Ref. Ares(2023)3777369 - 31/05/2023

Call: HORIZON-HLTH-2021-TOOL-06
Topic: HORIZON-HLTH-2021-TOOL-06-03
Funding Scheme: HORIZON Research and Innovation Actions (RIA)

Grant Agreement no: 101057062

AIDAVA

AI powered Data Curation & Publishing Virtual Assistant

Deliverable No. 1.2

Report from user survey with personas canvas

Approval by the European Commission Pending

Contractual Submission Date: 31/05/2023

Actual Submission Date: 31/05/2023

Responsible partner: P7- MUG

Funded by
the European Union

Grant agreement no.	101057062
Project full title	AIDAVA - AI powered Data Curation & Publishing Virtual Assistant

Deliverable number	**D1.2**
Deliverable title	**Report from user survey with personas canvas**
Type[1]	R
Dissemination level[2]	PU
Work package number	WP1
Work package leader	P8 - NEMC and P2 - b!loba
Author(s)	Carmen Zerner, Michaela Kargl, Monika Kindslehner (MUG) Kerli Norak, Mall Maasik (NEMC)
AIDAVA Reviewer(s)	Isabella Cina (EHN), Isabelle de Zegher (b!lo)
Keywords	User story, personas canvas, user survey

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HaDEA).

Neither the European Union nor the granting authority can be held responsible for them.

Document History

Version	Date	Description
V1	10.05.2023	Document ready for internal review
V2	22.05.2023	Document reviewed, commented and adapted
V3	31.05.2023	Final version

---

[1] **Type**: Use one of the following codes (in consistence with the Description of the Action):

R:	Document, report (excluding the periodic and final reports)
DEM:	Demonstrator, pilot, prototype, plan designs
DEC:	Websites, patents filing, press & media actions, videos, etc.

[2] **Dissemination level**: Use one of the following codes (in consistence with the Description of the Action)

PU:	Public, fully open, e.g. web
SEN:	Sensitive, limited under conditions of the Grant Agreement

List of Abbreviations and definitions

The abbreviations and definitions used in the deliverable are based on the AIDAVA Glossary[3].

Key definitions for this document:

Name	Definition
Information collection	Systematic gathering of data from various sources, such as surveys, interviews, observations, or existing databases. Information collection focuses on obtaining the necessary information required to address a research question, study, or analysis.
Data collection	Process of gathering or acquiring data from various sources for a specific purpose.
Data ingestion	Process of uploading or importing collected data into a data storage or processing system for further analysis or utilisation.
Data capture	Process of extracting information from any type of document or email and converting it into a format readable by a computer.

---

[3] https://www.aidava.eu/helpdesk/glossary

Table of contents

Executive summary

The AIDAVA project's Task 1.2 (T1.2) aimed to better understand the different user groups of the AIDAVA "AI-powered data curation and publishing virtual assistant" by involving two patient organisations, hospitals and health data intermediaries (HDIs). Also, T1.2 assessed patients' and citizens' interest and willingness to control and curate their personal health data.

To achieve this goal, the project team developed 8 personas based on 39 in-depth interviews, consisting of 2 patient personas, 2 data user personas, 2 data curator personas, and 2 third-party app developer personas. Foundation documents and persona canvases were created for each persona. Additionally, a survey was conducted with 250 participants to determine the general willingness of citizens to use a virtual assistant and what functionalities they would like the AIDAVA virtual assistant to have.

This information will be used to support the user-centred development of an AI-based data curation and publishing assistant. The personas will help developers empathise with different user groups, leading to more user-centric decisions. It will also serve as the foundation of the explainability and feedback layer for the user interface for patients - based on user profiles gathered when the user starts using the system for the first time. The personas will complement the business requirements specified in deliverable D1.3, and the default user profiles can be based on the main characteristics of the personas.

1 Introduction

This deliverable is the result of the work conducted during the first phase of Task 1.2 (user survey and definition of types of personas) in Work Package 1 of the AIDAVA project. Task 1.2 of the AIDAVA project is split into two phases: phase 1 runs from month 4 to month 8 and phase 2 runs from month 23 to month 24 in the project.

The objective of Task 1.2 was to enable broader participation across patient associations, hospitals and HDIs. The two main goals of phase 1 of Task 1.2 were:

- Creating user personas (i.e. archetypical user descriptions) for the AIDAVA "AI-powered data curation and publishing virtual assistant" as a basis for following human-centred design principles in the development of the AIDAVA prototype.
- Assessing - through a mix of literature survey, online survey, and direct interviews - patients/citizens interest and willingness to control and curate their data.

For the user survey, a web-based questionnaire and targeted interviews were used to better capture from the point of view of (potential) future AIDAVA users any important data curation & publishing aspects that may not have been included in existing literature and other projects. The aim was to enable the participation of a large number of patients, including members of patient associations and citizens working with HDIs. The survey also included clinical staff who currently maintain registries. During T1.2, for the different types of user groups of AIDAVA personas were created, which highlighted important factors related to the user's willingness to curate their data and which type of interaction and recommendations they expect to have with a virtual assistant.

1.1 Problem to be solved and aim of Task 1.2

To achieve user-centric design and development of the AIDAVA prototype, it is crucial to develop a clear understanding of the (potential) future users, their framework conditions, interests, needs and expectations.

The aim of Task 1.2 in the AIDAVA project is to enable broader participation across patient associations, hospitals and HDIs, to better understand these different user groups. Task 1.2 also covers the assessment of patients'/citizens' interest and willingness to control and curate their personal health data.

1.2 User groups identified for AIDAVA

For the AIDAVA prototype virtual assistant, four groups of potential users were identified:

1. Patients as the end users for AIDAVA who will curate their own data are one of the main user groups.
2. Expert data curators as hospital staff with medical knowledge/know-how in the healthcare field and with computer/data literacy who help patients to curate their data within AIDAVA; they will benefit from using AIDAVA in their everyday work if AIDAVA reduces their workload when curating patients´ data.

3. Data-users in medical or research facilities who use patients' data to do their everyday tasks in the field of medicine and/or research.
4. Third party app developers who develop applications which utilise AIDAVA's APIs (application programming interfaces)[4] either to provide health data to AIDAVA or to extract and reuse curated health data from AIDAVA.

---

[4] In the context of the prototype to be delivered during the project, all integration will be done through existing APIs and/or through asynchronous data transfer based on predefined data transfer specification to minimise disruption with productive systems. The full product is however expected to develop the relevant APIs.

2 Description of activities

To achieve the aims of Task 1.2, the first phase of T1.2 was dedicated to conducting in-depth interviews with representatives of the potential AIDAVA user groups, conducting an online survey, developing personas and deriving the content of user profiles. All these activities are described in detail in the following subsections.

2.1 In-depth interviews with potential AIDAVA users

The aim of the user interviews was to collect input from the previously described potential user groups of the AIDAVA prototype. The results of the interviews help to follow human-centred design principles and consider the users' needs, abilities, skills, constraints and preferences during the design and development of the AIDAVA system.

The interview results form the basis for the development of archetypical users, so-called personas. Personas is a method well-known in the human-computer interaction field [1]. It was introduced in this project to help designers and developers focus on the needs and goals of the target users throughout the product development process. These personas will form the basis for elaborating user profiles.

Specifically, this interview helped to better understand:

- Attitudes, goals, motivations, frustrations, challenges and pain points
- Aptitudes, competencies, knowledge, skills and experience
- Personal/work context and framework conditions
- Interest and willingness to control and curate health data
- Motivations to make health data accessible and available for research
- Vision of an “ideal” data curation and publishing tool

In order to conduct the interviews in the same way among all participants, guidelines were created for the interviewers to follow while preparing and conducting the interviews (see Annexes - Section 8.3). The guidelines included the selection criteria for interviewees, a reminder to collect a signed informed consent from the participants before starting (template also provided by the project team), a description of the aims of the interview and important notes to bear in mind while conducting the interviews. The guidelines also described how to consolidate and present the results.

The interviews were conducted in English or local language, depending on the interview partners. Each of the interviews took around 40-60 minutes and open-ended questions were used to collect answers to motivate the interview partners to tell their story. Whenever possible, 2 people conducted the interview where one asked the questions and interacted with the interviewee, while the other observed and took notes.
In total, 39 interviews were conducted: 4 with third party app developers, 9 with data curators, 12 with data users and 14 with patients.

The results of the interviews were transcribed and those that were not conducted in English were subsequently translated. All the results were anonymised and checked to ensure that there was no identifiable information. The interviewees were given their final transcript for review and approval before translating and uploading it to project files.

2.2 Online survey

The purpose of the survey was:

- to find out the approach of citizens regarding their personal health data;
- to discover the attitude and expectations towards future usage of an automatic data curation tool such as AIDAVA.

The survey was created in REDCap ("Research Electronic Data Capture")[5], a secure web application for building and managing online surveys and databases, and it was distributed by the project partners via an online link. The survey was targeted at citizens who are potential users of the AIDAVA virtual assistant in the future. It was mainly meant to confirm with a wider audience, the input we received from the 8 "patient consultants"[6] (4 cardiac patients from EHN and 4 breast cancers from ECPC) that helped us to define the patient perspective and requirements, during online calls and one face to face workshop in February 2023. The survey was shared therefore with the whole AIDAVA team, while the partners involved in T1.2 also shared it among their colleagues, family and friends. It was not intended to ask hospital patients in particular to fill in this survey nor to extend it broadly.

In order to cover all relevant topics, the questions were jointly developed by the partners in a shared document and then implemented in the survey in RedCap. For many of the questions, the response options consisted of selecting a numerical value from 0 (strongly disagree) to 10 (strongly agree). For example: The participants were asked to answer the question, "I am familiar with general medical terms." using the scale from 0 to 10 to indicate how much they agreed or disagreed with this statement.

Care was taken to minimise open-ended questions in the survey (as opposed to the interviews where open-ended questions were favoured) to prevent participants from sharing personal information or sensitive data in the open text fields. Therefore, the focus was placed on questions that could be answered using buttons, sliders or checkboxes. To keep the survey compact and clear, some questions were only displayed if certain answer options were selected. This was done using branching logic in RedCap.

Every question except the first question ("I agree that my answers to this questionnaire are processed and analysed as described above."), was defined as not mandatory.
Therefore, after agreeing that their answers can be processed and analysed, the participants were able to skip through the entire survey without filling out anything. This was done to motivate more people to participate in the survey, even if they do not want to answer specific questions.

Participants were not expected to spend more than 5-10 minutes to complete the survey.

---

[5] *REDCap* (Research Electronic Data Capture) https://www.project-redcap.org/

[6] Patient consultants are patients selected by the two patient organisations participating in the AIDAVA project (ECPC and EHN) - based on predefined "inclusion/exclusion criteria'' - who agreed to work throughout the project as consultant/ advisors representing the views from patients. These patients are not expected to share any personal data.

To make it more convenient for participants and to motivate more people to complete the survey, it was translated into several languages. The contents of the final version of the survey were translated from English into German, Dutch and Estonian with the help of a translation table (in the form of an .xls file). The translated content was implemented in RedCap using multi-language management.

The survey was open for one month. After this time, the results of the 250 participants were analysed. The results will be described in Section 3.1.

2.3 Development of personas

As already mentioned in Section 2.1, the *personas* method is well known in the field of human computer interaction (HCI) to support user-centred design. *Personas* are archetypes of users that make it easier for designers and developers to empathise with the users and to focus on the needs of the target user groups throughout the product development process.

For the development of personas, we followed the 5-steps approach suggested in *Personas for artificial intelligence (AI) [2]*.

Step 1 - identify user groups
Step 2 - collect information about users
Step 3 - consolidate and analyse information
Step 4 - create personas' foundation
Step 5 - visualise personas

In a first step - as described in Section 1.2 - four groups of people who would interact with a future AIDAVA system, were identified as potential future users.

In a second step, in-depth semi-structured interviews were conducted to collect information about these potential AIDAVA users, as described in detail in Section 2.1 of this document. The results of these interviews formed the basis for the development of personas.

2.3.1 Step 3. Consolidation and analysis of the interview results

The aim of step 3 in the development process of personas was threefold: a) to get an overview of the information collected, b) to distil the important findings from the collected information, and c) to decide which personas to develop.

To get an overview of the information collected, it is important to gather all relevant information in one place. To achieve this, all narrative texts resulting from the interviews were copied into a single spreadsheet for each user group and structured by the interview questions and topics.

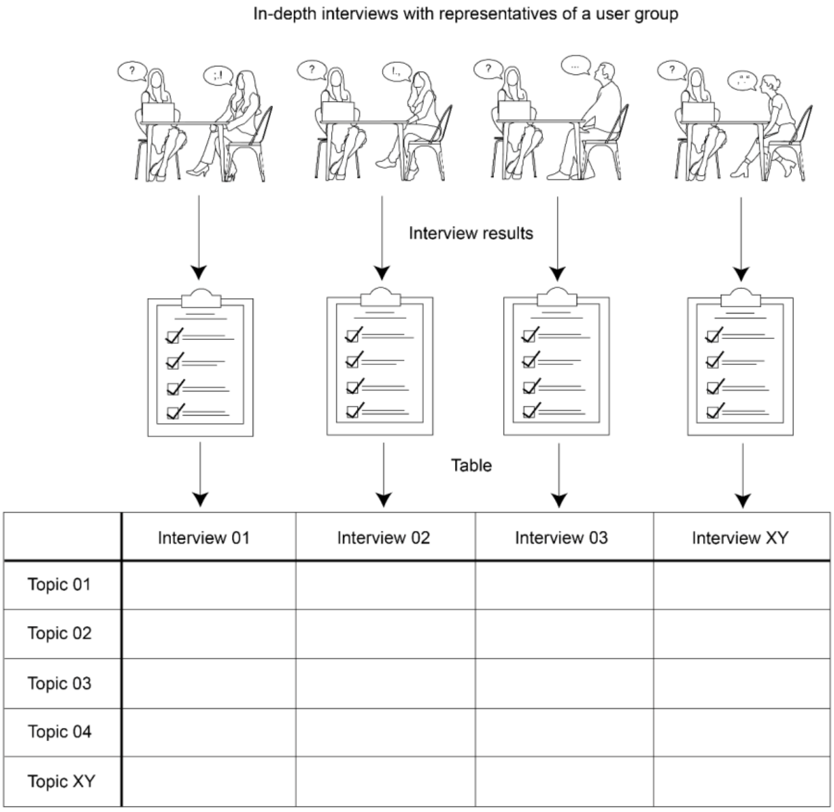
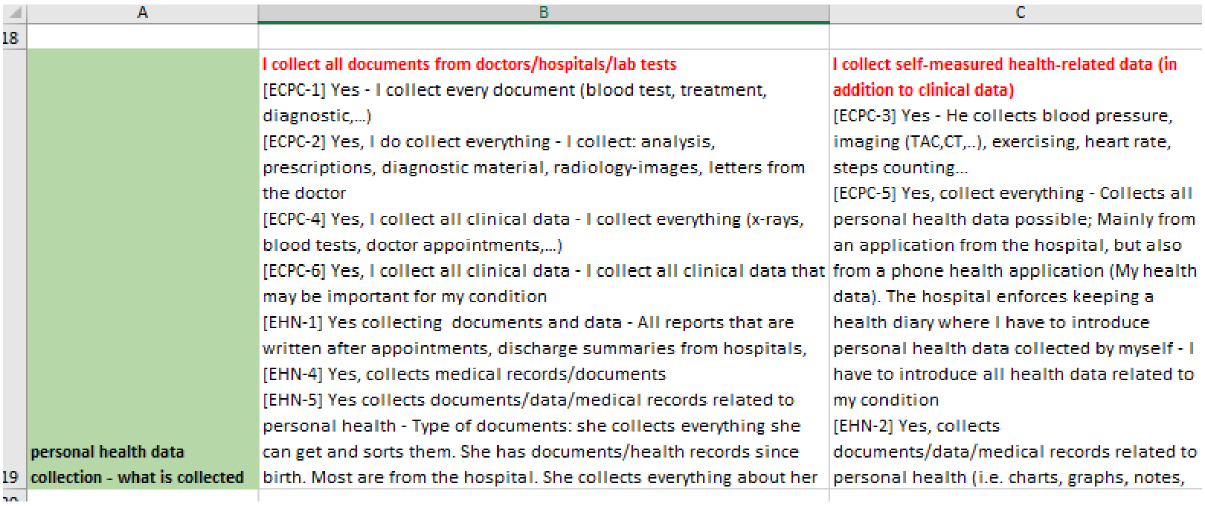


Figure 1. All interview results of a user-group were gathered in a spreadsheet structured by interview topics (source: MUG)

To distil from the collected information the important findings, i.e. those aspects which might be relevant for the usage of AIDAVA by the respective user group, we used the method of *affinity diagramming*. First, the narrative texts of the interview results were split into single aspects. Then, all these single pieces of information were grouped in clusters based on their content relationships. Finally, for each of these clusters, a summary was formulated that concisely described the aspect(s) contained in that specific cluster.



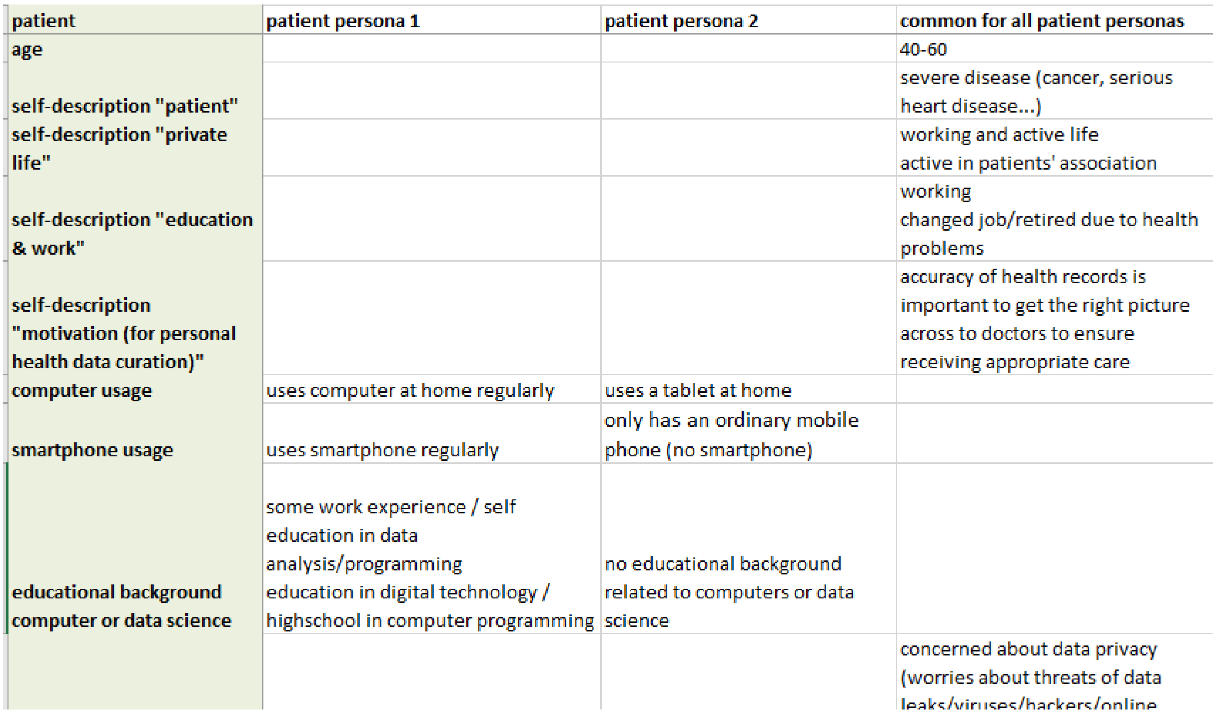
	A	B	C
18			
19	**personal health data collection - what is collected**	**I collect all documents from doctors/hospitals/lab tests** [ECPC-1] Yes - I collect every document (blood test, treatment, diagnostic,...) [ECPC-2] Yes, I do collect everything - I collect: analysis, prescriptions, diagnostic material, radiology-images, letters from the doctor [ECPC-4] Yes, I collect all clinical data - I collect everything (x-rays, blood tests, doctor appointments,...) [ECPC-6] Yes, I collect all clinical data - I collect all clinical data that may be important for my condition [EHN-1] Yes collecting documents and data - All reports that are written after appointments, discharge summaries from hospitals, [EHN-4] Yes, collects medical records/documents [EHN-5] Yes collects documents/data/medical records related to personal health - Type of documents: she collects everything she can get and sorts them. She has documents/health records since birth. Most are from the hospital. She collects everything about her	**I collect self-measured health-related data (in addition to clinical data)** [ECPC-3] Yes - He collects blood pressure, imaging (TAC,CT,..), exercising, heart rate, steps counting... [ECPC-5] Yes, collect everything - Collects all personal health data possible; Mainly from an application from the hospital, but also from a phone health application (My health data). The hospital enforces keeping a health diary where I have to introduce personal health data collected by myself - I have to introduce all health data related to my condition [EHN-2] Yes, collects documents/data/medical records related to personal health (i.e. charts, graphs, notes,

Figure 2. Digest of clustered interview results created by the method of 'affinity diagramming' with a concise summary of each cluster's content in red font

To decide which personas should be developed, a closer look was taken at the distilled findings to see if there were observable groupings across the interviewees' answers for some aspects. For each of those aspects, where it became apparent that the interviewees representing a user group clustered into subgroups, it was reflected whether that specific aspect might have an influence on the usage of the AIDAVA prototype. For all aspects, which were believed to probably have an influence on the usage of the AIDAVA prototype, it was taken care that the specific parameter values found in the subgroups of the interviewees are covered by personas.

2.3.2 Step 4. Creation of personas foundations

For each user group a personas' foundation document was created. The persona's foundation contains a complete list of all aspects derived from the interview results, which form the basis of a persona.



patient	**patient persona 1**	**patient persona 2**	**common for all patient personas**
age			40-60
self-description "patient"			severe disease (cancer, serious heart disease...)
self-description "private life"			working and active life active in patients' association
self-description "education & work"			working changed job/retired due to health problems
self-description "motivation (for personal health data curation)"			accuracy of health records is important to get the right picture across to doctors to ensure receiving appropriate care
computer usage	uses computer at home regularly	uses a tablet at home	
smartphone usage	uses smartphone regularly	only has an ordinary mobile phone (no smartphone)	
educational background computer or data science	some work experience / self education in data analysis/programming education in digital technology / highschool in computer programming	no educational background related to computers or data science	
			concerned about data privacy (worries about threats of data leaks/viruses/hackers/online

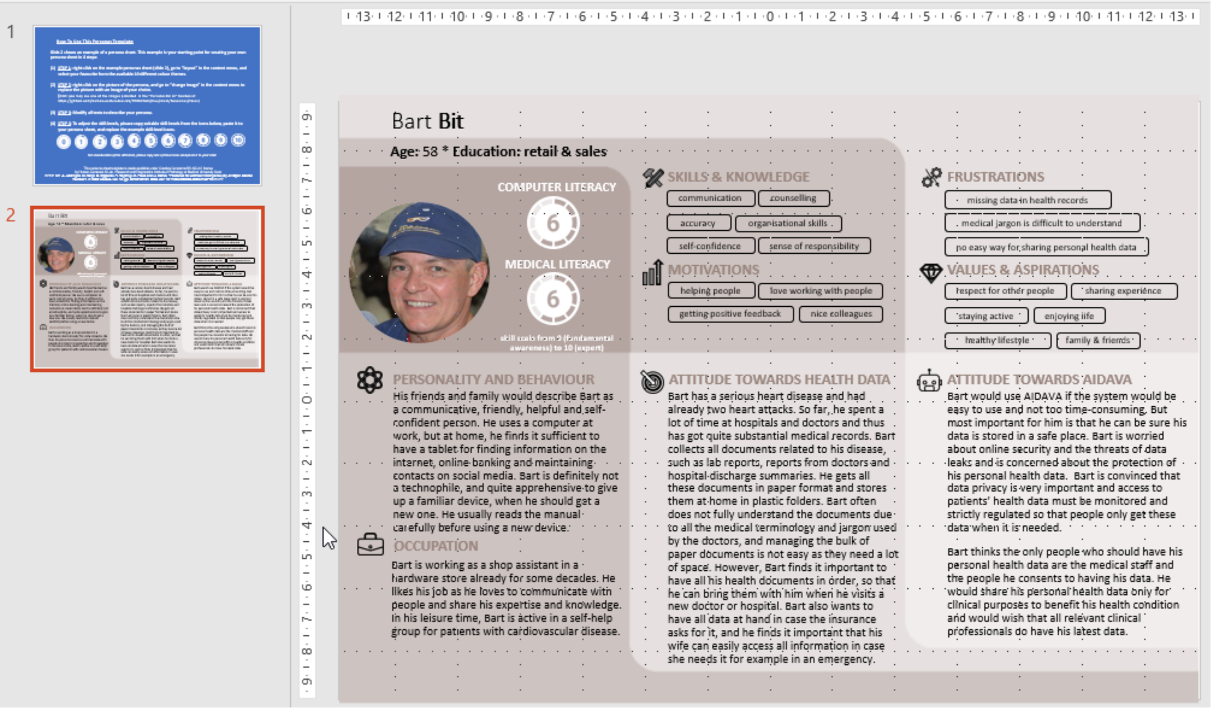
Figure 3. Detail from the patient personas foundations document

2.3.3 Step 5. Visualisation of the personas

To make it easier for people to empathise with the user represented by a persona, it is helpful to make the fictional persona a tangible realistic character. To achieve this, a persona is visualised in a 1-page layout, a so-called personas canvas, which includes a picture and name of the persona and a narrative text about the persona's interests, behavioural patterns and attitudes.

Most content of the personas canvas is based on the aspects included in the personas foundation. However, to make the persona more realistic, some fictional elements (e.g. a fictional picture, a fictional name, fictional hobbies, etc.) are added to the canvas. It is important that all these fictional elements are in line with the persona's characteristics as described in the persona's foundation so that the canvas conveys all key aspects from the personas foundation.

For the visualisation of the personas, we adapted a personas canvas template from *Personas for AI [3],* which is publicly available.



Bart Bit

Age: 58 * Education: retail & sales

COMPUTER LITERACY 6

MEDICAL LITERACY 6

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE

communication · counselling · accuracy · organisational skills · self-confidence · sense of responsibility

MOTIVATIONS

helping people · love working with people · getting positive feedback · nice colleagues

FRUSTRATIONS

missing data in health records · medical jargon is difficult to understand · no easy way for sharing personal health data

VALUES & ASPIRATIONS

respect for other people · sharing experience · staying active · enjoying life · healthy lifestyle · family & friends

PERSONALITY AND BEHAVIOUR

His friends and family would describe Bart as a communicative, friendly, helpful and self-confident person. He uses a computer at work, but at home, he finds it sufficient to have a tablet for finding information on the internet, online banking and maintaining contacts on social media. Bart is definitely not a technophile, and quite apprehensive to give up a familiar device, when he should get a new one. He usually reads the manual carefully before using a new device.

OCCUPATION

Bart is working as a shop assistant in a hardware store already for some decades. He likes his job as he loves to communicate with people and share his expertise and knowledge. In his leisure time, Bart is active in a self-help group for patients with cardiovascular disease.

ATTITUDE TOWARDS HEALTH DATA

Bart has a serious heart disease and had already two heart attacks. So far, he spent a lot of time at hospitals and doctors and thus has got quite substantial medical records. Bart collects all documents related to his disease, such as lab reports, reports from doctors and hospital discharge summaries. He gets all these documents in paper format and stores them at home in plastic folders. Bart often does not fully understand the documents due to all the medical terminology and jargon used by the doctors, and managing the bulk of paper documents is not easy as they need a lot of space. However, Bart finds it important to have all his health documents in order, so that he can bring them with him when he visits a new doctor or hospital. Bart also wants to have all data at hand in case the insurance asks for it, and he finds it important that his wife can easily access all information in case she needs it for example in an emergency.

ATTITUDE TOWARDS AIDAVA

Bart would use AIDAVA if the system would be easy to use and not too time-consuming. But most important for him is that he can be sure his data is stored in a safe place. Bart is worried about online security and the threats of data leaks and is concerned about the protection of his personal health data. Bart is convinced that data privacy is very important and access to patients' health data must be monitored and strictly regulated so that people only get these data when it is needed.

