects, information from trend reporting, relevant specialist or technical literature, databases and/or registers, information, including feedbacks and complaints, provided by users, distributors and importers (MDR, Annex III).

⁹⁶ The clinical evidence together with non-clinical data generated from non-clinical testing methods and other relevant documentation shall allow the manufacturer to demonstrate conformity with the general safety and performance requirements and shall be part of the technical documentation for the device in question. Both favourable and unfavourable data considered in the clinical evaluation shall be included in the technical documentation (MDR, Annex XIV, Part A, 1).

⁹⁷ The quality management system shall include, inter alia, ‘the procedures and techniques for monitoring, verifying, validating and controlling the design of the devices and the corresponding documentation as well as the data and records arising from those procedures and techniques.’ (MDR, Annex IX, Chapter I, 2.2(c)).

of the device and submitted documents is carried out by notified bodies - the third-party entities that are independent of manufacturers and that are designated to verify if a device complies with the requirements imposed by the MDR.⁹⁸ After a device is authorised, providers have to continuously gather information about the device's real-life use and promptly react to any occurred incidents, side-effects, or deviations from a device's intended performance. Overall, the existing legal framework provides rather elaborated procedures to control medical devices, even the most innovative and challenging – AI medical devices. In addition, the existing framework will be complemented by the requirements of the AI Act, once the legislative proposal is adopted and comes into force.⁹⁹

*In scenario 2, when an AI application, for example, aids in calculating the social security allowance to be allocated to an individual, it might be covered by the future framework suggested by the EC Proposal for the AI Act. Similar to the law applicable to medical devices, 'the aim of this framework is to ensure that AI systems placed on the EU market and used in the Union are safe and respect existing law on fundamental rights and Union values.'*¹⁰⁰

The legislative proposal establishes that AI systems used for evaluating the eligibility of individuals for public assistance benefits and services¹⁰¹ are considered to be high-risk AI applications. For this type of system, providers should also perform conformity assessment to verify compliance with the AI Act requirements: high quality of data;¹⁰² drawing up of technical documentation;¹⁰³ record-keeping;¹⁰⁴ transparency and provision of information to users,¹⁰⁵ human oversight,¹⁰⁶ robustness, accuracy and cybersecurity.¹⁰⁷ 'Importantly,

⁹⁸ MDR, art. 2 (41) and 2(42).

⁹⁹ EC Proposal for AI Act, art. 43 (3).

¹⁰⁰ Council of the EU, 'Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights,' (Press release, December 06, 2022), available at: < <https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/>> accessed December 07, 2023.

¹⁰¹ In addition to that it covers AI systems to as well as to grant, reduce, revoke, or reclaim social security benefits and services (EC Proposal for the AI Act, art. 6 and Annex III).

¹⁰² Training, validation and testing data sets shall be relevant, representative, free of errors and complete and be subject to appropriate data governance and management practices (EC Proposal for the AI Act, art. 10).

¹⁰³ The technical documentation of a high-risk AI system shall be drawn up before that system is placed on the market or put into service and shall be kept up-to date. The documentation shall be drawn up in such a way to demonstrate that the high-risk AI system complies with the requirements set out in the AI Act (EC Proposal for the AI Act, art. 11).

¹⁰⁴ High-risk AI systems shall be designed and developed with capabilities enabling the automatic recording of events ('logs') while the high-risk AI systems is operating (EC Proposal for the AI Act, art. 12).

¹⁰⁵ 'High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately <...> Such systems shall also be accompanied by instructions for use in an appropriate format that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to users.' (EC Proposal for the AI Act, art. 13).

¹⁰⁶ 'High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use.' (EC Proposal for the AI Act, art. 14).

¹⁰⁷ EC Proposal for the AI Act, art. 15.

most of these requirements must be embedded in the design of the high-risk AI system since the beginning phases of the development process.¹⁰⁸

4.2.3. Safety and quality control of AI applications in social security and work capacity assessment: clarification, unification and enhancement of applicable procedures are required

The described legal frameworks provide a rather similar set of procedures and requirements, but a more elaborated approach to regulate AI applications in the social security and work capacity assessment is desired. Clarification, unification and enhancement of applicable procedures are necessary.

- *Clarification* is necessary for defining how AI applications with combined functionalities should be classified. For example, if an AI application both assesses medical conditions and calculates social security allowance, clear guidance on applicable rules should be developed.
- *Unification* is required to set the bar for controlling the process of all the AI solutions to be applied in social security and work capacity evaluation. For example, AI medical devices (scenario 1) usually require an external conformity assessment, while AI systems to allocate social security allowance (scenario 2) may need only an internal one. Such discrepancies may result in misuse of describing AI functionality to adjust it to the preferable procedure, which in the end may result in the lack of trust from professionals and their clients.
- *Enhancement* of the existing procedures is essential to address the AI risks specifically relevant to the work capacity assessment. For example, in the said area the risks of discrimination and stigmatization may be particularly high. Although the proposed AI Act provides a solid framework to control biases in high-risk AI systems, additional guidance directed at AI developers on the sector-specific manifestations of unfairness and discrimination may be necessary. 'People with impairments are excluded from full participation and equality and this constitutes a form of discrimination that is institutionalised throughout society.'¹⁰⁹ When developing AI solutions, providers should consult with social security professionals on how to address the interests of all the groups of individuals who may be affected by algorithmic recommendations.

