# Giving DIAnA More TIME – Guidance for the Design of XAI-Based Medical Decision Support Systems

Enrico Bunde[1](✉), Daniel Eisenhardt[1], Daniel Sonntag[2], Hans-Jürgen Profitlich[2], and Christian Meske[1]

[1] Ruhr-University Bochum, Bochum, Germany
{enrico.bunde,daniel.eisenhardt,christian.meske}@rub.de
[2] German Research Center for Artificial Intelligence (DFKI), Oldenburg University, Oldenburg, Germany
{sonntag,profitlich}@dfki.de

**Abstract.** Future healthcare ecosystems integrating human-centered artificial intelligence (AI) will be indispensable. AI-based healthcare technologies can support diagnosis processes and make healthcare more accessible globally. In this context, we conducted a design science research project intending to introduce design principles for user interfaces (UIs) of explainable AI-based (XAI) medical decision support systems (XAI-based MDSS). We used an archaeological approach to analyze the UI of an existing web-based system in the context of skin lesion classification called DIAnA (Dermatological Images – Analysis and Archiving). One of DIAnA's unique characteristics is that it should be usable for the stakeholder groups of physicians and patients. We conducted the in-situ analysis with these stakeholders using the think-aloud method and semi-structured interviews. We anchored our interview guide in concepts of the Theory of Interactive Media Effects (TIME), which formulates UI features as causes and user psychology as effects. Based on the results, we derived 20 design requirements and developed nine design principles grounded in TIME for this class of XAI-based MDSS, either associated with the needs of physicians, patients, or both. Regarding evaluation, we first conducted semi-structured interviews with software developers to assess the reusability of our design principles. Afterward, we conducted a survey with user experience/interface designers. The evaluation uncovered that 77% of the participants would adopt the design principles, and 82% would recommend them to colleagues for a suitable project. The findings prove the reusability of the design principles and highlight a positive perception by potential implementers.

**Keywords:** Design Science Research · Design Principles · Explainable Artificial Intelligence · Medical Decision Support Systems · Healthcare

## 1 Introduction

For future developments in the healthcare sector, human-centered ecosystems that integrate artificial intelligence (AI) are becoming indispensable [1]. Novel healthcare technologies like AI-based medical decision support systems (MDSS) can promote good


A. Gerber and R. Baskerville (Eds.): DESRIST 2023, LNCS 13873, pp. 107–122, 2023.
https://doi.org/10.1007/978-3-031-32808-4_7

health, well-being, and support access to healthcare globally [2]. In addition, such AI-based technologies can help make medical analysis more efficient [3] and outperform human experts in tasks like classifying pigmented skin lesions [4]. However, AI also introduces unique challenges. For example, Tschandl et al. [5] concluded that faulty AI could mislead a spectrum of clinicians, including experts. Additionally, many modern AI approaches are black boxes and thus not interpretable, which leads to questions regarding accountability, liability, fairness, and explainability [6].

Scholars from the field of explainable AI (XAI) work on such challenges and aim to introduce transparent AI models or techniques that explain black box models [7]. XAI-based MDSS have already entered clinical practice, and researchers have investigated them in contexts like in-hospital mortality, intensive care unit admissions, or incidences of leukemia and cancer [8]. When designing such systems, involving relevant stakeholder groups like physicians or patients to consider human factors is highly important [9]. Well-designed XAI-based MDSS can facilitate the acceptance of AI predictions, support detecting errors, or aid in establishing appropriate levels of trust toward the system [7, 10, 11]. However, there is only little prescriptive design knowledge that guides practitioners when developing XAI-based MDSS and their user interfaces (UIs) as the first entry point of the Human-XAI interaction [9, 10]. We address this research gap by establishing the following research question:

***RQ:*** *Which design principles for UIs can be derived from an in-situ analysis of an XAI-based MDSS for skin-related diagnoses that consider multiple stakeholders?*

We have conducted a design science research (DSR) project to answer the research question. Following an archaeological approach [12], we have analyzed an existing XAI-based MDSS in-situ called *Dermatological Images – Analysis and Archiving* (DIAnA). The unique characteristic of DIAnA is that it aims to be accessible and usable by multiple stakeholders, including physicians and patients [13]. Therefore, we involved two physicians, four physicians in training, and six patients during the analysis to consider their individual information needs [10, 14]. For the analysis, we used think-aloud protocols and semi-structured interviews [15]. Since the Theory of Interactive Media Effects (TIME) formulates UI features as causes and user psychology as effects [16], it provided a valuable knowledge base for deriving our applied interview guide. Based on the insights, we derived 20 design requirements (DRs) associated with the needs of physicians, patients, or hybrid nature if they address the needs of both stakeholder groups. We addressed these DRs with nine design Principles (DPs), which were formalized according to the scheme of Gregor et al. [17]. We anchored the DPs in concepts of TIME, which explains the effect of technologies on humans in terms of affordances and thus provides valuable insights on perceived system properties that could drive users to operate the system in nuanced ways [16, 18]. We evaluated the DPs qualitatively with four software developers and quantitatively with 66 user interface and user experience designers regarding their reusability [19]. The evaluation ensured that the DPs introduced are applicable in practice and revealed a positive perception.

We present the DSR project in the following structure. First, in Sect. 2, we provide the background. Then, we describe our research design and the methodologies used in Sect. 3. Afterward, we derive and present the prescriptive design knowledge in Sect. 4, followed by a presentation of the evaluation and the results in Sect. 5, which we discuss in Sect. 6. We conclude the article with Sect. 7.

