



Ethical dimensions of AI-supported fact-checking in elderly health care: a qualitative exploratory study

Haline Maia^{1,2,4} · Sérgio Nunes^{3,4} · Patrícia Cordeiro^{5,3} · Carolina Vila-Chã⁶ · Helena Lima^{3,7}

Received: 5 January 2026 / Accepted: 23 March 2026
© The Author(s) 2026

Abstract

This article examines the ethical dimensions of AI-supported fact-checking tools in the context of elderly health care, focusing on how caregivers interpret trust, credibility, and cultural relevance when engaging with AI-mediated information. The study adopts a qualitative, exploratory methodology grounded in Value Sensitive Design and Human–Machine Communication. Empirical data were collected through a co-design workshop, focus group discussions, empathy map exercises, and descriptive questionnaires with elderly caregivers in a rural Portuguese context. Rather than aiming for statistical generalization, the research prioritizes contextual understanding and value articulation. The findings reveal that caregivers evaluate AI-supported fact-checking solutions not only in terms of informational accuracy, but also through ethical considerations such as protection, inclusion, trustworthiness, and perceived legitimacy. Differences between traditional journalistic approaches, automated AI solutions, and visual or animated formats were less salient than participants' interpretations of how these tools aligned with their lived experiences and communicative norms. By foregrounding the perspectives of caregivers in a socioeconomically constrained setting, this study contributes empirically grounded insights to debates in AI ethics and responsible AI design. It highlights the importance of culturally sensitive, user-centered approaches when deploying AI technologies in health communication and misinformation management, particularly for vulnerable populations.

Keywords Artificial intelligence · AI ethics · Fact-checking · Health communication · Value sensitive design · Misinformation

1 Introduction

In the rapidly evolving landscape of journalism, generative artificial intelligence is enabling new types of content creation, verification, and dissemination. The expansion of health-related misinformation during the COVID-19

infodemic intensified the need for structured fact-checking initiatives capable of countering misleading claims [1, 2]. Within this broader transformation, our case study adopts an analytical lens to examine a fact-checking chatbot tailored for caregivers of older people, a group especially susceptible to misinformation and structural vulnerabilities in

✉ Haline Maia
haline.maia@uab.pt

Sérgio Nunes
sergio.nunes@fe.up.pt

Patrícia Cordeiro
pcordeiro@morecolab.pt

Carolina Vila-Chã
cvilacha@sal.ipg.pt

Helena Lima
hllima@letras.up.pt

¹ Universidade Aberta, Lisbon, Portugal

² CIRCLE - Centre for Innovation Research - Lund University, Lund, Sweden

³ Faculty of Engineering, University of Porto, Porto, Portugal

⁴ Institute for Systems and Computer Engineering of Porto, Porto, Portugal

⁵ MORE Colab, Bragança, Portugal

⁶ Research Center in Sports Sciences, Health Sciences and Human Development, Vila Real, Portugal

⁷ Faculty of Arts and Humanities, University of Porto, Porto, Portugal

access to reliable health information. This research was conducted as part of a dissemination activity of the EducAGE project, which took place in Macedo de Cavaleiros, Portugal, and sought to counteract misinformation regarding healthy ageing while addressing the educational challenges faced by caregivers, often women, in socioeconomically disadvantaged regions.

Against this background, this study asks which values, ethical expectations, and interpretive tensions emerge when elderly caregivers make sense of AI-mediated fact-checking in health communication contexts. By shifting the analytical focus from technical performance to value articulation and meaning-making, the research situates AI-supported fact-checking within broader debates in AI ethics and responsible design. Drawing on Value Sensitive Design [3, 4] and Human–Machine Communication [5], the study approaches AI tools as sociotechnical artifacts whose legitimacy depends on alignment with users' lived experiences, communicative norms, and moral expectations.

Empirically, the study is grounded in a co-design workshop conducted with elderly caregivers in Macedo de Cavaleiros within the framework of the EducAGE project. This European initiative focuses on healthy ageing and digital empowerment in socioeconomically constrained regions [6]. Informal caregivers represent a substantial proportion of long-term care providers in Europe and frequently operate without structured digital training, despite increasing reliance on online health information [7]. In this setting, AI-supported fact-checking tools were explored as potential mechanisms to address misinformation related to healthy ageing.

Participants interacted with multiple fact-checking prototypes that ranged from non-functional interface models to iterations incorporating AI-generated scripts, avatar-based dissemination, and text and audio solutions. Rather than evaluating whether artificial intelligence improves fact-checking outcomes in abstract terms, the study examines how caregivers interpret these dissemination formats and how such formats intersect with expectations of authority, protection, familiarity, inclusion, and credibility.

Methodologically, the research adopts an exploratory qualitative design that combines co-design practices, focus group discussions, empathy mapping exercises, and descriptive questionnaire data. Quantitative elements, including Likert-scale responses, are used descriptively to structure participant perceptions and to support qualitative interpretation rather than to enable statistical inference.

2 Fact-checking, AI and the ethical stakes of health misinformation

To understand the problem, we will draw on the literature to organize the discussion around the main tensions that motivated our research project, particularly the gap between verification and dissemination in fact-checking, the challenge of operationalizing ethical principles in AI design, and the underrepresentation of vulnerable populations in empirical work on AI ethics.

2.1 Fact-checking beyond the verification phase

We define fact-checking as a rigorous process dedicated to validating factual accuracy and pursuing investigative clarity, moving beyond simple assessments of credibility [1, 8]. The Duke Reporters' Lab's annual fact-checking census [2] noted increased active fact-checking initiatives, underscoring a global expansion in fact-checking efforts with nearly 400 professionals addressing misinformation across 105 nations.

Fact-checking distinguishes itself from traditional journalism through its laser focus on precision, prioritizing factual accuracy over interpretative narratives [9]. Given the multifaceted nature of this challenge, our initial focus has been on the relatively underexplored domain of automated fact-checking, which presents a significant opportunity for innovation. Research by Full Fact Org [10] delineates four automated stages of fact-checking: monitoring, spotting, verifying, and disseminating findings. Maia [11] conducted a systematic review highlighting the predominant emphasis on the verification phase, followed by identification and monitoring, with relatively few studies delving into the nuances of dissemination technologies used for fact-checking.

This imbalance is an important turning point, as the effectiveness of fact-checking depends not only on accurate assessment, but also on whether corrections reach and are accepted by the target audience. This gap becomes more pressing as AI technologies, such as large language models, increasingly mediate the dissemination phase.

Given this, this study seeks to understand ethical ways of using AI in the dissemination phase of results already verified by fact-checking agencies, as pointed out to be a prominent area of study and focusing on verifying misinformation in health.

2.2 Value sensitive design and human–machine communication

AI, particularly in spreading fake news, underscores the necessity of incorporating ethical considerations into its

development. Leading experts also advocate for frameworks integrating values into decision-making processes [12]. For this purpose, ethical complexities such as balancing source protection and transparency in algorithmic decisions must be addressed [13]. Helberger and Diakopoulos [14] also elucidate the complex ethical and regulatory landscapes in AI governance in media and journalism.

Some policies have been created to help generate innovation and technologies within ethical challenges. Various models support this integration, including the value survey model [15], the value pie model [16], the theory of fundamental values [17], and Value Sensitive Design—VSD [4]. This study explores the ethical dimensions of computing employing VSD as a foundational framework, which systematically incorporates human values into technology design [3]. The VSD framework posits that technology and human values are intrinsically linked, each shaping the other and reflecting the values of its creators and implementers, intentionally or not.