Overall, the legislative framework provides foundational procedures for controlling the safety and quality of AI applications in social security and work capacity assessment.

¹⁰⁸ Katerina Demetzou, 'Introduction to the Conformity Assessment Under the Draft EU AI Act, And How it Compares To DPIAs' (Future of Privacy Forum, August 12, 2022), available at: < <https://fpf.org/blog/introduction-to-the-conformity-assessment-under-the-draft-eu-ai-act-and-how-it-compares-to-dpias/> > accessed December 07, 2023.

¹⁰⁹ Priestley (n 86), 26.

However, the sector-specific implementation guidelines are necessary. The guidelines should be developed in collaboration with experts from various fields, including social security specialists, ethicists, legal professionals, tech developers and affected individuals. This collaborative approach ensures that the guidelines encompass a comprehensive perspective, integrating technical excellence with ethical considerations and real-world experiences. To be enforceable, the guidelines should be verified and adopted by the relevant authorities, such as a new European AI Office within the European Commission or national competent market surveillance authorities. The guidelines would provide necessary clarification and unification of applicable procedures and tailor them to the specific challenges presented in the sector.

4.3. Transparency is the key factor for users' acceptance

Since the very first promising results of AI, the “black-box” problem¹¹⁰ - opacity of how AI comes to decisions¹¹¹ - became a cornerstone of its successful application. For this type of models, even their creators have the limited capacity to trace how the algorithmic input turned into the specific output and what factors contributed to it.¹¹² This feature of the technology makes the detection of errors and latent biases in algorithmic recommendations rather challenging. The lack of transparency is the fundamental issue of the safe and trustworthy AI implication, but in the sensitive domains such as healthcare, social security and work capacity assessment it is paramount due to the direct impact on people's lives and health.

‘Transparency is one of the core values promoted by the EU for the development, deployment, and use of AI systems. Since the start of the policy process to regulate AI, all the relevant documents¹¹³ included transparency in the ethical or legal frameworks they respectively suggested.’¹¹⁴ The EC Proposal for the AI Act also includes transparency requirement: ‘an AI system is deemed sufficiently transparent if it enables its users to interpret the AI's system output and apply it appropriately.’¹¹⁵ The Program of the Belgian Presidency of the Council of the European Union outlines that a *human-centered approach* to digital transformation requires particular attention to *algorithmic transparency*.¹¹⁶

¹¹⁰ The detailed description of the ‘black-box’ problem is provided by one of the most prominent scholars in data science focusing on algorithmic interpretability, C. Rudin: “In machine learning, these black box models are created directly from data by an algorithm, meaning that humans, even those who design them, cannot understand how variables are being combined to make predictions. Even if one has a list of the input variables, black box predictive models can be such complicated functions of the variables that no human can understand how the variables are jointly related to each other to reach a final prediction’ (Rudin, C., & Radin, J. (2019). Why Are We Using Black Box Models in AI When We Don’t Need To? A Lesson From an Explainable AI Competition. *Harvard Data Science Review*, 1(2). <https://doi.org/10.1162/99608f92.5a8a3a3d>).

¹¹¹ Linardatos, P., Papastefanopoulos, V., and Kotsiantis, S. (2021). Explainable AI: A review of machine learning interpretability methods. *Entropy* 23:18. doi: 10.3390/e23010018

¹¹² Kiseleva A, Kotzinos D and De Hert P (2022) Transparency of AI in Healthcare as a Multilayered System of Accountabilities: Between Legal Requirements and Technical Limitations. *Front. Artif. Intell.* 5:879603. doi: 10.3389/frai.2022.879603.

¹¹³ See: AI HLEG (2019). Ethics Guidelines for Trustworthy AI. Brussels, European Commission, available at: < <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>>, accessed December 04, 2023 (AI-HLEG Ethics Guidelines for Trustworthy AI); European Commission (2020). White Paper ‘On Artificial Intelligence - A European Approach to Excellence and Trust,’ available at: < https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en>, accessed December 04, 2023 (EC White Paper on AI), and European Parliament (2020). Report with Recommendations to the Commission on a Framework of Ethical Aspects of Artificial Intelligence, Robotics and Related Technologies (2020/2012(INL), October 08, 2020) (EP Report on AI Framework).

¹¹⁴ A. Kiseleva, ‘MAKING AI’S TRANSPARENCY TRANSPARENT: notes on the EU Proposal for the AI Act,’ (European Law Blog, July 29, 2021). Available at: <https://europeanlawblog.eu/2021/07/29/making-ais-transparency-transparent-notes-on-the-eu-proposal-for-the-ai-act/>, accessed December 03, 2023.