# 2 Background

## 2.1 Explainable Artificial Intelligence in Medical Decision Support Systems

The pervasiveness of AI has significantly increased across domains, which also applies to the healthcare industry [7, 10]. AI can outperform human experts in medical contexts like the classification of pigmented skin lesions and can reduce the time required in processes of medical analysis [3, 5]. The influence of AI in a medical context is so powerful that surgeons or physicians may even change their decisions [20]. However, the lack of transparency and explainability is still a major obstacle to the usage of AI in a high-risk context like health [10]. Therefore, the relevance of XAI for AI-based MDSS becomes steadily more relevant [8]. Nevertheless, the sheer presence of explanations may not improve human-computer interaction or interpretation of AI-based results [7, 9]. Although XAI has entered health-related research streams, there is still much need for research on XAI-based MDSS to generate user benefits [8, 10]. Especially the design of XAI-based MDSS and their UIs is under-researched [9, 10].

## 2.2 The Multi-stakeholder Perspective

When designing MDSS that are efficient and supportive, it is highly relevant to involve individual stakeholder groups and take a human-centered perspective [8, 9]. Different stakeholder groups can use MDSS, like physicians or patients [10]. Prior research uncovered clinicians prefer systems to quickly gain an overview and identify critical information immediately [21, 22]. Because MDSSs for patients are no longer viewed as a repository [23], new requirements for this specific stakeholder group need to be derived. For example, from the user experience perspective, patients prefer essential information on conditions and treatments to better evaluate their health status and physician decisions [24]. Due to patients' high possibility of misinterpretation of information, it is necessary to understand how information can be optimally communicated to them [25]. Design-oriented studies can therefore improve the accessibility of novel healthcare technologies and may reduce the barriers for patients to participate in their treatment [2].

## 2.3 Theory of Interactive Media Effects (TIME)

TIME considers two individual routes of affordances, the action route and the cue route [16]. The action route considers the psychological impacts on users' knowledge, attitudes, and behaviors. In contrast, the cue route focuses on the presence of individual cues in the UI and their effect on users' perception. Overall, TIME includes four models to explain its propositions regarding the predictors, mediating variables, and outcomes. The models are *Interactivity Effects Model*, *Agency Model*, *Motivation Technology Model*, and the *Modality-Agency-Interactivity-Navigability (MAIN) Model*. The first three models explain the action route of TIME, and the MAIN model the effects of the cue route. Since TIME consists of four underlying models focusing on different psychological facets [16], we adopted suitable elements for designing UIs of XAI-based MDSS, including interactivity, navigability, perceptual bandwidth, or interface and content perception.

TIME further posits that UIs provide interactive features which influence user engagement in nuanced ways. They involve various technological affordances and psychological variables [16]. To consider these factors, we grounded our DPs in TIME, which predicts affordances that affect user psychology [16]. Transferred to an AI context, afforded user actions can influence user engagement and experience through the provided interactions with the system [18]. Similarly, the features of UIs influence content perception [26]. Moreover, a well-designed interaction between users and AI systems can increase trust in the system [27]. TIME looks at different aspects of the interaction between users and the medium and explores how these influences can affect the behavior of users toward the UI [18].

## 3 Method: The Design Science Research Process

For our DSR project, we adapted the framework of Kuechler and Vaishnavi [28]. Figure 1 provides a concise overview of the process steps, methods used, activities conducted, and outputs. The first step, *awareness of problem*, was initiated by a literature review and the in-situ artifact analysis. In the second step, *suggestion*, we investigated existing literature to identify a suitable theoretical basis for the DPs. During the third step, *development*, we conceptualized the DPs and integrated insights from TIME. The *evaluation* is the fourth step. Finally, we evaluated the reusability of the proposed qualitatively and quantitatively. By reflecting and interpreting the proposed DPs in combination with the evaluation results, we complete the last step of the *conclusion*.

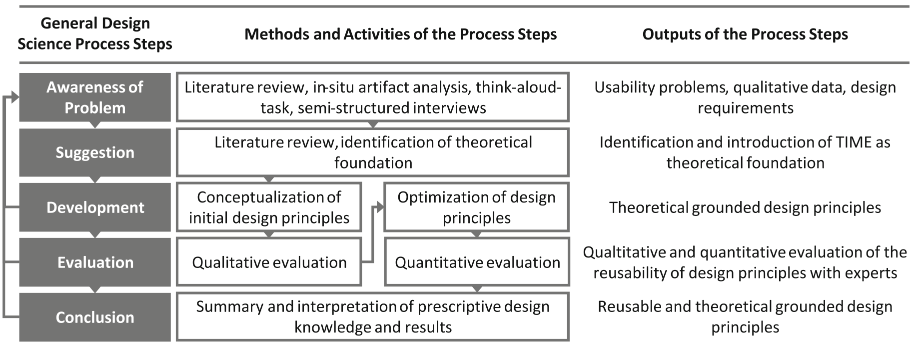
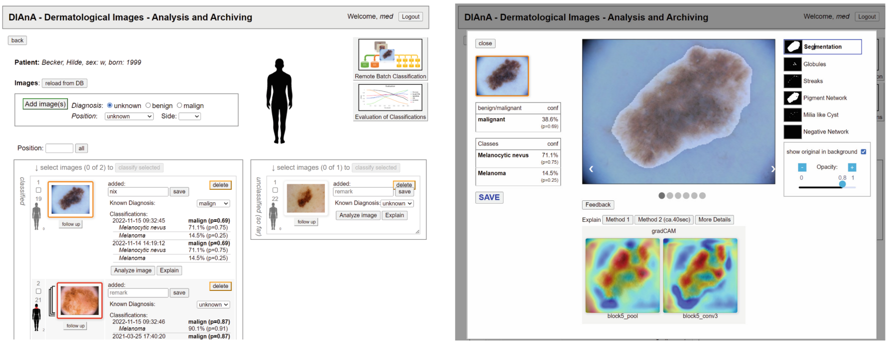