Bart thinks the only people who should have his personal health data are the medical staff and the people he consents to having his data. He would share his personal health data only for clinical purposes to benefit his health condition and would wish that all relevant clinical professionals do have his latest data.

Figure 5. Personas Canvas Template [3]

Validation of the personas

Visualisations of personas bear the risk of reinforcing stereotypes instead of providing a realistic picture of the users. Therefore, it is strongly recommended to validate the personas' canvases by asking people from the respective user group whether they feel plausibly represented by those personas. To obtain feedback from potential future AIDAVA users, we shared the personas canvas sheets to the respective interview partners with whom we conducted the in-depth interviews. We asked these people to tell us, (1) whether they think these personas are realistic and (2) whether in their point of view, any important aspects are missing in these personas.

The user feedback from this step was considered in the final versions of the AIDAVA personas. For example, we set the computer literacy of one of the patient personas to a lower level in response to the feedback from interview partners representing the user group of patients, who stated that one of the patient personas should be representative for people who are not very versed with computers.

The results of the whole process of personas development are described in Section 3.2.

4 Results

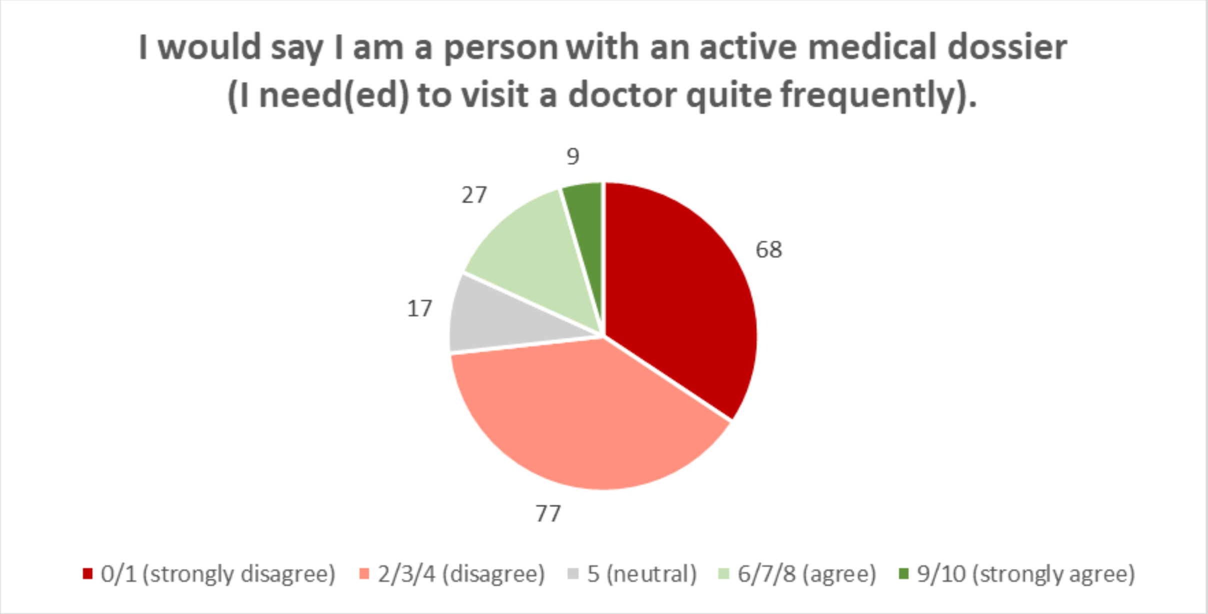
In the following subsections, we describe in detail the results of the activities in the first phase of the Task 1.2 of the AIDAVA project, covering first the survey results and afterwards, in Section 3.2, the personas as the results of the interviews.

4.1 Summary of the survey results

4.1.1 Result description

Every question in the online survey was non-mandatory to answer, except the first one, thus the amount of answers for each question might be different. From the 250 participants who filled out our survey at least partly, we obtained about 180-220 responses per question.

Of all the participants we reached, almost 50% (108 out of 218) were between 18 and 40 years old, while only 11% (25) were over the age of 60 years.

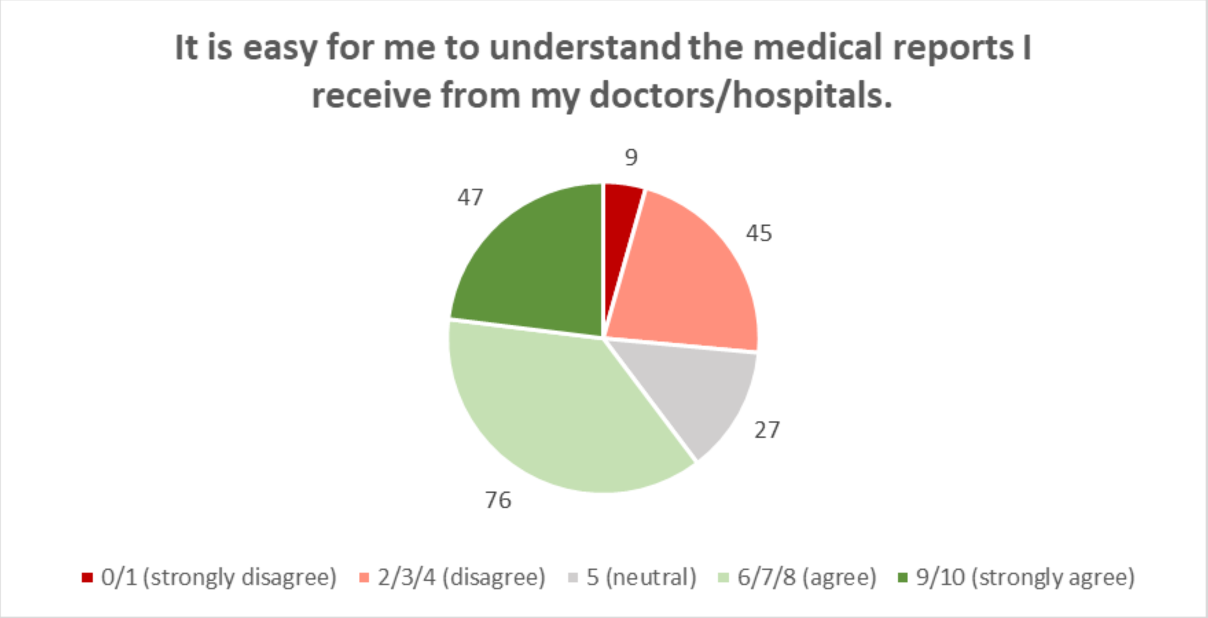


Results of the question "I would say I am a person with an active medical dossier (I need(ed) to visit a doctor quite frequently)."

- Explanation:
  - The scale to answer this question goes from 0 (strongly disagree) to 10 (strongly agree). For getting a better picture, we grouped them to 0/1 (strongly disagree), 2/3/4 (disagree), 5 (neutral), 6/7/8 (agree) and 9/10 (strongly agree).
  - A total number of 198 participants answered this question. The number in the pie chart indicates the number of participants that chose the respective answer option.
- Key message: 73% of the participants disagree or strongly disagree with the statement that they are a person with an active medical dossier; while 18% have an active dossier. We can consequently consider that the majority of the population that answered the survey are citizens in reasonable health.

Of the 216 participants that answered the question whether they have an educational background related to healthcare or medicine, 144 participants (66% of the total responses) do not have an educational background related to medicine or healthcare. This seems to indicate a bias in the surveyed population toward citizens better informed on health issues than the normal citizen population.

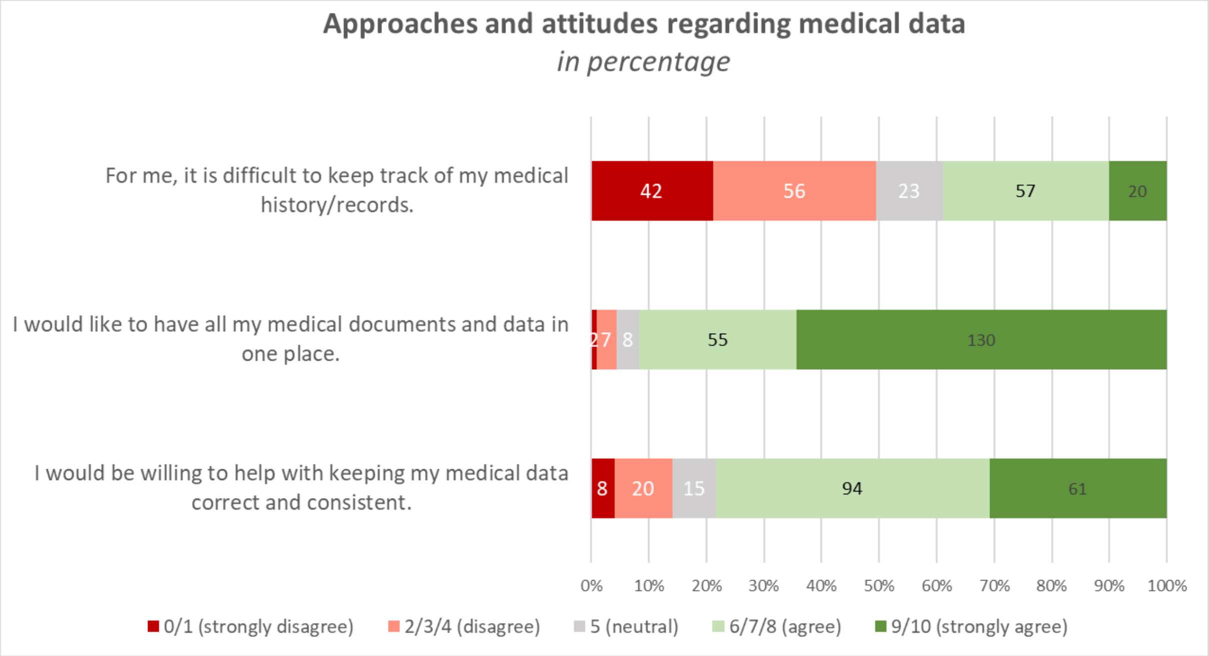
This bias is also supported by the following pie chart, which shows how easy the participants perceive it to understand reports they receive from their doctors / hospitals.



Results of the question "It is easy for me to understand the medical reports I receive from my doctors/hospitals"

- Explanation:
    - The scale to answer this question goes from 0 (strongly disagree) to 10 (strongly agree). For getting a better picture, we grouped them to 0/1 (strongly disagree), 2/3/4 (disagree), 5 (neutral), 6/7/8 (agree) and 9/10 (strongly agree).
    - A total number of 204 participants answered this question. The number in the pie chart indicates the number of participants that chose the respective answer option.
- Key message: The majority (60%) of the participants agrees or strongly agrees that they find it easy to understand the medical reports they receive from their doctors or hospitals. This seems to confirm the bias in the population who answered the survey toward citizens with some health literacy.

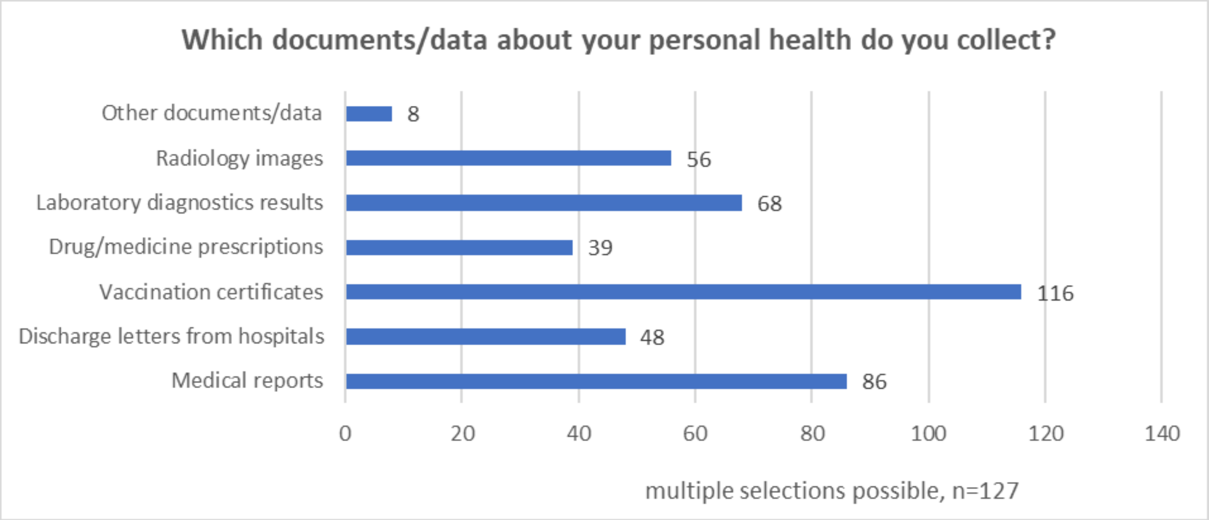
Additionally, below the results of three more important statements are shown. While the participants don't seem to agree whether it is difficult to keep track of their medical history and reports or not, they very clearly agree on wanting to have all their medical documents and data in one place. Furthermore, there is a clear willingness to help in keeping their medical data correct and consistent. Especially the results of the last question draw a very favourable picture for the usage of AIDAVA in the future.



Results of specific statements in "Approaches and attitudes regarding medical data", in percentage

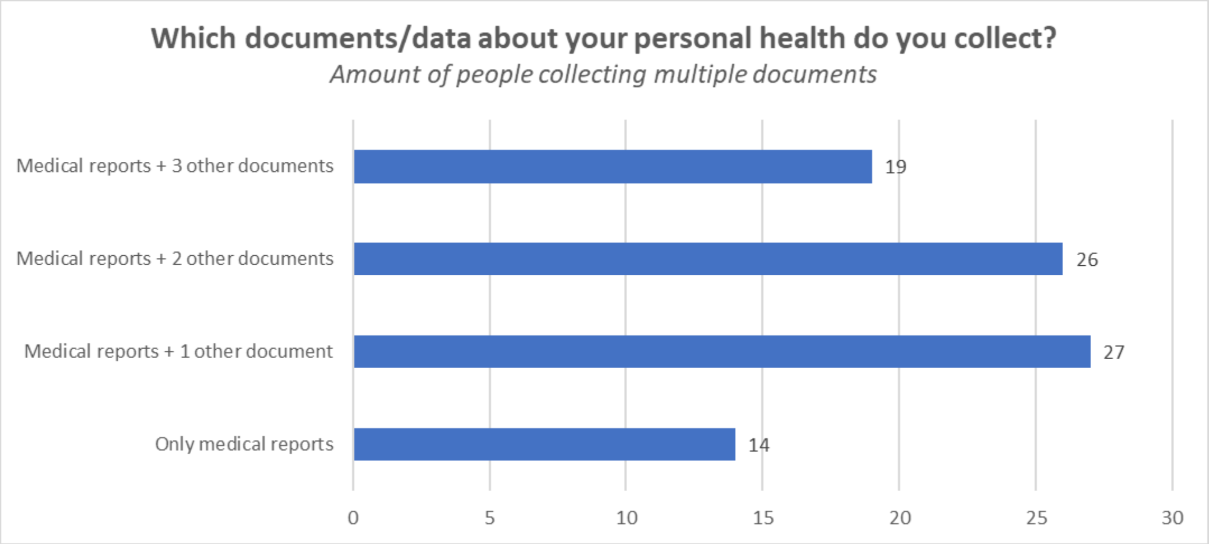
- Explanation:
  - The scale to answer these questions goes from 0 (strongly disagree) to 10 (strongly agree). For getting a better picture, we grouped them to 0/1 (strongly disagree), 2/3/4 (disagree), 5 (neutral), 6/7/8 (agree) and 9/10 (strongly agree).
  - A total number of 198 participants answered the first and the third question. A total number of 202 participants answered the second question. The numbers in the bar diagram indicate the number of participants that chose the respective answer options.
- Key message: No clear conclusion can be made on whether the participants as a whole find it difficult to keep track of their medical history or records. Meanwhile, over 91% of the participants agree or strongly agree that they would like to have all their medical documents or data in one place, 4,4% of the participants disagree or strongly disagree on that. More than 78% of the participants agree or strongly agree that they would be willing to help to keep their medical data correct and consistent.

Of all 203 participants that answered the question whether they collect documents/data about their personal health, 127 participants answered with yes. Currently, of those 127 people, the participants mainly collect vaccination certificates and medical reports, and over half of the participants also collect laboratory diagnostic results and radiology images. The high amount of vaccination certificates collected by respondents could be related to the Covid-19 vaccination policy in their country of residence.



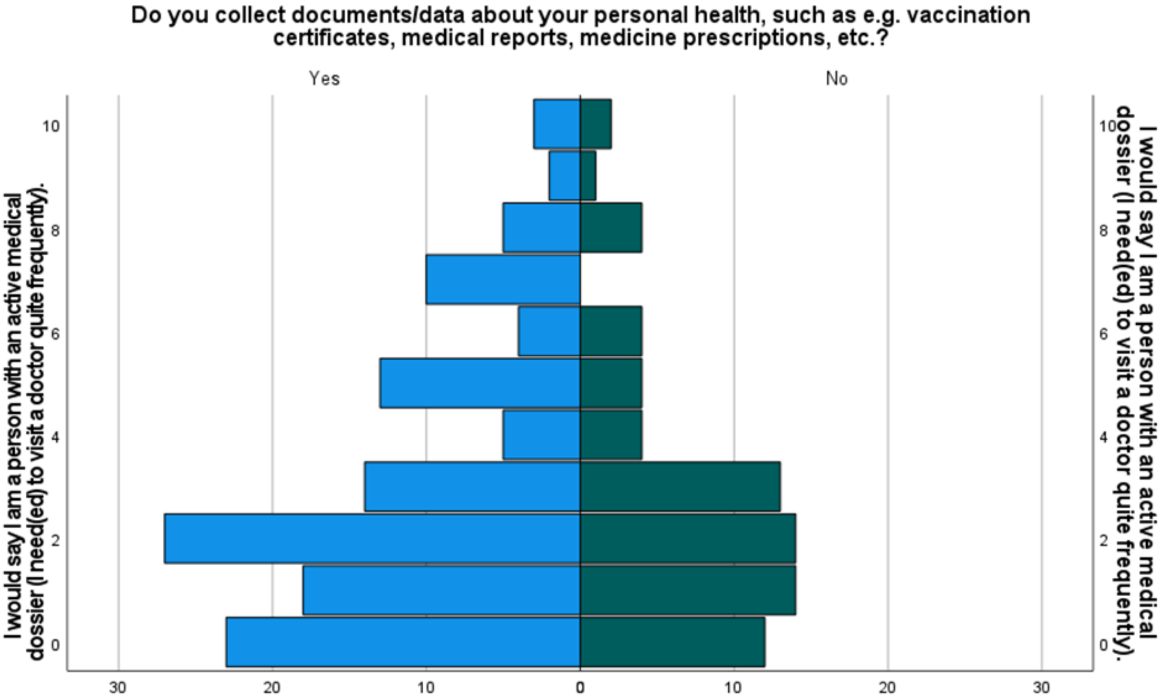
Results of the question "Which documents/data about your personal health do you collect?"

The bar chart below shows how many participants collect more than just one document, based on the medical reports that were most frequently collected. For the purpose of removing the Covid-19 bias, neither the vaccination certificates are considered in the following bar chart, nor are the radiology images.



Extracted number of participants collecting multiple documents

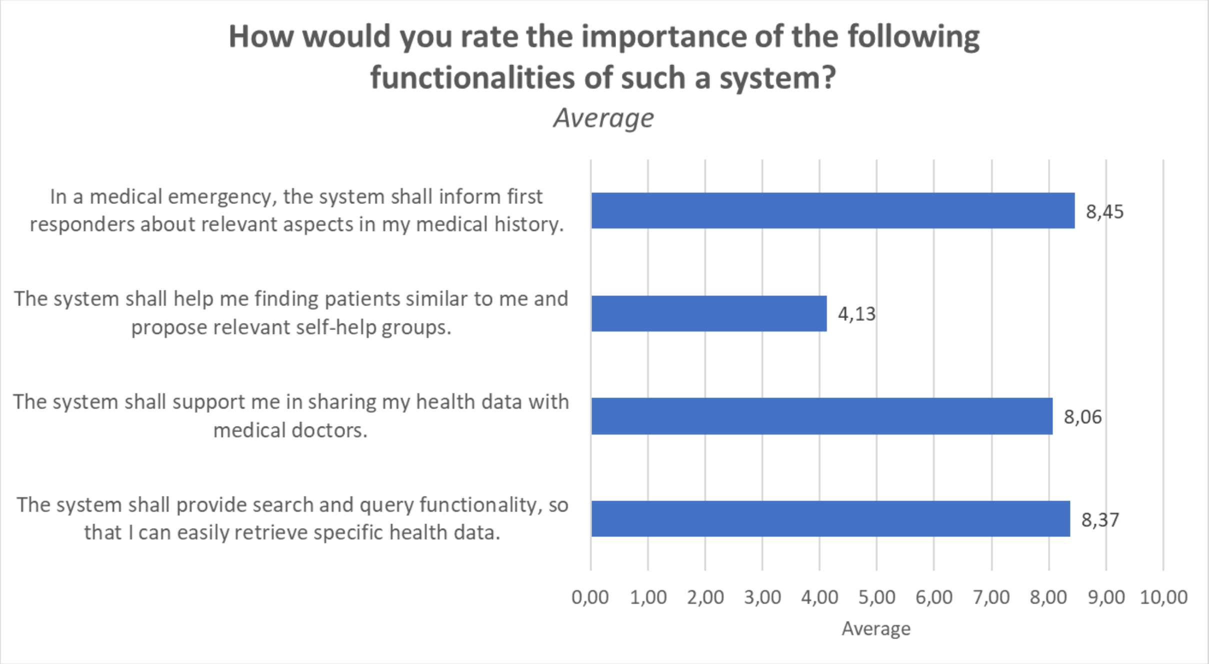
Despite this, having an active medical dossier does not seem to influence the decision of the participants to collect data about their personal health or not.



Comparison of the participants who collect / don't collect their health data with how the participants rated themselves in terms of having an active medical dossier

- Explanation
    - Scale of vertical axis (medical dossier): 0 = strongly disagree, 10 = strongly agree
    - Scale of horizontal axis: number of participants selecting the respective answer
- Key message: The answer on the question whether participants collect documents or data about their personal health does not show any significant correlation (p=0,107) to their answer on having an active medical dossier. Participants, who have an active medical dossier are not more likely to collect medical data than participants, who do not have an active medical dossier.

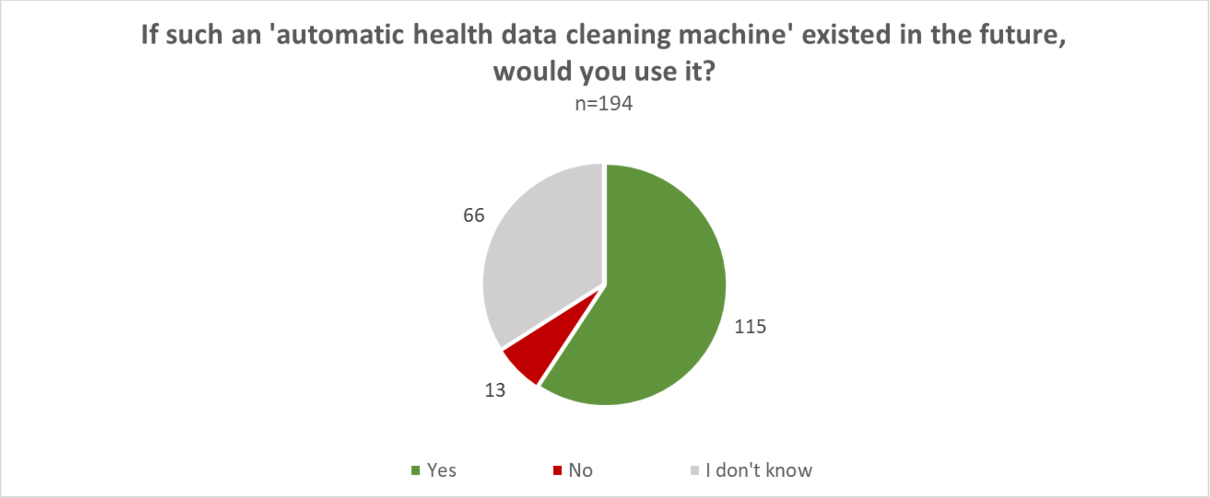
The participants were also asked that if there would be an "intelligent health data system" in the future, how would they rate the importance of the listed functionalities of such a system? The clearest answers are summarised as follows.



Average results of specific statements in "How would you rate the importance of the following functionalities of such a system?"

- Explanation
    - Scale: 0 = strongly disagree, 10 = strongly agree
- Key message: The participants are most interested in having the system inform first responders (i.e. treating physicians) in a medical emergency about relevant aspects of their medical history, provide search and query function and support them with sharing their health data with medical doctors. On the other hand, most participants do not find it important for the system to help them find patients similar to them and propose relevant self-help groups. This corresponds to the feedback we received from the patient consultants - supporting definition of patient related requirements for AIDAVA - that the first reason for them to use a tool like AIDAVA would be to ensure their treating physician has access to the correct information - anytime, anywhere.

When participants were asked if they would use an “automatic health data cleaning machine” with the functionality of AIDAVA, 115 out of 194 (59%) selected “Yes”, and only 13 selected “No”.



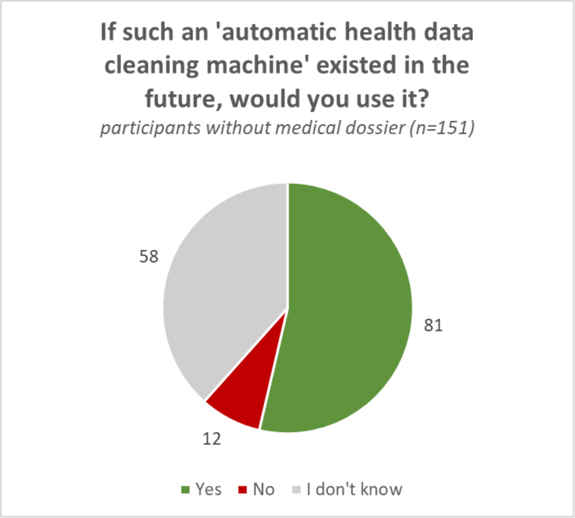
Results of the question “If such an ‘automatic health data cleaning machine’ existed in the future, would you use it?”

As 59% does not allow to make a valid conclusion, we decided to split the participants in two groups:

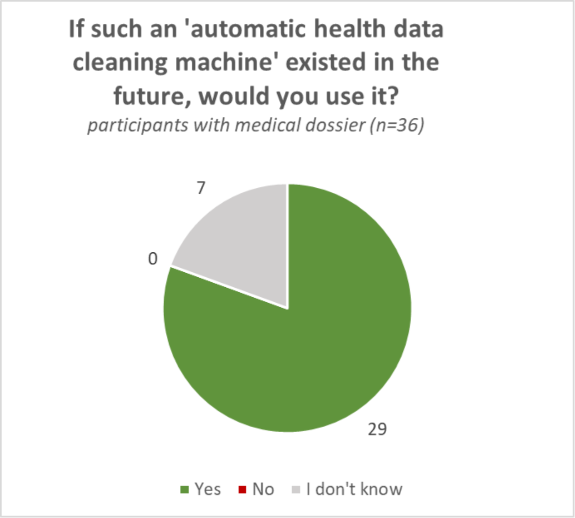
- Those who disagree or strongly disagree with having an active medical dossier, further summarised as “WITHOUT medical dossier” (73%) in pie chart “I would say I am a person with an active medical dossier”
- Those who agree or strongly agree having an active medical dossier, further summarised as “WITH medical dossier”

The participants (n=7) who filled out the question on whether they would use a tool like AIDAVA, but did not answer the question of the active medical dossier, are not included in the following pie charts.

The percentage of participants who would agree to use a tool like AIDAVA is strikingly different between the 2 populations: 54% for the people WITHOUT medical dossier; 81% for the people WITH medical dossier (and therefore people with active medical conditions). This nicely demonstrates that people with a real medical problem would be more interested in a tool like AIDAVA.