¹¹⁵ EC Proposal for the AI Act, art 13(1).

¹¹⁶ Official Program of the Belgian Presidency of the Council of the European Union (from the 1st of January to the 30th of June 2024), ‘Protect, Strengthen, Prepare’, available at: < https://belgian-presidency.consilium.europa.eu/media/3kajw1io/programme_en.pdf> accessed December 15, 2023.

Similarly, the existing legal framework, the Medical Devices Regulation, highlights the importance of transparency:

The Medical Devices Regulation, recital 43:

'transparency and adequate access to information, appropriately presented for the intended user, are essential in the public interest, to protect public health, to empower patients and healthcare professionals and to enable them to make informed decisions, to provide a sound basis for regulatory decision-making and to build confidence in the regulatory system.'

Importantly, the law recognises users of AI applications as the key addressees of transparency – they have to be provided with all the necessary information in the appropriate form so that can make informed decisions. Insurance physicians, medical advisors, other professionals who assess medical conditions make the final decisions regarding the health state and work capacity of individuals, which might affect the scope of social security benefits available to them. If the specified professionals apply AI tools in their decision-making process, they evaluate whether an algorithmic recommendation is reliable and whether it should be implicated in the specific case of their client. To make such a complex decision, users of AI applications have to be properly guided and informed by providers of the technology.

Finding 4 of the survey confirms that both users (insurance physicians, medical advisors, other relevant professionals) and deployers (organisations where users are employed) should become more empowered in the informational circle relevant to AI applications in the social security sector.

The scope of the information to be provided by AI developers to users can be grouped based on the sub-decisions that the users of the technology should take:

- **Sub-decision 1**: deciding if one of the deployed in the organisation AI applications should be applied in a certain case of a client:
 - *'Motivational' informational nodes*: information about clinical benefits, performance (accuracy, stability, and precision), and safety (risks and measures to mitigate them).
 - *'Justificatory' informational nodes*: intended purpose, indications, contra-indications, patients' target groups, users' groups, limitations, warnings, precautions, measures to be taken in the event of the changes in algorithmic performance that may affect safety.

- Sub-decision 2: Examination of the instructions to use an AI application to see if a professional in the social security sector understands them well and if she is equipped with everything to follow them:
 - *Usage materials*: instructions for the appropriate use, information allowing assessment of the application's suitability, specifications a user requires to use the AI application appropriately, requirements for special facilities, special training, or particular qualifications of a user.
 - *Elimination or reduction of use errors*: AI providers should take the relevant measures and consider the technical knowledge, experience, education, training, use environment, and the medical and physical conditions of the intended user.
 - *Usability engineering and testing*: to identify and minimise use errors and reduce use-associated risks, developers of AI applications have to conduct testing of the instructions for use with representative users to mitigate three sources of use errors: the difficulty of perception, the difficulty of cognition, and the difficulty of manipulation.
 - *Data governance guidelines for users*: to avoid data shifts, users have to be instructed on what input data are considered to be relevant and appropriate for the AI system in question.

- Sub-decision 3: Receival of the recommendation generated by AI, deciding if the recommendation should be followed, and assessing the risks associated with it.
 - *Actionable explanations of algorithmic outcomes should be provided*: professionals in the social security sector – users of AI applications – should be able to assess the correctness of algorithmic suggestions and carry out their own risk assessments.

Information to be provided to users should be contextualised for the intended purpose, area of use, and functionality of the AI application and tailored to the technical knowledge, experience, education, or training of users. Instructions should be written in terms readily understandable and when needed, be supplemented with drawings and diagrams – they should be sufficient and understandable so that users do not seek additional consultations or materials to apply instructions. If users have to spend a considerable amount of time investigating instructions and understanding algorithmic recommendations, any time savings offered by AI applications might become void.

Only when insurance physicians and other relevant professionals are sufficiently informed regarding AI application, they can guide their clients to enable them making informed choices. As declared by the UNESCO Universal Declaration on Bioethics and Human Rights of 2005: ‘The autonomy of persons to make decisions, while taking responsibility for those decisions and respecting the autonomy of others, is to be respected.’¹¹⁷ To take and hold this responsibility, patients shall be armoured with a very strong informational shield that shapes their understanding of what they sign up for (or not). Such shield is created for patients by their physicians. Guided by the requirement of the medical informed consent,¹¹⁸ healthcare professionals define what is necessary to disclose in the specific case of their client, observe the values, knowledge and character of a patient, step-by-step build trust with him or her and decide on how to better exchange the information with that specific patient.

Overall, ‘transparency shall be viewed as a continuous process accompanying the whole life cycle of an AI system, from development to implementation of the technology.’¹¹⁹ It shall become a “way of thinking” rather than just a single activity to tick the box of the legal obligation.’¹²⁰

¹¹⁷ UNESCO Universal Declaration on Bioethics and Human Rights, October 19, 2005, Paris, France <<https://en.unesco.org/about-us/legal-affairs/universal-declaration-bioethics-and-human-rights>>, art. 5.