This method, supported by Freier [18], emphasizes the significance of mock-ups, prototypes, and field deployments in exploring the value implications of technologies at early stages, a notion further explored by Yoo et al. [19]. The subsequent phase focused on co-design and empathy principles highlighted by Dimitrakopoulou and Lewis [20]. This stage is crucial for fostering empathy, identifying values and ethical considerations, and using technological probes and ideation techniques for innovative solution-seeking, as Portugal et al. [21] discussed.

Our analysis is aligned with Human–Machine Communication (HMC), a field focused on creating meaning between humans and machines [5]. In this study, VSD and HMC work complementarily at different levels. VSD guides the definition of fundamental values for the functioning of the proposed technology; HMC informs how we interpret participants' responses in different AI-mediated formats [22]. Together, they allow us to examine how the defined values are well accepted and integrated into norms through interaction with technological probes that simulate AI scenarios used for fact-checking.

For the purposes of this analysis, we define four key concepts: trust, credibility, values and inclusion. Trust refers to users' willingness to rely on information from a technological system, shaped by perceptions of competence, transparency, and benevolence [4]. Credibility refers to the perceived believability of information, influenced by source characteristics, format, and alignment with users' prior knowledge. Values, following Schwartz [17] and VSD, are trans-situational goals that guide people's lives and become embedded in sociotechnical systems. Inclusion denotes the extent to which a technology accommodates diverse needs, particularly of marginalized users [4].

In methodological terms, VSD is operationalized in this study as a qualitative and interpretive framework rather than as a prescriptive checklist of values. The co-design workshop functions as an empirical strategy to elicit participants' value articulations, ethical concerns, and interpretive frameworks when engaging with AI-based fact-checking solutions. This approach aligns with prior work that frames VSD as an exploratory process focused on identifying value tensions, contextual meanings, and normative assumptions embedded in sociotechnical systems [4, 19].

2.3 Vulnerable populations in AI ethics research

Another aspect that deserves attention is the limited empirical research on vulnerable populations in AI ethics research. As Furtado et al. [7] noted, elderly populations and communities in economically disadvantaged regions remain underrepresented. In our context, caregivers face distinct barriers to engaging with digital fact-checking tools, including limited digital literacy and communication norms that differ from those assumed by mainstream AI design. This gap was also evident during the COVID-19 infodemic, which disproportionately affected populations with less access to reliable digital information sources [1].

These three areas of work reveal an unaddressed gap. Fact-checking research neglects the communicative dimensions of AI-mediated dissemination. VSD and HMC offer analytical tools, but they have rarely been applied to vulnerable populations in resource-limited contexts. This study seeks to understand not only technological usability and efficiency, as is usual in applied technology studies, but also whether these proposed solutions can respond to our assumption that it is possible to learn from the disinformation to create antidotes that are as eye-catching and captivating as the disinformation itself. The central research question is the following: how can the integration of artificial intelligence promote ethically responsible practices in health fact-checking, with an explicit focus on the dissemination of verified content?

Faced with these challenges, we sought to examine participants' feedback on principles such as Protection, Trust, Engagement, Familiarity, Frustration, and Fear, as well as additional aspects such as Information, Speed, Resonance, Inclusion, Character Appearance, and Novelty. Our research seeks to reveal insights into the design and effectiveness of digital solutions in combating health misinformation. These principles were derived from the literature on VSD and refined through phases of our previous research [23] in which workshops and VSD methods helped identify the most relevant value dimensions for participants in this context.

2.4 Caregivers as direct stakeholders in a situated context

Within the framework of Value Sensitive Design, identifying directly affected stakeholders is a key step in VSD. In this study, elderly caregivers in Macedo de Cavaleiros, Portugal, were approached as direct stakeholders whose everyday practices and constraints shape the ethical viability of AI-supported fact-checking tools.

The empirical context is linked to the broader EducAGE initiative, a European programme aimed at strengthening competencies related to healthy ageing and digital engagement among caregivers in underrepresented regions [6]. EducAGE was not the object of evaluation; it provided access to a community of caregivers whose profiles reflect broader structural challenges identified in European long-term care systems. Informal caregivers represent a substantial proportion of care provision across Europe and frequently operate without structured digital training, despite increasing reliance on online health information [7].

Positioning caregivers from this context as stakeholders enables the study to examine AI ethics not in abstract terms, but in relation to lived experiences of care and the challenges of accessing reliable health information. The focus is therefore not on evaluating the EducAGE programme itself, but on understanding how AI-mediated fact-checking is interpreted within a community characterized by limited digital readiness and high responsibility for others' well-being.

3 Value-oriented co-design workshop

When applying VSD, our focus when connecting with an audience of elderly caregivers was directed toward participation in an existing activity of this project in a complementary phase. As the project was already involved in health initiatives, our methodology was included in this context as a complementary addition. That being said, the recruitment was conducted in collaboration with the Senior Citizens University in Macedo de Cavaleiros (Macedo [24]), which regularly hosted EducAge activities for older adults and caregivers. Participation was voluntary, following a public flyer invitation aligned with the dissemination activities of the EducAGE initiative. The aim was not statistical representativeness but contextual depth, focusing on individuals who are actively engaged in caregiving practices and who are likely to encounter health-related information challenges in their daily lives.

The workshop was designed as a qualitative inquiry aimed at eliciting caregivers' interpretations, values, and ethical expectations regarding AI-mediated health information. Rather than evaluating system performance, the

activities focused on how participants made sense of the proposed solutions, how they articulated trust, protection, and inclusion, and how these values were constructed collectively. The data generated through questionnaires, group discussions, empathy maps, and focus group conversations were therefore treated as qualitative material, analyzed through thematic and interpretive lenses. By working with this specific community, the study prioritizes depth of understanding over sample size. This workshop offered the following activities aimed at improving the impact and relevance of design:

- a. *Questionnaire evaluation* participants interact with various chatbot prototypes, providing detailed feedback on values and effectiveness. A questionnaire with 13 questions in line with Kitchenham and Pfleeger [25] recommendations will be explained further in the next part of the research. The responses are on a Likert scale, ranging from 1 (do not agree at all) to 5 (completely agree). Participants can rate which proposed prototype solution is their favorite.
- b. *Brainstorming and empathy map* participants engaged in co-design activities to collaboratively define the persona and characteristics of the chatbot avatar. The Empathy Map entails developing a visual map to consolidate insights from the users into sections like seeing, hearing, and feeling, facilitating the creation of detailed personas. These personas offer a comprehensive understanding of user motivations and challenges, guiding user-centered decision-making in the design process. In this workshop, participants were divided into three groups to carry out the activity, and after 30 min, they presented their creations to one another.
- c. *Focus group discussion* focus groups, as defined by Krueger [26], are structured group debates designed to probe a series of issues, where the researcher engages with at least six participants in face-to-face discussions. Leveraging the values brought to light, participants will collaboratively forge a normative framework to guide the ethical dimensions of the design process. For this research, we held a session lasting around 30 min, with a pre-established script that aimed to provide greater clarity about the participants' preferences for the technologies and other details about their evaluations of the values presented.