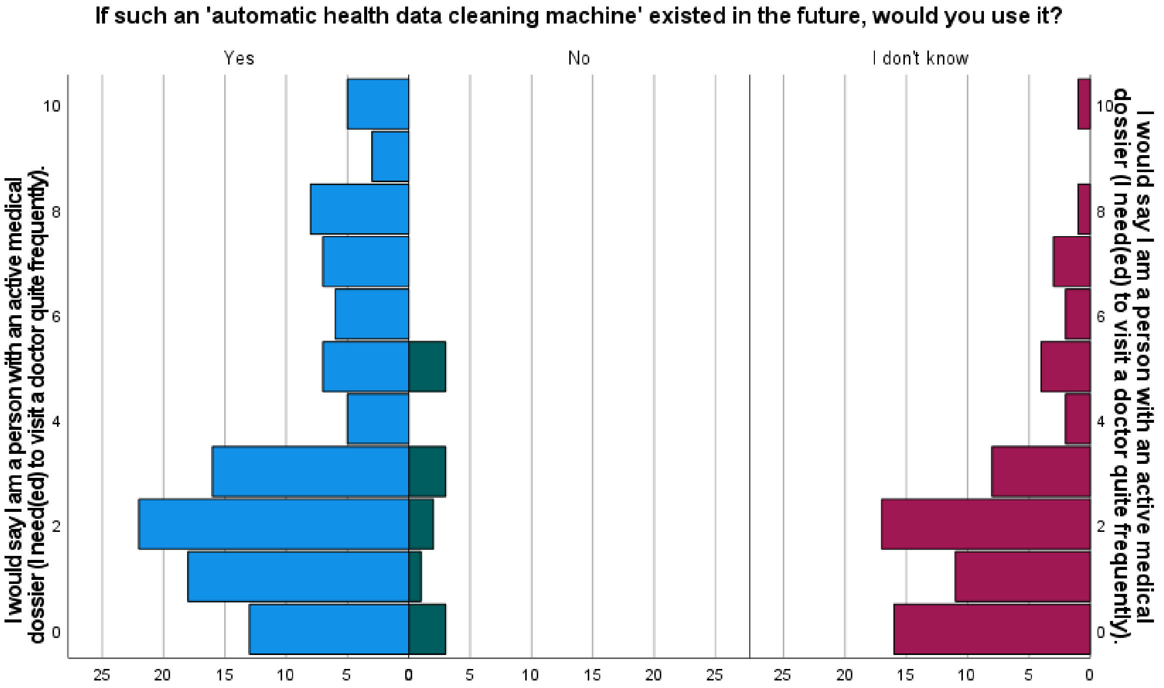


Participants (n=151) WITHOUT active medical dossier who agreed to potentially use AIDAVA.



*Participants (n=36) **WITH** active medical dossiers who agreed to potentially use AIDAVA.*

The same pattern can also be seen in the following distribution of answers below, again comparing how the participants having an active medical dossier would use a tool like AIDAVA. It can be seen that the participants who consider themselves persons with an active medical dossier tend to be more open to using AIDAVA in the future.



Comparison of the participants who would use an "automatic health data cleaning machine" with how the participants rated themselves in terms of having an active medical dossier.

4.1.2 Conclusions

The population who answered the survey seems to be more representative of healthy citizens - with some level of health literacy - than patients with an active disorder (and medical dossier). The majority of the survey respondents is interested in keeping their dossier in a central place (91%), and to increase the quality of this dossier (78%) - independently of the fact that they have an active medical dossier or not. The main reason for doing so is to ensure the treating physicians have access to the appropriate information, certainly in emergency situations. These findings align perfectly with the input we received from the 8 patient consultants.

For AIDAVA, it is interesting to mention that currently, 59% of the survey respondents would use a tool like AIDAVA to support them. When only considering the participants having an active medical dossier, this number increases to 81%. It indicates that the participants are more concerned about risks of such a system like security and reliability, when they are not active patients, therefore not having the need for a tool like AIDAVA. This means that AIDAVA would profit from ensuring high security and reliability of the system, and also from offering user training covering both data literacy as well as the risks and benefits of a system like AIDAVA.

4.2 Personas

In total, eight personas were developed across the four groups of people, who were identified as potential future AIDAVA users in Section 1.3. In the following subsections, the persona's foundation as well as the persona's visualisation is presented for each of these eight personas.

3.2.1. Patient personas

Foundation for patient personas

All aspects mentioned in this personas foundations table are based on the interview results.

The aspects written in red font were used for the creation of the respective persona's canvas;

	patient persona 1	**patient persona 2**	**common for all patient personas**
age			40-60
health condition			severe disease (cancer, serious heart disease...);
private life			active life (hobbies…) active in patients' association
education & work			working changed job/retired due to health problems
motivation for personal health data curation			accuracy of health records is important to get the right picture across to doctors to ensure receiving appropriate care
computer usage	uses computer at home regularly proficient computer user	uses a tablet at home	
smartphone usage	uses smartphone regularly	only has an ordinary mobile phone (no smartphone)	
educational background computer or data science	some education in digital technology (school/work experience or self-education)	no educational background related to computers or data science	
data rights and privacy aspects		data privacy is important, but having to give consent every time when you go to the doctor is cumbersome; the only people who should have their data are the medical staff and the people they consent to having their data	concerned about data privacy (worries about threats of data leaks/viruses/hackers/online insecurity) and personal right to access own data; like to be the owner of their data and to know who holds what data about them/their family and what is happening with their data; do not want their data to be used commercially without their knowledge and consent
reading manual for new device	usually dives right into when using a new device (reads manual only when encountering an issue -	usually reads the manual carefully before using a new device	

	patient persona 1	patient persona 2	common for all patient personas
	prefers visual instruction cause reading is boring)		
feelings about getting a new device	finds changing to new devices exciting, loves new gadgets and likes exploring them	apprehensive to give up familiar device, but then happy to have new functions and possibilities	
level of specific medical knowledge regarding own disease	8-10 (I google and read everything about my condition to better understand it)	6-7 (I do not understand everything but have a good baseline knowledge; I know more than general population)	
level of familiarity with general medical terms	8-9 (I know many but not all medical terms)	5-7 (mostly familiar with terms regarding my own personal health condition)	
educational background medicine / healthcare	educational background in medicine/healthcare	no educational background in medicine or healthcare	
personal health data collection - what is collected	collects in addition to clinical data also self-measured personal health data *(blood pressure, physical activity, heart rate, steps counting, weight…* - from hospital application & phone health-application)	collect all medical documents related to their health condition *(lab test results, treatment, diagnostic material, prescriptions, letters from doctors, imaging (radiology images, CT,...), discharge summaries from hospitals,...)*	
format of collected data	most health data are digital (pdf), but some on paper	receives only paper documents (and medical images on CD)	
how is the collected data stored	stores digital health documents on computer at home; stores paper documents in folder/box and keeps scanned copy on computer self-measured personal health-related data is kept in health app	stores paper documents in folder/box at home	
motivation for data collection		I collect everything in paper format, because I can read it and I like to insert comments on the side of the document on paper	I collect the health data for myself (to better understand/monitor my condition...); I collect the health data to share it (with doctors, insurance...)

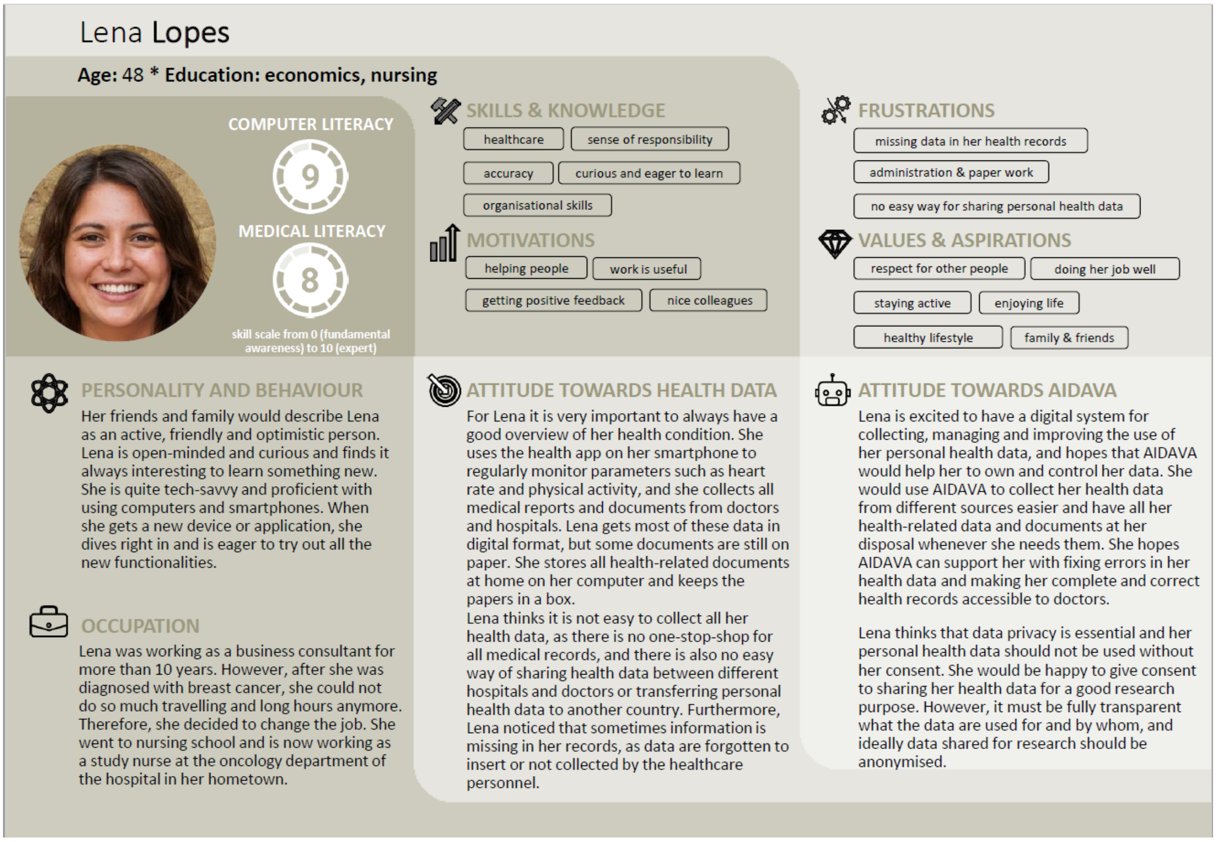
	patient persona 1	patient persona 2	common for all patient personas
challenges and issues regarding health data collection		paper documents need a lot of space to store, are difficult to manage, and not environmentally friendly; I do not understand the reports I get because of the medical jargon/terminology	I cannot see images/MRI scans at home since I do not have the necessary software to read these files; health information is often not complete (parts of information are forgotten to insert/upload and health information coming from other sources is not collected by the health personnel); there is no easy way for sharing personal health data from different locations / countries with each other; health data retrieval is not easy (*no one-stop-shop for all medical records, difficult to retrieve data from the past, complicated to receive doctor's notes, lengthy and tedious formal consent procedures needed to get own health data from hospital, difficult to know whom to ask for the documents...*)
monitoring health-related parameters at home	regularly monitoring health-related parameters *(e.g. blood pressure, heart rate, physical activity, oxygen saturation, weight)* at home, but sharing with doctor (print documents) only if there is an issue; regularly monitoring health-related parameters at home and sharing data with doctor through app; measuring health-related parameters at home only when feeling not ok and sharing data (as pdf) with doctors only if something alarming;	measuring no health-related parameters at home	
using health/fitness app or device	I use a health app / fitness tracker device	I do not use a fitness-tracker app or device	

	patient persona 1	patient persona 2	common for all patient personas
sharing personal health data	would consent to share data for good research purpose; must be fully transparent for which purpose data are used - would consent on a case-by-case basis; no commercial use; data shared for research should be anonymised; there must be consent and strict regulation; people should not be overly cautious and understand how helpful data can be;	would share data only for clinical purposes if benefit for my health - all relevant clinical staff should have my latest data, but it has to be monitored and regulated strictly so that people only access it when they need it	
Reasons for AIDAVA usage - expectations towards AIDAVA	I would use AIDAVA to own and control my own health data; I would use AIDAVA cause I am excited to have a digital system for collecting, managing and improving the use of my health data	I would use AIDAVA if it is easy to use and if I can be sure that my data is stored in a safe place	I would use AIDAVA... ...to collect my health data from different sources easier and faster ...to have all my health data (from different sources) with me all the time ...to have a clear picture of my health ...to make my complete, correct health data accessible to doctors ...to curate my health data (fix errors, remove unnecessary data) ...since I hope that the system takes care of my health

Additional aspects distilled from interview results, which are not covered by the current patient personas:

- user is not the patient, but a relative (e.g. parent of a child with severe (congenital) disease)
- all medical records are online accessible in electronic health system (national health system or local health maintenance organisation / health data intermediary), and thus there is no need to collect medical documents personally; problems mentioned by the interviewees: nobody else than me can access my medical records in my local health data account; there is no line of communication/access between hospitals and the health maintenance organisation / health data intermediary

Visualisations of the patient personas



Lena Lopes

Age: 48 * Education: economics, nursing

COMPUTER LITERACY 9

MEDICAL LITERACY 8

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE

- healthcare
- sense of responsibility
- accuracy
- curious and eager to learn
- organisational skills

MOTIVATIONS

- helping people
- work is useful
- getting positive feedback
- nice colleagues

FRUSTRATIONS

- missing data in her health records
- administration & paper work
- no easy way for sharing personal health data

VALUES & ASPIRATIONS

- respect for other people
- doing her job well
- staying active
- enjoying life
- healthy lifestyle
- family & friends

PERSONALITY AND BEHAVIOUR

Her friends and family would describe Lena as an active, friendly and optimistic person. Lena is open-minded and curious and finds it always interesting to learn something new. She is quite tech-savvy and proficient with using computers and smartphones. When she gets a new device or application, she dives right in and is eager to try out all the new functionalities.

OCCUPATION

Lena was working as a business consultant for more than 10 years. However, after she was diagnosed with breast cancer, she could not do so much travelling and long hours anymore. Therefore, she decided to change the job. She went to nursing school and is now working as a study nurse at the oncology department of the hospital in her hometown.

ATTITUDE TOWARDS HEALTH DATA

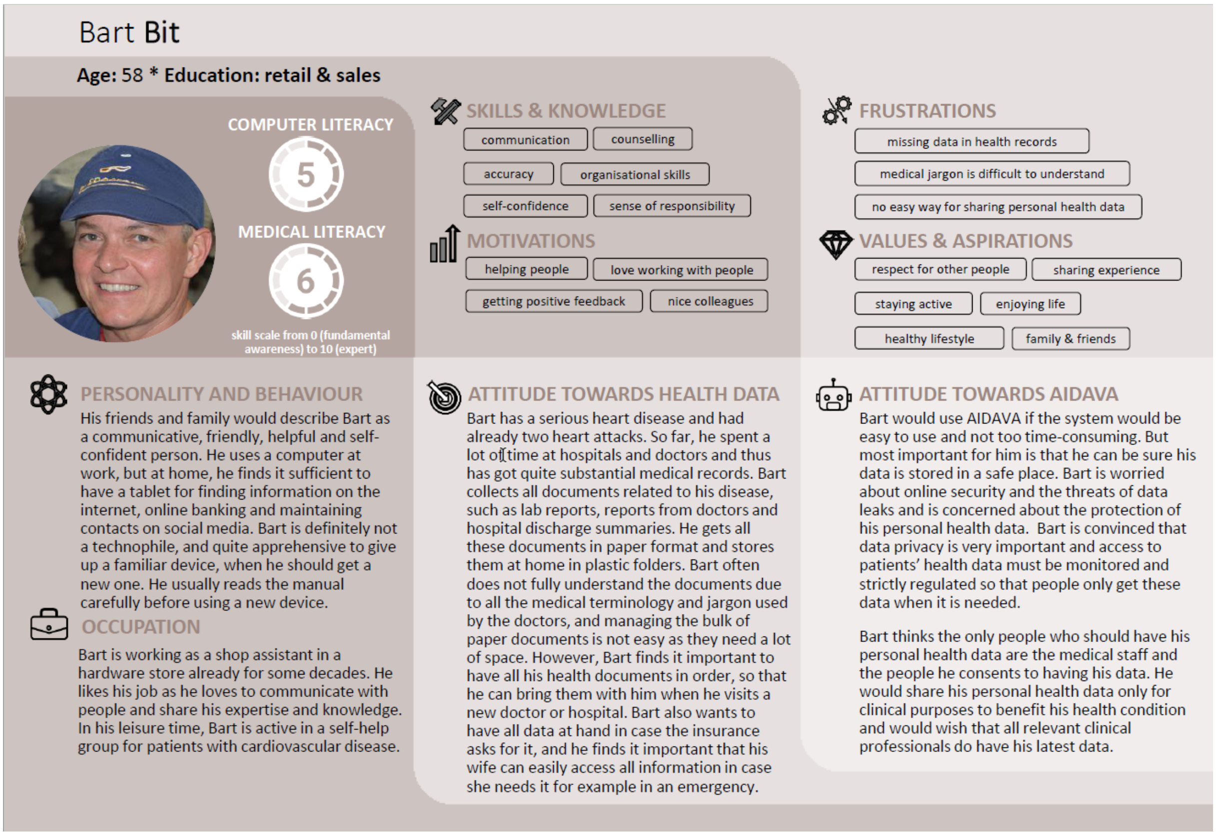
For Lena it is very important to always have a good overview of her health condition. She uses the health app on her smartphone to regularly monitor parameters such as heart rate and physical activity, and she collects all medical reports and documents from doctors and hospitals. Lena gets most of these data in digital format, but some documents are still on paper. She stores all health-related documents at home on her computer and keeps the papers in a box.

Lena thinks it is not easy to collect all her health data, as there is no one-stop-shop for all medical records, and there is also no easy way of sharing health data between different hospitals and doctors or transferring personal health data to another country. Furthermore, Lena noticed that sometimes information is missing in her records, as data are forgotten to insert or not collected by the healthcare personnel.

ATTITUDE TOWARDS AIDAVA

Lena is excited to have a digital system for collecting, managing and improving the use of her personal health data, and hopes that AIDAVA would help her to own and control her data. She would use AIDAVA to collect her health data from different sources easier and have all her health-related data and documents at her disposal whenever she needs them. She hopes AIDAVA can support her with fixing errors in her health data and making her complete and correct health records accessible to doctors.

Lena thinks that data privacy is essential and her personal health data should not be used without her consent. She would be happy to give consent to sharing her health data for a good research purpose. However, it must be fully transparent what the data are used for and by whom, and ideally data shared for research should be anonymised.



Bart Bit

Age: 58 * Education: retail & sales

COMPUTER LITERACY 5

MEDICAL LITERACY 6

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE

- communication
- counselling
- accuracy
- organisational skills
- self-confidence
- sense of responsibility

MOTIVATIONS

- helping people
- love working with people
- getting positive feedback
- nice colleagues

FRUSTRATIONS

- missing data in health records
- medical jargon is difficult to understand
- no easy way for sharing personal health data

VALUES & ASPIRATIONS

- respect for other people
- sharing experience
- staying active
- enjoying life
- healthy lifestyle
- family & friends

PERSONALITY AND BEHAVIOUR

His friends and family would describe Bart as a communicative, friendly, helpful and self-confident person. He uses a computer at work, but at home, he finds it sufficient to have a tablet for finding information on the internet, online banking and maintaining contacts on social media. Bart is definitely not a technophile, and quite apprehensive to give up a familiar device, when he should get a new one. He usually reads the manual carefully before using a new device.

OCCUPATION

Bart is working as a shop assistant in a hardware store already for some decades. He likes his job as he loves to communicate with people and share his expertise and knowledge. In his leisure time, Bart is active in a self-help group for patients with cardiovascular disease.

ATTITUDE TOWARDS HEALTH DATA

Bart has a serious heart disease and had already two heart attacks. So far, he spent a lot of time at hospitals and doctors and thus has got quite substantial medical records. Bart collects all documents related to his disease, such as lab reports, reports from doctors and hospital discharge summaries. He gets all these documents in paper format and stores them at home in plastic folders. Bart often does not fully understand the documents due to all the medical terminology and jargon used by the doctors, and managing the bulk of paper documents is not easy as they need a lot of space. However, Bart finds it important to have all his health documents in order, so that he can bring them with him when he visits a new doctor or hospital. Bart also wants to have all data at hand in case the insurance asks for it, and he finds it important that his wife can easily access all information in case she needs it for example in an emergency.

ATTITUDE TOWARDS AIDAVA

Bart would use AIDAVA if the system would be easy to use and not too time-consuming. But most important for him is that he can be sure his data is stored in a safe place. Bart is worried about online security and the threats of data leaks and is concerned about the protection of his personal health data. Bart is convinced that data privacy is very important and access to patients' health data must be monitored and strictly regulated so that people only get these data when it is needed.

Bart thinks the only people who should have his personal health data are the medical staff and the people he consents to having his data. He would share his personal health data only for clinical purposes to benefit his health condition and would wish that all relevant clinical professionals do have his latest data.

3.2.2. Data user personas

Foundation for data user personas

All aspects mentioned in this persona's foundations table are based on the interview results. The aspects written in red font were used for the creation of the respective persona's canvas.

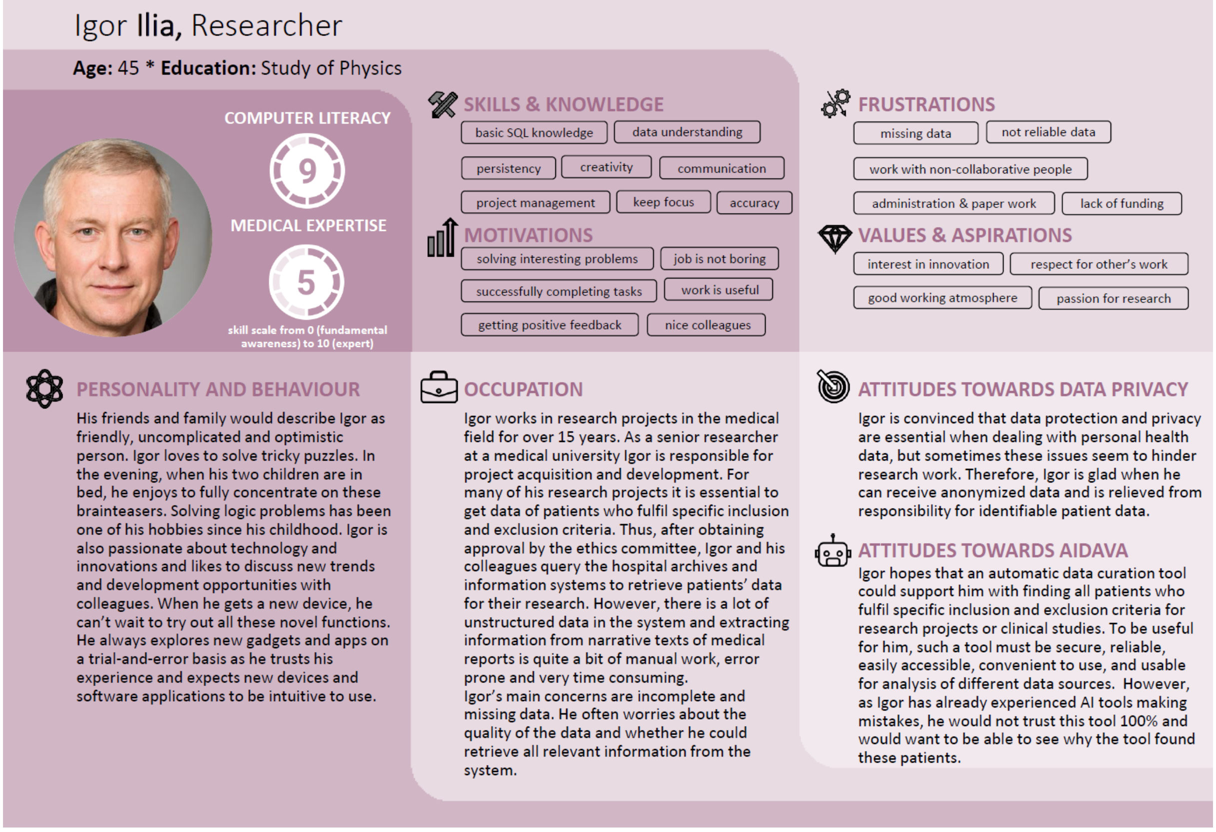
	data-user persona 1 ("researcher")	**data-user persona 2 ("medical doctor")**
age	40-60	40-60
self description	friendly (4x), likeable, always good mood, eloquent, happy, optimistic, enthusiastic, hard working, decision maker, straight forward, honest, uncomplicated, rather precise, careful, often very indecisive parent of 2 children	hard working (3x), conscientious, precise, quick, stubborn, relatively good stress tolerance in working under pressure, experienced, very calm person, likes chocolate very much
computer / data literacy	computer-user skills: 9-10 (4x) - educational background in data science (3x)/ no education in data science but acquainted knowledge in course of work and attended training courses (1x)	self-assessment of computer-user skills: 6-7 - no education in computer/data science; can manage everything work-related / everyday work is not a problem / work quite proficiently with programs I often use
medical/health literacy	familiarity with medical terms: 4-6; no medical education; when I need medical terms I know where to look them up; acquired knowledge of medical terms during career;	familiarity with medical terms: 8-10; education: medical doctor; you cannot know everything and medical terminology also changes over time
attitude towards data protection and privacy (at work)	data protection and data privacy is very important especially when it comes to personal health data, but sometimes it is bothersome as it hinders us in many surveys and research projects	data protection and data privacy is very important especially when it comes to personal health data, but sometimes it is a complex issue and an ethical dilemma (e.g. when inadvertently opening the wrong person's records or when discussing patient-related matters with another doctor and there may be another person in the room...) Also in my personal life I do think about data protection rules and data protection
attitude towards new technologies - getting a new smartphone...	usually not reading user manuals, trying out new stuff on a trial-and-error basis, expect new devices to be self-explanatory and intuitive to use; I trust my experience on technology. if I get a new device, I'm excited and eager to explore the new functionalities;	usually reading user manuals, but prefers that someone shows how it works (that's always easier); getting new technology is not a big thrill, I'm not the kind of person to immediately buy a new thing and start researching on my own how to use it; usually I choose the same brand that I had before when I need to get a new device - they are similar and there is no problem for me to handle the new one then.
job / tasks	working in research projects in the medical field; initiating clinical studies and medical research projects;	medical doctor in a hospital (diagnostics, patient treatment); pathologist (histopathology diagnostics);

	data-user persona 1 ("researcher")	**data-user persona 2 ("medical doctor")**
important skills for the job	computer & data-science related skills (*basic SQL knowledge and data understanding; ability and willingness to learn dealing with hospital information system, specific databases, ETL and evaluation tools*); communication and social skills (ability to communicate with professionals in different languages, coordinate and integrate different experts); management & research skills (keeping an overall picture and avoid getting lost in details, project management, problem solving, read & write scientific articles and conduct research);	medical knowledge; basic computer user skills, and willingness to keep learning different information systems that are constantly being updated; communication and social skills (ability to communicate with different types of people, making things understandable to people with different levels of education, day-to-day interaction with patients and colleagues, psychological knowledge);
personality traits for the job	self-motivation, focus, persistence acting self-reliant and autonomous calm, patience, excellent self-control interest in innovation, desire to develop (your specialty), innovative, creative, curious open to teamwork, collaboration	empathy, being considerate of others and not too self-centred, friendly, ability to accept different opinions and tolerate various behaviours and situations; resilience to stress, stable nervous system, able to work under pressure; accuracy, conscientiousness; being smart, speed in thinking and acting, analytical thinking; calm, patience, excellent self-control; interest in innovation, desire to develop (your specialty), innovative, creative, curious, teamwork, collaboration;
motivations / pleasure in job	successful completion of a task; getting positive feedback; freedom, independence, creativity; passion about future research, curiosity job is not boring - different things to do, opportunity for constant development, difficult and interesting cases (solving problems); work is useful, beneficial and valuable *(I can help people with my skills, and bring in / put into practice good ideas...)*; good team and working atmosphere *(respect and appreciation for colleagues and their work)*	physical work/hands-on approach in the operating room job is not boring - there are always different things to do, opportunity for continuous development, difficult and interesting cases; my work is useful, beneficial and valuable *(I can help people with my skills, I can bring in and put into practice good ideas...)*; good team and working atmosphere *(respect and appreciation for colleagues and their work)*
challenges / frustrations in job	missing data and not reliable data *(often documentation is poor, don't know whether data items are up-to-date)*; having to work with / depend on non-collaborative people; funding politics; administration and paperwork;	work overload and shortage of support staff - need to do a massive amount of extra hours. administration and paperwork; need to keep up with fast changes but no time for learning new things; lack of mental rewards;
purpose of health data retrieval	retrieve (many) patients' health data for research projects and clinical studies (need to find all patients that fulfil specific inclusion and exclusion criteria)	retrieve a specific patient's health data for diagnostics and treatment of this patient (need to get an overview of the patient's previous medical history)

	data-user persona 1 ("researcher")	**data-user persona 2 ("medical doctor")**
current procedure for health data retrieval	access to patients' data for research purposes always needs the approval of the ethics committee and is not possible without a project	to access patients' health records authentication is required; access to patients' data is logged in the system (doctors should look only at "their" patient's data, whom they are treating)
	retrieve data via SQL queries from the hospital archive system, or specify which data you need and ask a specific department at the university (data clearing site) to retrieve these data from hospital information system; researcher is relieved from responsibility for identifiable patient data since they get pseudonymised data from data-clearing site; data cleaning and pre-processing for research is a lot of manual work	go through the health record documents of the patient in the hospital information system and find the necessary information in heterogeneous data (*including also a lot of unstructured data such as narrative report texts*)
current challenges in data retrieval	bad / unclear data quality (*you don't know if that is the latest status for a specific entry/document in the system*); analysing unstructured data and narrative texts (which often include words like "rather" and "probably"…); currently no structured data entry into the hospital information system, no structured reports; no access to external patient data *(e.g. from GP or from other (private) clinics)*; missing and incomplete data, due to a) mistakes in data retrieval process *(not all available data from system retrieved)* b) data was not entered into the system *(someone forgot to input this information, or this information is not recorded - there is no field for that information in the system)*	I always need to justify why I have looked at someone's data (*if someone asks long time after it has happened I may not even remember the reason anymore…*); sometimes I cannot find data/information I need (e.g. if layout of the digital system has changed); difficult to find information in long medical records (some people's medical history is many pages long and going through them takes time…); always feel that I´m missing some medical data. The lack of data *(e.g. if someone has forgotten to input information on a medical test conducted for a patient)* affects my work, wastes time and may lead to repetitive tests and procedures. missing and incomplete data, due to a) mistakes in data retrieval process b) data was not entered into the system *(someone forgot to input this information, or this information is not recorded)*
attitudes/expectations towards automatic health data curation	would use such a tool for patient recruitment in clinical studies or for finding patients who fulfil specific inclusion and exclusion criteria for research projects	such a tool should help with finding/identifying important parts of the information in a patient's medical records; such a tool should extract information from report texts and fill in forms (e.g. order forms for analyses or concilium…) automatically

	data-user persona 1 ("researcher")	data-user persona 2 ("medical doctor")
to be useful, such a system must…	be safe and secure; be easily accessible & easy to configure; be accurate; be interoperable; be usable for analysis of different data sources; recognise automatically if something is wrong; check data quality automatically; deliver data I need in perfect quality; be up-to-date and fast; be user friendly, convenient/easy to use; be reliable (provide reliable data)	provide a good overview of the medical record (show only important things) provide data so that they can be easily processed be up-to-date and fast be user friendly, convenient/easy to use be reliable (provide reliable data)
trust in such an AI-based tool	the tool should show me why it found these patients/data, I want to be able to understand that; I would not trust it 100% as I've already seen AI-tools making mistakes…	if the company says this works, then I do not check; from this tool I would expect to get all data - so I would only make some random sample checks;