¹¹⁸ Convention for the Protection of Human Rights (1997). Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine. Oviedo, 4.IV.1997. Council of Europe. European Treaty Series - No. 164 (Oviedo Convention), art. 5; Charter of Fundamental Rights of the European Union (2012). OJ C 326, p. 391–407 (EU Charter of Fundamental Rights or CFR). The general requirement is established by the international treaties – the Oviedo Convention on Human Rights and Biomedicine as of April 4, 1997 (art. 5) and the EU Charter of Fundamental Rights (art. 3.2(a)) – and detailed by national legislations.

¹¹⁹ Kiseleva, Kotzinos, De Hert (n 112).

¹²⁰ Kiseleva, Kotzinos, De Hert (n 112).

4.4. Reliable implementation of AI shall be based on cooperative efforts

Cooperation among different stakeholders is vital for implicating AI solutions in social security that are technically robust, ethically sound, legally compliant, and, most importantly, beneficial to the individuals and communities they aim to support. Partnership should be ongoing, spanning across sectors and regions.

Collaboration should be facilitated at all the stages of the AI lifecycle: from the development to validation, reliable implementation and ongoing monitoring of the technology. All the stakeholders should participate: tech developers, private and public entities in social security, medical advisors, insurance physicians and other professionals in the sector, citizens, regulatory agencies and political institutions. Finally, efforts should be cooperated and united among stakeholders from different Member States, thus strengthening the EU single market and harmonising the protection of individuals at the Union level.

- *During the development and validation phase*, it is necessary to engage the representatives of users for the assessment of AI solutions. Such engagement would facilitate user acceptance and transparency, as well as enable AI providers to benefit from the users' expertise in the domain of applying the AI solution. As finding 5 confirmed, the EUMASS community participating in the survey expressed their strong demand and willingness to participate in pre-market assessment and verification of AI applications.

To facilitate user acceptance, AI developers should carry out usability engineering – the process intended to identify and minimise use errors and reduce use-associated risks.¹²¹ Prevention of user errors is critical in the AI context because the quality of outcomes highly depends on the quality of the input. In addition, user interfaces as the means to lead Human-AI conversations should be tested during the development and validation stages.

To facilitate transparency, AI providers should ensure that the information, instructions and explanations accompanying AI is fully and easily comprehensible by potential users. The described materials should be evaluated by users: representatives of professionals should participate in the device's testing and verification procedures to assess material accompanying AI applications. This process is especially important in the AI context, where the

¹²¹ The International Electrotechnical Commission (IEC), IEC 62366-1:2015(en), Medical devices — Part 1: Application of usability engineering to medical devices, art. 3.17, available at: <https://www.iso.org/obp/ui/#iso:std:iec:62366:-1:ed-1:v1:en>
>

complexity of the technology might require much more attention to the transparency needs of users.

Engagement of users in the development and validation may substantially increase the quality of AI applications. Insurance physicians, medical advisors and other relevant professionals may share with AI developers their unique sector-specific expertise and insights, translate the concerns and expectations of their clients, and outline the factors crucial for adopting the developed solution (for example, reliability metrics, time constraints).

- *During the implementation and monitoring phase*, cooperation and communication between users, deployers and developers of AI applications should be facilitated in two directions: support from developers and feedback from users. This bidirectional communication is crucial for refining the application, addressing issues, and continuously improving its functionality based on real-world usage and experiences.

Developers should offer support to users and deployers, ensuring that they comprehend the application's functionalities and aiding in resolving any issues that arise during implementation. Such support is especially critical at the initial stages of technological implementation, as it might impact the motivation of users to accept an AI solution. As the survey confirms, users demand to have a mechanism for contacting AI providers when necessary.

AI providers should also build a structured system for gathering feedback from users. 'Learning from new data, AI algorithms constantly adapt and change.'¹²² These changes can enhance the algorithm's performance by leveraging a larger pool of available training data. However, if not properly governed, real-life algorithmic self-learning could cause data shifts and subsequently reduce the reliability of a medical device incorporating such algorithms. Users and deployers, as the main subjects monitoring how AI functions in real-life settings, should be provided with all the mechanisms to report AI providers about any incidents or anomalies encountered during the utilisation of AI in real-life settings. This feedback loop is essential for ensuring that providers are informed about real-time issues, enabling them to take corrective measures swiftly and enhance the overall performance and safety of the AI systems deployed.

In addition to the cooperation during AI development and implementation, general efforts in supporting innovations in the sector and raising awareness should be commonly taken by all stakeholders, including regulators and political decision-makers.

¹²² Kiseleva, A. (2020a). AI as a medical device: is it enough to ensure performance transparency and accountability? *European Pharmaceutical Law Review* 4, 5–16. doi: 10.21552/eplr/2020/1/4.