The choice of these methods is part of a broader doctoral research context, in which similar workshops were used in earlier phases [27]. These earlier phases helped us refine the methods within the VSD and establish values and standards that correspond to certain fundamental principles in the use of automated fact-checking. Given these iterations

and feedback from participants, these principles were established as our guidelines: Protection, Trust, Engagement, Familiarity, Frustration, and Fear; and additional questions about Information, Speed, Resonance, Inclusion, Character Appearance, and Novelty.

In these workshops, the author acted as principal investigator and lead facilitator. To mitigate potential biases and ensure balanced participation, facilitation was carried out in collaboration with other project leaders associated with the EducAGE initiative, who supported the moderation of discussions and the execution of workshop tasks. This shared facilitation approach aimed to encourage diverse viewpoints and reduce the influence of a single researcher's interpretation.

Analytically, the data generated was analyzed using thematic analysis [28]. A first round of open coding considered all values explicitly mentioned by participants in previous workshop exercises inductively. In a second phase of iterative coding, these values were compared across data generated to identify recurring patterns, convergences, and tensions in a more deductive manner. The codes were then grouped into higher-order thematic categories, reflecting the dimensions of core values relevant to AI-mediated fact-checking. This process allowed us to move from descriptive elicitation of values to an interpretive mapping of the core value orientations that shape participants' assessments.

4 Disclosing iterated technological probes

An important part was the presentation of technological prototypes to the public. For this part, focusing on developing an automatic fact-checking verifier for a web portal, we tested a few advanced technologies, particularly emphasizing the use of large language models (LLMs) and their differentiated applications and software.

Beyond their technical implementation, these probes were designed as analytical instruments rather than finished products. Each configuration intentionally varied dimensions

such as institutional authority, character embodiment, communicative tone, and degree of automation to elicit stakeholders' value judgments. The goal was not to test technological performance per se, but to examine how different representational and communicative forms of AI-mediated fact-checking shape perceptions of trust, legitimacy, protection, and inclusion. In this sense, the probes functioned as value-elicitation devices aligned with the study's central research question. The overall process of creating the technological probes is illustrated in Fig. 1.

In the initial step, prompts were crafted to specify the instructions for generating the solutions to be automated by APIs. One solution used for comparison was a video produced without AI, following the traditional formats of SIC, a journalistic company in Portugal that operates a fact-checking agency known as Polígrafo and conducts special fact-checking reports. The prompts also delineated the content to be summarized automatically, bearing in mind that in a functional system, there would be a filter from the datasets to be consulted with data from recommended fact-checking agencies. In this instance, each prompt specified a different text from a fact-checking agency.

Additionally, the prompt indicated that those texts should be summarized according to the specific unique characteristics of each character, with a text limit imposed. This phase is critical as it defined features and personality traits in our prompts, such as “respond with the seriousness of a traditional journalist,” “respond with the grumpiness of a well-informed cat,” or “respond with the kindness of a beloved pet.” Depending on each proposed solution, the summarized text automatically generates audio or text in the subsequent stage. Simultaneously, the video is produced in the next step according to the necessary simulation.

In total, we presented seven solutions for experimentation. A detailed description of the prototypes, including APIs and software tools used, is available in the project repository [29]. The characters were developed to showcase a range of approaches, from traditional to disruptive. These included well-established journalists in Portugal represented through deepfake technology to a grumpy cat, which was a solution suggested by participants in our previous workshop (Table 1).

Each technological setup was selected to explore LLMs' distinct capabilities and integration challenges in varying contexts. The configurations were explicitly chosen to demonstrate the potential of AI in enhancing user engagement, which is crucial for an effective fact-checking system within a digital platform. A non-functional prototype (Fig. 2) demonstrated how this fact-checking search engine would be integrated into the website. Subsequently, each probe was showcased within this prototype using a projector to facilitate the continuation of the experiments.

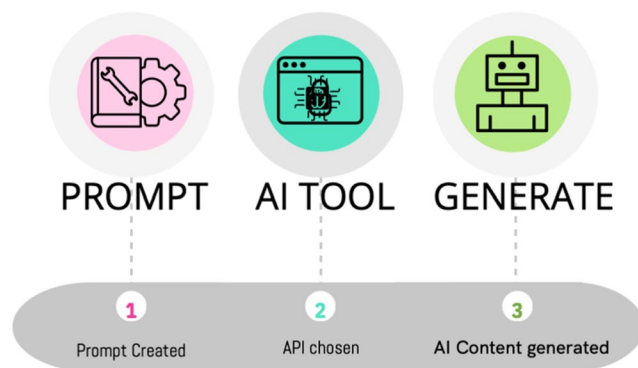


Fig. 1 Iterated flowchart for creating fact-checking videos

Table 1 Technological probes presented at the educage workshop

Solution	Description	Analytical dimension explored
Sônia Rebelo	Disclosure of Fact-checking using Full Fact data and Deepfake character. The character's name was created by us	Journalistic authority
AIDA	Disclosure of Fact-checking using Full Fact data and 3D character in Portuguese language from Globo	Innovative approach within 3D character
Dina Aguiar	Disclosure of Fact-checking using Full Fact data and deepfake from a Portuguese journalist known to the population	Familiar Journalistic authority
Melinda (female cat)	Disclosure of fact-checking using data from Full Fact and a 3D character of a friendly cat. The character's name was created by us	Embodiment and relational tone
Bigodão (male cat)	Disclosure of fact-checking using data from Full Fact and a 3D character of a grumpy cat. The character's name was created by us	Incorporation of irony and humor
Text and audio only	Disclosure of fact-checking using data from Full Fact and text only	Content clarity without embodiment
SIC Polígrafo	Report shown on the a television channel in Portugal with data from a fact-checking agency belonging to the same group	Non-AI baseline comparison

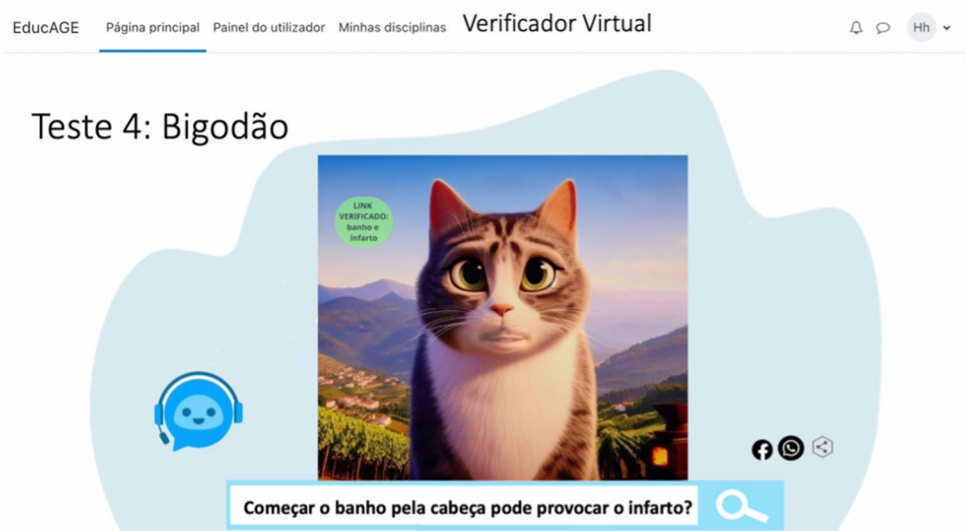
Analytically, the variation across probes allowed us to explore specific dimensions of the research question. Journalist-based deepfakes tested the role of institutional credibility and recognizable authority. Animated characters examined how embodiment and emotional tone influence perceptions of seriousness and trustworthiness. Text-and-audio formats isolated content clarity from visual mediation. The inclusion of a traditional non-AI journalistic

Table 2 Demographic profile and identifications of study participants

id	SEX	Age	Education
PQ1	male	64	Secondary education
PQ2	female	80	Basic education
PQ3	female	55	Secondary education
PQ4	female	67	Secondary education
PQ5	female	23	Higher education
PQ6	female	49	Higher education
PQ7	female	55	Secondary education
PQ8	male	69	Basic education
PQ9	female	67	Secondary education
PQ10	female	77	Basic education
PQ11	female	43	Higher education
PQ12	female	62	Secondary education

report provided a comparative baseline to assess whether AI mediation itself shaped evaluations. By systematically contrasting these formats, the workshop was structured to surface implicit criteria through which caregivers assess AI-supported dissemination.