**Fig. 1.** Overview of the adapted DSR process.

We applied an archaeological perspective with an in-situ analysis inspired by Chandra Kruse et al. [12]. We approached the in-situ analysis by combining methods from human-computer interaction and usability research. In the in-situ analysis, we chose an archaeological perspective to take a human-centered approach and involve the relevant stakeholder groups (i.e., physicians and patients) [8, 9]. While we took the archaeological perspective, we analyzed dimensions like the aesthetic and symbolic one. Moreover, we considered intended and unintended effects, which XAI-based MDSS can trigger, like misleading decision support [5, 20]. The analyzed XAI-based MDSS, DIAnA [13], provides features like classifying skin lesions, generating XAI-based explanations, and managing patient-related data. DIAnA is currently available online via https://iml.dfki.de/demos/diana/index.html (2023/01/23). The following Fig. 2 depicts two exemplary screens of DIAnA.




(a) (b)

**Fig. 2.** Screenshots of DIAnA: (a) overview for a patient and (b) XAI-based diagnosis.

We started with think-aloud protocols since they allow users to communicate usability problems freely since we, as scholars, exert no influence on participants. Overall, we involved 12 participants in the in-situ analysis: two physicians, four physicians in training, and six patients. After the 12 interviews, we did not identify new insights, reached theoretical saturation, and did not conduct further interviews. The analysis started with a think-aloud part based on the cognitive task analysis [15]. The participants had to work through simple tasks to get familiar with the XAI-based MDSS while verbalizing their thoughts and actions during the interaction. After the think-aloud part, we conducted the semi-structured interviews guided by our TIME-inspired interview guide. Finally, we conducted follow-up semi-structured interviews to discuss concepts from TIME, focusing on UI features and their psychological effects on users [16; 18]. We followed an interview guide developed based on the before-mentioned aspects and dimensions of TIME, supplemented by questions focusing on the XAI features. The resulting data were analyzed using the thematic analysis of Braun and Clarke [29]. The codes were inductively derived using a descriptive coding method [30].

To formalize the DPs according to the status quo in DSR, we applied the scheme for specifying DPs for information technology-based artifacts in sociotechnical systems by Gregor et al. [17]. During the evaluation, we used the reusability dimensions and the associated template provided by Ivari et al. [19]. In the first and qualitative evaluation, we derived an interview guide based on the template. Four experienced software developers participated in the qualitative evaluation. For the second quantitative evaluation, we measured the reusability dimensions using a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). Within the quantitative evaluation, we recruited 66 experienced user interface and user experience designers through Prolific and applied statistical analysis using *R*.

# 4 The Development of Prescriptive Design Knowledge

## 4.1 The Derivation of Design Requirements Based on Empirical Insights

The qualitative data analysis laid out the foundation to derive the DRs. Table 1 provides an overview of the individual DRs and the associated codes, supplemented by exemplary quotes. We have named the DRs, according to the results from the interviews with physicians and physicians in training, as expert DRs (E-DRs). DRs that resulted from patient interviews were named patient DRs (P-DRs). We named hybrid DRs (H-DRs) for requirements that address both stakeholder groups. For traceability reasons, we added the thematic specification to the requirements as XAI-focused (H-XAI-DRs), AI-focused (H-AI-DRs), or a focus on MDSS (H-MDSS-DRs). We used the same logic for naming the codes so that expert codes are named E-XAI-C, the patient codes P-AI-C, and hybrid codes H-MDSS-C.

**Table 1.** Overview of the defined DRs and associated codes with exemplary quotes.

| Requirements | Codes and Exemplary Quotes |
|---|---|
| **E-AI-DR1**: If a MDSS is based on artificial intelligence, the system should provide the medical staff with information about the performance of the artificial intelligence in order to enable an evaluation of the system's recommendation | **E-AI-C1 System Information:** *"I might need more data on that, like what is the success rate of the AI, or how often has the AI been wrong."* (Interviewee 5) |
| **E-MDSS-DR2**: If the development of the patient-related object of analysis (e.g., melanoma) over time is critical to the diagnosis, the system should provide the possibility to compare historical patient data | **E-MDSS-C2 Diagnostic Timeline:** *"Then of course that would be great if you can compare old photos, see the analysis, let's say 2020 the system says 20% malignant, 2023 it says 60% malignant and 2026 it's going to scare the hell out of me, and we have to cut that out now."* (Interviewee 4) |
| **E-MDSS-DR3**: If multiple patients are managed by one MDSS, it should provide the medical staff with the ability to uniquely identify each patient | **E-MDSS-C3 Patient Identification:** *"The patient identification is done by the date of birth and the name, […]. Often there are several Becker Hilde which are born perhaps in 1990, therefore it would be important to display all information […]."* (Interviewee 5) |