Collective responsibility involves fostering an environment conducive to innovation through supportive policies, funding initiatives, and regulatory frameworks. Joint advocacy for research and development grants or incentives can stimulate innovation in AI technologies specifically tailored to the demands of the social security sector. Encouraging collaboration between academia, industry, and governmental bodies can accelerate the development of cutting-edge solutions addressing the sector's unique requirements. By uniting efforts, stakeholders can create a conducive environment for responsible and effective AI integration in social security and work capacity assessments, ensuring alignment with societal needs and ethical considerations.

Appendix 1. Introduction to the survey

'Dear Participant,

Thank you for your interest in participating in this survey.

Background and Purposes

In recent years, the sphere of Artificial Intelligence (AI) has catapulted to the forefront of innovation, promising to revolutionize certain areas and hugely impact others. The primary aim of this survey is to identify the current positioning of AI in the area of social security. We gather insights and opinions from stakeholders with diverse perspectives and experiences related to the integration of AI technologies in social security.

Accelerated interest in the implementation of AI creates the need to develop policies that can simultaneously ensure the safety and quality of the technology, guarantee the fundamental rights of individuals and at the same time unleash the opportunities and benefits that the technology can bring. The findings from this survey will contribute to the development of policy recommendations for the trustworthy integration of AI in social security. Your participation will help us to tailor these recommendations to the expectations and experiences represented in the sector.

The survey is designed by the European Union of Medicine in Assurance and Social Security (EUMASS) in collaboration with the Belgian Ministry (Federal Public Service FPS) of Social Affairs.

Duration and Instructions

Completing this survey will take approximately 20-25 minutes of your valuable time. In simple choice questions (YES/NO), we do not expect you to add comments, we only welcome providing them if you have any.

Privacy and Confidentiality

Your identity details are not required to fulfill the survey. It is optional for you to specify your name and your organisation. Other questions do not require providing the data that can directly or indirectly identify you. In any case, your responses are confidential and will be used for the statistical and research purposes defined above. This means that you cannot be identified in any way in reports or publications. We only report/publish aggregated and anonymised results.

If provided, your personal data will be processed strictly in accordance with the European General Data Protection Regulation 2016/679 of April 27, 2016, in force since May 25, 2018 (GDPR) based on your consent and the privacy policy (see below).

We kindly request you not to provide any personally identifiable patient information when completing the questionnaire.

Contact information

If you have any questions, concerns, or require further clarification about this survey, please do not hesitate to contact us (Anastasiya Kiseleva, anastasiya.kiseleva@vub.be). Your inquiries will be addressed promptly and thoroughly.

We extend our sincere gratitude for your valuable contribution.

The EUMASS website: <https://eumass.eu>

This survey is anonymous.

The record of your survey responses does not contain any identifying information about you unless a specific survey question explicitly asked for it.

If you used an identifying access code to access this survey, please rest assured that this code will not be stored together with your responses. It is managed in a separate database and will only be updated to indicate whether you did (or did not) complete this survey. There is no way of matching identification access codes with survey responses.'

Appendix 2. List of questions posed to the potential users of AI

	Question	Response
1. Information about the participant		
1.1.	Full name (optional)	
1.2.	Your organisation (optional)	
1.3.	Country of your organisation	
1.4.	Your role	
1.5.	Your specialty	
1.6.	Your experience	1-5 years/6-15 years/16-30 years/more than 30 years
2. Assessment of the intent to start using AI		
2.1.	How would you rate your overall intent to start using AI tools in your practice?	0 to 10 Where 0 is: I do not have any intentions to start using AI, 5 – I am neutral, will be fine with using AI, but do not have an intense demand for it; and 10 – I am sure AI application will greatly benefit my work and I want to start using it as soon as possible
2.2.	If you have a strong intent to start using AI in your practice (you answered from 6 to 10), could you please describe the reasons for that and specify the benefits you expect AI to bring?	Description/not applicable
2.3.	If you do not have strong intentions to start using AI in your practice (you answered from 0 to 4), could you please describe the reasons for that and specify your concerns?	Description/not applicable
2.4.	If you are neutral about the use of AI in your practice (you rated your intent as 5), could you please elaborate on your answer?	Description/not applicable
2. Assessment of the challenges regarding applying AI		
2.1.	Do you see any challenges to the adoption and implementation of AI in your field of work?	YES/NO
2.2.	If you see the challenges to the adoption and implementation of AI in your field of work, could you please describe them? <i>Please provide your initial thoughts, more detailed questions follow.</i>	Description