The event occurred on the 5th of December at the Senior Citizens University from 9 am to 12 pm, in Macedo de Cavaleiros, Bragança. The proposed agenda for the event was aligned with the overarching goals of the EducAGE project and publicized. The research involved 12 participants from the town of Macedo de Cavaleiros, ten women and two men, with an average age of 59 (Table 2). The workshop was attended voluntarily by students from the Senior Citizens University, which provides courses and regular activities in the city. The elderly caregivers' educational levels also vary, with some having primary education, others having secondary education, and a few having higher education. PQ refers to Participants from the Questionnaire phase.

Fig. 2 High-fidelity non-functional prototype with probes for the project site

5 Descriptive patterns supporting qualitative analysis

The questionnaire data are presented here as descriptive patterns that support the qualitative interpretation of participants' perceptions. Given the exploratory nature of the study and the small, context-specific sample, the Likert-scale responses are not used for statistical generalization. Instead, they serve as analytical cues that help identify recurring themes, contrasts between solutions, and areas of ethical sensitivity, which are further examined through qualitative data from the co-design activities and focus groups.

The questionnaire was printed, and response symbols were on a Likert scale, ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). The following sections show data responses from the 12 participants (PQ1-PQ12) for the solutions Sônia Rebelo, Dina Aguiar, Melinda, Bigodão, text and audio only, and Polígrafo da SIC. To avoid fragmentation across multiple dimensions, we synthesize the questionnaire results into two visual summaries: (1) comparative scores in core trust-related categories and (2) overall preference ranking. These figures serve as entry points for cross-cutting analysis rather than isolated reporting.

5.1 Preferred solutions

The final chart (Fig. 3) presents a compilation of user preferences across various health information solutions, with participants asked to rank these from their most to least favorite. In an assessment of overall satisfaction, Sônia Rebelo and Dina Aguiar emerged as the top choices, each accumulating 48 points across all participants. This outcome suggests that these two solutions were equally favored among the users, resonating well with their needs and preferences.

Sônia Rebelo shows diverse rankings but notably secures the top spot (6 points) from PQ1, PQ6, and PQ9, indicating extreme preferences from specific individuals. Meanwhile, Dina Aguiar appears to be a consistent favorite, achieving the highest ranking (6 points), notably from PQ4, PQ7, PQ9, and PQ10, showing a broad appeal among the participants. In conclusion, while Sônia Rebelo and Dina Aguiar stood out as the most appreciated solutions for health information, indicating their effectiveness in user engagement and content delivery, there were varied preferences across the board. This diversity in user rankings underscores the importance of offering various solutions to cater to any user base's needs and tastes.

This ranking shows that Sônia Rebelo and Dina Aguiar emerged as the most preferred solutions, accumulating the highest total scores across participants. Rather than interpreting this as mere popularity, this pattern indicates that formats associated with recognizable journalistic authority tend to function as anchors of perceived legitimacy. When read alongside focus group discussions, this preference aligns with participants' explicit emphasis on scientific references, traceable sources, and institutional credibility. In this sense, the ranking reflects not simply aesthetic preference but a deeper evaluative mechanism in which institutional recognition serves as a proxy for trustworthiness. However, the distribution of top rankings also reveals individual variation, suggesting that legitimacy is constructed rather than uniformly assigned.

5.2 Descriptive patterns supporting qualitative analysis

To complement the preferred solutions of questionnaire results, Table 3 provides a consolidated overview of mean scores per solution across all evaluative dimensions. This

Fig. 3 Chart presenting character ranking

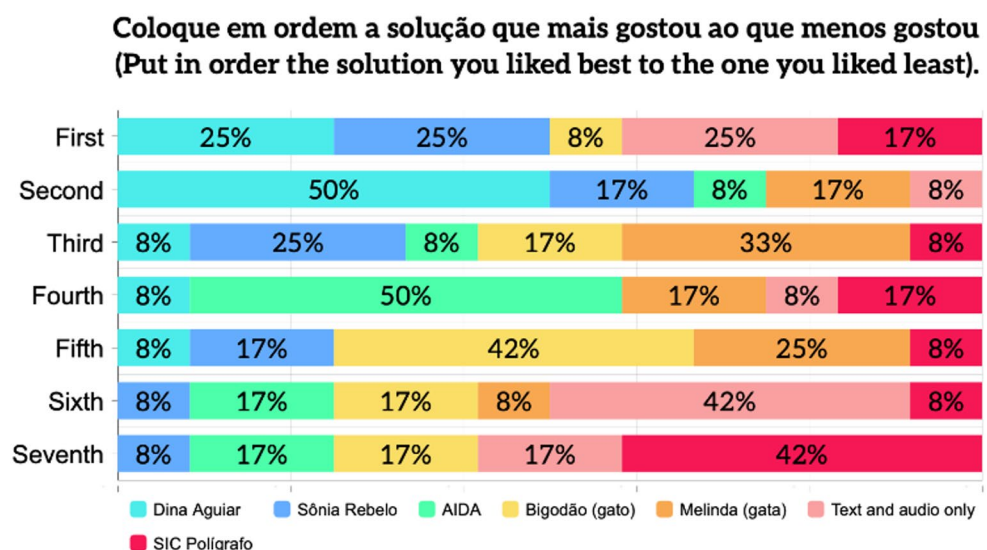


Table 3 Mean scores per dimension and solution

Dimension	Sônia Rebelo	AIDA	Dina Aguiar	Melinda	Bigodão	Text & Audio Only	Polígrafo da SIC
Protection	4.08	3.58	3.83	3.58	3.50	3.92	3.67
Trust	3.75	3.58	3.92	3.58	3.50	3.83	3.58
Engagement	3.83	3.58	3.92	3.67	3.33	3.83	3.67
Familiarity	4.08	3.92	4.00	3.42	3.42	3.50	3.92
Frustration	2.25	1.92	2.00	2.17	1.83	2.58	2.50
Frightening	2.25	2.42	2.08	1.92	2.00	2.50	2.33
Information	4.17	3.50	3.92	3.83	3.33	3.58	3.50
Promptness	3.67	3.75	3.92	3.50	3.50	3.83	3.58
Resonance	3.75	3.42	3.83	3.42	3.25	3.58	3.42
Inclusion	3.67	3.50	3.58	3.58	3.33	3.67	3.42
Beauty	3.67	3.33	3.75	3.58	3.17	3.75	3.75
Aesthetics	3.50	3.33	4.00	3.67	3.33	3.75	3.58
Novelty	3.58	3.58	3.83	3.75	3.50	3.92	3.58

table allows for a comparison across dimensions, showing where solutions scored similarly and where they differed.