(*continued*)

**Table 1.** (*continued*)

| Requirements | Codes and Exemplary Quotes |
|---|---|
| **E-MDSS-DR4**: If the medical staff feels the need to adjust the appearance and accessibility of task-related functionalities, the system should provide the individual with the ability to customize the interface to optimize the workflow | **E-MDSS-C4 Customization:** *"But that would be cool, of course, if you could really drag your most popular function right where you want it. Or you could set a favorite. That's a good way to save time."* (Interviewee 5) |
| **H-MDSS-DR5**: If diagnosis-relevant patient data is displayed (e.g., images or tabular data), the system should provide the user with the ability to customize the presentation of corresponding data (e.g., zooming in/out of an image) in order to enhance the cognitive processing of information | **H-MDSS-C5 Visual Adjustment***: "So some photos are very big, and some photos are very small probably it's the original file size I don't know. In some cases, I cannot see everything equally well."* (Interviewee 4) |
| **H-MDSS-DR6**: If the MDSS is used by individuals with different linguistic backgrounds (e.g., German, English), the system should be able to display user interfaces in different languages in order to increase its comprehensibility | **H-MDSS-C6 Language Barrier:** *"If my mother had to use the system, she doesn't know much about English. I think she would be overwhelmed. Maybe it would be cool to offer it in different languages."* (Interviewee 6) |
| **H-MDSS-DR7**: If the MDSS provides information within separate, modal windows and stakeholders are able to adjust these windows with regards to their size (e.g., to full screen), the user interface should be designed in a responsive way so that the presentation of data is automatically adjusted (e.g., enlargement of images) | **H-MDSS-C7 Automatic Scalability:** *"It's kind of difficult because you don't know which button to press here and what the buttons belong to, whether it's unknown benign or malignant."* (Interviewee 4) |
| **E-MDSS-DR8**: If the medical staff feels the need to get a second opinion, the system should provide the medical staff with the ability to communicate with other experts in order to enable a professional exchange | **E-MDSS-C8 Communication Between Experts:** *"[…] it is nice when you could contact a colleague for a second opinion […]"* (Interviewee 7) |
| **H-MDSS-DR9**: If the MDSS is used by multiple stakeholders (e.g., patients/layperson and professionals/experts), the system should provide the stakeholders with the ability to communicate with each other in order to enable a corresponding discourse | **H-MDSS-C9 Expert-Patient Communication:** *"I didn't think of that at first, but that would certainly be an interesting function as well. It would certainly be a good function for patients."* (Interviewee 9) |

(*continued*)

**Table 1.** (*continued*)

| Requirements | Codes and Exemplary Quotes |
|---|---|
| **H-MDSS-DR10**: If the MDSS offers options to communicate between multiple stakeholders (e.g., patients/laymen and professionals/experts), the system should enable the stakeholders to reference specific (patient) data for a precise and unambiguous communication | **H-MDSS-C10 Communication clarification:** *"If there are multiple lesions. Yes, that would be the most important to me that you can refer very specifically to one. Maybe like on WhatsApp, you can click specifically on a message and then answer the quote, so to speak."* (Interview 2) |
| **P-MDSS-DR11**: If the MDSS processes patient-related data (e.g., images or tabular data), the system should provide the patient with information on where the data is stored, how the data is processed and by whom or what the data is analyzed in order to increase the transparency of the data handling | **P-MDSS-C11 Information on data processing:** *"I think I might also need contextual information about whom I'm actually sending this to and what kind of person it is from which institute, so contextual information about who is processing it."* (Interviewee 2) |
| **H-MDSS-DR12**: If the MDSS presents information and functionalities on different subpages, the system should provide the user with an efficient and consistent navigation in order to increase the user's orientation when interacting with the system | **H-MDSS-C12 Navigational consistency***:* *"I've pressed the wrong button several times now when trying to close the window because I'm always looking for an X somewhere, but you have to press it [the window] again."* (Interviewee 4) |
| **H-MDSS-DR13**: If a MDSS uses separate, modal windows to present information, the system should provide the windows with a standardized graphical design (e.g., same buttons or icons and their locations as on the main page) in order to have a consistent presentation of information and interface functions | **H-MDSS-C13 Continuity of System Display:** *"It is questionable to me in several places how things are arranged."* (Interviewee 10) |
| **H-MDSS-DR14**: If the MDSS provides different kind of data (e.g., tabular and/or image data), its presentation should have a clear structure with easy to interpret labels so that users can quickly gain an overview of the content | **H-MDSS-C14 Over Information:** *"Partly, you are overwhelmed with information in the different subpages and functions which are perhaps not so important."* (Interviewee 9) |
| **H-MDSS-DR15**: If the MDSS is used by multiple stakeholders with varying levels of experience with such systems, the MDSS should provide an on-demand walkthrough on how to operate it in order to increase the stakeholders' independency and effectivity of their system usage | **H-MDSS-C15 Introduction to the System:** *"My feeling is […] somewhat overwhelming, so at first glance when you see so much data, what it could mean and why so many data is given, and everything is just on top of each other. […] Maybe I would be happy, and the feeling would be better if you were eased into it a little friendlier."* (Interviewee 2) |

(*continued*)