2.2.1.	In your opinion, what are the main <u>technical</u> challenges to the adoption of AI in your field of work?	Description
2.2.2.	In your opinion, what are the main <u>organisational</u> challenges to the adoption of AI in your field of work?	Description
2.2.3.	In your opinion, what are the main <u>ethical/legal</u> challenges to the adoption of AI in your field of work?	Description
3. Assessment of the information demanded for the confident use of AI		
3.1.	To start using an AI application, do you think you need to be additionally educated and trained for that?	YES/NO (comment)
3.2.	If you use an AI application, what do you think you need to know about the application itself to use it confidently? <i>Please share your initial thoughts, more detailed questions follow.</i>	Description
3.2.1.	Do you think you need to be informed on the nature, functions and intended purpose of an AI application?	YES/NO (comment)
3.2.2.	Do you think you need to be informed about the performance characteristics, accuracy level and stability of an AI application?	YES/NO (comment)
3.2.3.	Do you think you need to be informed on its target groups and contra-indications of an AI application?	YES/NO (comment)
3.3.	If you use an AI application, what do you think you need to be instructed on to use it confidently? <i>Please share your initial thoughts, more detailed questions follow.</i>	Description
3.3.1.	Do you think you need to be instructed on how to handle/filter the input data to be added to the AI application?	YES/NO (comment)
3.4.	If you use an AI application, what do you think you need to know to rely on the advice suggested by the application? <i>Please share your initial thoughts, more detailed questions follow.</i>	Description
3.4.1.	Do you think you need instructions on how to assess the correctness of the outcomes/suggestions provided by the AI application (through embedded explanations, for example)?	YES/NO (comment)
3.4.2.	Do you think you need to be informed on how the quality and safety of the AI application were tested and verified?	YES/NO ((comment)
3.4.3.	Do you think you need to be informed on how the training, testing and validation data was selected and handled to ensure data quality and sufficient representativeness?	YES/NO (comment)

3.4.4.	Do you think it should have embedded explanations for the provided suggestions of an AI application?	YES/NO (comment)
3.5.	If you use an AI application, what do you think you need to know about the benefits of an AI application? <i>Please share your initial thoughts, more detailed questions follow.</i>	Description
3.5.1.	Do you think you need to be informed on how an AI application performs in comparison to non-AI systems?	YES/NO (comment)
3.5.2.	Do you think you need to be informed on how an AI application performs in comparison to human specialists?	YES/NO (comment)
3.5.3.	Do you think you need to be informed on the claimed accuracy level of an AI application?	YES/NO (comment)
3.6.	If you use an AI application, what do you think you need to know about the risks of an AI application? <i>Please share your initial thoughts, more detailed questions follow.</i>	Description
3.6.1.	Do you think you need to be informed on the risk of algorithmic opacity?	YES/NO (comment)
3.6.2.	Do you think you need to be informed on the risk of shifts in the accuracy level/data shifts?	YES/NO (comment)
3.6.3.	Do you think you need to be informed on the limited ability to verify the correctness of the specific AI outcome?	YES/NO (comment)
3.6.4.	Do you think you need to be informed on the risk of latent biases?	YES/NO (comment)
3.6.5.	Do you think you need to be informed on how the AI provider handles the limitations, minimises and manages the risks?	YES/NO (comment)
3.7.	If you use an AI application, what do you think is necessary to inform patients/individuals about?	Description
4. Assessment of AI awareness		
4.1.	Have you heard about AI applications being used in your field of work?	YES/NO
4.2.	If you heard about the specific AI application being used in your field of work, could you please name and describe it? If possible, could you please describe the experience of the application's users (overall impression, benefits and challenges)?	Description/not applicable
5. Assessment of the expected support from AI providers		
5.1.	If you use an AI application, do you think you need to have the mechanism to contact AI providers/seek their support if you have any questions or doubts about the use of AI and the implication of AI outcomes?	YES/NO

5.2.	If you need to have a mechanism to contact AI providers, could you please describe how regularly, if it should be a direct contact or organised centrally from your organisation, could you please share your general thoughts on it?	Description/not applicable
6. Assessment of the need and willingness to cooperate with AI providers		
6.1.	Do you think it is necessary to involve the representatives of users (healthcare and social security professionals) in the assessment and verification of AI applications before they are placed on the market?	YES/NO (comment)
6.2.	If you think that users' involvement is necessary, will you be willing and able to participate in the assessment and verification of AI applications before they are placed on the market?	YES/NO (comment)
6.3.	Besides users' involvement, how else do you see the forms of cooperation between AI providers and users regarding AI applications?	YES/NO (comment)
7. Regulatory and Ethical Considerations		
7.1.	Do you believe there should be specific regulations or ethical guidelines governing the use of AI in social security?	YES/NO
7.1.1.	Could you please elaborate on your answer regarding the need to have regulations or ethical guidelines governing the use of AI in social security systems?	Description
7.2.	Do you believe AI applications in social security should be internally and externally verified before being used toward individuals?	YES/NO
7.2.1.	Could you please elaborate on your answer regarding the need for internal and external verification of AI applications in social security?	Description
7.2.2.	If you think that AI applications should be internally and externally validated, what are the elements to be controlled and verified?	Description
7.3.	In your view, what ethical principles should AI systems in social security adhere to?	Description
8. Feedback and Suggestions		
8.1.	Do you have any suggestions or recommendations for facilitating the integration of AI in healthcare within the social security system?	Description/not applicable
9. Final Thoughts		
9.1.	Please share any additional comments or insights regarding your expectations and concerns for the integration of AI in healthcare within the social security system.	Description