5.2.1 Trust, protection and credibility as convergent dimensions

As shown in Table 3, dimensions related to *Protection* and *Trust* display consistently higher mean scores across most solutions, particularly for Sônia Rebelo and Dina Aguiar. These formats score close to or above 3.8 in trust-related items, reinforcing the interpretation that recognizable journalistic formats function as anchors of perceived legitimacy.

This convergence suggests that feelings of protection are closely linked to institutional cues, communicative seriousness, and familiarity with established media figures. Rather than being driven by technological sophistication or novelty, trust appears to be anchored in symbolic signals of authority.

These descriptive patterns gain additional significance when read alongside the qualitative data. During the workshop discussions, several participants emphasized the importance of traceable sources, scientific backing, and credibility markers. Those preferences correspond to the comparatively stronger performance of journalist-like formats in Table 3.

At the same time, the empathy mapping exercise revealed a parallel desire for warmth, listening, and relational tone. This creates a tension: caregivers appear to value relational sensitivity, yet hesitate to associate it with formats that may weaken institutional authority. Rather than reflecting contradiction, this tension illustrates a dual evaluative framework through which AI-mediated fact-checking is assessed. Participants simultaneously seek institutional reliability and communicative proximity.

5.2.2 Limited negative affect and calibrated evaluation

The dimensions *Frustration* and *Frightening* show consistently lower mean scores across all solutions, generally clustering around values close to 2.0 (Table 3). This pattern indicates that participants did not experience the AI-mediated formats as significantly irritating or anxiety-inducing.

The relative absence of strong negative affect is relevant. It suggests that ethical concerns surrounding AI-mediated dissemination are not expressed primarily through emotional rejection. Instead, participants' evaluations appear calibrated and differentiated. Differences between solutions emerge mainly in dimensions related to credibility, familiarity, and representational alignment rather than in outright technological resistance.

The data indicate that caregivers are not rejecting AI as such. Instead, they are evaluating whether it fits the seriousness of a health context. Their concerns focus on whether the tools feel credible, clear, and adequate for the care situations they face.

5.2.3 Inclusion and communicative alignment

Patterns related to *Inclusion* and *Representation* show moderately positive scores across formats, with slightly stronger performance for institutional and journalist-like solutions (Table 3). When interpreted alongside focus group discussions about digital literacy, accessibility, and clarity of language, these results suggest that inclusion is understood less in terms of aesthetic diversity and more in terms of communicative alignment and recognizability.

Participants appear to associate inclusion with being addressed clearly, respectfully, and through formats that resonate with familiar communicative norms. This reinforces the broader interpretation that AI-supported fact-checking is evaluated not merely on informational grounds but within a relational and contextual framework.

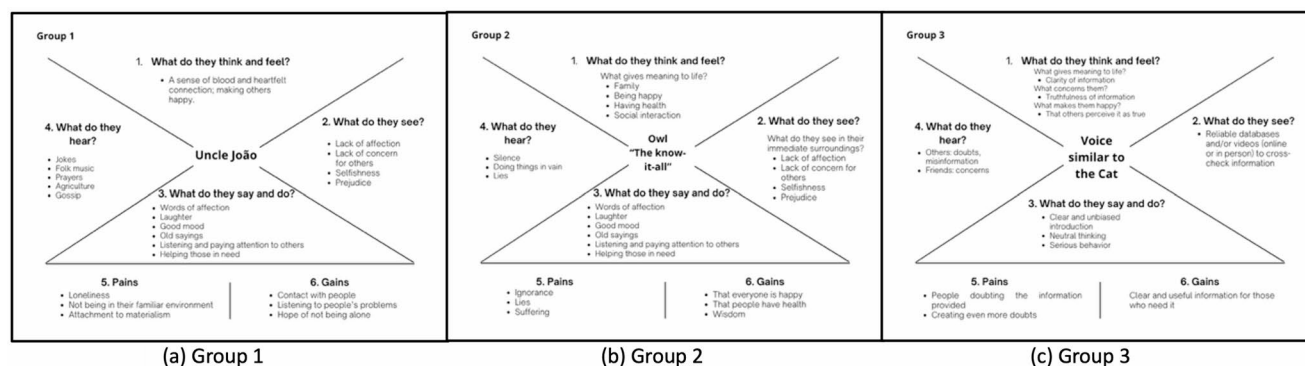


Fig. 4 Three paper sheets of work resulting from the workshop subgroup's exercise

Taken together, the cross-dimensional patterns summarized in Table 3 reinforce the central theoretical argument of this study. Dissemination format alone does not determine trust. Rather, AI-mediated fact-checking is assessed through layered expectations involving institutional credibility, relational tone, communicative familiarity, and perceived protection. When read alongside the qualitative workshop findings, the descriptive patterns support a value-oriented understanding of how legitimacy is formed in vulnerable health communication contexts.

5.3 Empathy map brainstorming

The empathy mapping exercise represents a crucial interpretive stage of the workshop, allowing participants to articulate the values, expectations, and concerns that shape their engagement with AI-mediated health information. Instead of serving merely as a creative activity, this phase functioned as a structured elicitation of normative orientations, revealing how caregivers conceptualize trust, authority, listening, and protection in communicative contexts. The characters developed during this activity provide insight into the relational and moral dimensions through which technological tools are evaluated. These value expressions later inform the interpretation of questionnaire patterns, enabling a more integrated understanding of how descriptive ratings correspond to deeper expectations about legitimacy and care. In the empathy map part of the workshop, three distinct groups (Fig. 4) developed unique characters, each reflecting varied nuances of life experiences in Macedo de Cavaleiros.

The first group, consisting of PQ4, PQ12, PQ10, and PQ6, created the character Ti João, inspired by a local radio journalist. Ti João finds meaning in life by bringing joy to others, standing out for his ability to listen and meet the needs of others. He lives in a world he perceives as lacking affection and solidarity, marked by disinterest and prejudice. Expressing himself through words of affection and laughter, Ti João is concerned with loneliness and materialism, longing for deep and genuine human connections. From the

Table 4 Group 1 participants explain their choices

Thematic code	Participant	asserts example
Reliability and consistency	PQ6	So, we chose a character inspired by radio announcers like Rui Costa and Uncle João from Bragança. They can reach out further in terms of communication and speaking with them
Participation and involvement	PQ6	There is a difference compared to the characters of this project: They can listen. That could be a tip to leave that it could listen to people's problems. Because throughout all the things we were adding, listening was a fundamental aspect

audio transcription of this group presentation from PQ6, we could highlight in Table 4:

The second group (PQ8, PQ2, PQ3, PQ9) introduced the Owl, a character symbolizing wisdom and perception. The Owl values family, happiness, and health but is tormented by inequality and loneliness. Visually, this character may reflect the wisdom traditionally attributed to owls, with a keen observation of its surroundings, notably similar to the scenario perceived by Ti João. Studious and reflective, the Owl seeks to convey knowledge and love while fighting against ignorance and aspiring to a world where everyone is happy and wise. These are the highlights (Table 5):