Visualisations of the data user personas



Igor **Ilia,** Researcher

Age: 45 * **Education:** Study of Physics

COMPUTER LITERACY 9

MEDICAL EXPERTISE 5

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE
- basic SQL knowledge
- data understanding
- persistency
- creativity
- communication
- project management
- keep focus
- accuracy

MOTIVATIONS
- solving interesting problems
- job is not boring
- successfully completing tasks
- work is useful
- getting positive feedback
- nice colleagues

FRUSTRATIONS
- missing data
- not reliable data
- work with non-collaborative people
- administration & paper work
- lack of funding

VALUES & ASPIRATIONS
- interest in innovation
- respect for other's work
- good working atmosphere
- passion for research

PERSONALITY AND BEHAVIOUR

His friends and family would describe Igor as friendly, uncomplicated and optimistic person. Igor loves to solve tricky puzzles. In the evening, when his two children are in bed, he enjoys to fully concentrate on these brainteasers. Solving logic problems has been one of his hobbies since his childhood. Igor is also passionate about technology and innovations and likes to discuss new trends and development opportunities with colleagues. When he gets a new device, he can't wait to try out all these novel functions. He always explores new gadgets and apps on a trial-and-error basis as he trusts his experience and expects new devices and software applications to be intuitive to use.

OCCUPATION

Igor works in research projects in the medical field for over 15 years. As a senior researcher at a medical university Igor is responsible for project acquisition and development. For many of his research projects it is essential to get data of patients who fulfil specific inclusion and exclusion criteria. Thus, after obtaining approval by the ethics committee, Igor and his colleagues query the hospital archives and information systems to retrieve patients' data for their research. However, there is a lot of unstructured data in the system and extracting information from narrative texts of medical reports is quite a bit of manual work, error prone and very time consuming.

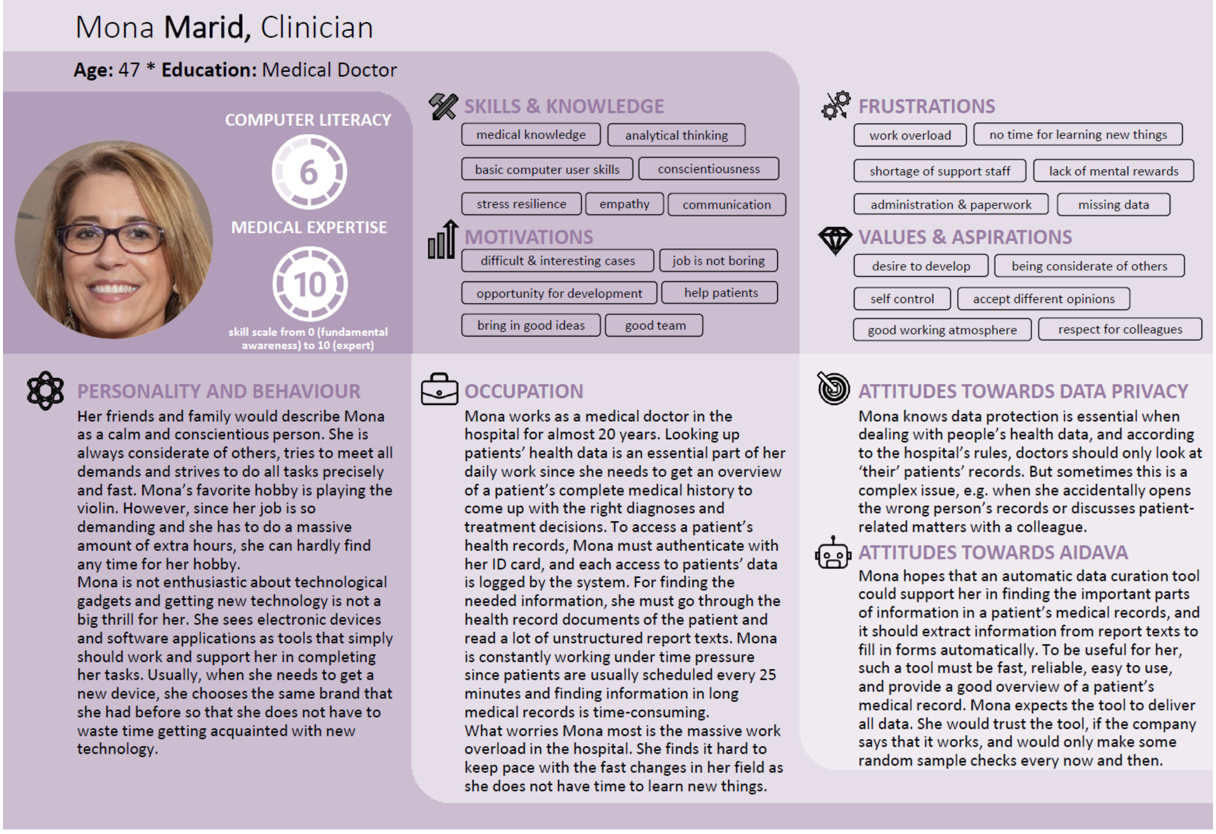
Igor's main concerns are incomplete and missing data. He often worries about the quality of the data and whether he could retrieve all relevant information from the system.

ATTITUDES TOWARDS DATA PRIVACY

Igor is convinced that data protection and privacy are essential when dealing with personal health data, but sometimes these issues seem to hinder research work. Therefore, Igor is glad when he can receive anonymized data and is relieved from responsibility for identifiable patient data.

ATTITUDES TOWARDS AIDAVA

Igor hopes that an automatic data curation tool could support him with finding all patients who fulfil specific inclusion and exclusion criteria for research projects or clinical studies. To be useful for him, such a tool must be secure, reliable, easily accessible, convenient to use, and usable for analysis of different data sources. However, as Igor has already experienced AI tools making mistakes, he would not trust this tool 100% and would want to be able to see why the tool found these patients.



Mona **Marid,** Clinician

Age: 47 * **Education:** Medical Doctor

COMPUTER LITERACY

6

MEDICAL EXPERTISE

10

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE

medical knowledge | analytical thinking | basic computer user skills | conscientiousness | stress resilience | empathy | communication

MOTIVATIONS

difficult & interesting cases | job is not boring | opportunity for development | help patients | bring in good ideas | good team

FRUSTRATIONS

work overload | no time for learning new things | shortage of support staff | lack of mental rewards | administration & paperwork | missing data

VALUES & ASPIRATIONS

desire to develop | being considerate of others | self control | accept different opinions | good working atmosphere | respect for colleagues

PERSONALITY AND BEHAVIOUR

Her friends and family would describe Mona as a calm and conscientious person. She is always considerate of others, tries to meet all demands and strives to do all tasks precisely and fast. Mona's favorite hobby is playing the violin. However, since her job is so demanding and she has to do a massive amount of extra hours, she can hardly find any time for her hobby.

Mona is not enthusiastic about technological gadgets and getting new technology is not a big thrill for her. She sees electronic devices and software applications as tools that simply should work and support her in completing her tasks. Usually, when she needs to get a new device, she chooses the same brand that she had before so that she does not have to waste time getting acquainted with new technology.

OCCUPATION

Mona works as a medical doctor in the hospital for almost 20 years. Looking up patients' health data is an essential part of her daily work since she needs to get an overview of a patient's complete medical history to come up with the right diagnoses and treatment decisions. To access a patient's health records, Mona must authenticate with her ID card, and each access to patients' data is logged by the system. For finding the needed information, she must go through the health record documents of the patient and read a lot of unstructured report texts. Mona is constantly working under time pressure since patients are usually scheduled every 25 minutes and finding information in long medical records is time-consuming.

What worries Mona most is the massive work overload in the hospital. She finds it hard to keep pace with the fast changes in her field as she does not have time to learn new things.

ATTITUDES TOWARDS DATA PRIVACY

Mona knows data protection is essential when dealing with people's health data, and according to the hospital's rules, doctors should only look at 'their' patients' records. But sometimes this is a complex issue, e.g. when she accidentally opens the wrong person's records or discusses patient-related matters with a colleague.

ATTITUDES TOWARDS AIDAVA

Mona hopes that an automatic data curation tool could support her in finding the important parts of information in a patient's medical records, and it should extract information from report texts to fill in forms automatically. To be useful for her, such a tool must be fast, reliable, easy to use, and provide a good overview of a patient's medical record. Mona expects the tool to deliver all data. She would trust the tool, if the company says that it works, and would only make some random sample checks every now and then.

3.2.3. Data curator personas

Foundation for data curator personas

All aspects mentioned in this persona's foundations table are based on the interview results. The aspects written in red font were used for the creation of the respective persona's canvas.

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
computer/data literacy	Very good/expert computer user with good computer skills. (rating on 0-10 scale; 7; 10; 10)	Self-rated proficiency as computer & smartphone user: 7-8 (2); 9; 6;7; 6-7 , average 7,25 Can handle everyday tasks related to IT, yet there is always room for improvement; uses a computer daily; weak in graphs. Enjoys learning new things - keeps mind fresh	
attitude towards data protection and privacy	Believes that data protection and privacy aspects are underestimated and underappreciated within research; necessary for every research project, although it complicates things for certain research goals - you have to take a lot of approvals from different people to acquire data. It's highly important to deal with data rights and data privacy aspects for patient needs to trust the researchers who deal with this data.	Data rights and privacy are very important(2) and also a risk in AIDAVA (e.g. problem if the employer receives those data or insurance companies utilize these data for their purpose - Not everybody needs to know what is wrong with patients´ health). However, for research, it is important to have high-quality, correct and complete data. Sometimes data protection & privacy is overstrained and this causes also problems as the related bureaucracy is extremely high. Data protection rules are very strict.	Data rights and privacy very/highly important

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
attitude towards new technologies	It depends on the device whether to read the manual because it can be a waste of time or it can save a lot of time. Sometimes I like to read manuals because it saves so much time. It's like a construction of an IKEA furniture. Excited to explore the new functionalities of the new device and I don't have a problem with the old device because it's taken care of by technology due to the restore to factory settings option.	Usually, I do not read manuals (2). I consult the manual only when I get stuck or it is a complicated device (3) I always read the manual (3) Sometimes I ask a colleague or expert (2) to help with/show how to use a new device. When I would get a new smartphone, I would look forward to it/gladly use it (2) If a new solution or application makes my work easier then I feel fine. There are so many new devices, equipment and systems coming out all the time so I do not get frustrated. Not all new solutions are helpful.	Depends on the device whether to read a manual. No problem to use new device
medical/health literacy	On a scale of 0-10 (10, 8, 5) some educational background related to medicine and healthcare generally; familiar with medical terms because dealing with medical data. Does not consider themselves an expert, but not a person who doesn't have any idea about medical terms either.	Self-rated familiarity with medical terms: 8; 10; 7 (2); 6-7, 7-8, average 7.7	

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
Job	Working as a data engineer ; doing a PhD in clinical data science ; programming in SQL and MATLAB; some administrative stuff as well; postdoctoral researcher (2) within a clinical data science department; supervision of PhD students (2) within the department and coding; preparing the data for researchers and transforming them into a machine readable format.	Patient care, including patient education, examination, guiding through chemotherapy, aftercare and discussing the reports – a lot of my work is talking to the patients. In addition: organizational management of the study centre, science and research activities Medical secretary work (2) - Filling in the data to registries (2) is an added task for me but my main tasks are also related to entering patient data (including myocardial infarction notice, treating physicians’ information, patients´ data and care records, test results, procedures, medications and billing/invoice input). Main tasks: organizing the work of the nurses, ordering, work schedules; all personnel related issues. Make sure all the input is correct, running tasks include helping the head nurse. I also help with patient care and cover shifts.	

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
important skills for the job	Skills in programming(3), coding(2) and semantic web technologies that transform data into machinery that will format according to a specific set of principles called FAIRdata principles. Communication skills (3) with people in your job in order to understand and execute tasks and needs of a specific project; patience (2) is one of the important traits a clinical data scientist should have. Project management skills; development of applications; statistical analysis. You should also be open minded. Dealing with a half dozen research projects, international and National Research projects as well. You have to be precise and extremely careful with the medical data you are using and you have to clean them in a specific way in order to be usable.	Most important for doing the job well: you have to like this job and this work. A person should be able to think for themselves – work independently and stay calm. The skills you need to do this job are acquired over time. You will just remember everything at one point. A Prerequisite to working here is a general knowledge of what goes on in our field. Language skills needed include (local language + largest ethnic minorities, tourists´ languages and English, sometimes medical Latin) Skills needed in this job: proper education, computer skills (4), procurement procedures, administrative skills. Also medical skills(3) – need to also take shifts and care of patients(2). Traits needed: patience, dealing with emotions (your own and others), there are a lot of people and (sometimes difficult) personalities (2) around you – In general you need to have a pretty thick skin to do this job. Conscientiousness (2), precise,	computer skills, communication skills/dealing with different personalities, precise, administrative skills

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
		have a wide understanding of many things, **friendly** (2) and helpful, empathic, multitasking;	
I like most in job	Communication with colleagues in the office and out of the office as well, enjoy collaborating with people and have face to face contact with them; working together on different projects. Improve patient's lives treatments. Implementation of new technology within medicine.	Combination of the fact that the work is professionally demanding and interesting on the one hand, and on the other hand, you can also look after and support the patients over a longer period. The job itself is also interesting, there is not a lot of routine. I like working with patients because I like practical work. I enjoy my colleagues, the field of	communication/working with colleagues, improve patients' lives

	data-curator persona 1 ("researcher")	**data-curator persona 2 ("clinical assistant")**	**common aspects for both personas**
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
	Independence by working from wherever and when I want. Love helping people to make decisions based on data; I like the application of programming in healthcare. So the new technology, that's the new era of artificial intelligence, brings us. Application of technology in healthcare and therefore the ability to improve patients' lives.	work I'm in, teamwork, fast paced work and decisions, variation of administrative and medical tasks. I enjoy working /communicating with my colleagues and if you get your job done and people are pleased with you. The room I work in is nice and warm, equipment (to work with) are ok.	
challenge in job	Collaborating with people because they usually have different time schedules and can be difficult (personality wise?). Commuting to work - prefer working from home. Not enough legal standardised framework to use patients' data within our organization - due to that dealing with laws, lawyers and different data privacy officers to acquire data and it's extremely difficult. I don't really like debugging in my coding. Every day I have to debug my code. I hate politics in my world because you and	Sometimes I have the feeling that I can't finish my work/time factor (2). Patients have to wait so long to be examined, there are no more patient beds, you don't know where to admit patients. Increasing bureaucracy in every respect, technical solutions are urgently needed to support with these bureaucracy tasks. New programs you need to use and learn that appear quite often. Health insurance checks for the patients -there used to be a phone number you could call to check but that does not exist anymore. Often the system info	Time pressure/not enough time to finish tasks, increasing bureaucracy/dealing with laws/politics

	data-curator persona 1 ("researcher")	**data-curator persona 2 ("clinical assistant")**	**common aspects for both personas**
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
	requiring approvals. Need to build some personal or other interests from a company and this is really annoying to me. Hate time pressure and deadlines. Problems with the extraction of medical data from hospital- time consuming part and it's really stressful. Lot of meetings during my daily work.	is not correct. Main challenge with this job is communication with people (2). New devices and programs are usually not an issue. No day is the same.	

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
current approach for health data curation	I use and retrieve patients' health data because I want to implement prediction modelling within the healthcare domain and specifically radiotherapy. I mainly use EHR data that are collected within the different data storage systems of the hospitals In my daily routine, I mainly use SPARQLqueries (2) to retrieve the data from the hospital systems.Use some statistical analysis tools like R Studio to clean and get the data ready for research. My main task is to develop prediction models based on deep learning techniques. Furthermore, I use patient data to implement some algorithms that delineate medical images using deep learning algorithms. I mainly deal with CT and MRI images and I receive these images in DICOM format.	When collecting data from a patient to create a plan, that actually includes everything: Data entry, curation, and use. I currently do all of that manually. The findings have to be collected from everywhere: from the national electronic health records, from the evaluating institute, sometimes the patients bring reports from other physicians, sometimes we get a CD, sometimes the medical information is in another language and we need translation. It is always different and very diverse. Some data is put into the hospital system but some are not. Sometimes we ask the patient themselves or check the ambulance card/notes. I get the data both from the information system and from the national system.	significant amount of labour to clean data

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
	They are coming from the EHR systems of hospitals and I have this data stored in SQL Server. I retrieve patients' health data in order to implement artificial intelligence algorithms for prediction modelling. I mainly transform this data into a machinery that will format using fair data principles and semantic web technologies.		
current challenges in data curation	The main challenges are missing data and data that don't make sense. Cleaning the data and having a standardised way of format/ different formats that the data are given to me. Keeping in track of several projects because I'm not good at	Sometimes the information doesn't match and then you have to look more closely. Sometimes, you can only rely on the image data. Doctors are urgently calling for measures to reduce workload. So that patients don't all come to us for aftercare, but the general	missing data

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
	project management.	practitioners and technology should take over. The main challenges I have with data is that I cannot access all needed data and some of the data is missing. It is extra work to look for and enter data to registries	

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
attitude towards automatic health data cleaning	AIDAVA has to be secure- (2) because this tool will deal with patient data, so it's extremely significant to have a security framework within this system.Second, it has to be user friendly(2). The users do not have to spend a lot of time to find all the different aspects of this system. It's important to be collectively owned. It has to be public, open to the people who are interested in data curation and data collection of patients.This system has to be a server based system and what I mean is that this system should be something like a cloud based system. Furthermore it has to be machine readable using fair data principles and API, in other words it has to be user friendly. The ideal system consists of an automation of the errors detection that may exist in a data set. The privacy preserving data exchange or data modelling exchange within the different	The laboratory checks should be automatically uploaded to the system. And perhaps with a warning system, for example, if the leukocytes are below the necessary value. The AIDAVA tool would have to be simple and self-explanatory, and should be able to filter the right information for me. It should ask the patient correctly everything that is needed, and also support by telling the patient what to do when this or that arises. We need a simplification of communication at the interfaces and fewer patient-hospital contacts. AIDAVA should be less work and not more. Also some background info about the patient could be automatically added from the system (phone number, home address, contact info.) AIDAVA could collect all the needed data from different	user friendly/self-explanatory,;secure/data is protected; correct/error detection

	data-curator persona 1 ("researcher")	data-curator persona 2 ("clinical assistant")	common aspects for both personas
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
	participants is also important so I would highly consider it. I will say that the recommendation is really important to this tool. Standard terminologies have to be used in order to achieve a standardisation framework of the data sets and the data that are used within this tool. The ideal data curation and publishing machine would include standardised terminologies and extract common entities from it.	places (2). Diagnosis, test results, procedures information. Some of this text is in free text, some are structured. It would also be nice to know whether they have allergies, smokes or not (and old is this information – maybe they quit smoking). What has the home treatment been, what medications are they on. What about the data on paper? Speech to text function would also be helpful. It would be great if there could be a machine that could collect all the necessary data and put it in a register. At the moment, there are so many clicks and ticks to be ticked in the information system. You have to make sure you don't forget anything. It's those clicks and clicks that have made our work very busy today. AIDAVA could check or keep track so that registry fields that have already been checked and don't have to be checked again each time. Somehow it should be possible to put a mark on the field so that the information doesn't change again. AIDAVA	

	data-curator persona 1 ("researcher")	**data-curator persona 2 ("clinical assistant")**	**common aspects for both personas**
age	18-40	18-40 (1); 40-60(3); >60(2)	
self-description	Social person with a lot of friends, likes spending time with colleagues outside of work; likes spending time with friends and family; passionate about new technologies and computers. Easy going person, hard working. Serious and friendly person.	Likes their work; Interested in people and strives for the well-being of the patients; open minded and curious; enthusiastic doctor who is annoyed by all the things that distract from the medical doctor's work; conscientious, correct, quick/fast, proper, friendly, empathic, precise, innovative, loyal, organized	friendly
Educational background	Bachelor in medicine and Master in data science, but not a clinical person; educational background in computer and data science;	medical secretary/assistant courses (if no-medical previous education)	
		could e.g. point out if a field is blank and ask why it is blank.	

Visualisations of the data curator personas

Helga Johanson, Clinical assistant

Age: 50* Education: clinical assistant, nurse

COMPUTER LITERACY 7

MEDICAL EXPERTISE 8

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE

- solution orientated
- language skills
- administrative skills
- communication
- accuracy
- computer skills
- medical knowledge

MOTIVATIONS

- improve patient's lives
- nice colleagues
- good working conditions
- independence
- continuous development
- interesting job

FRUSTRATIONS

- time pressure
- information system errors
- people’s different time schedules
- increasing bureaucracy
- missing data

VALUES & ASPIRATIONS

- professionally demanding
- modern equipment
- good working atmosphere
- new technology

PERSONALITY AND BEHAVIOUR

Helga is interested in people and strives for the well-being of patients. She is open-minded and curious. Colleagues say that Helga is a hardworking, conscientious, correct, quick, empathetic, innovative, loyal, organized and friendly person and a good communicator.
The main challenges at her work are the amount of new programs she needs to learn to use and dealing with relationships between people with various personalities.
If Helga gets a new device it is not difficult for her to start using it. She usually consults the manual first, and if that does not help, she asks for advice from a specialist.

OCCUPATION

Helga has been working in the medical field for over 20 years. Her main tasks are patient care, including patient education, guiding the treatment process and discussing reports. Entering data into registries is an additional task for Helga, but her main tasks are also related to entering patient data into the hospital's information system.
Every morning, Helga goes through all the system/patient care areas which she is responsible for and makes sure that all entries are correct. If not, she corrects them. Helga is very conscientious and precise in her work. She is quick and independent in her tasks.
She likes her job because it is different every day and offers many learning opportunities. The most annoying part of her work is when the information systems don't work properly.

ATTITUDES TOWARDS DATA PRIVACY

Helga feels that data privacy is very important. It is beneficial for science and patient care to gather all health data in one place, but it is a risk if e.g. insurance companies use this data for their own purposes or if the employer receives this data. Access to data must be very precisely restricted.

ATTITUDES TOWARDS AIDAVA

Helga believes the AIDAVA tool should be simple, self-explanatory and user-friendly. Most important part is that it has to be secure because this tool will deal with patient data. AIDAVA should have automatic error detection that may exist in data sets. Background information about the patient should be automatically added from different sources (phone number, home address, contact info). It’s important that the curated information doesn't change multiple times. AIDAVA should also have an alert system - for example, if some health indicators are abnormal.

Jonas Thompson, Researcher

Age: 33* Education: PhD clinical data science

COMPUTER LITERACY 9

MEDICAL EXPERTISE 8

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE

- Programming/coding
- Good communicator
- Precise
- Patience
- Open-minded
- Semantic web technologies
- Project management

MOTIVATIONS

- Ability to improve patient lives/treatments
- Independence in tasks
- collaboration
- Enjoy the application of technology in healthcare

FRUSTRATIONS

- Legal paperwork
- Commuting
- Debugging
- work with non-collaborative and busy people
- Non-standardized data
- Time pressure and deadlines
- Missing data

VALUES & ASPIRATIONS

- Being the boss of your own time
- Values face to face interactions
- Enjoys helping people
- passion for research

PERSONALITY AND BEHAVIOUR

People describe Jonas as social, tech-savvy and hard-working person. In his spare time Jonas loves to spend time with his colleagues, family and friends.
He loves learning new things and assembling IKEA furniture or adapting to new technologies is no problem for him.
When Jonas is currently not programming or coding, he is helping to supervise PhD students. He is always happy to help out in administrative tasks if needed. He really enjoys helping people to make decisions based on data.
When he gets a new device, he is excited to explore all the new possibilities and does not worry about data getting lost when changing smartphones. He has no problems navigating new devices without a manual but depending on the device reading the manual first can save a lot of time.

OCCUPATION

Jonas works in research projects in the medical data field and supervises PhD students. As a data engineer and researcher at a medical university Jonas´s main tasks are programming, coding and development.
A big part of his work is to transform and clean the data into machine readable or standardized form. Jonas really hates all the time consuming legal and ethical paperwork that comes with processing medical data. It substantially slows down the work process.
Jonas likes it very much to be the boss of his own time and have the ability to work from home as well. He also enjoys to apply his abilities in programming to healthcare and see it used to improve patient lives or treatment.

ATTITUDES TOWARDS DATA PRIVACY

Jonas believes that data protection and privacy issues are underestimated and underappreciated within research - especially in healthcare. Privacy aspects are very necessary in every research although it sometimes complicates things for certain research goals since it is so time consuming to get permissions.

ATTITUDES TOWARDS AIDAVA

One of the most important factors for Jonas is that this tool is secure since it deals with patient data. Also it needs to be user friendly to save time for all users. Jonas finds it important for AIDAVA to be collectively owned – it has to be a public server based (cloud) system and available to all who are interested in data collection and curation. The ideal system should in addition to automation have error detection. It should use standardized frameworks of the data sets and extract common entities from it.