Appendix 3. List of questions posed to the current users of AI

	Question	Response
1. Information about the participant		
1.1.	Full name (optional)	
1.2.	Your organisation (optional)	
1.3.	Country of your organisation	
1.4.	Your role	
1.5.	Your specialty	
1.6.	Your experience	1-5 years/6-15 years/16-30 years/more than 30 years
2. Information about the AI application		
2.1.	Name the AI application* <i>*If you have experience with more than one AI application, could you please fill the separate survey for every application that you use?</i>	
2.2.	Describe the AI application	<i>For example (if known): intended purpose, functionality, target population, type of algorithms, accuracy level, website of the AI application and/or the provider</i>
2.3.	Do you have the information about the mechanisms for certification that are applied for the AI application you use?	YES/NO
2.3.1.	If you have the information about the mechanisms for certification that are applied for the AI application you use, could you please describe them?	Description/not applicable
2.3.2.	Is the AI application certified as a medical device?	YES/NO/NOT AWARE
3. Assessment of the experience with the AI application		
3.1.	How would you rate your overall experience with the use of the AI application in your practice?	1 to 10 where 1 is: it made it much worse and it would be better without AI, 5 – it didn't change anything and 10 is the AI application greatly benefited your work and now you cannot imagine your work life without it

3.2.	If the AI application benefitted your work (you answered from 6 to 10), could you please describe the benefits?	Time savings, increased accuracy, and other
3.3.	If the AI application challenged your work (you answered from 1 to 4), could you please describe the challenges?	Too complex, too much time to be trained to use it, the lack of transparency, data shifts, other
3.4.	If the use of the AI application did not change anything in your practice (you rated your experience as 5), could you please elaborate on your experience and describe challenges and benefits (if any)?	Description/not applicable
4. Assessment of general instructing and informing users regarding the AI application		
4.1.	Have you been informed on the nature, functions and intended purpose of the AI application?	YES/NO (comment)
4.2.	Have you been informed about the performance characteristics, accuracy level and stability of the AI application?	YES/NO (comment)
4.3.	Have you had a general training and educational courses on how to use the AI application?	YES/NO (comment)
4.4.	Have you been provided with the instructions on how to handle/filter the input data to be added to the AI application?	YES/NO (comment)
4.5.	Have you been provided with instructions on how to assess the correctness of the outcomes/suggestions provided by the AI application (through embedded explanations, for example)?	YES/NO (comment)
4.6.	Have you been informed on how the quality and safety of the AI application was tested and verified?	YES/NO (comment)
4.7.	Have you been informed on how the training, testing and validation data was selected and handled to ensure data quality and sufficient representativeness?	YES/NO (comment)
4.8.	Have you been informed on the target groups and contra-indications for the use of the AI application?	YES/NO (comment)
5. Assessment of informing users on the benefits of the AI application		
5.1.	Have you been informed of the benefits of the AI application?	YES/NO (comment)
5.1.1.	If you have been informed of the benefits of the AI application, did they include a comparison with non-AI systems?	YES/NO (comment)
5.1.2.	If you have been informed of the benefits of the AI application, did they include a comparison with human performance?	YES/NO (comment)

5.1.3.	If you have been informed of the benefits of the AI application, did they include the claimed accuracy level?	YES/NO (comment)
6. Assessment of informing users on the limitations of the AI application		
6.1.	Have you been informed on the limitations and risks of the AI application?	YES/NO (comment)
6.1.1.	If you have been informed of the limitations and risks of the AI application, did they include algorithmic opacity?	YES/NO (comment)
6.1.2.	If you have been informed of the limitations of the AI application, did they include shifts in the accuracy level/data shifts?	YES/NO (comment)
6.1.3.	If you have been informed of the limitations and risks of the AI application, did they include issues with the ability to verify the correctness of the specific AI outcome (due to opacity)?	YES/NO (comment)
6.1.4.	If you have been informed on the limitations and risks of the AI application, did they include issues with latent biases?	YES/NO (comment)
6.1.5.	If you have been informed of the limitations of the AI application, have you been also informed on how the AI provider handles the limitations, minimises and manages the risks?	YES/NO (comment)
7. Assessment of the ability to explain, verify, or contest the suggestions provided by the AI application		
7.1.	Have you had cases where the suggestion of the AI application differed from your preliminary analysis?	YES/NO
7.1.1.	If you had cases where the suggestion of the AI application differed from your preliminary analysis, did you have tools to verify the (non-)correctness of the AI advice (such as explanations or instructions from AI providers)?	YES/NO (comment)
7.2.	If the AI application that you use has embedded explanations, could you please describe them? For example, what was explained (how algorithms work, the received outcomes (to the possible extent), how the AI application was trained, etc.) What type of explanations: textual, visual, or combination? If you have more technical details on explanations, do not hesitate to add them (for example, in-model or post-hoc explanations, contrafactual or feature relevance explanations).	Description/not applicable
8. Assessment of AI performance		
8.1.	Did you experience the evident errors in the suggestions of the AI application?	YES/NO