**Table 1.** (*continued*)

| Requirements | Codes and Exemplary Quotes |
|---|---|
| **H-MDSS-DR16**: If the MDSS is used by multiple stakeholders with varying levels of experience with such systems, it should provide the stakeholders with possibilities to inform themselves (e.g., help functions) about the provided functionalities so that the stakeholders can get familiar with the system | **H-MDSS-C16 Information on Demand:** *"A wiki, a FAQ section or something like that, […]"* (Interviewee 2) |
| **H-MDSS-DR17**: If the MDSS uses imaging techniques to support the diagnostic process, the system should enable the stakeholders to compare images from other cases so that they can expand their knowledge base by identifying relevant similarities or differences across the analyses of image data | **H-MDSS-C17 Comparison Options:** *"In addition to filtering options, it would, of course, be interesting to be able to compare cases with each other and different classifications, I have not seen this function so far."* (Interviewee 9) |
| **P-AI-DR18**: When the MDSS communicates the results of its analysis, the system should provide recommendations for action (e.g., to export the data and contact a specialist) so that the patients are assisted in planning the next steps | **P-AI-C18 Recommendation for Action:** *"Because the website doesn't currently tell me what I have to do now, i.e., a recommendation for action or something similar."* (Interviewee 6) |
| **H-AI-DR19**: If the MDSS provides automated, diagnostic recommendations to support the analytical process, it should integrate techniques to generate explanations for the given recommendations so that stakeholders can comprehend and evaluate the output, and to develop an appropriate level of trust into the system | **H-AI-C19 Explanation Necessity:** "*Of course, it is always difficult to rely on a program. But if you can now see why, it has decided this way, that's good. It is also possible that something was misinterpreted or that something strange was marked, which could be an error in the system. That's why I think it's important and makes sense."* (Interviewee 7) |
| **H-XAI-DR20**: If the MDSS provides explanation functionalities for the automated recommendations, the system should provide guidance on how to interpret the specific explanation so that stakeholders can develop an accurate mental model of the system's output | **H-XAI-C20 XAI Interpretation:** *"It confused me now because it was called Heatmap, and I thought it had something to do with the blood flow and temperature. But after an additional explanation, it was clear with what it correlates, and then I also found the XAI-tool clear."* (Interviewee 1) |

## 4.2 The Development of Theoretically Grounded Design Principles

The developed DRs build the foundation for the initial set of DPs, which we derive and describe in the following. Figure 3 provides an overview of the relationship between the derived DRs and DPs.

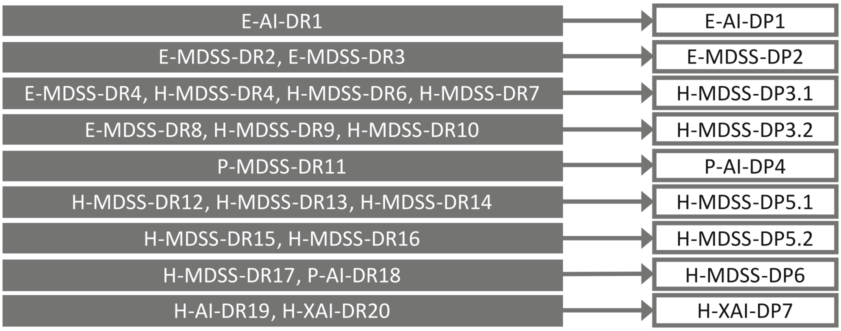



**Fig. 3.** Overview and summary of the relationship between DRs and DPs.

Transparency and explainability are important aspects of AI-based MDSS, which can influence users' acceptance [10]. AI performance metrics are essential information that can further increase transparency [9]. Providing competencies as information in a UI can lead to a positive perception of the underlying system and positively influence trust [18]. Consequently, we establish **E-AI-DP1:** *To provide an AI-MDSS, which clearly communicates its performance in the context of health-related diagnoses, the system should provide AI-related performance metrics as well as the probability of its classification in an easy-to-understand way, so that medical professionals without knowledge in computer science can appropriately interpret the system's recommendation.*

TIME proposes completeness and level of detail as essential quality criteria of information systems [16]. Since participants of the interviews communicated that a specific amount of information is necessary to identify patients and work effectively with the system, it should provide selected information for the user. Presenting relevant information will also influence the ease of use and the identification of critical information immediately [21, 22]. Furthermore, during the interviews, it was mentioned that time-related information like timespans is vital for a holistic overview. Consequently, we establish **E-MDSS-DP2**: *To provide an MDSS, which supports unique identification of patients and provides a trackable overview of their health status for medical professionals in the context of health-related diagnoses, the system should provide unique patient-related information as well as timestamps of previous diagnoses, so that the experts can quickly gain an overview of a specific patient's medical history.*

Prior research has proven that customizing systems to individual needs can strengthen information intake [31]. Customization can also enhance users' ability to manage tasks [16]. Moreover, customization can be valuable as it enhances self-efficacy beliefs and increases learning performance [32]. Consequently, we establish **H-MDSS-DP3.1**: *To provide an MDSS, which allows the customization of the user interface regarding the appearance, the interaction with information, and system level settings in the context of health-related diagnoses, the system should provide intuitive customization features, so that the system can be adjusted on a graphical and functional level, provide filter options, and options to interactively explore the information, so that all stakeholders can customize the system according to their cognitive needs and individual workflows.*