3.2.4. Third party app developer personas

Foundation for 3rd party app developer personas

	app developer persona 1 - user persona ("software developer")	**app developer persona 2 - customer persona ("manager")**
education	study of medical informatics and communication science	study of psychology / molecular biology / computer science;
career	worked for several years in hospital administration and health-related communication;	co-founder and board member of digital health / software development company; worked for 20 years as CTO in software industry; work profile evolved from doing lots of programming in earlier days to now having mainly organisational and management tasks; working in health care system policy consulting for 20 years;
current job	developing software for health-related apps	CTO, leading and coordinating the developer teams of a digital health company - responsible for project management and timelines; board member of digital health / software development company;
main motivation in job	likes logical and structured thinking; enjoys solving problems and debugging - i.e. finding solutions for stuff that does not work;	
main frustration / challenge in job	in another project they do not have a real test environment and the partner's implementation is not available for testing 24/7	The highest hurdle is the reimbursement system: even if you have evidence for the functioning of a health IT solution, it's not clear who pays for a health IT solution.
data literacy		very high data literacy (*consultant for government in data privacy/protection issues; expert in state of the art software technology*)
medical/health literacy		familiarity with medical terms is very high (*consultant for government and co-lead of health-related national research program; expert in HL7-FHIR - currently developing a connection with the national health terminology server*)

	app developer persona 1 - user persona ("software developer")	**app developer persona 2 - customer persona ("manager")**
attitude towards health-data privacy	data protection in health-related field is extremely important, as you cannot make it undone when data is leaked; health-data should be accessible in a way that patients can understand, and the patients should be able to decide with whom they want to share their data; *(Current problem: legally the data owners are the patients, but practically the real owners are medical doctors)*	health data privacy is very important - when you want to enter DIGA or CE, you have to have all the data privacy and data protection aspects covered. It's pretty complicated. We follow the required privacy and security guidelines from the ministry of health (*like ISO 27001 and NEN 7510 certification and a yearly penetration test on the software done by an external party*). We enter into an era of 2nd use of health data and need to be transparent with what we do with the data. You must make sure that a patient understands but not overestimates the implication of sharing data; emphasizing the giant value for society in having this kind of data that allows to build decision support systems, that will ultimately benefit all patients;
expectations/needs towards AIDAVA	important that API is well documented; need a test environment that's available 24/7; it must be transparent what we can expect from the tool; ideally, if I have to go to the hospital I want the doctor to see a report of two pages about my health that contains everything relevant and correct;	In an AIDAVA setting, our company could put data in. We don't see our platform as a standalone solution. We integrate into other platforms depending on the use case. The ideas in AIDAVA are something where we as a company want to go - we are eager to learn about the possibilities of AI. The phase where you put in PROMs with your finger is going away. There will be more NLP with all the PROM tools. Our company's use case in healthcare: we focus on the collection of health information for the individual - not only medical data but also the individual lifestyle (e.g. personalized nutrition, activity and other lifestyle activities, like alcohol and tobacco use) in collaboration with research and pharma institutes.

	app developer persona 1 - user persona ("software developer")	app developer persona 2 - customer persona ("manager")
attitude towards automatic health-data curation and publishing		It is important to have good data curation in healthcare. To achieve this, you need a top down approach from the ministry of health (by law healthcare providers need to deliver the medical data in a FHIR format), and when this works we have a solution for structured data, but not yet for unstructured, and for an ideal world it will take several years. Key issue is that lots of the data in clinical records is plain unstructured text - thus it would make sense feeding textual data to AIDAVA and reading back structured, interpreted data we can use in visualisation; One of the key advantages AIDAVA would be able to offer: you have meaningful structured data from unstructured data and you can compare it to a population and use it for a decision support system. Another valuable use case: Let's say you have three hospitals, and they record the same data and they do it structured, but they do it in different structures because there is no overarching agreement on what the exact structure should be. So if AIDAVA would be able to get from all these hospitals these data fields and also some form of metadata on what those fields mean in the context of registration of that hospital, then translating the values from each of these hospitals into some knowledge graph that is compatible with all of these representations that would be very valuable ;
attitudes towards AI	It must be transparent when data are aggregated where the data came from and which sources are used; Many people are sceptical about whether AI can be as good as humans in data curation - the possibility to see what exactly the system did to generate the curated data is essential for trustworthiness;	Key problem in AI and machine learning: how to verify the validity of the output of the machine learning algorithm? In a registry there's no way, especially when anonymized, to ever know whether the automatically curated data was correct. I think you're not going to get around, in the medical context at least, having a person verify the output. If you would want to replace an unstructured piece of data with a structured piece of data, as a physician you can't use that structured data to base your decision on. AI has to improve the efficiency and quality of care. *(You cannot motivate clinical staff to do more work)*

Visualisations of the app developer personas

Didi **Drdlik,** Software Developer

Age: 37 * **Education:** Study of Medical Informatics

COMPUTER LITERACY

10

MEDICAL LITERACY

5

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE

programming | data understanding | persistence | creativity | logical thinking | structured working

MOTIVATIONS

solving interesting problems | job is not boring | completing tasks | finding bugs and solutions

FRUSTRATIONS

unclear specifications | unrealistic deadlines | missing documentation in third party software | insufficient test environment

VALUES & ASPIRATIONS

good work/life balance | rejoice in his work | self-determined working hours and workflow

PERSONALITY AND BEHAVIOUR

His friends and family would describe Didi as an easygoing and optimistic person.
Didi has been interested in technical gadgets since his childhood. His friends frequently ask for his advice with respect to computers, smart watches or mobile phones since Didi is always well-informed about new technologies and knows the state-of-the-art products.
In his leisure time, Didi likes playing computer games or going out with his friends to play a round of billiards.

OCCUPATION

After finishing his university studies in medical informatics, Didi worked for some years in the hospital administration and health-related communication sector before he started his current job as a software developer in a digital health company. As part of a team of eight software developers, Didi is responsible for the maintenance and implementation of new features in the company's health-related apps. Tasks and issues are distributed according to a plan among the developers in the team. Most of the time, the software developers work on the assigned tasks on their own. However, they are in close contact with each other, and each day a short meeting with the teammates is scheduled. Didi aims to fulfil the tasks assigned to him always on time and in good quality.

ATTITUDES TOWARDS DATA PRIVACY

Didi thinks that data protection in the health sector is extremely important as you cannot make it undone when health data is leaked. He believes that health data should be accessible in a way patients can understand, and the patients should be empowered to decide with whom they want to share their data.

ATTITUDES TOWARDS AIDAVA

When it comes to cooperation with AIDAVA, as a software developer Didi would want to know in detail what he can expect from that tool. He would hope that the API is well documented and a test environment is available 24/7.
Regarding AI-based data curation, Didi finds it extremely important for trustworthiness to have the possibility to see what exactly the system did to generate the curated data and which sources are used.

Alec **Aarn,** Chief Technical Officer

Age: 61 * **Education:** Study of Molecular Biology

COMPUTER LITERACY

10

MEDICAL LITERACY

7

skill scale from 0 (fundamental awareness) to 10 (expert)

SKILLS & KNOWLEDGE

management | data understanding | coordination | networking | communication | policy consulting | business sense

MOTIVATIONS

fruitful cooperation | staying on the ball | development of the company | business opportunity

FRUSTRATIONS

reimbursement system in healthcare | finance issues | bureaucracy | uncooperative people

VALUES & ASPIRATIONS

value for society | rejoice in his work | collaboration | leadership

PERSONALITY AND BEHAVIOUR

His friends and family would describe Alec as a correct, communicative, solution-oriented and always forward-looking person with a great sense of humour. Alec loves to discuss about politics and society and is also very interested in technological trends and innovations.
Alec thinks that it is very important to never rest on one's oars but always keep learning. Thus, he often spends after-work hours reading newspapers and technical journals to keep up with societal as well as technological developments.

OCCUPATION

Alec is a co-founder and board member of a digital health company, which is developing software solutions and applications focusing on the collection of health information for the individual. As CTO of the company, Alec is leading and coordinating the developer teams and has mainly got organizational and management tasks.
In addition, Alec has been working in healthcare system policy consulting for over 20 years. He is a consultant for the government in health-data issues and is actively involved with health-related national research programs.
According to Alec, currently, the biggest challenge for IT companies in the healthcare sector is the reimbursement system, as it is not clear who pays for a health IT solution.

ATTITUDES TOWARDS DATA PRIVACY

Alec believes that data privacy is very important in the health sector and the company needs to strictly follow the respective privacy & security guidelines from the national ministry. However, Alec thinks for societal benefits it is also essential to ensure that patients understand but do not overestimate the implications of sharing data.

ATTITUDES TOWARDS AIDAVA

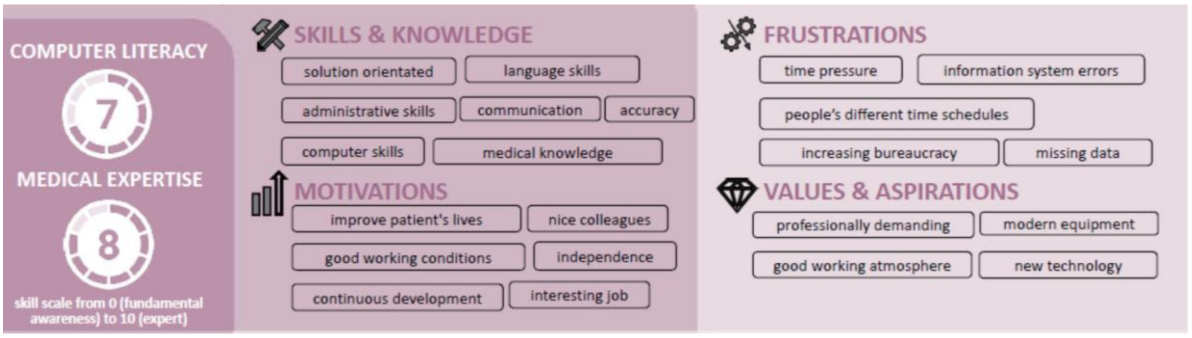
Alec knows that having good data curation in healthcare is important as lots of the data in clinical records are plain unstructured text. Thus, Alec is eager to learn about the possibilities of AI and the AIDAVA system, as the ideas in AIDAVA fit with where the company wants to go.
However, as Alec knows that verifying the output of a machine learning algorithm is still the key problem in AI, he believes that in a medical context, you will still need to have a person verifying the automatically curated data.

4.3 User profiles

A user profile is a collection of information and settings associated with an individual user within a system, application, or online platform. It represents a unique identity for a user and typically includes personal attributes (e.g. user's name, age, gender, …), and preference-related attributes, such as for example language, time zone, theme, notification preferences, or privacy settings.

User profiles serve several purposes, including personalization, customization, and user management. They allow users to maintain their unique identity, tailor their experience, and access personalised content or features within the system. User profiles also enable system administrators and operators to manage user accounts, track user behaviour, and provide targeted services or recommendations.

In AIDAVA, the patients' user profiles will also include attributes derived from the analysis of the personas that are relevant to customise the human-computer interaction (HCI) tools to be developed mainly in Task 5.3. Building on the highlights extracted from the personas - see below - we expect to use at minimum two attributes.



- Medical expertise (scale 1 to 10). This attribute will be used in the AIDAVA prototype to decide if a question that occurs in the curation workflow can be sent to the patient or should be sent to the expert curator
- Computer literacy (scale 1 to 10). This attribute can be used in the AIDAVA prototype to provide guidance and information to the user in a customised level of technical complexity

In addition, we will consider using other attributes such as skillset, motivation, frustration and aspirations (all based on a set of predefined values) when developing adapted human interaction and explanations in Task 5.3.

All attributes related to the user profiles will be stored in a dedicated data store ("User Directory"), as specified in the Deliverable 2.3 - Solution Design.

5 Summary and next steps

In conclusion, based on 39 in-depth interviews and a survey of 250 participants the AIDAVA project's Task 1.2 successfully developed eight personas: 2 patient personas, 2 data user personas, 2 data curator personas and 2 third party app developer personas. For these personas, we formulated foundation documents and visualised each persona also in the form of a personas canvas.

From the survey, we could conclude that nearly all (91%) of the survey respondents are interested in keeping their medical dossier in a central place, and the majority (78%) is furthermore willing to increase the quality of this dossier. The main reason for doing so is to ensure the treating physicians have access to the appropriate information, particularly in emergency situations.
Noticeable in the context of AIDAVA is that only 59% of the survey respondents would - currently - be interested to use a digital tool to support them. Interestingly, this number increases to 81% for the people with an active medical dossier. This seems to demonstrate that when there is no real need, people put more consideration to the risks than to the benefits. Thus, this finding supports the need to ensure high security and reliability of the system, together with user training covering both data literacy as well as the risk and benefits of a system like AIDAVA.

This is also supported by the interviews and the resulting personas, where we found that especially people who are not so tech-savvy are rather anxious regarding online insecurity and worried about threats of data leaks, viruses, and hackers. These people are also rather reluctant to share their medical data with anybody else but their treating medical staff. It is important to take these findings into account when developing training materials and user guidance. For example, it must be taken care that security aspects of the AIDAVA system are not only described in technical detail to cater for the needs of a system administrator, but are also well explained in layman language addressing the concerns of (not so tech-savvy) users.

This deliverable and the associated material and results created during Task 1.2. will support a user-centred development of the AIDAVA “AI-based data curation & publishing assistant”. The personas will help the developers to emphasise with the different user groups so that they can take decisions with the users in mind. This should lead to more user-centric decisions. In this sense, the personas complement the business requirements specified in D1.3.
For example, attributes derived from main characteristics of the personas, such as computer literacy and medical expertise, will be included in the default user profiles. These will be used in Task 5.3 to customise the HCI tools and user interface of the explainability and feedback layer for patients. In addition, we will consider using other attributes such as skillset, motivation, frustration and aspirations (all based on a set of predefined values) when developing adapted human interaction and explanations in Task 5.3.

The next steps to ensure that the material and results created during Task 1.2 will be used effectively for user-centred development of the AIDAVA “AI-based data curation & publishing assistant” include:

- Persona communication: We will share the personas (canvas and foundation documents) with the project partners and development teams and make sure everyone understands the personas and their characteristics relevant in shaping the AIDAVA prototype’s design and features.

- Feature prioritisation: To ensure that the most critical features and functionalities (from users' point of view) are developed first, we will focus on addressing the pain points/frustrations, goals and preferences of the personas, who represent the target users of the AIDAVA prototype.
- Customised explanations and guidance: We will consider attributes derived from the personas (such as computer literacy, skillset, motivation…), to ensure that human interaction and explanation components of the AIDAVA prototype as well as user guidance elements and training procedures are developed according to the needs of the target groups.
- User testing: We will test the G1 prototype of AIDAVA with representative users and gather their feedback. By comparing the users' experiences and feedback with the personas we can identify any gaps and areas for improvement, and update and refine the personas accordingly.

7 References

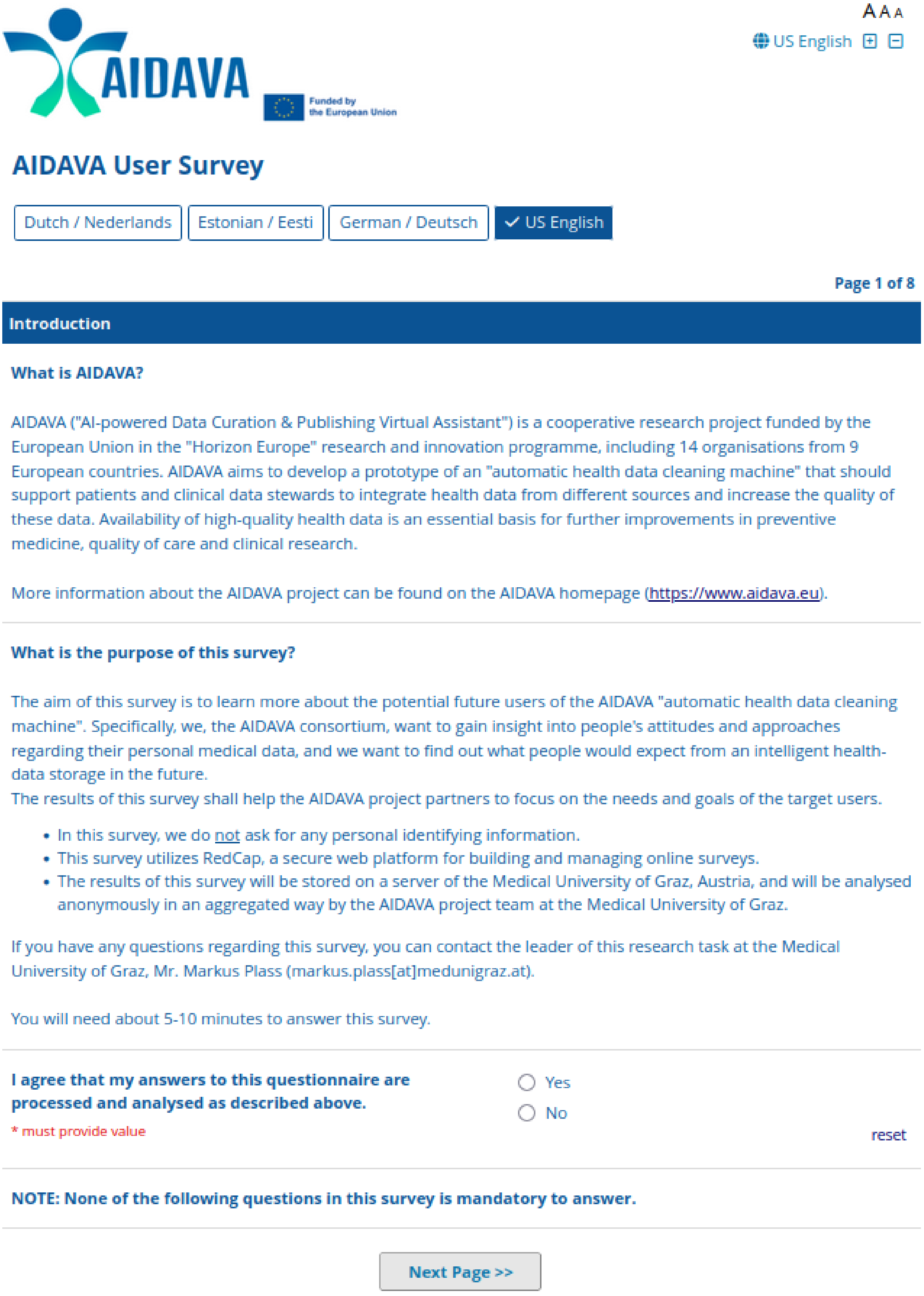
[1] T. Adlin and J. Pruitt, *The Essential Persona Lifecycle: Your Guide to Building and Using Personas*. Morgan Kaufmann, 2010.

[2] A. Holzinger, M. Kargl, B. Kipperer, P. Regitnig, M. Plass, and H. Müller, “Personas for Artificial Intelligence (AI) an Open Source Toolbox,” *IEEE Access*, vol. 10, pp. 23732–23747, 2022.

[3] *PERSONAS*. Github. Accessed: May 17, 2023. [Online]. Available: https://github.com/human-centered-ai-lab/PERSONAS

8 Annexes

8.1 Survey



AIDAVA User Survey

Introduction

What is AIDAVA?

AIDAVA ("AI-powered Data Curation & Publishing Virtual Assistant") is a cooperative research project funded by the European Union in the "Horizon Europe" research and innovation programme, including 14 organisations from 9 European countries. AIDAVA aims to develop a prototype of an "automatic health data cleaning machine" that should support patients and clinical data stewards to integrate health data from different sources and increase the quality of these data. Availability of high-quality health data is an essential basis for further improvements in preventive medicine, quality of care and clinical research.

More information about the AIDAVA project can be found on the AIDAVA homepage (https://www.aidava.eu).

What is the purpose of this survey?

The aim of this survey is to learn more about the potential future users of the AIDAVA "automatic health data cleaning machine". Specifically, we, the AIDAVA consortium, want to gain insight into people's attitudes and approaches regarding their personal medical data, and we want to find out what people would expect from an intelligent health-data storage in the future.
The results of this survey shall help the AIDAVA project partners to focus on the needs and goals of the target users.

- In this survey, we do not ask for any personal identifying information.
- This survey utilizes RedCap, a secure web platform for building and managing online surveys.
- The results of this survey will be stored on a server of the Medical University of Graz, Austria, and will be analysed anonymously in an aggregated way by the AIDAVA project team at the Medical University of Graz.

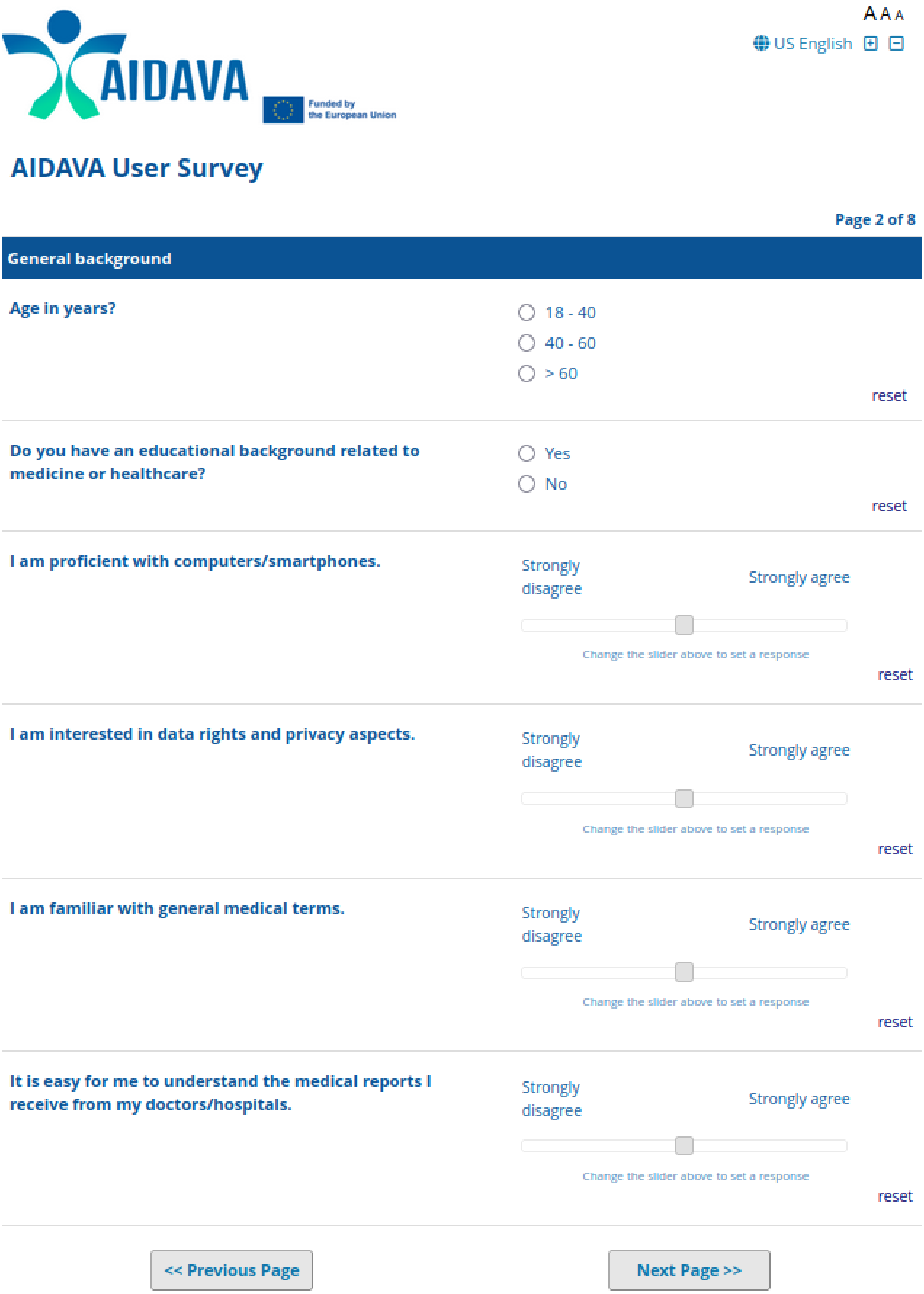
If you have any questions regarding this survey, you can contact the leader of this research task at the Medical University of Graz, Mr. Markus Plass (markus.plass[at]medunigraz.at).

You will need about 5-10 minutes to answer this survey.

I agree that my answers to this questionnaire are processed and analysed as described above.
* must provide value

○ Yes
○ No

NOTE: None of the following questions in this survey is mandatory to answer.



AIDAVA User Survey

Page 2 of 8

General background

Age in years?

○ 18 - 40
○ 40 - 60
○ > 60

reset

Do you have an educational background related to medicine or healthcare?

○ Yes
○ No

reset

I am proficient with computers/smartphones.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset

I am interested in data rights and privacy aspects.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset

I am familiar with general medical terms.

Strongly disagree — Strongly agree

Change the slider above to set a response

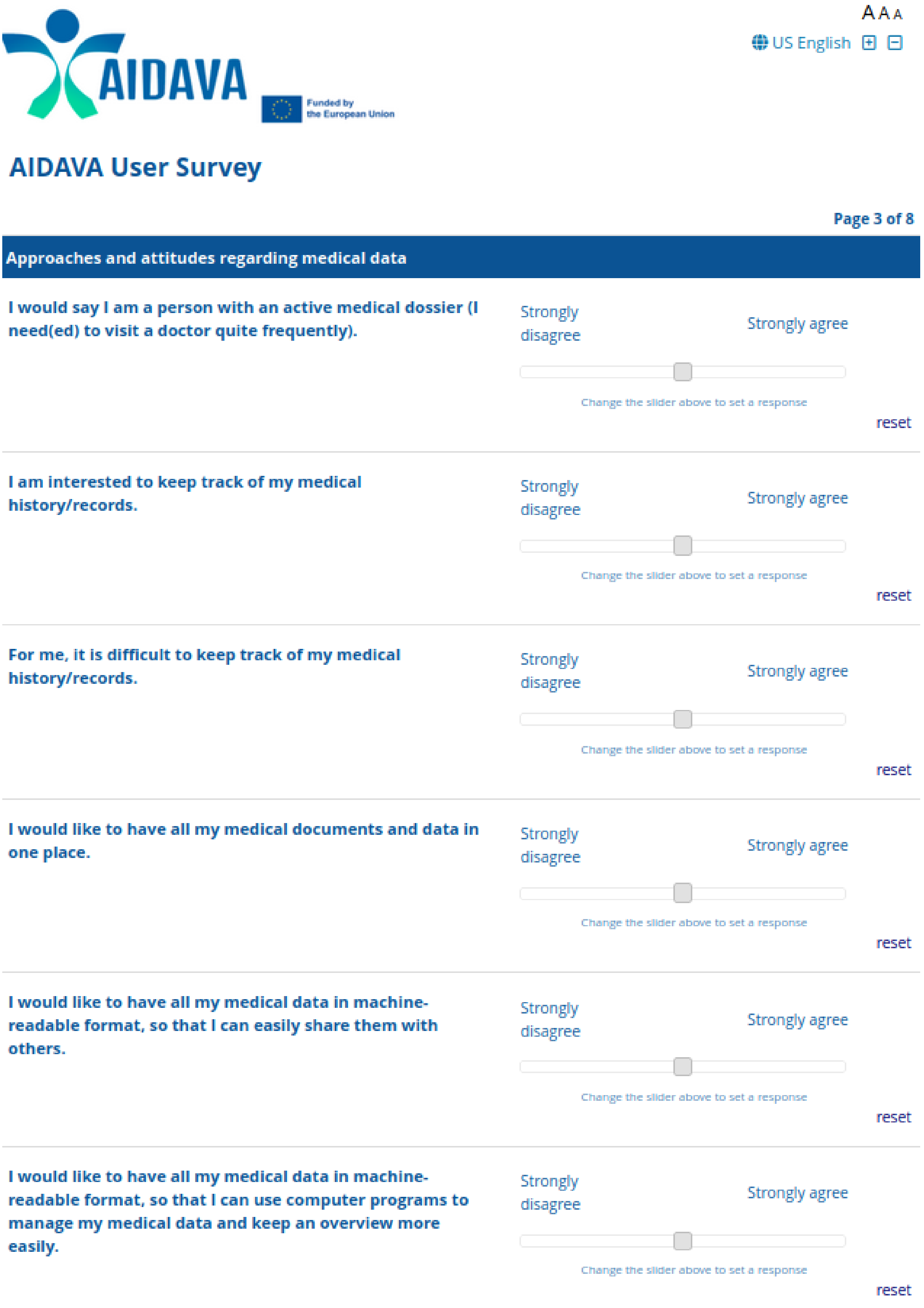
reset

It is easy for me to understand the medical reports I receive from my doctors/hospitals.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset



AIDAVA User Survey

Page 3 of 8

Approaches and attitudes regarding medical data

I would say I am a person with an active medical dossier (I need(ed) to visit a doctor quite frequently).