8.2.	Did you experience errors in the suggestions of the AI application that could be detected only after additional verification?	YES/NO
8.3.	Did you experience the errors in the suggestions of the AI application that were detected only after the errored suggestion was implicated?	YES/NO
8.4.	If you experienced any type of errors specified above, could you please describe them and how often they occurred?	Description/not applicable
8.5.	Could you please describe how the errors were handled?	Description/not applicable
9. Assessment of informing patients		
9.1.	Do you inform patients ¹²³ about the use of the AI application in their certain cases?	YES/NO (if you answered NO, please go to the next section)
9.1.1.	If you inform patients about the use of the AI application in their certain case, do you do it in written or orally?	In written/orally
9.1.2.	If you inform patients about the use of the AI application in their case, what do you usually describe to them?	Description/not applicable
9.1.3.	Did you have challenges with informing patients about the AI application?	YES/NO
9.1.4.	If you had challenges with informing patients about the AI application, could you please describe them?	Description/not applicable
9.1.5.	Did you face hesitations of patients regarding the use of the AI application?	YES/NO
9.1.6.	If you faced hesitations from patients regarding the use of the AI application, could you describe them?	Description/not applicable
9.1.7.	If you faced hesitations of patients regarding the use of the AI application, could you please describe if you could address the hesitations and what did you do for that?	Description/not applicable
9.1.8.	When you inform patients, do you use explanations provided with the AI application (if any)?	YES/NO (comment)
10. Assessment of the support from AI providers		
10.1.	Do you have the mechanism to contact AI providers/seek for their support if you have any questions or doubts about the use of AI and about AI outcomes?	YES/NO
10.1.1.	If you can contact AI providers/seek for their support, could you please describe how regularly, if you can contact directly or it is organised centrally inside your organisation? Is the communication usually helpful, easily carried out and addresses your concern?	Description/not applicable

¹²³ We can use the term 'individuals' or other similar if not all the clients are covered by the term 'patient'.

11. Assessment of the need and willingness to cooperate with AI providers		
11.1.	Do you think it is necessary to involve the representatives of users (healthcare and social security professionals) in the assessment and verification of AI applications before they are placed on the market?	YES/NO (comment)
11.2.	If you think that users' involvement is necessary, will you be willing and able to participate in the assessment and verification of AI applications before they are placed on the market?	YES/NO (comment)
11.3.	Besides users' involvement, how else do you see the forms of cooperation between AI providers and users regarding AI applications?	Description
12. Feedback and Suggestions		
12.1.	Do you have any suggestions or recommendations for improving the integration of AI in healthcare within the social security system?	Description/not applicable
13. Final Thoughts		
13.1.	Please share any additional comments or insights regarding your experiences with the AI application you use and in general on AI systems in social security.	Description

Appendix 4. Table on AI transparency expectations and experiences

Transparency experience of users of AI applications designed for the specific purposes concerning their professional activities in social security

Question	Percentage of positive replies			
	ID68 ¹²⁴	ID72 ¹²⁵	ID89 ¹²⁶	
General information about an AI application				
Nature, functions and intended purpose	+	+	+	100%
Performance characteristics, accuracy level and stability	+	+	+	100%
Target groups and contra-indications	+	+	+	100%
Instructions				
How to handle/filter the input data to be added to the AI application	-	-	+	33,33%
How to assess the correctness of the outcomes/suggestions provided by the AI application	+	-	+	66,66%
How the quality and safety of an AI application were tested and verified	+	+	+	100%
How the training, testing and validation data was selected and handled to ensure data quality and sufficient representativeness	+	+	+	100%
Embedded explanations of the algorithmic recommendations	+	-	+	66,66%
Benefits of an AI application				
How an AI application performs in comparison to non-AI systems	-	-	+	33,33%
How an AI application performs in comparison to human specialists	+	+	+	100%
Claimed accuracy level of an AI application	+	+	+	100%
Risks of an AI application				
Risk of algorithmic opacity	+	-	-	33,33%
Risk of shifts in the accuracy level/data shifts	+	-	-	33,33%
Limited ability to verify the correctness of the specific AI outcome	+	-	-	33,33%
Risk of latent biases	-	+	-	33,33%
Information about how an AI provider handles the limitations, minimises and manages the risks	-	+	-	33,33%

¹²⁴ The experience with the AI application is rated 5 – neutral.

¹²⁵ The experience with the AI application is rated 8 – very positive. The respondent shares that the practically reduced workload is particularly evident.

¹²⁶ The experience with the AI application is rated 5 – neutral.