A valuable feature with beneficial aspects is the interactivity of exchange between users that use the same system [33]. Interactivity with other users positively affects users' perception of the content quality and the value of the information [16]. Participants of the interviews also communicated the desire to exchange with other users. For example, physicians may want a second opinion and a feature to reference specific information to avoid miscommunication. Consequently, we establish **H-MDSS-DP3.2**: *To provide an MDSS, which allows an unambiguous communication between medical experts or between medical experts with patients in the context of health-related diagnoses, the system should provide the possibility for professional exchange, which enables the stakeholders to reference specific data points of analyses, so that an effective discussion can take place.*

According to TIME, trustworthiness is an important aspect when users evaluate the credibility of systems [16]. Therefore, influencing and appropriately calibrating trust is another integral objective pursued by XAI [6]. During the interviews, we uncovered specific aspects that users would desire and could influence their credibility assessment and trust towards the AI-based MDSS, for example, regarding information for data storage or information processing methods used in the underlying system. Consequently, we establish **P-AI-DP4**: *To provide an AI-MDSS, which includes transparent information of the collected and analyzed data for patients in the context of health-related diagnoses, the system should provide a separate feature for patients to retrieve information regarding the storage, processing, and analysis of their personal data, so that they understand how their data is handled.*

To ensure the usability and efficiency of AI-MDSS, the UI of the system must be consistently designed [18; 33]. Designing functionalities and consistently navigating through a UI leads to usability enhancements, for example, by guiding users through the system [16]. We discussed these aspects also with both stakeholder groups during the interviews. Consequently, we establish **H-MDSS-DP5.1**: *To provide an MDSS, which includes an efficient navigational flow, concise labels, and descriptions in the context of health-related diagnoses, the system should maintain a consistent structure as well as the appearance of the graphical user interface elements, so that multiple stakeholders can navigate intuitively as well as effectively through the system resulting in a satisfying user experience.*

An easy way to navigate through systems and additional information that support the user in getting familiar with the system can influence the user's judgment regarding its credibility [16]. In TIME it is a helper heuristic. Both stakeholder groups also addressed these aspects during the interviews. Moreover, guiding users through actions can trigger the scaffolding heuristic, which can positively impact users' perception of systems [16]. We include both heuristics. Consequently, we establish **H-MDSS-DP5.2**: *To provide an MDSS, which includes assistance regarding the operation of the system's functionalities and information presented in the context of health-related diagnoses, a series of walkthroughs in combination with a help function to get detailed information about the system should be integrated, so that multiple stakeholders can independently and effectively work with the system.*

Clinicians prefer to learn case-based and the integration of practical cases. It involves comparing or recognizing essential characteristics, which is more beneficial for them than just being confronted with problems that need to be solved [10]. This aspect is also represented in TIME as self as source which describes that engaging users enable them to critically reflect on decisions and draw conclusions by themselves [16]. This authority heuristic is vital for patients since it requires recommendations and guidance for initiating appropriate actions. Especially if they lack the expertise to understand and analyze medical cases alone. Consequently, we establish **H-MDSS-DP6**: *To provide an MDSS, which extends the stakeholders' knowledge base regarding the specific diagnostic context, and enables an informed decision-making in the context of health-related diagnoses, the system should provide a functionality to compare and connect information across diagnostic cases as well as to recommend appropriate actions within the corresponding medical scope that a patient can initiate, so that the stakeholders can operate the system in a self-effective manner and initiate informed follow-up actions.*

Following TIME, trustworthiness and understandability are essential to perceiving a system's credibility [16]. Similarly, XAI can help to improve the credibility assessment, trustworthiness, and understandability of AI-based systems [6; 34; 35]. Furthermore, the role of explanations is also crucial for the acceptance of AI-based systems in clinical practice [10]. Moreover, we discussed these aspects during the interviews, and both stakeholder groups communicated the need for additional information to interpret the explanations correctly. Consequently, we establish **H-XAI-DP7**: *To provide an AI-MDSS, which includes explanations for its diagnostic recommendations and that enables an accurate interpretation of these explanations in the context of health-related diagnoses, the system should provide supplementary information on how to interpret the explanations, so that the stakeholders can develop an understanding as well as an appropriate level of trust towards the system.*

## 5 Evaluation and Results

As described, we have conducted two evaluation cycles with potential implementers focusing on the reusability of the proposed DPs. The four software developers in the first qualitative evaluation had experience in versatile domains, including web development, mobile app development, and machine learning. We started with a practitioner-oriented introduction to the concept of DPs and the type of system they are intended for (i.e., XAI-based MDSS). Subsequently, we worked through the different dimensions of the reusability evaluation. All participants perceived the proposed DPs as valuable and perceived them positively. We uncovered only little potential for optimization, which focused on the wording rather than the content. Two exemplary quotes from the qualitative evaluation to highlight the positive perception: *"Overall, the design principles are very comprehensible and also understandable."* (Interview 1); *"I can imagine that such design principles can have a positive impact on my productivity."* (Interview 2).