Strongly disagree — Strongly agree

Change the slider above to set a response

reset

I am interested to keep track of my medical history/records.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset

For me, it is difficult to keep track of my medical history/records.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset

I would like to have all my medical documents and data in one place.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset

I would like to have all my medical data in machine-readable format, so that I can easily share them with others.

Strongly disagree — Strongly agree

Change the slider above to set a response

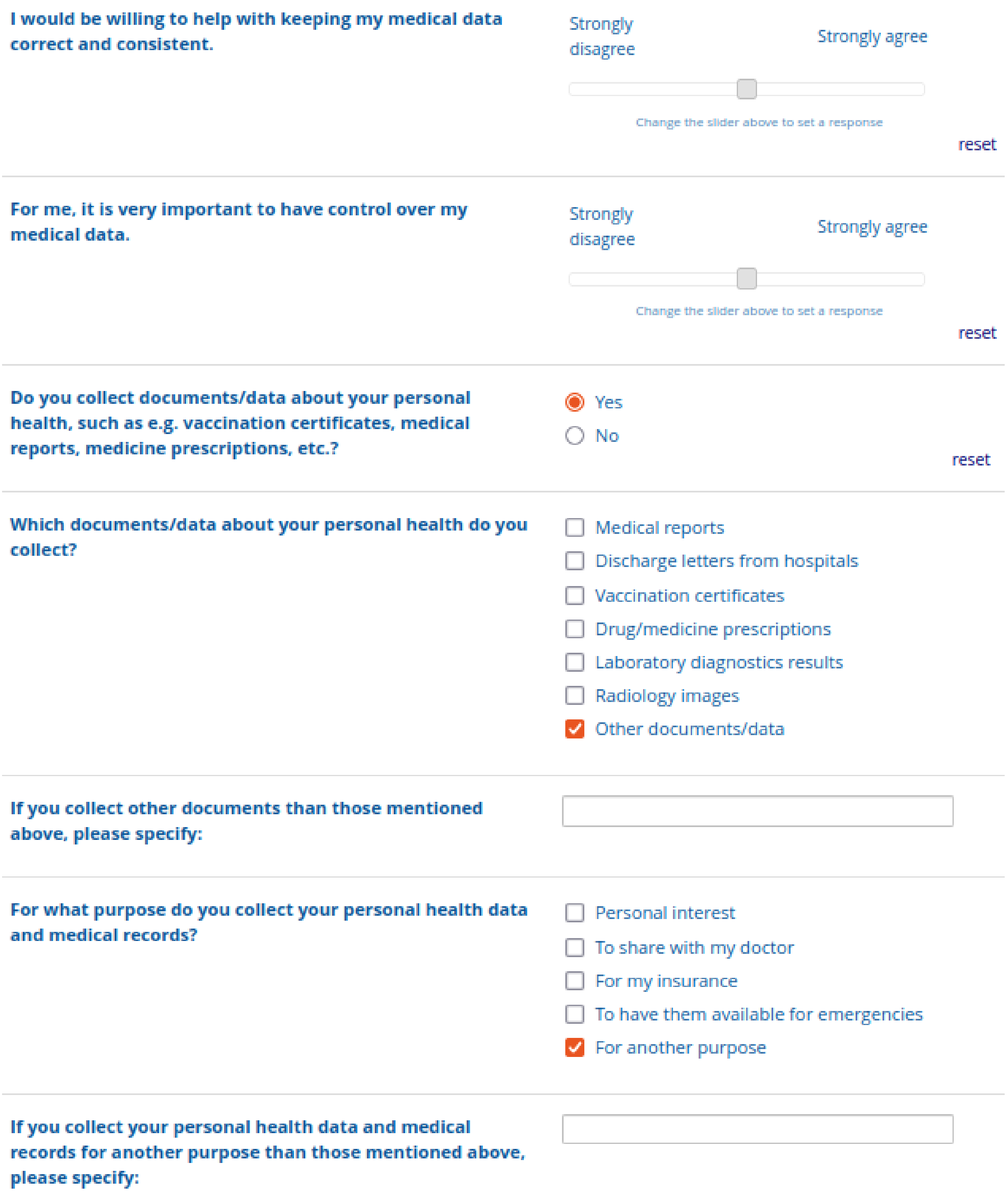
reset

I would like to have all my medical data in machine-readable format, so that I can use computer programs to manage my medical data and keep an overview more easily.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset



I would be willing to help with keeping my medical data correct and consistent.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset

For me, it is very important to have control over my medical data.

Strongly disagree — Strongly agree

Change the slider above to set a response

reset

Do you collect documents/data about your personal health, such as e.g. vaccination certificates, medical reports, medicine prescriptions, etc.?

● Yes
○ No

reset

Which documents/data about your personal health do you collect?

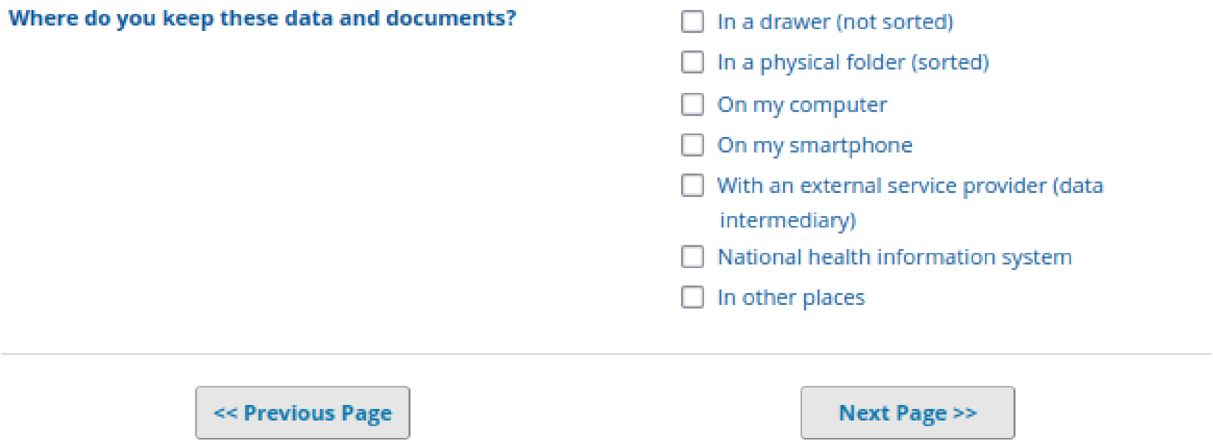
☐ Medical reports
☐ Discharge letters from hospitals
☐ Vaccination certificates
☐ Drug/medicine prescriptions
☐ Laboratory diagnostics results
☐ Radiology images
☑ Other documents/data

If you collect other documents than those mentioned above, please specify:

For what purpose do you collect your personal health data and medical records?

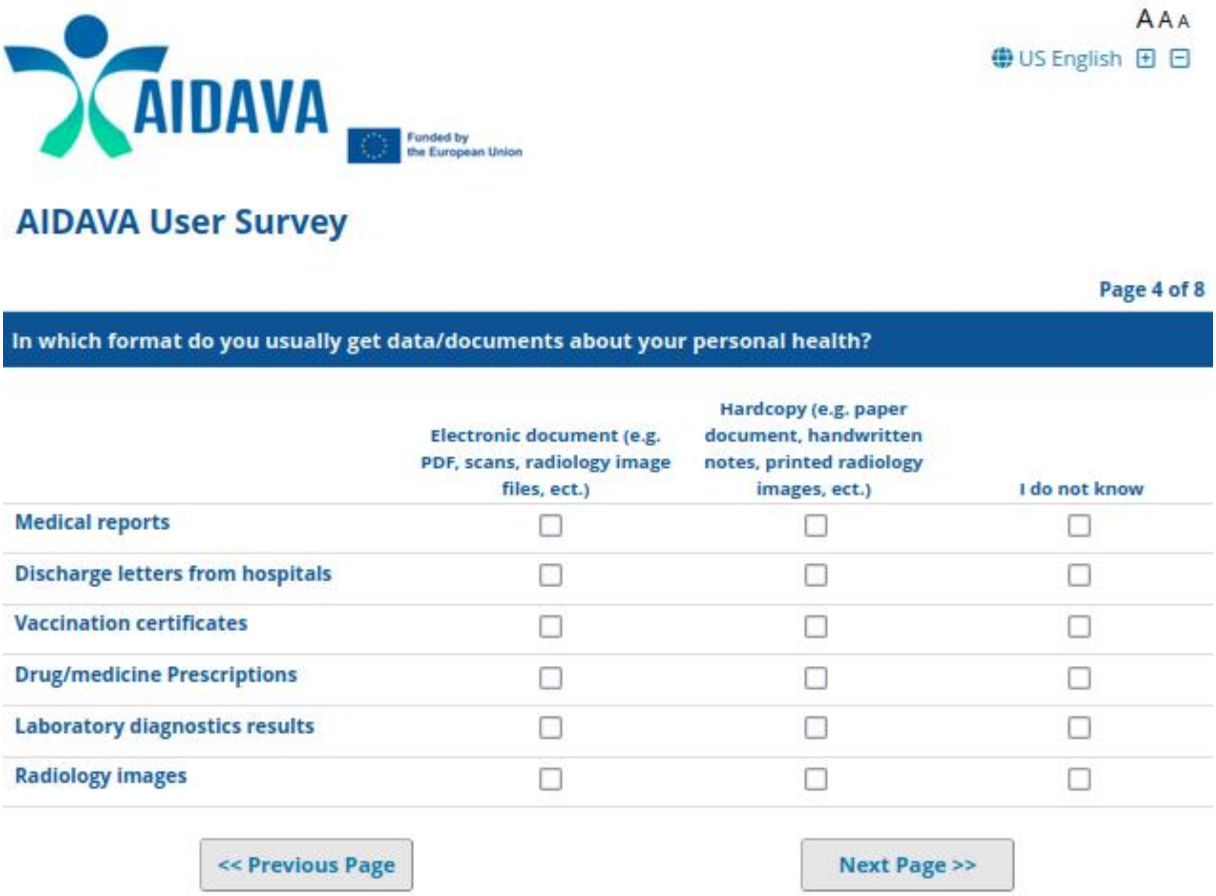
☐ Personal interest
☐ To share with my doctor
☐ For my insurance
☐ To have them available for emergencies
☑ For another purpose

If you collect your personal health data and medical records for another purpose than those mentioned above, please specify:



Where do you keep these data and documents?

☐ In a drawer (not sorted)
☐ In a physical folder (sorted)
☐ On my computer
☐ On my smartphone
☐ With an external service provider (data intermediary)
☐ National health information system
☐ In other places

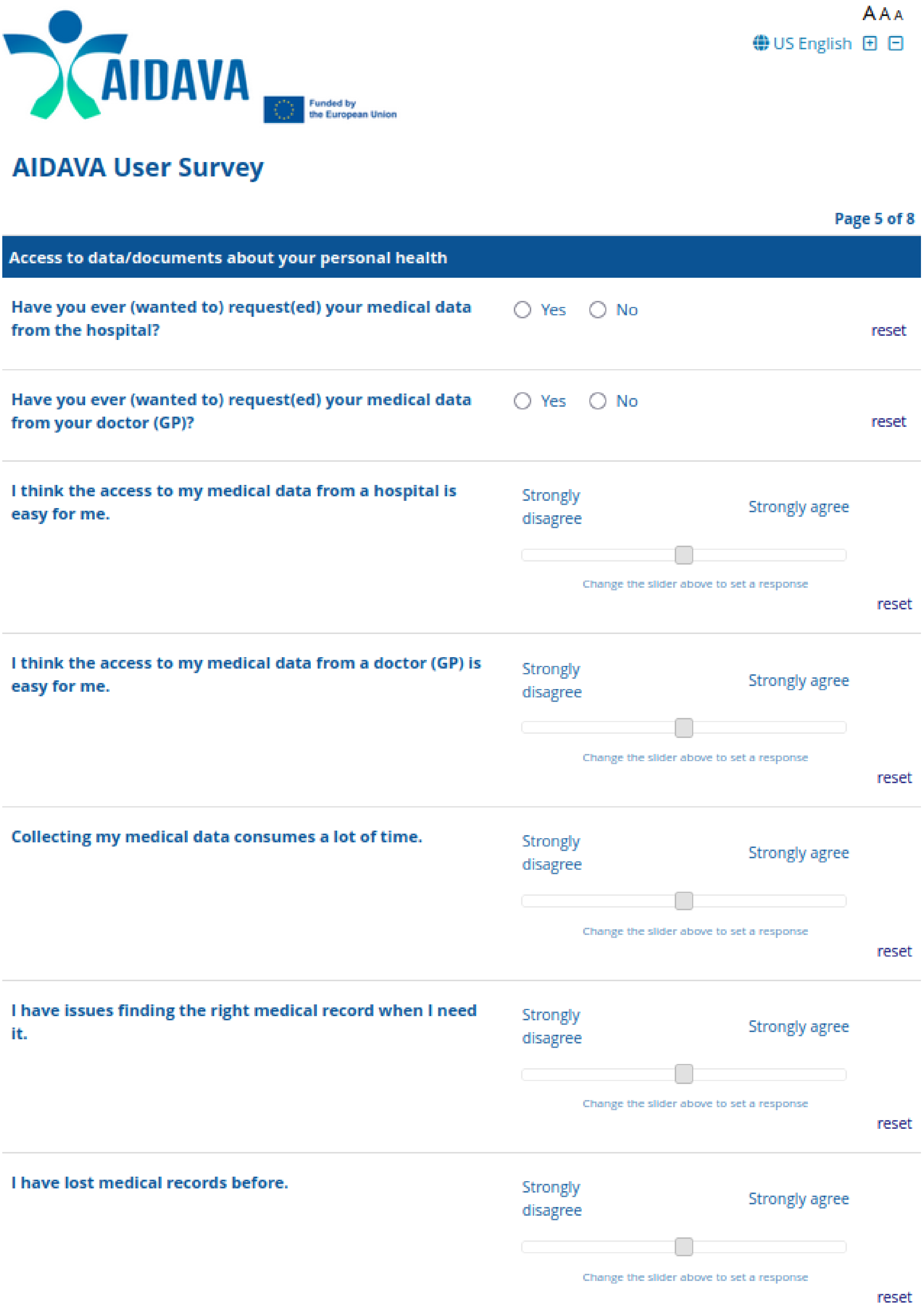


AIDAVA User Survey

Page 4 of 8

In which format do you usually get data/documents about your personal health?

	Electronic document (e.g. PDF, scans, radiology image files, ect.)	Hardcopy (e.g. paper document, handwritten notes, printed radiology images, ect.)	I do not know
Medical reports	☐	☐	☐
Discharge letters from hospitals	☐	☐	☐
Vaccination certificates	☐	☐	☐
Drug/medicine Prescriptions	☐	☐	☐
Laboratory diagnostics results	☐	☐	☐
Radiology images	☐	☐	☐



AIDAVA User Survey

Page 5 of 8

Access to data/documents about your personal health

Have you ever (wanted to) request(ed) your medical data from the hospital? ○ Yes ○ No reset

Have you ever (wanted to) request(ed) your medical data from your doctor (GP)? ○ Yes ○ No reset

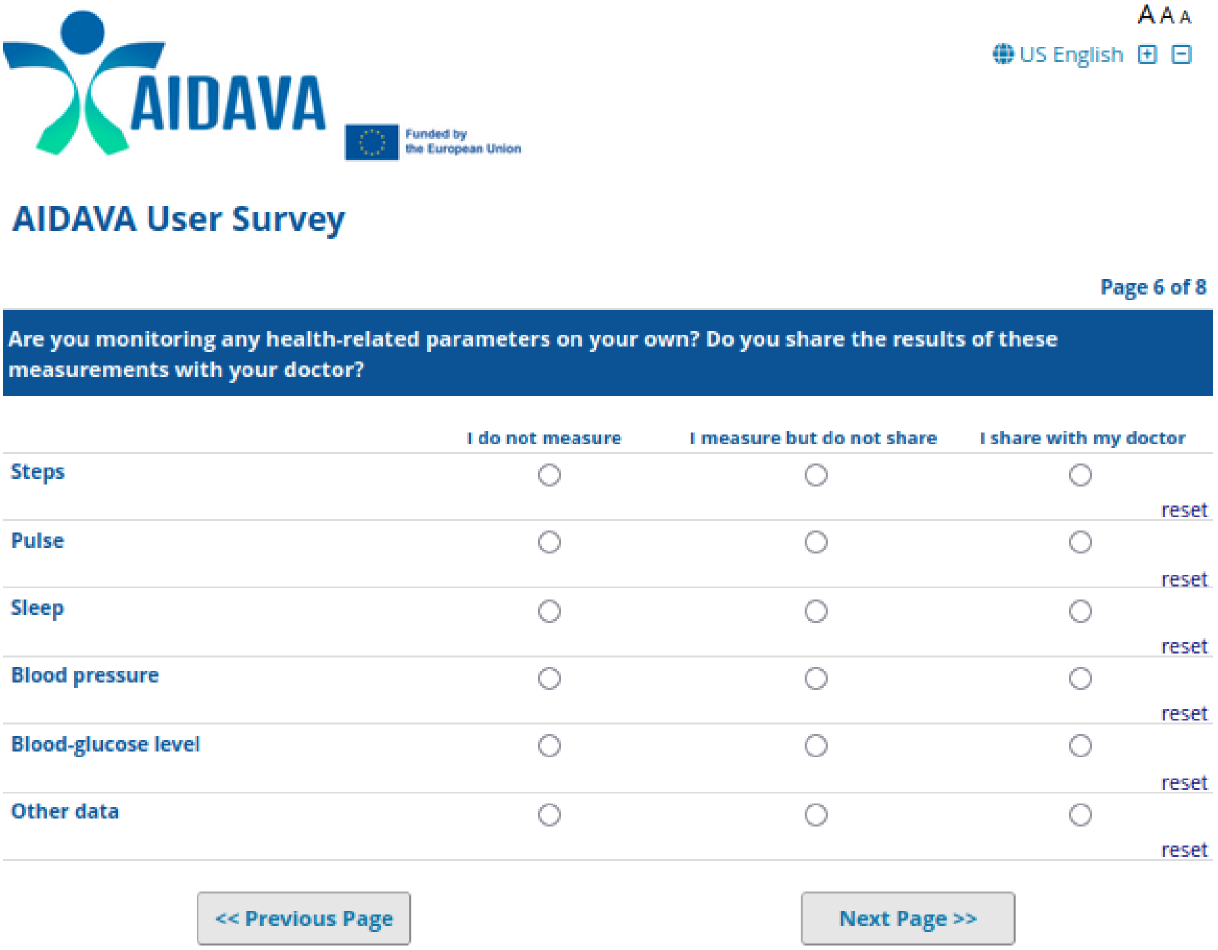
I think the access to my medical data from a hospital is easy for me. Strongly disagree — Strongly agree
Change the slider above to set a response
reset

I think the access to my medical data from a doctor (GP) is easy for me. Strongly disagree — Strongly agree
Change the slider above to set a response
reset

Collecting my medical data consumes a lot of time. Strongly disagree — Strongly agree
Change the slider above to set a response
reset

I have issues finding the right medical record when I need it. Strongly disagree — Strongly agree
Change the slider above to set a response
reset

I have lost medical records before. Strongly disagree — Strongly agree
Change the slider above to set a response
reset

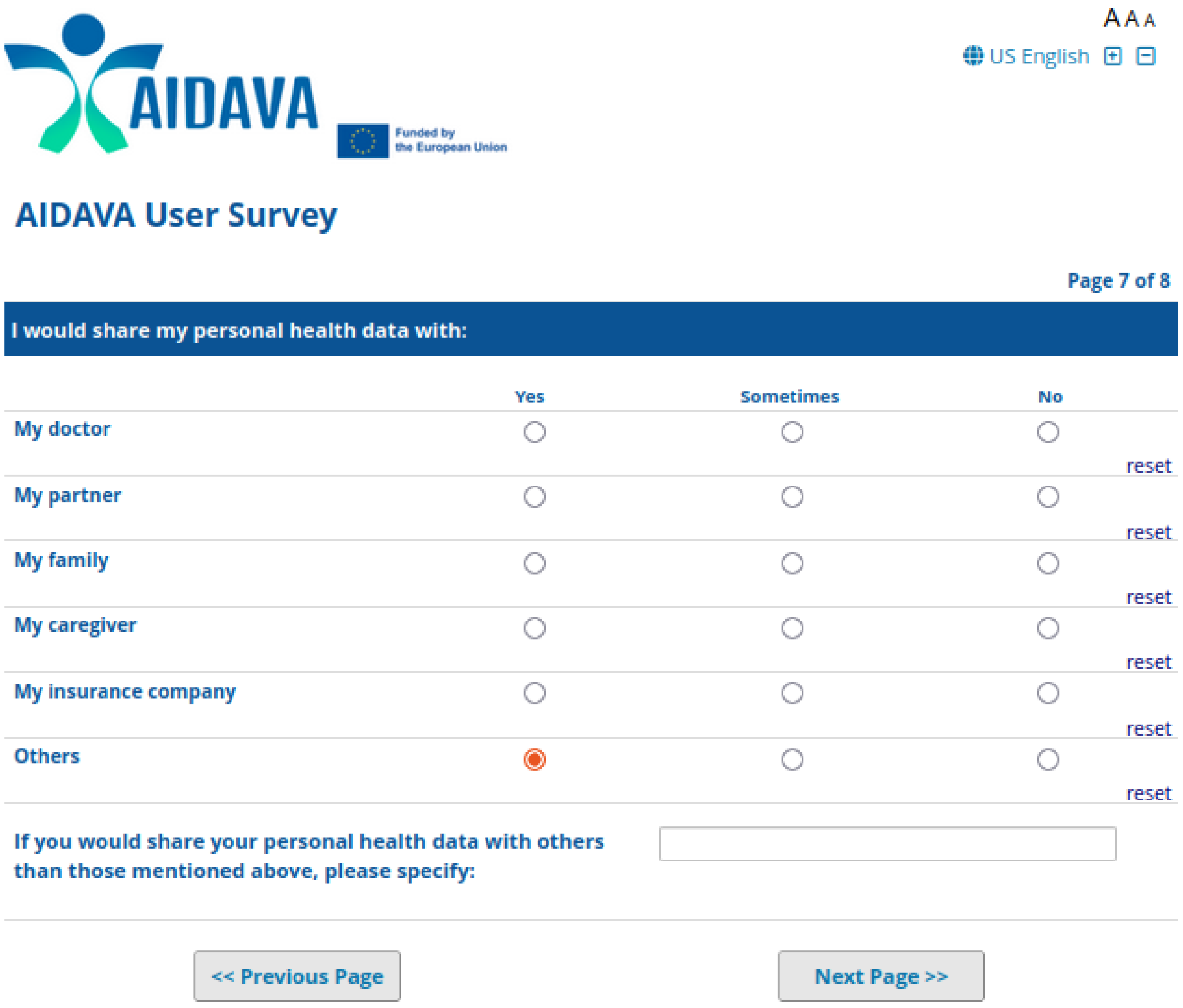


AIDAVA User Survey

Page 6 of 8

Are you monitoring any health-related parameters on your own? Do you share the results of these measurements with your doctor?

	I do not measure	I measure but do not share	I share with my doctor	
Steps	○	○	○	reset
Pulse	○	○	○	reset
Sleep	○	○	○	reset
Blood pressure	○	○	○	reset
Blood-glucose level	○	○	○	reset
Other data	○	○	○	reset



AIDAVA User Survey

Page 7 of 8

I would share my personal health data with:

	Yes	Sometimes	No	
My doctor	○	○	○	reset
My partner	○	○	○	reset
My family	○	○	○	reset
My caregiver	○	○	○	reset
My insurance company	○	○	○	reset
Others	●	○	○	reset

If you would share your personal health data with others than those mentioned above, please specify:

<< Previous Page

Next Page >>

AIDAVA User Survey

Page 8 of 8

Expectations and attitude towards "intelligent health data system"

If there would be an "intelligent health data system" in the future, how would you rate the importance of the following functionalities of such a system?

The system shall provide search and query functionality, so that I can easily retrieve specific health data.

Not important — Very important

Change the slider above to set a response

reset

The system shall inform me about any inconsistencies in my health data.

Not important — Very important

Change the slider above to set a response

reset

The system shall inform me about risk factors found/deducted from my health data.

Not important — Very important

Change the slider above to set a response

reset

The system shall support me in sharing my health data with medical doctors.

Not important — Very important

Change the slider above to set a response

reset

The system shall help me finding patients similar to me and propose relevant self-help groups.

Not important — Very important

Change the slider above to set a response

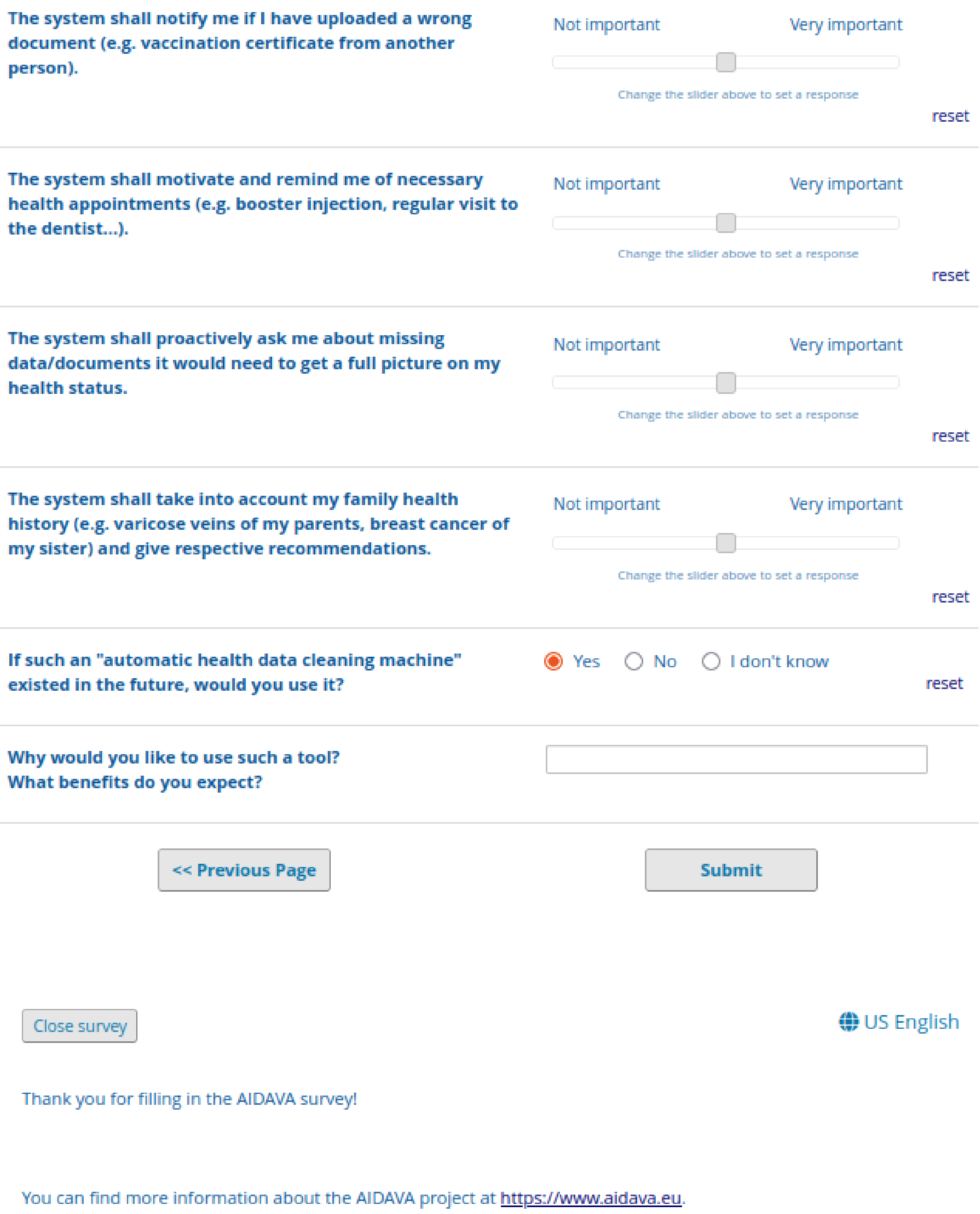
reset

In a medical emergency, the system shall inform first responders about relevant aspects in my medical history.