Finally, participants PQ1, PQ11, PQ5, and PQ7, with the third group (Table 6), focused on their character's voice without reaching a consensus on its visual appearance. The chosen voice is similar to that of a cat presented earlier, suggesting a gentle yet assertive approach. This character, valuing clarity and truthfulness of information, is deeply concerned with the authenticity and accuracy of the data shared. Their voice would be used to express neutral and impartial thoughts, concerned with the reception and impact of their information, always in pursuit of elucidating truths and dispelling doubts. Here are the highlights of this group:

Table 5 Group 2 participants explain their choices

Thematic code	Participant	Asserts example
Availability and accessibility	PQ6	So, the main thing is family. We have someone from the family to help us be happy and healthy, with coexistence, union, guilt, inequality, loneliness, abandonment, and lack of love
Cultural sensitivity	PQ6	And now here, what do they hear? Silence. Often, we need to correct them and remain silent
clarity and comprehensibility	PQ3	What do they say and do? The owl. What do they allow us? She studies, sees everything, hears, shares knowledge, and speculates for certainty, love, and affection

Table 6 Group 3 participants explain their choices

Thematic code	Participant	Asserts example
Protection	PQ6	We did not choose any character because we chose different characters... However, we recommend that the tone of voice or the audio must always accompany the text or character, which creates some closeness. So, the tone of voice we chose was that of Melinda, the cat, who, in times of difficulty, can help you in this passage of information. Therefore, we chose a voice with the audio, so you do not have those feelings as if it were a person
Resonance	PQ6	The image makes little difference. The voice, tone, or audio perhaps creates closeness and verifies what we hear

In summary, each group created a character with a distinct personality and perspective, ranging from Ti João, who embodies the human warmth and compassion of a community journalist, to the reflective and wise Owl, and the cautious and accurate voice of the third character, all reflecting different aspects of the human desire to connect, understand, and communicate effectively. The activity was publicized on the city's official advertising channels on Facebook, demonstrating how relevant this action is for community engagement.

Beyond the descriptive richness of the characters created, the empathy mapping exercise reveals how caregivers conceptualize AI-mediated fact-checking not merely as an informational tool, but as a relational and moral actor. The figures of Ti João, the Owl, and the disembodied yet careful voice reflect distinct value orientations: community-based listening and solidarity, wisdom and reflective authority,

and cautious neutrality. These constructions resonate with the Value Sensitive Design perspective by surfacing stakeholders' implicit value hierarchies, particularly around reliability, participation, cultural sensitivity, and protection. From a Human–Machine Communication standpoint, the exercise demonstrates that participants attribute communicative roles and ethical expectations to technological artefacts, imagining them as listeners, educators, companions, or guardians of truth. The empathy maps therefore do more than illustrate creativity; they expose the expectations that shape how AI-supported health fact-checking is perceived as legitimate, trustworthy, and socially appropriate.

5.4 Focus group analysis

The focus group reveals a nuanced discussion on health information solutions among caregivers (Appendix D). Some participants expressed preferences for solutions like Dina Aguiar and text and audio formats, highlighting the importance of clear and direct information delivery, and an AI solution for co-existence in both the real and virtual worlds. Some found the characters and virtual elements less appealing, suggesting a divide between those who favor a more traditional, straightforward approach and those who may appreciate innovative, interactive methods. Concerns were raised about the relatability of virtual characters and the effectiveness of humorous approaches in serious health contexts. Participants also discussed the accessibility of digital tools for older populations and the varying levels of internet proficiency among caregivers, emphasizing the need for inclusive solutions that cater to a range of digital literacy levels. The conversation pointed towards the potential of technology to aid in disseminating information and the desire for trustworthy, easily verifiable sources, reflecting a collective awareness of the challenges in navigating health information in the digital age.

During the focus group, participants engaged in a rich discussion about the various health information solutions they experienced. PQ8 expressed a preference for realism over virtual elements, stating, “I prefer something more real than virtual... The voice and information don't present as negative or false. But a virtual image does raise more doubts for me about how the response might be” (translated from Portuguese). This sentiment was echoed by PQ11 who preferred text and audio solutions, suggesting that a serious health topic might not be best served by a playful character like a cat, noting, “If it was just to create an icon and have it in audio and text, that would be fine.”

PQ4 and PQ3 highlighted the importance of access to reliable information, with PQ4 saying, “I always look for scientific sites. And that there is a reference.” PQ3 pointed out the digital divide affecting older individuals, “From the

age of 60 and 70, many people don't know how to use the internet. They are left out of this solution."

The focus group also touched on the potential of these health information tools to assist caregivers in their roles. PQ10 voiced the need to educate caregivers on accessing such digital tools, saying, "I would want to teach caregivers how to access this." While PQ6 raised a critical change in the educational background between the elderly and their caregivers, suggesting that caregivers are now more informed and often have higher education levels than the people they are caring for.

PQ2 shared a personal anecdote about caring for her husband, illustrating the practical need for accessible health information among caregivers, "I didn't have training. But I had to care for my husband." The group also touched on the broader implications of such tools for elderly care, with PQ9 suggesting that the country needs to be more attentive and PQ11 advocating for self-empowerment, "If we want the elderly to be treated differently, or better, we ourselves have to do something about it."

PQ3 recounts her experience in a Swiss care home, emphasizing the autonomy and choices offered to the residents there: "I worked in a nursing home for elderly women in Switzerland, in Lausanne... It was the best because there were three houses in the same place... those who could look after themselves. And they had television. Everyone had television. Everyone had their own will. And in another sector, we had bedridden people. And they had problems. Few problems. And in another sector was psychological problems. I worked with people who had psychological problems. And the interesting thing there is that they could stay up much later together. Here in our homes no, because the elderly have to be all in bed. And they don't have access to television. There I realized that they want the television. They want, for example, the soap operas. They love to watch the soap operas. They want to see the programs..." (translated from Portuguese). This participant's observation highlights the importance of entertainment and personal choice in care facilities, suggesting that such amenities can significantly impact the well-being of the elderly.

PQ12's comment touches on the challenges in providing care for individuals with Alzheimer's: "But, in a situation of Alzheimer's, sometimes even the doctors don't give good advice." Similarly, PQ4 adds, "Exactly. You have to have your own knowledge. Then, realizing what is good for him, what is bad. How is our health system? Sometimes there is no doctor. Because the doctors, you know? They don't inform the patient; they don't say what should be. For example, they need a specialty. They don't have it because they don't even know the specialist doctor." These remarks underscore a gap in the healthcare system, where caregivers often need to become self-educated advocates to navigate

the complexities of care, especially in cases like Alzheimer's where nuanced understanding and support are crucial.

PQ8 shared a personal technique used to engage Alzheimer's patients: "So, for five years I transported sick people. I had two experiences with some Alzheimer's people. One, known. I used music in the car every day. And even recorded music. And we played guessing the singer. They said the name. Look, my strength isn't in singing. And they sang the song. Music greatly improved people with Alzheimer's, the man already transformed." This highlights the therapeutic potential of music and personalized interaction in Alzheimer's care, emphasizing non-traditional methods that can offer significant benefits.