Afterward, we used the questionnaire template for a quantitative evaluation, which we implemented in a survey via Prolific. We measured each reusability dimension with a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). We recruited 75 participants with experience in either user interface design or user experience design. From these 75 participants, we excluded nine participants for speeding. We included the responses of the remaining 66 participants in the analysis. Overall, the positive perception of the qualitative evaluation episode was confirmed in the quantitative evaluation. The analysis yielded the following results: accessibility ($\alpha = 0.83$; $M = 4.9$; $SD = 1.2$); importance ($\alpha = 0.88$; $M = 4.5$; $SD = 1.4$); novelty and insightfulness ($\alpha = 0.64$; $M = 5.1$; $SD = 1.1$); actability and appropriate guidance ($\alpha = 0.82$; $M = 5.2$; $SD = 0.84$); effectiveness ($\alpha = 0.82$; $M = 5.2$; $SD = 0.97$). 77% of the 66 participants indicated they would adopt the DPs, and 82% would recommend them to a colleague for a suitable project.

# 6 Discussion

## 6.1 Summary of Findings

Through our DSR project, we provide in-depth insights into the perception and design of XAI-based MDSS and their UIs for skin-related diagnosis. Research characterizes this research area still as under-researched [9, 10]. Therefore, we aimed to take a human-centered approach to analyze DIAnA by involving relevant stakeholder groups using an archaeological approach from a user perspective [12]. In doing so, we considered the individual information needs of versatile stakeholder groups vital to achieving a high acceptance of XAI-based systems and AI predictions [7; 10; 14]. The in-situ analysis led us to 20 empirically grounded DRs for XAI-based MDSS for skin-related diagnosis. By applying concepts of TIME and using the scheme of Gregor et al. [17], we introduced nine DPs for guiding the design of XAI-based MDSS. Through our qualitative and quantitative reusability evaluation with experienced practitioners, we ensured that the proposed DPs are transferrable into practice, a desirable characteristic of DPs [19]. Therefore, practitioners and scholars can adopt suitable DPs for their use cases and particular application context. Following Gregor and Hevner's [36] DSR contribution framework, we consider our contribution an improvement. We used a human-centered archaeological approach to derive prescriptive design knowledge (i.e., design principles) to improve the human-centered design of solution artifacts (i.e., XAI-based MDSS) [36]. Our findings can therefore support the development of XAI-based MDSS and align with related research that highlighted the high relevance of human-centeredness when designing XAI-based systems [7; 9; 10; 35].

## 6.2 Limitations and Opportunities for Future Research

Like any other research project, ours is not without its limitations. At the same time, this also creates a range of opportunities for future research. For example, the derivation of design features, their instantiation in an MDSS, and their evaluation were not part of this project. Therefore, our future research aims to instantiate the theoretically grounded DPs and evaluate the resulting MDSS in a controlled laboratory or real-world setting. Other areas that were not part of our research are subjects like data protection or regulation of such classes of MDSS. Therefore, scholars from these disciplines could investigate how to protect the privacy of patient-related data or the regulation of such MDSS in the real world. We have also not focused on how to integrate the investigated MDSS in human-centered health ecosystems [1], which could be a flourishing area of research in the future. Consequently, many research opportunities arise around MDSS, especially with a multi-stakeholder perspective and human-centeredness.

## 7 Conclusion

In our DSR project, we investigated the design of XAI-based MDSS. These systems can be an influential driving force for future human-centered health ecosystems. During different stages of the DSR project, we involved multiple stakeholders, including physicians, physicians in training, patients, software developers, user interface, and user experience designers. We summarize the results of our project as empirically grounded DRs, which we address with a set of reusable DPs that potential implementers positively perceive. Consequently, we propose a set of reusable DPs, grounded in empirical insights and theoretical concepts of TIME.

**Acknowledgements.** This research is partly funded by the pAItient project (BMG, 2520DAT0P2).

## References

1. Bulc, V., Hart, B., Hannah, M., Hrovatin, B.: Society 5.0 and a human centred health care. In: Simini, F., Bertemes-Filho, P. (eds.) Medicine-Based Informatics and Engineering. LNB, pp. 147–177. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-87845-0_9
2. Rojas, C.N., Penafiel, G.A.A., Buitrago, D.F.L., Romero, C.A.T.: Society 5.0: a Japanese Concept for a super intelligent society. Sustainability **13**, 1–16 (2021)
3. Sahoo, P.R., Chatterjee, S.M.: Threats and challenges of artificial intelligence in healthcare industry. In: Zhang, Y.D., Senjyu, T., So-In, C., Joshi, A. (eds.) Smart Trends in Computing and Communications, LNNS, vol. 396, pp. 761–770. Springer, Singapore (2023)
4. Tschandl, P., et al.: Comparison for the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: an open, web-based, international, diagnostic study. Lancet oncology **20**(7), 938–947 (2019)
5. Tschandl, P., et al.: Human-coputer collaboration for skin cancer recognition **26**, 1229–1234 (2020)
6. Arrieta, A.B., et al.: Explainable Artificial Intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf. Fusion **58**, 82–115 (2020)
7. Adadi, A., Berrada, M.: Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access **6**, 52138–52160 (2018)
8. Payrovnaziri, S.N., et al.: Explainable artificial intelligence models using real-world electronic health record data: a systematic scoping review. J. Am. Med. Inform. Assoc. **27**(7), 1173–1185 (2020)
9. Schoonderwoerd, T.A.J., Jorritsma, W., Neerincx, M.A., van den Bosch, K.: Human-centered XAI: developing design patterns for explanations of clinical decision support systems. Int. J. Hum. Comput. Stud. **154**, 1–25 (2021)
10. Barda, A.J., Horvat, C.M., Hochheiser, H.: A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. BMC Med. Inform. Decis. Mak. **20**, 1–16 (2020)
11. Meske, C., Bunde, E.: Transparency and trust in human-AI-interaction: the role of model-agnostic explanations in computer vision-based decision support. In: Degen, H., Reinerman-Jones, L. (eds.) HCII 2020. LNCS, vol. 12217, pp. 54–69. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-50334-5_4