Not important — Very important

Change the slider above to set a response

reset



The system shall notify me if I have uploaded a wrong document (e.g. vaccination certificate from another person).

The system shall motivate and remind me of necessary health appointments (e.g. booster injection, regular visit to the dentist…).

The system shall proactively ask me about missing data/documents it would need to get a full picture on my health status.

The system shall take into account my family health history (e.g. varicose veins of my parents, breast cancer of my sister) and give respective recommendations.

If such an "automatic health data cleaning machine" existed in the future, would you use it?

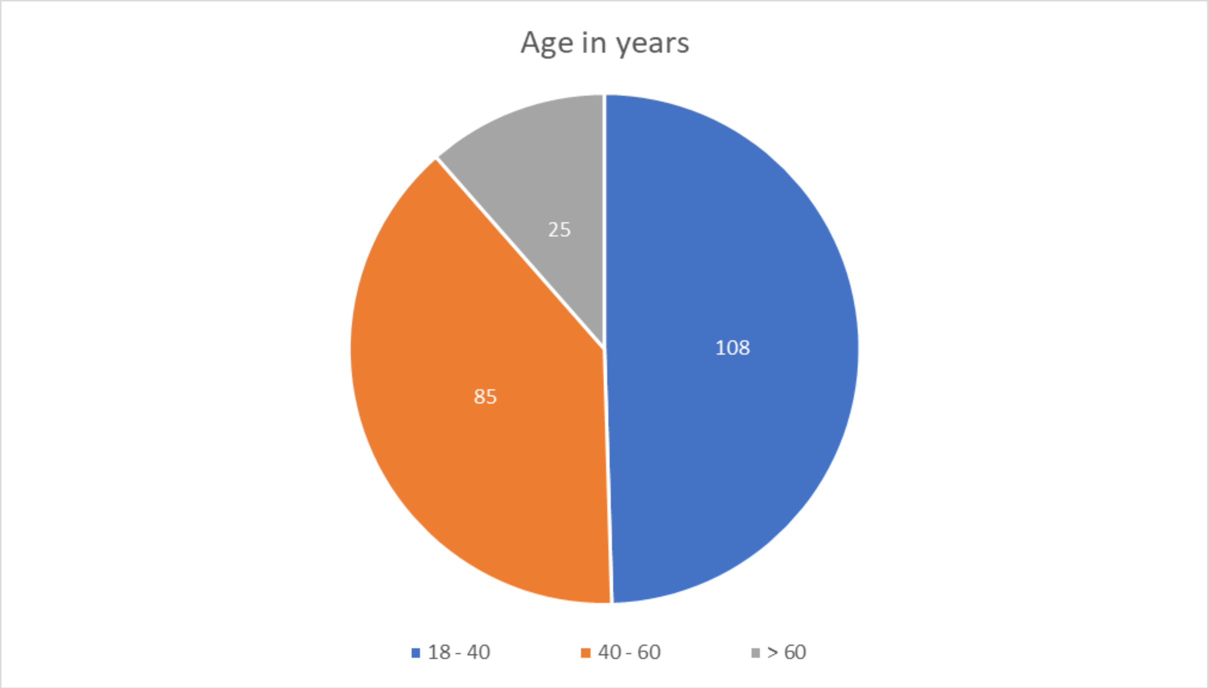
● Yes ○ No ○ I don't know

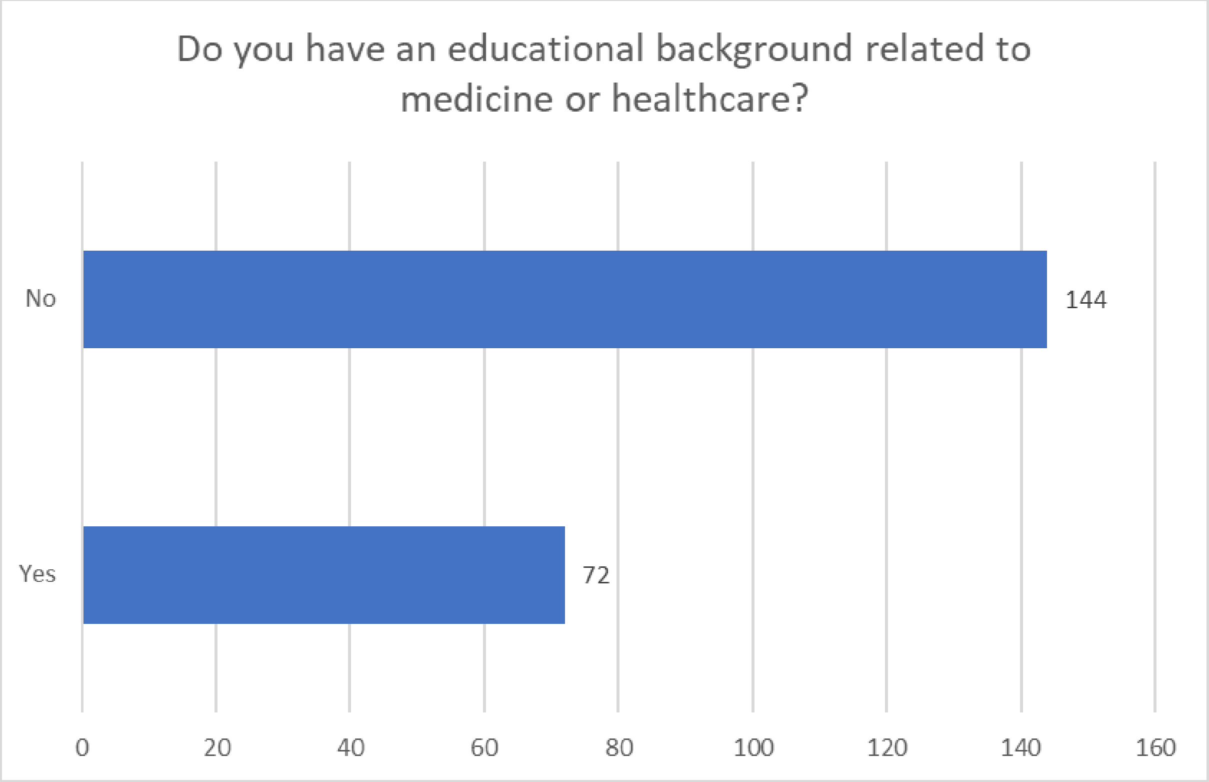
Why would you like to use such a tool?
What benefits do you expect?

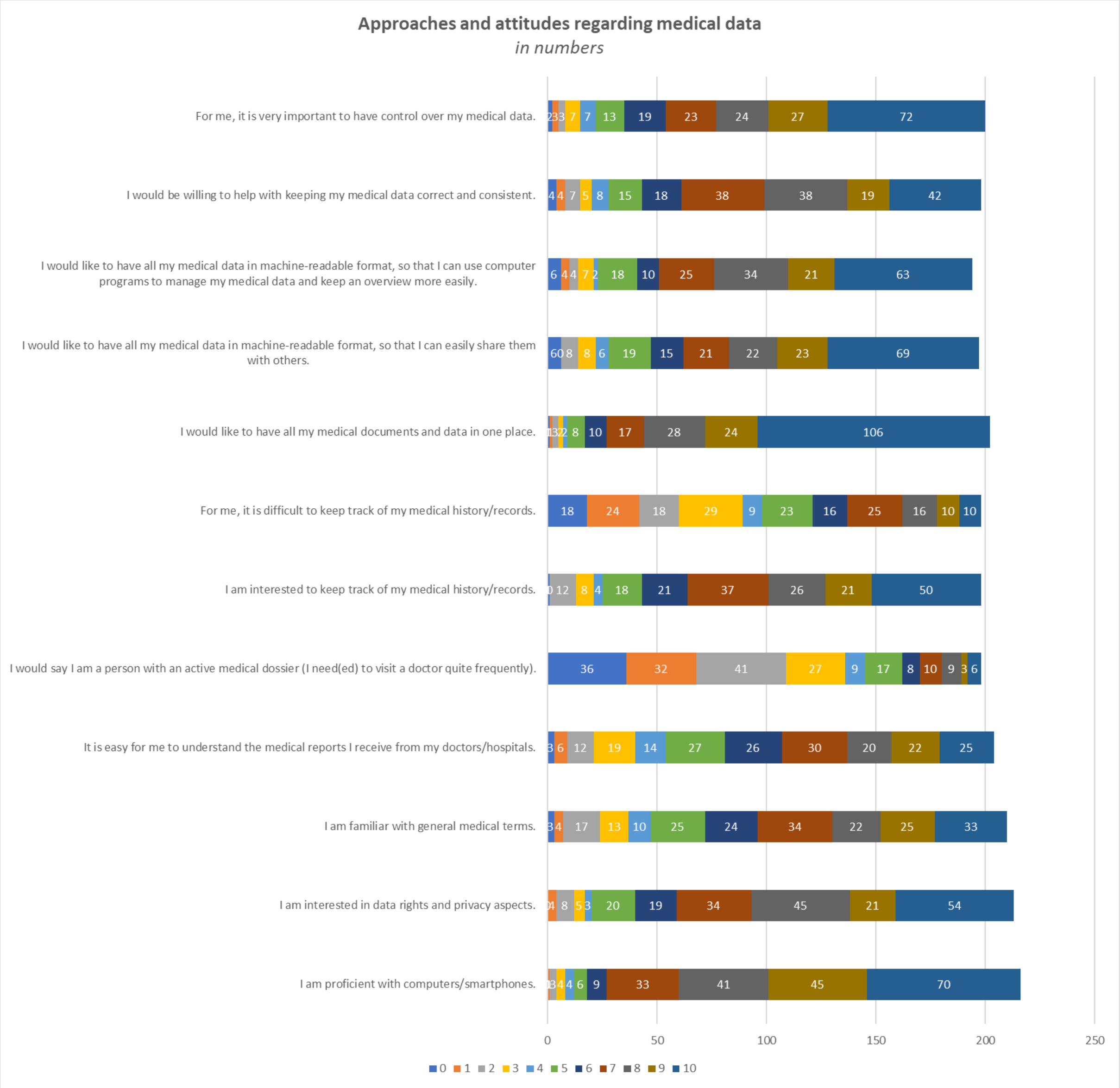
Thank you for filling in the AIDAVA survey!

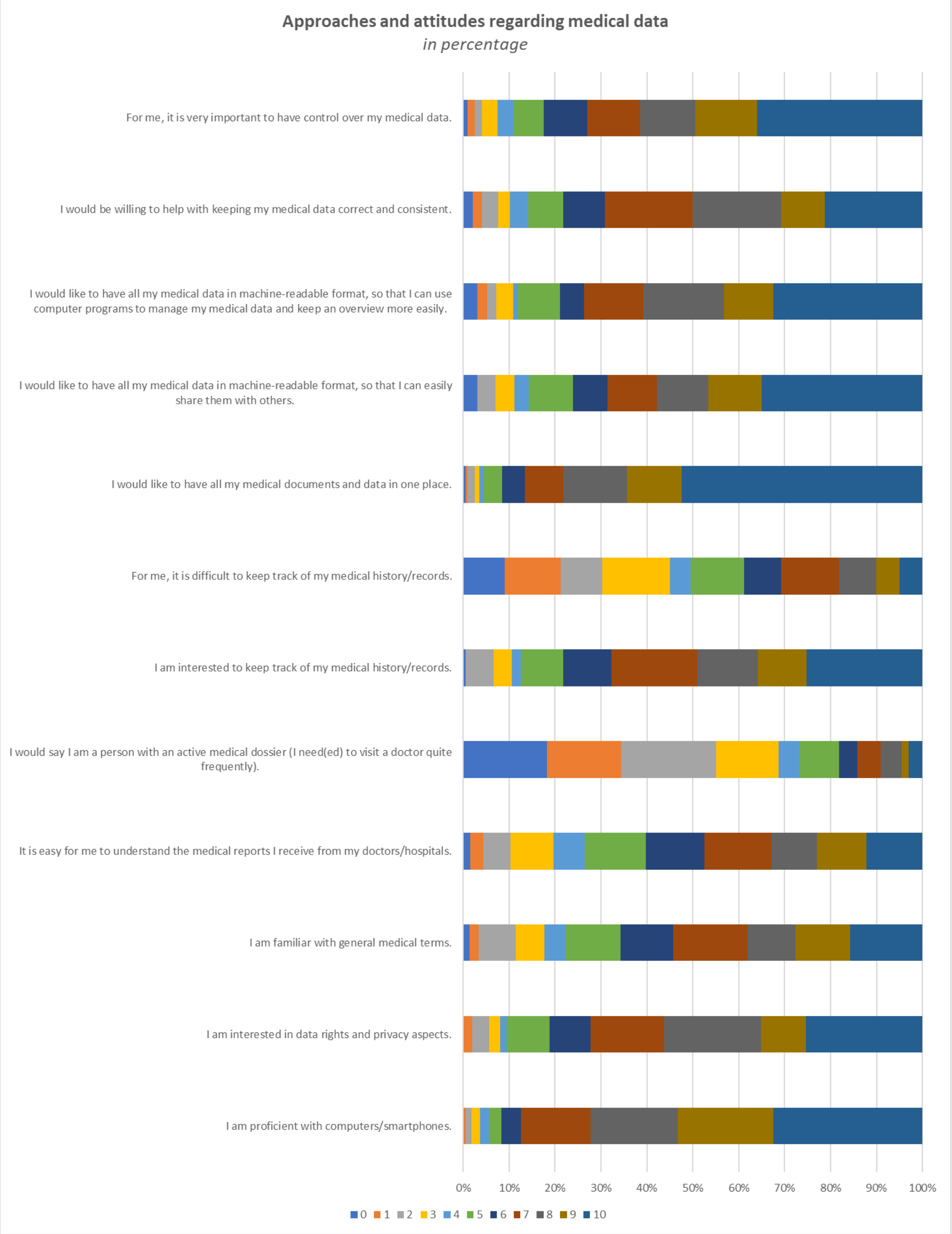
You can find more information about the AIDAVA project at https://www.aidava.eu.

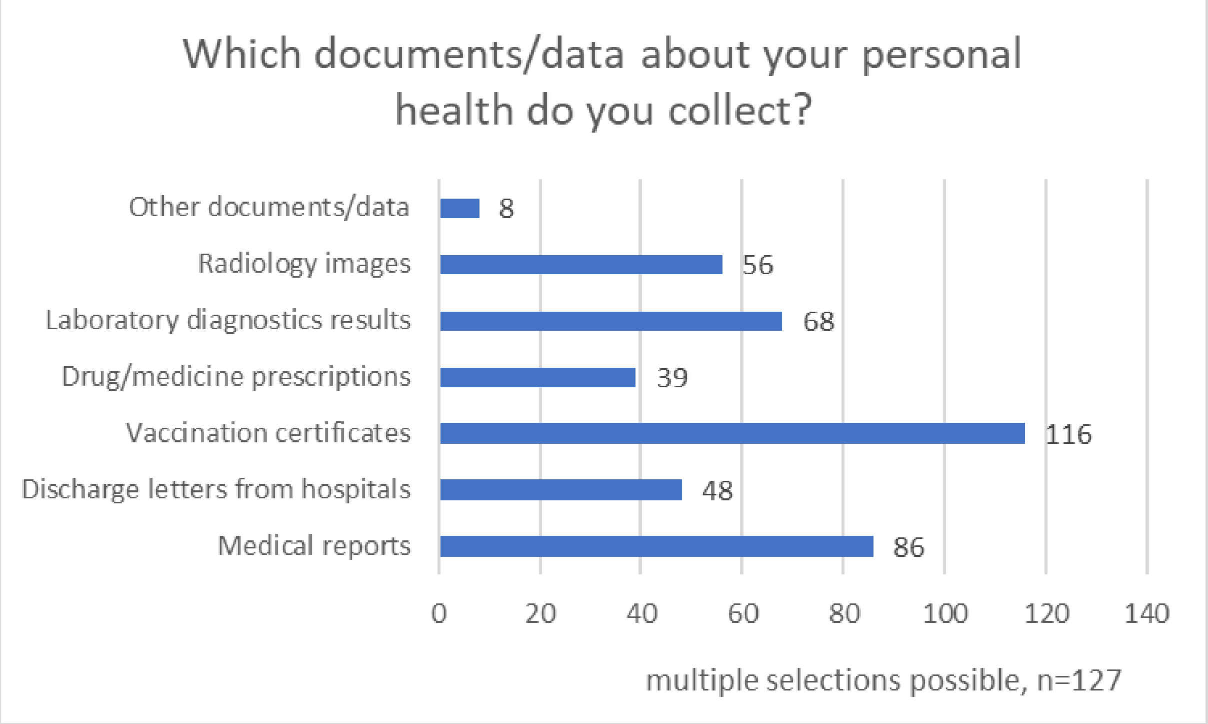
8.2 Further results from the survey

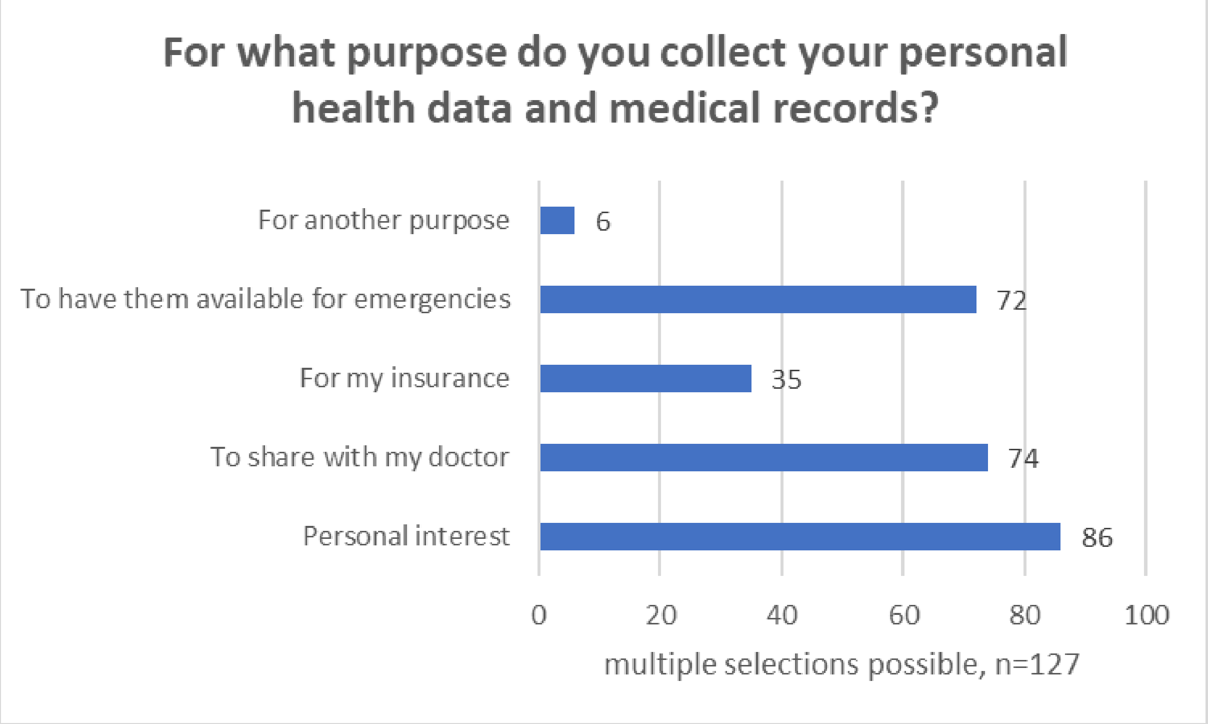


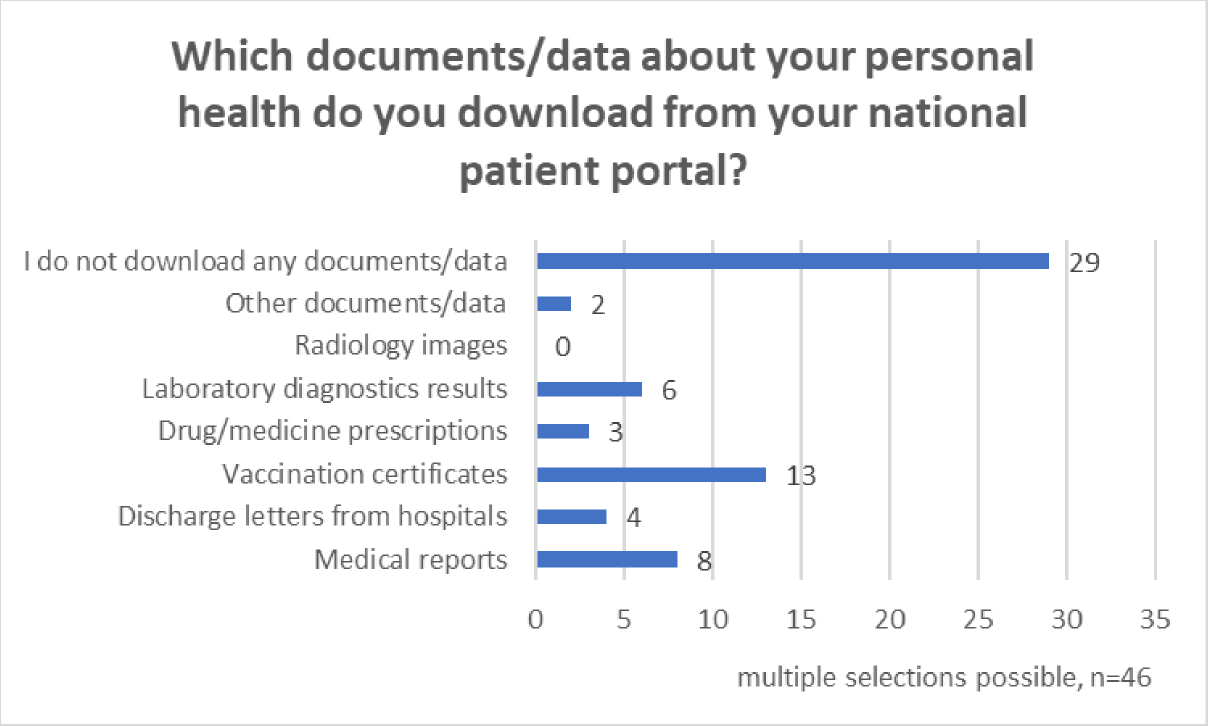


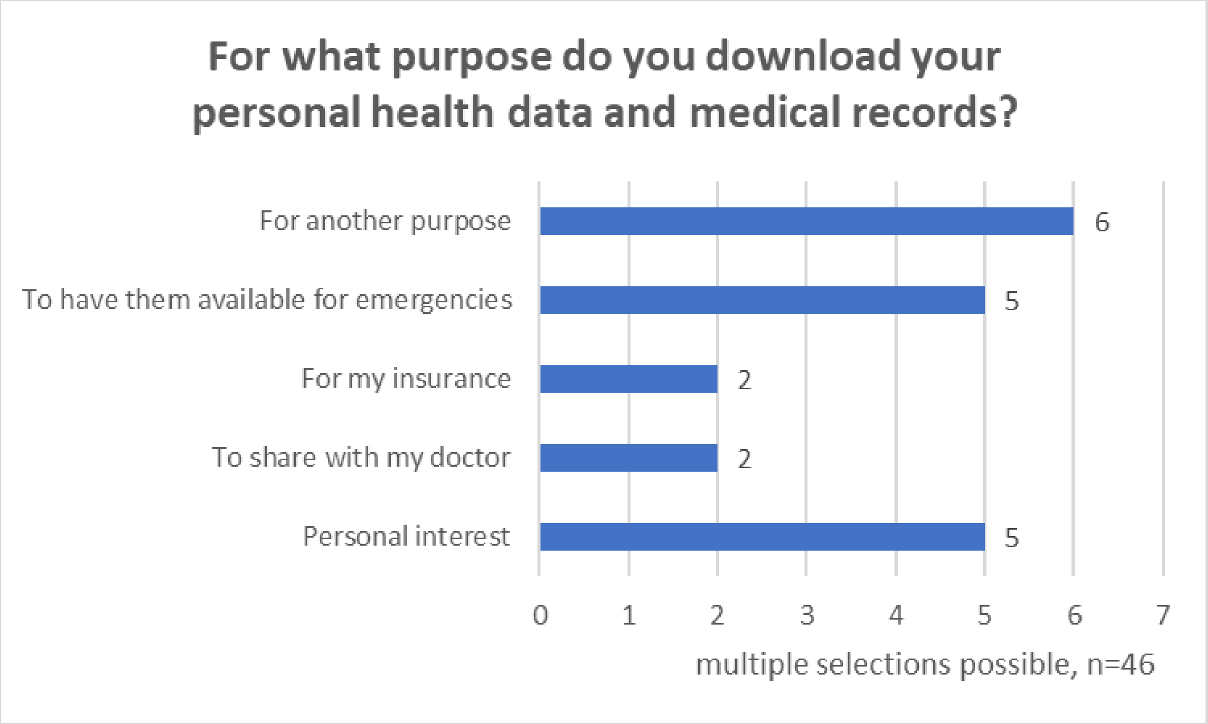

Approaches and attitudes regarding medical data
in numbers
For me, it is very important to have control over my medical data.
2 3 3 7 7 13 19 23 24 27 72
I would be willing to help with keeping my medical data correct and consistent.
4 4 7 5 8 15 18 38 38 19 42
I would like to have all my medical data in machine-readable format, so that I can use computer programs to manage my medical data and keep an overview more easily.
6 4 4 7 2 18 10 25 34 21 63
I would like to have all my medical data in machine-readable format, so that I can easily share them with others.
6 0 8 8 6 19 15 21 22 23 69
I would like to have all my medical documents and data in one place.
1 3 2 2 8 10 17 28 24 106
For me, it is difficult to keep track of my medical history/records.
18 24 18 29 9 23 16 25 16 10 10
I am interested to keep track of my medical history/records.
0 12 8 4 18 21 37 26 21 50
I would say I am a person with an active medical dossier (I need(ed) to visit a doctor quite frequently).
36 32 41 27 9 17 8 10 9 3 6
It is easy for me to understand the medical reports I receive from my doctors/hospitals.
3 6 12 19 14 27 26 30 20 22 25
I am familiar with general medical terms.
3 4 17 13 10 25 24 34 22 25 33
I am interested in data rights and privacy aspects.
0 4 8 5 3 20 19 34 45 21 54
I am proficient with computers/smartphones.
0 3 4 4 6 9 33 41 45 70
0
50
100
150
200
250
0
1
2
3
4
5
6
7
8
9
10