These excerpts from the focus group discussion reveal the caregivers' insights into the nuanced needs of the elderly and those with cognitive challenges. The participants advocate for respectful, informed, and personalized approaches to care, stressing the need for accessible health information and the role of caregivers in ensuring the well-being of their charges. They underscore the diverse opinions and needs within the caregiver community regarding health information solutions. The conversation reveals a clear preference for realistic, direct, and accessible information, highlighting the importance of considering the end-user's background, preferences, and environment when developing and implementing such tools.

While the focus group documents varied preferences regarding formats and characters, its theoretical significance lies in how participants articulate legitimacy criteria for AI-mediated health communication. The discussion consistently returns to concerns about credibility, traceability of sources, realism, and accessibility, suggesting that trust is shaped through institutional alignment instead of technological novelty. Participants' skepticism toward overly playful avatars in serious health contexts shows that they draw a line between warmth and seriousness. At the same time, references to digital literacy gaps, caregiving burdens, and healthcare limitations situate AI-supported tools within broader problems of access and care. These discussions illustrate that acceptance of AI-mediated fact-checking depends not only on usability but on how these tools fit into the relationships and responsibilities caregivers already manage. The findings therefore contribute to AI ethics debates by showing that caregivers evaluate these technological through their own experiences of responsibility, limited support, and trust in institutions.

6 Discussions and conclusions

This study explored how elderly caregivers interpret and evaluate AI-supported fact-checking tools within a specific socio-cultural context. Rather than assessing technological performance or efficiency, the research examined how participants negotiate values such as trust, protection, credibility, and inclusion when interacting with AI-mediated health communication. The findings suggest that the legitimacy of AI-supported fact-checking does not emerge automatically from technological functionality but is formed through culturally embedded expectations about authority, care, and communicative responsibility.

Across the workshop activities, questionnaire responses, and focus group discussions, participants consistently evaluated the proposed solutions through credibility-oriented criteria. Participants repeatedly emphasized the importance of traceable sources, scientific references, and institutional legitimacy. Formats associated with recognizable journalistic authority were therefore perceived as more trustworthy. This reinforces existing research on fact-checking and media credibility, which shows that verification systems gain trust not merely through technological automation but through their connection to institutional credibility structures. In this sense, the findings suggest that AI-supported fact-checking tools are most likely to be accepted when they extend, rather than replace, established journalistic authority.

At the same time, the empathy mapping exercise revealed that participants' expectations of such tools extend beyond the traditional boundaries of fact-checking. Two groups explicitly identified loneliness and emotional isolation as significant contextual factors affecting older adults. The personas developed during the exercise reflect this concern. For instance, the character inspired by a local radio journalist emphasized listening, warmth, and companionship, while the owl persona symbolized attentive guidance and wisdom. These representations suggest that participants imagine information tools not only as verification mechanisms but also as communicative actors capable of providing interpretive support and emotional reassurance.

However, these expectations were not shared uniformly. A third group expressed clear reservations about the idea of virtual characters functioning as communicative agents in serious health contexts. Their preference for a neutral voice-based interface and emphasis on clarity and factual reliability reflects a more traditional understanding of fact-checking as an informational service. This divergence indicates the presence of at least two distinct interpretive orientations among participants. Some caregivers primarily value factual verification and authority, while others seek communicative tools that combine reliable information with relational interaction.

This divergence suggests that AI-supported fact-checking tools may need to address different audience needs rather than assume a single model of interaction. For some users, the primary function of such systems is the rapid verification of claims and access to credible sources. For others, particularly caregivers navigating emotionally demanding situations, informational verification may need to be embedded within more relational communicative interactions. The workshop results, therefore, indicate that future systems may benefit from flexible design approaches that allow users to select different modes of interaction depending on their informational and emotional needs.

The focus group discussions further contextualize these expectations within participants' lived experiences as caregivers. Conversations about institutional care practices and challenges related to Alzheimer's disease initially appear to extend beyond the immediate scope of fact-checking technologies. However, these discussions reveal how participants situate informational tools within broader care infrastructures. Caregivers responsible for individuals with complex health conditions must regularly interpret medical information, evaluate conflicting advice, and make practical decisions in situations characterized by uncertainty. Within this context, access to reliable and understandable information becomes a critical resource embedded in everyday care practices.

Participants' reflections on Alzheimer's care illustrate this dynamic particularly clearly. Caregivers described situations in which professional medical advice was insufficient or unavailable, forcing them to rely on personal judgement and informal information networks. In such circumstances, a fact-checking tool is not simply a verification service but a potential support mechanism that helps caregivers navigate uncertainty and conflicting information. Similarly, discussions about institutional care environments highlight how informational needs intersect with broader concerns about autonomy, dignity, and quality of life for older adults.

These observations align closely with the principles of Value Sensitive Design. The workshop activities allowed participants to articulate values such as listening, protection, credibility, and accessibility that are central to their everyday caregiving practices. Rather than treating technological design as a purely technical challenge, the findings demonstrate how AI-supported systems must incorporate these social and moral expectations to be perceived as legitimate and useful. The co-design process, therefore, reveals how stakeholder values shape the interpretation and acceptance of technological solutions.

The findings also resonate with perspectives from Human–Machine Communication research. Participants consistently attributed communicative roles and social expectations to the proposed AI systems. Rather than

perceiving them simply as tools, participants described them in terms that implied social agency, such as listening, guiding, informing, and protecting. This suggests that AI-supported fact-checking tools function not only as information interfaces but also as communicative actors.

An important insight emerging from the study is the absence of strong technological resistance among participants. While certain formats were preferred over others, participants did not reject AI-mediated solutions outright. Instead, their evaluations were nuanced and context-sensitive. Concerns were primarily directed toward communicative appropriateness, credibility cues, and accessibility rather than toward the technology itself. This suggests that ethical debates surrounding AI in journalism and health communication should focus less on technological acceptance and more on how systems align with existing trust structures and communicative expectations.

From a broader perspective, the study contributes to ongoing discussions on AI ethics in journalism and health communication by highlighting the importance of culturally grounded interpretations of trust and legitimacy. In contexts characterized by uneven digital literacy and limited technological infrastructure, the ethical deployment of AI-supported information systems cannot rely solely on technical innovation. Instead, responsible design requires engagement with local communicative practices, institutional trust relationships, and the everyday realities of caregivers and vulnerable populations.

Methodologically, the study demonstrates the value of combining co-design workshops, empathy mapping exercises, focus group discussions, and descriptive questionnaires to capture these interpretive dynamics. The questionnaire results were not used to produce statistically representative conclusions but served as descriptive indicators that helped identify patterns emerging from qualitative interactions. This triangulated approach allowed the research to capture how participants interpret the ethical implications of AI-supported communication technologies in practice.

The study is necessarily limited by its exploratory design and small sample size. The findings, therefore, cannot be generalized to broader populations. Instead, they offer context-specific insights into how elderly caregivers articulate ethical expectations regarding AI-supported fact-checking tools. Future research could extend these findings by examining similar systems in different cultural contexts, exploring how different interface designs influence trust and credibility, or comparing how various user groups interpret AI-mediated verification tools.

The central contribution of this study lies in demonstrating that AI-supported fact-checking tools are interpreted not only as mechanisms for verifying information but also as

communicative actors situated within broader networks of care, authority, and social meaning. The legitimacy of such systems, therefore, depends not only on informational accuracy but also on their ability to align with users' expectations of credibility, empathy, and communicative responsibility. By foregrounding the perspectives of elderly caregivers, this research provides new insight into how ethical considerations shape the reception of AI-mediated journalism and health communication technologies.