12. Chandra Kruse, L., Seidel, S., vom Brocke, J.: Design archaeology: generating design knowledge from real-world artifact design. In: Tulu, B., Djamasbi, S., Leroy, G. (eds.) DESRIST 2019. LNCS, vol. 11491, pp. 32–45. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-19504-5_3
13. Sonntag, D., Nunnari, F., Profitlich, H.-J.: The Skincare project, an interactive deep learning system for differential diagnosis of malignant skin lesions. DFKI Tech. Rep. **H2020**, 1–20 (2020)
14. Meske, C., Bunde, E., Schneider, J., Gersch, M.: Explainable artificial intelligence: objectives, stakeholders, and future research opportunities. Inf. Syst. Manag. **39**(1), 53–63 (2022)
15. Fernandez, A., Insfran, E., Abrahao, S.: Usability evaluation methods for the web: a systematic mapping study. Inf. Softw. Technol. **53**, 789–817 (2011)
16. Sundar, S.S., Jia, H., Waddell, T.F., Huang, Y.: The Handbook of Psychology of Communication Technology, 1st edn. Wiley, Hoboken (2015)
17. Gregor, S., Chandra Kruse, L., Seidel, S.: Research perspectives: the anatomy of a design principle. J. Assoc. Inf. Syst. **21**(6), 1622–1652 (2020)
18. Sundar, S.S.: Rise of machine agency: a framework for studying the psychology of Human-AI Interaction (HAII). J. Comput.-Mediat. Commun. **25**, 74–88 (2020)
19. Ivari, J., Hansen, M.R.P., Haj-Bolouri, A.: A proposal for minimum reusability evaluation of design principles. Eur. J. Inf. Syst. **30**(3), 286–303 (2021)
20. Dwivedi, A., Dwivedi, S.S., Tariq, M.R., Qiu, X., Hong, S., Xin, Y.: Scope of artificial intelligence in medicine. J. Res. Med. Dent. Sci. **8**(3), 137–140 (2020)
21. Westerbeek, L., et al.: Barriers and facilitators influencing medication-related CDSS acceptance according to clinicians: a systematic review. Int. J. Med. Inform. **152**, 1–14 (2021)
22. Westerbeek, L., de Bruijn, G.-J., van Weert, H.C., Abu-Hanna, A., Medlock, S., van Weert, J.C.M.: General practitioners' needs and wishes for clinical decision support systems: a focus group study. Int. J. Med. Inform. **168**, 1–7 (2022)
23. Fuji, K.T., Abbot, A.A., Galt, K.A., Drincic, A., Kraft, M., Kasha, T.: Standalone personal health records in the United States: meeting patient desires. Health Technol. **2**, 197–205 (2012)
24. Attfield, S.J., Adams, A., Blandford, A.: Patient information needs: pre- and post-consultation. Health Inform. J. **12**(2), 165–177 (2006)
25. Jefford, M., Tattersall, M.H.: Informing and involving cancer patients in their own care. Lancet Oncol. **3**(10), 629–637 (2002)
26. Hwang, A.H.-C., Oh, J.: Interacting with background music engages E-Customers more: the impact of interactive music on consumer perception and behavioral intention. J. Retail. Consum. Serv. **54**, 1–15 (2020)
27. Sun, Y., Sundar, S.S.: Exploring the effects of interactive dialogue in improving user control for explainable online symptom checkers. In: Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems, pp. 1–7. Association for Computing Machinery, New Orleans, LA, USA (2022)
28. Kuechler, W., Vaishnavi, V.: A framework for theory development in design science research: multiple perspectives. J. Assoc. Inf. Syst. **13**(6), 395–423 (2012)
29. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. **3**(2), 77–101 (2006)
30. Xu, W., Zammit, K.: Applying thematic analysis to education: a hybrid approach to interpreting data in practitioner research. Int. J. Qual. Methods **19**, 1–9 (2020)
31. Kreuter, M.W., Strecher, V.J., Glassman, B.: One size does not fit all: the case for tailoring print materials. Ann. Behav. Med. **21**, 276–283 (1999)
32. Ku, O., Hou, C.-C., Chen, S.Y.: Incorporating customization and personalization into game-based learning: a cognitive style perspective. Comput. Hum. Behav. **5**, 359–368 (2016)

33. Sundar, S. S., Oh, J., Bellur, S., Jia, H., Kim, H.-S.: Interactivity as self-expression: a field experiment with customization and blogging. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 395–404. Association for Computing Machinery, Austin, Texas, USA (2012)
34. Bunde, E., Kühl, N., Meske, C.: Fake or credible? towards designing services to support users' credibility assessment of news content. In: Proceedings of the 55th Hawaii International Conference on System Sciences, pp. 1883–1892 (2021)
35. Meske, C., Bunde, E.: Design principles for user interfaces in AI-based decision support systems: the case of explainable hate speech detection. Inf. Syst. Front. 1–31 (2022)
36. Gregor, S., Hevner, A.R.: Positioning and presenting design science research for maximum impact. MIS Q. **37**(2), 337–355 (2013)