Taken together, these findings underscore the importance of user-centered and culturally sensitive approaches in the development of AI-supported fact-checking tools. Rather than positioning AI as a standalone solution to misinformation, the study illustrates how caregivers interpret such technologies through existing communicative norms, moral expectations, and local references. This perspective contributes to a more grounded understanding of how AI-mediated journalism and health communication are received in contexts where digital literacy and access to technological infrastructures may be uneven.

The research further demonstrates how exploratory, small-scale qualitative studies can inform ethical reflection on AI deployment in underrepresented regions. By focusing on the situated needs of older women and caregivers, the project highlights how AI applications can support media literacy and health education without assuming technological neutrality or universal applicability. While the findings are necessarily context-specific, they offer empirically grounded insights that may inform future research and practice on the ethical integration of AI in journalism, particularly in domains that require sensitivity to vulnerability, trust, and cultural meaning-making.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s43681-026-01112-4>.

Acknowledgements The authors thank the participants of the EducAGE project in Macedo de Cavaleiros, Portugal, for their valuable contributions to the workshops.

Author contributions Conceptualization: Haline Costa Maia; Methodology: Haline Costa Maia; Formal analysis and investigation: Haline Costa Maia; Writing – original draft preparation: Haline Costa Maia; Writing—review and editing: Haline Costa Maia, Helena Lima, Patrícia Cordeiro, Carolina Vila-Chã, Sérgio Nunes; Supervision: Helena Lima, Sérgio Nunes; Funding acquisition: Haline Costa Maia.

Funding Open access funding provided by FCT|FCCN (b-on). Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2021.05707.BD.

Data availability <https://doi.org/10.25747/TKJR-EV33>

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Posetti, J., Bontcheva, K.: Infodemic: disinformation and media literacy in the context of COVID-19. *Internet Sect. Overv.* **3**(13), 1–21 (2021)
- Stencel, M., Ryan, E., & Luther, J. (2022). Fact-checkers extend their global reach with 391 outlets, but growth has slowed. *Duke Reporters' Lab*.
- Borning, A., Friedman, B., & Kahn, P. (2004). Designing for human values in an urban simulation system: Value sensitive design and participatory design. *Proceedings From the Eighth Biennial Participatory Design Conference*,
- Friedman, B., & Hendry, D. G. (2019). *Value sensitive design: Shaping technology with moral imagination*. Mit Press.
- Guzman, A.L.: What is human-machine communication, anyway. *Hum. Mach. Commun. Rethink. Commun. Technol. Ourselves* **117**, 1–28 (2018)
- Ortega Morán, J. F., Vila-cha, C., Zamillo, D., Curaj, A., Sánchez Margallo, F. M., & Pagador, J. B. (2022). Validation of an App for the Awareness of Active Ageing Linked to Cultural Heritage. *Proceedings at International Workshop on Gerontechnology*.
- Furtado, G.E., Vaz, C., Bovolini, A., Marques, E., Serra, N., Costa-Brito, A.R., Vila-Chã, C.: The impact of physical activity levels and functional fitness status on the quality of life perceived by older adults living in rural and urban areas: The Portuguese inland case. *Healthcare* **10**(7), 1266 (2022)
- Miller, I. (2020). *Research Guides: Journalism - Fact-Checking Sites*. (Report from Eastern Washington University., Issue.
- Lamprou, E., Antonopoulos, N., Anomeritou, I., Apostolou, C.: Characteristics of fake news and misinformation in Greece: The rise of new crowdsourcing-based journalistic fact-checking models. *Journal. Media* **2**(3), 417–439 (2021)
- Babakar, M., & Moy, W. (2016). The state of automated fact-checking. *Full Fact*, 28.
- Maia, H. (2022). Disinformation and Health: A systematic review study on health automatic fact-checking during the Pandemic. *Proceedings at 17TH International AAAI Conference on Web and Social Media*,
- Spiekermann, S.: *Ethical IT Innovation: A Value-Based System Design Approach*. CRC Press (2015)
- Dörr, K.N., Hollnbuchner, K.: Ethical challenges of algorithmic journalism. *Digit. J.* **5**(4), 404–419 (2017)
- Helberger, N., & Diakopoulos, N. (2022). The European AI act and how it matters for research into AI in media and journalism. *Digit. Journal.*, 1–10.
- McCoy, J.L.: Rokeach. beliefs, attitudes and values: a theory of organization and change” (book review). *Rural. Sociol.* **34**(4), 597 (1969)
- Pereira, R., & Baranauskas, M. C. C. (2014). Value pie: a culturally informed conceptual scheme for understanding values in design. *Human-Computer Interaction. Theories, Methods, and Tools: 16th International Conference, HCI International., Crete, Greece*.
- Schwartz, S.H.: An overview of the Schwartz theory of basic values. *Online Read. Psychol. Cult.* **2**(1), 2307–0919.1116 (2012)
- Freier, N. G. (2008). Children attribute moral standing to a personified agent. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 343–352.
- Yoo, D., Hultgren, A., Woelfer, J. P., Hendry, D. G., & Friedman, B. (2013). A value sensitive action-reflection model: evolving a co-design space with stakeholder and designer prompts. *Proceedings of the SIGCHI conference on human factors in computing systems*.
- Dimitrakopoulou, D., Lewis, S.C.: The generative dialogue framework and the pursuit of better listening by journalists: A design-centered approach for more constructive conversations with audiences. *Digit. J.* **11**(3), 547–568 (2023)
- Portugal, R., Wilczek, B., Eder, M., Thurman, N., & Haim, M. (2023). *Design Thinking for Journalism in the AI age: Towards an Innovation Process for Responsible AI Applications* Paper presented at the The Joint Computation + Journalism European Data & Computational Journalism Conference, Zurich.
- Shneiderman, B., Plaisant, C., Cohen, M. S., & Jacobs, S. M. (2009). Designing the user interface: strategies for effective human-computer interaction.
- Maia, H., Ariel, P., Nunes, S.: Adding human values on the deepfake: co-designing fact-checking solutions to combat misinformation. *AI Ethics* **5**(3), 3035–3050 (2025)
- Macedo De Cavaleiros. (2024). Retrieved 5 de abril from <https://www.cm-macedodecavaleiros.pt/pages/262>
- Kitchenham, B., Pfleeger, S.L.: Principles of survey research Part 4: questionnaire evaluation. *ACM SIGSOFT Softw. Eng. Notes* **27**(3), 20–23 (2002)
- Krueger, R. A. (2014). *Focus groups: A practical guide for applied research* (5 ed.). Sage publications.
- Maia, H. C. (2025). *Artificial Intelligence and Infodemic: a study on fact-checked Health Communication and Synthetic Media* Faculty of Engineering of University of Porto]. Porto. <https://repositorio-aberto.up.pt/bitstream/10216/165758/2/714041.pdf>
- Braun, V., Clarke, V.: What can “thematic analysis” offer health and wellbeing researchers? *Int. J. Qual. Stud. Health Well-being* **9**(1), 26152 (2014)
- Maia, H. (2024). *Artificial Intelligence and Infodemic: Video Dataset for Fact-Checked Health Communication and Synthetic Media*. (INESC TEC) <https://doi.org/10.25747/TKJR-EV33>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.