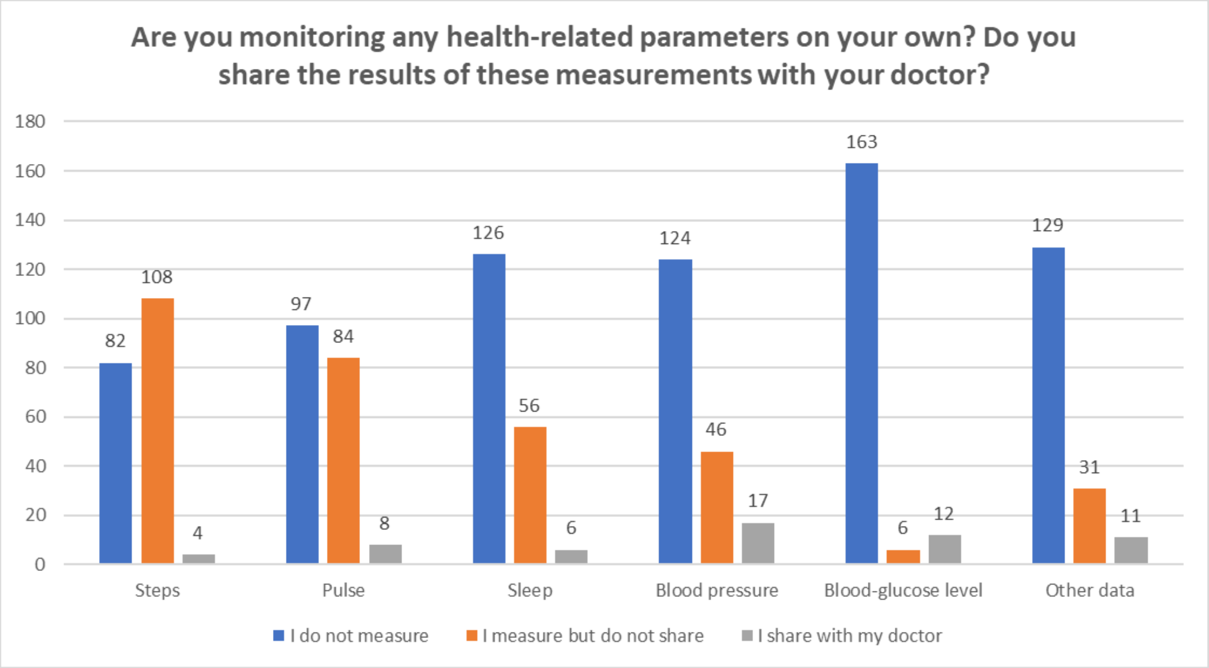

Approaches and attitudes regarding medical data
in percentage
For me, it is very important to have control over my medical data.
I would be willing to help with keeping my medical data correct and consistent.
I would like to have all my medical data in machine-readable format, so that I can use computer programs to manage my medical data and keep an overview more easily.
I would like to have all my medical data in machine-readable format, so that I can easily share them with others.
I would like to have all my medical documents and data in one place.
For me, it is difficult to keep track of my medical history/records.
I am interested to keep track of my medical history/records.
I would say I am a person with an active medical dossier (I need(ed) to visit a doctor quite frequently).
It is easy for me to understand the medical reports I receive from my doctors/hospitals.
I am familiar with general medical terms.
I am interested in data rights and privacy aspects.
I am proficient with computers/smartphones.
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
0
1
2
3
4
5
6
7
8
9
10

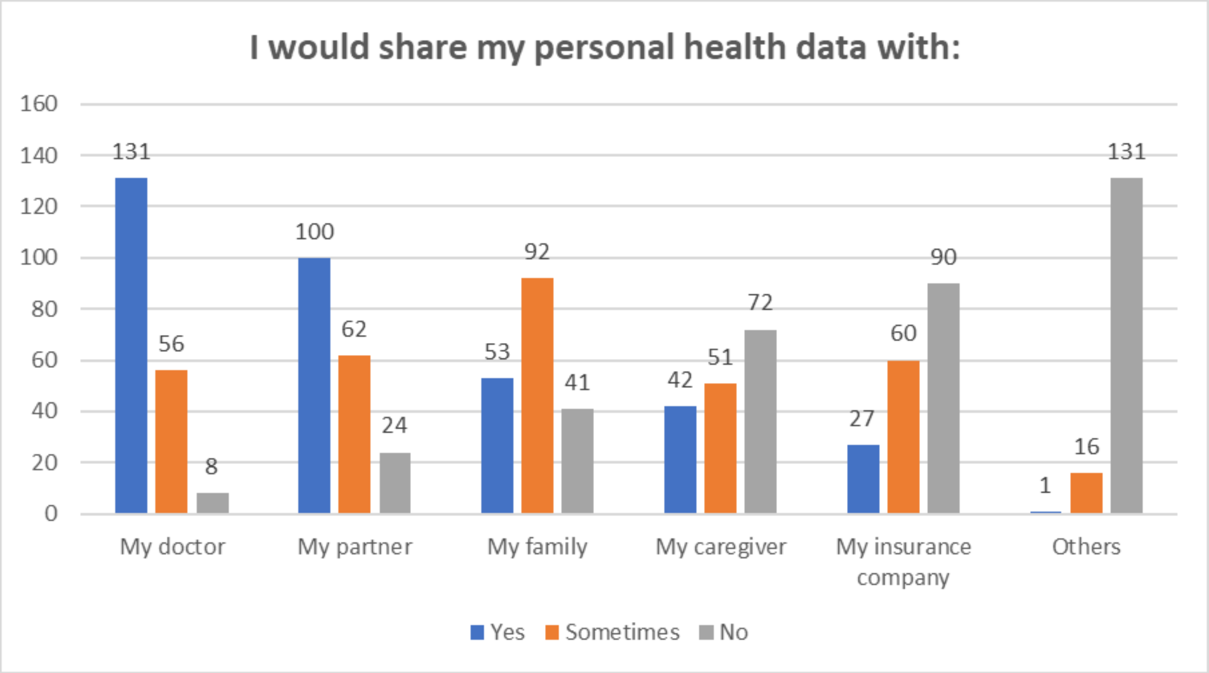

Which documents/data about your personal health do you collect?
Other documents/data
8
Radiology images
56
Laboratory diagnostics results
68
Drug/medicine prescriptions
39
Vaccination certificates
116
Discharge letters from hospitals
48
Medical reports
86
0
20
40
60
80
100
120
140
multiple selections possible, n=127


For what purpose do you collect your personal health data and medical records?
For another purpose
6
To have them available for emergencies
72
For my insurance
35
To share with my doctor
74
Personal interest
86
0
20
40
60
80
100
multiple selections possible, n=127

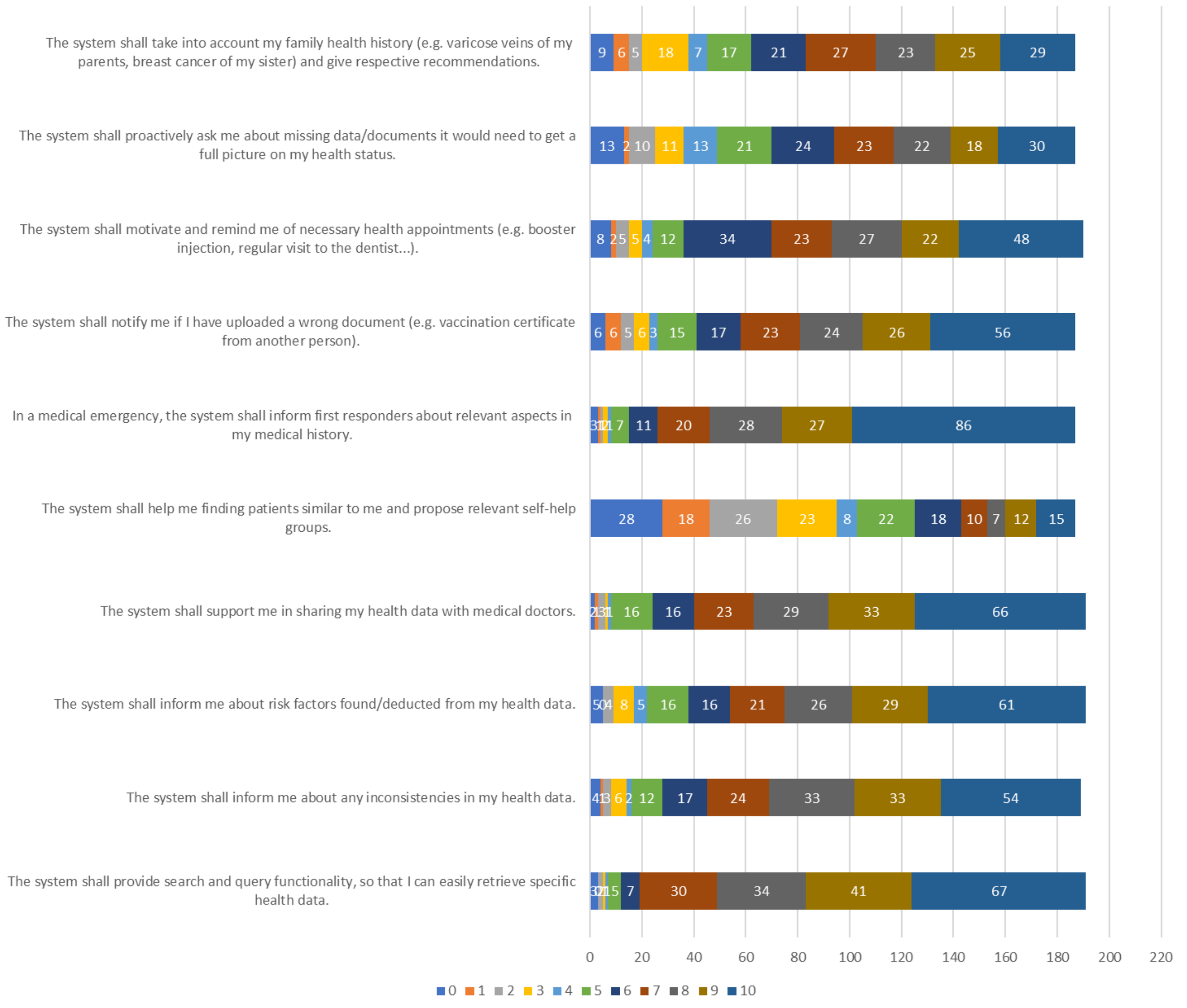

Which documents/data about your personal health do you download from your national patient portal?
I do not download any documents/data
29
Other documents/data
2
Radiology images
0
Laboratory diagnostics results
6
Drug/medicine prescriptions
3
Vaccination certificates
13
Discharge letters from hospitals
4
Medical reports
8
0
5
10
15
20
25
30
35
multiple selections possible, n=46

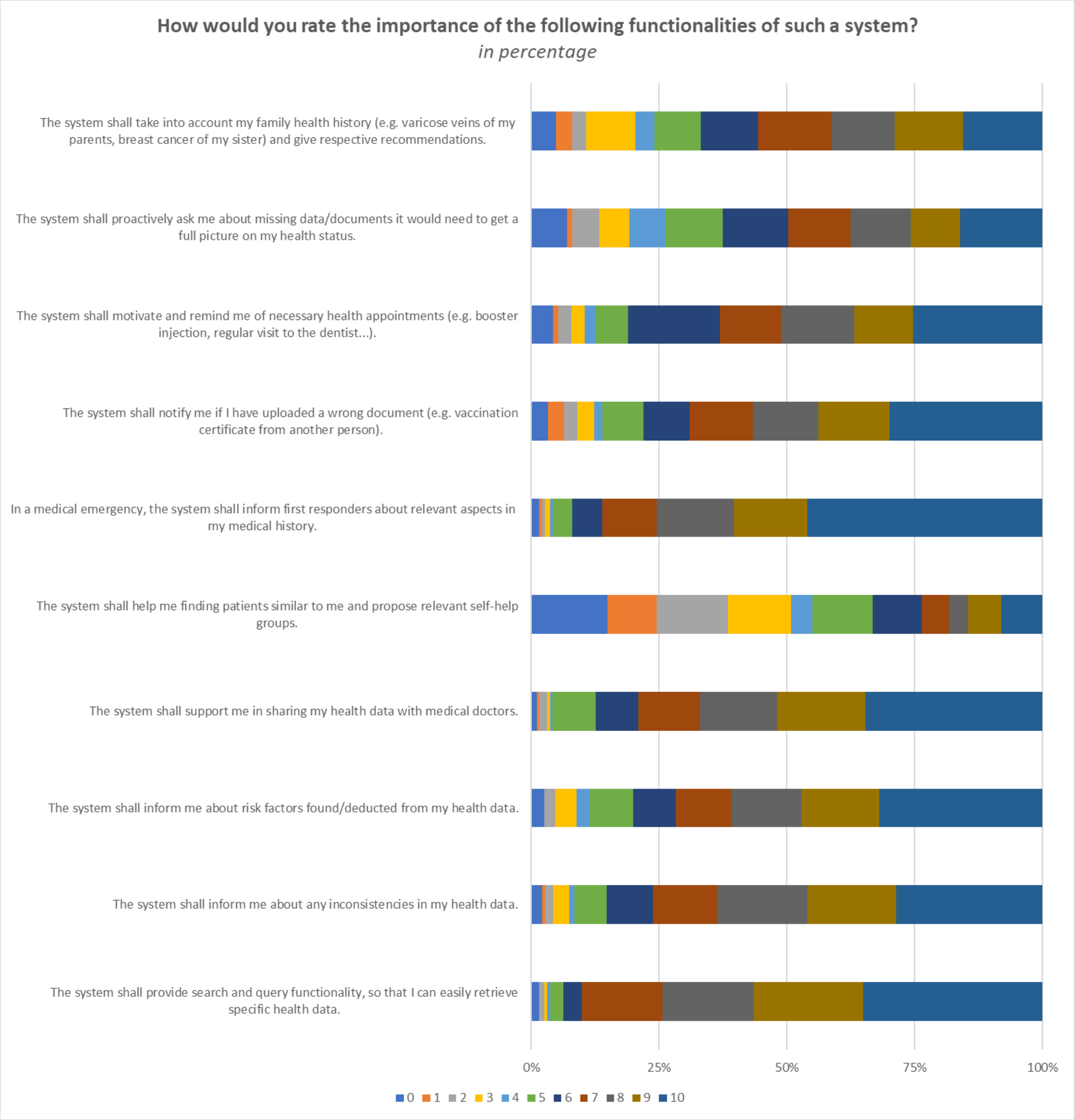

For what purpose do you download your personal health data and medical records?
For another purpose
6
To have them available for emergencies
5
For my insurance
2
To share with my doctor
2
Personal interest
5
0
1
2
3
4
5
6
7
multiple selections possible, n=46

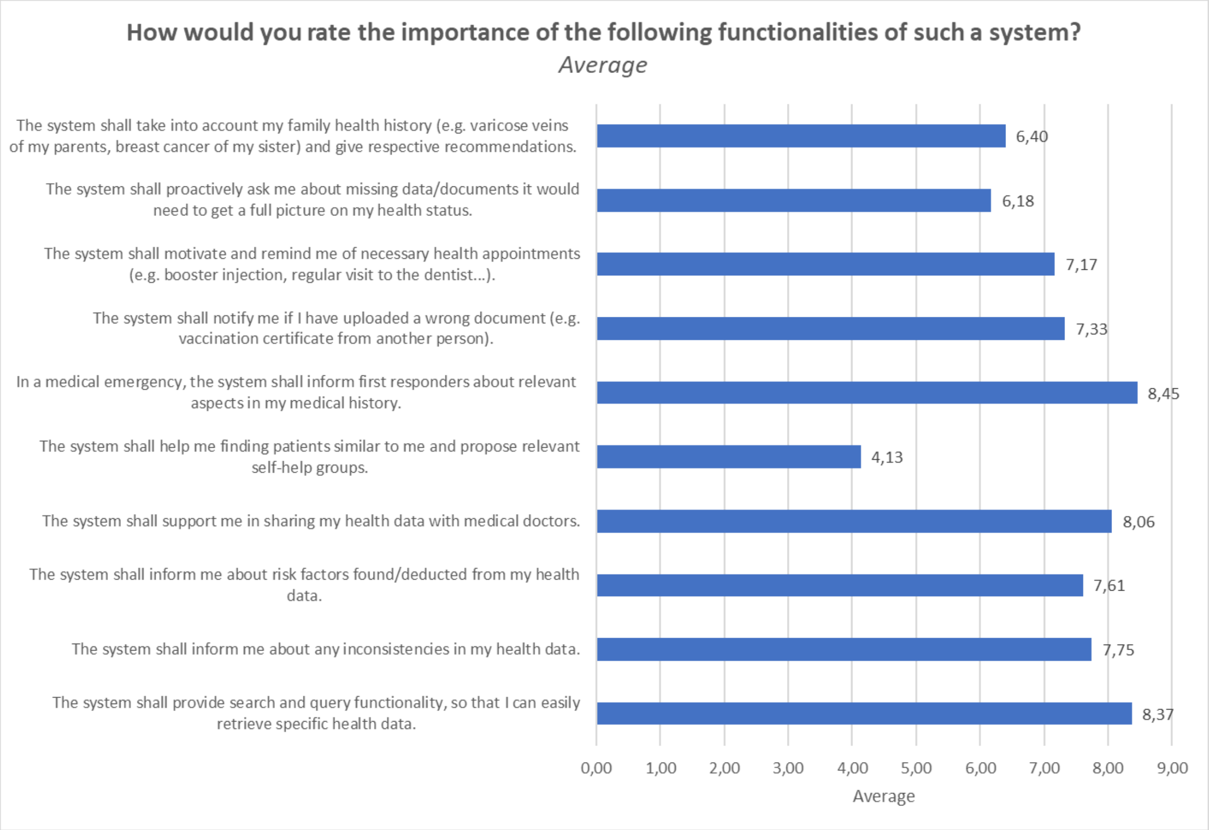

Are you monitoring any health-related parameters on your own? Do you share the results of these measurements with your doctor?
180
160
140
120
100
80
60
40
20
0
Steps
82
108
4
Pulse
97
84
8
Sleep
126
56
6
Blood pressure
124
46
17
Blood-glucose level
163
6
12
Other data
129
31
11
■ I do not measure
■ I measure but do not share
■ I share with my doctor

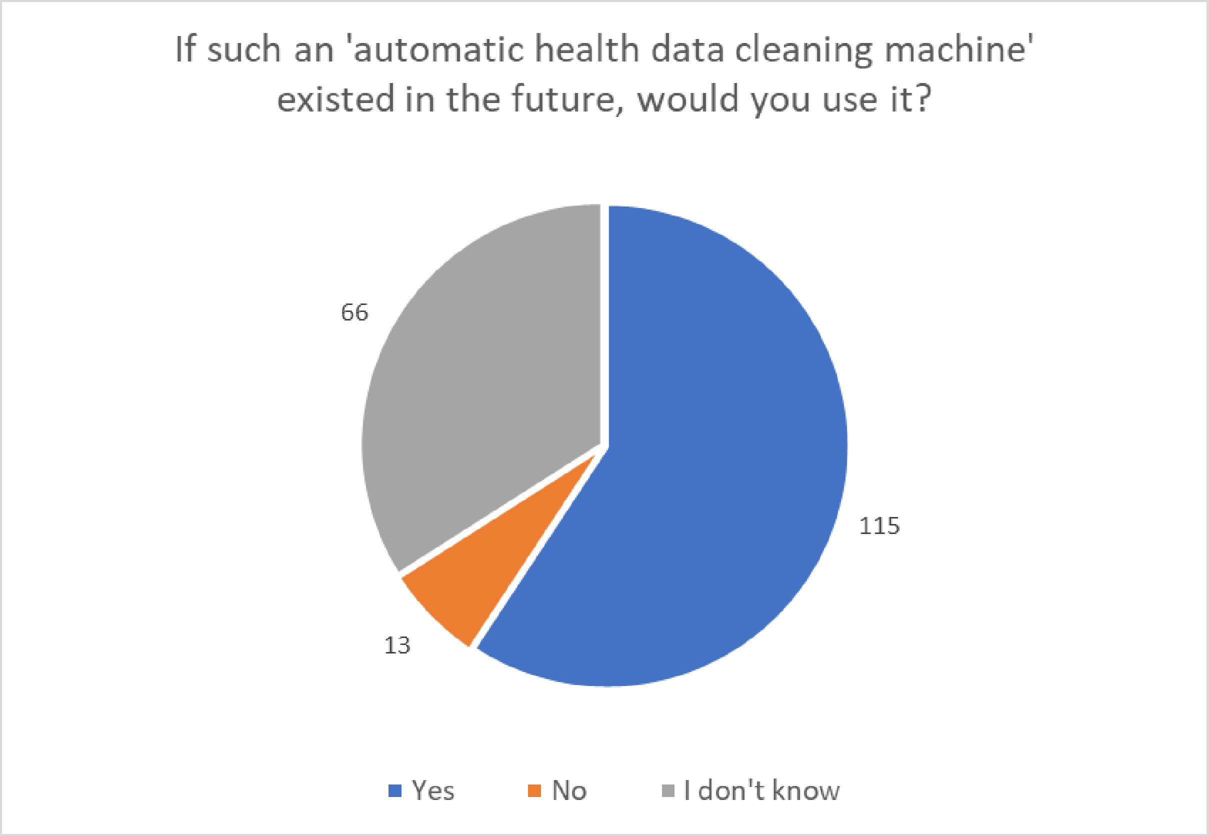

I would share my personal health data with:
160
140
120
100
80
60
40
20
0
My doctor
131
56
8
My partner
100
62
24
My family
53
92
41
My caregiver
42
51
72
My insurance company
27
60
90
Others
1
16
131
■ Yes
■ Sometimes
■ No

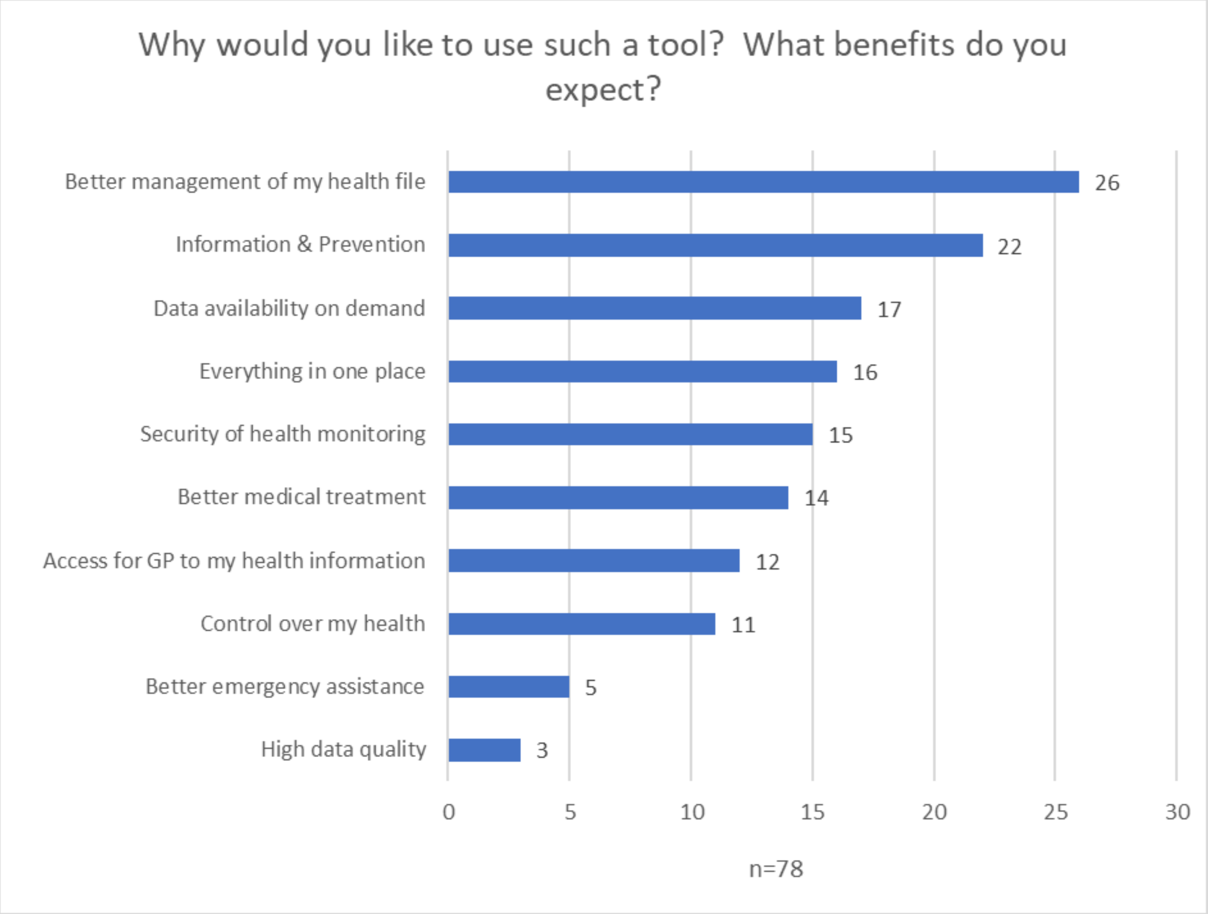
How would you rate the importance of the following functionalities of such a system?
in numbers



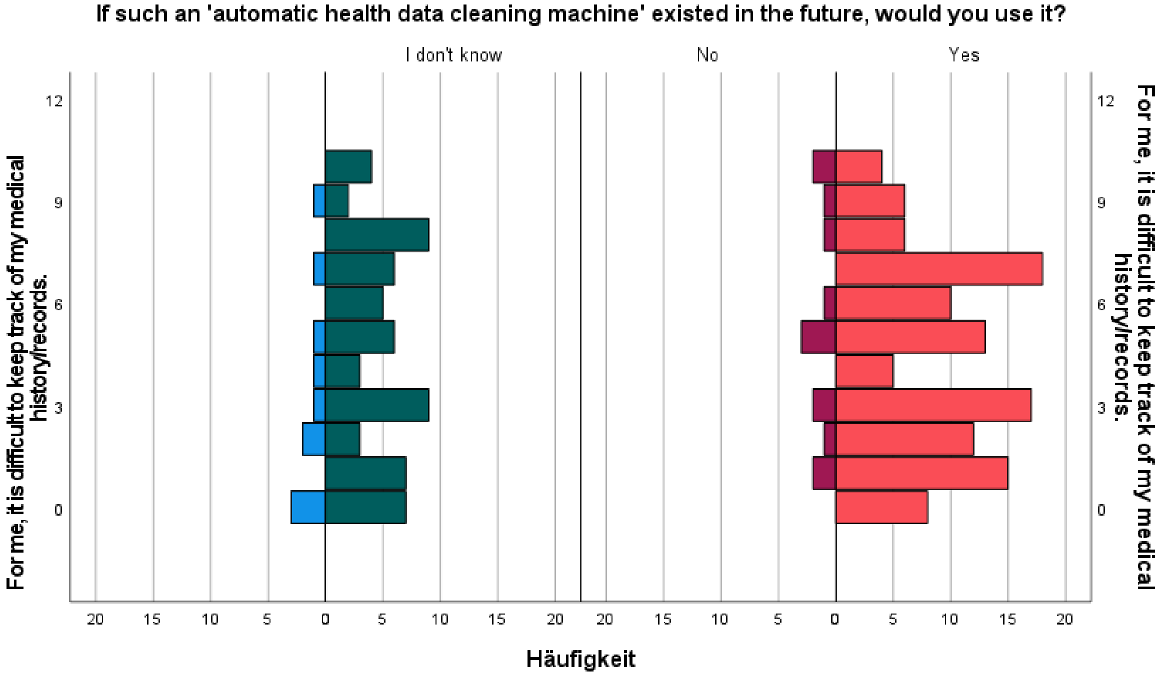

How would you rate the importance of the following functionalities of such a system?
in percentage
The system shall take into account my family health history (e.g. varicose veins of my parents, breast cancer of my sister) and give respective recommendations.
The system shall proactively ask me about missing data/documents it would need to get a full picture on my health status.
The system shall motivate and remind me of necessary health appointments (e.g. booster injection, regular visit to the dentist...).
The system shall notify me if I have uploaded a wrong document (e.g. vaccination certificate from another person).
In a medical emergency, the system shall inform first responders about relevant aspects in my medical history.
The system shall help me finding patients similar to me and propose relevant self-help groups.
The system shall support me in sharing my health data with medical doctors.
The system shall inform me about risk factors found/deducted from my health data.
The system shall inform me about any inconsistencies in my health data.
The system shall provide search and query functionality, so that I can easily retrieve specific health data.
0%
25%
50%
75%
100%
0
1
2
3
4
5
6
7
8
9
10


How would you rate the importance of the following functionalities of such a system?
Average
The system shall take into account my family health history (e.g. varicose veins of my parents, breast cancer of my sister) and give respective recommendations.
6,40
The system shall proactively ask me about missing data/documents it would need to get a full picture on my health status.
6,18
The system shall motivate and remind me of necessary health appointments (e.g. booster injection, regular visit to the dentist...).
7,17
The system shall notify me if I have uploaded a wrong document (e.g. vaccination certificate from another person).
7,33
In a medical emergency, the system shall inform first responders about relevant aspects in my medical history.
8,45
The system shall help me finding patients similar to me and propose relevant self-help groups.
4,13
The system shall support me in sharing my health data with medical doctors.
8,06
The system shall inform me about risk factors found/deducted from my health data.
7,61
The system shall inform me about any inconsistencies in my health data.
7,75
The system shall provide search and query functionality, so that I can easily retrieve specific health data.
8,37
0,00
1,00
2,00
3,00
4,00
5,00
6,00
7,00
8,00
9,00
Average

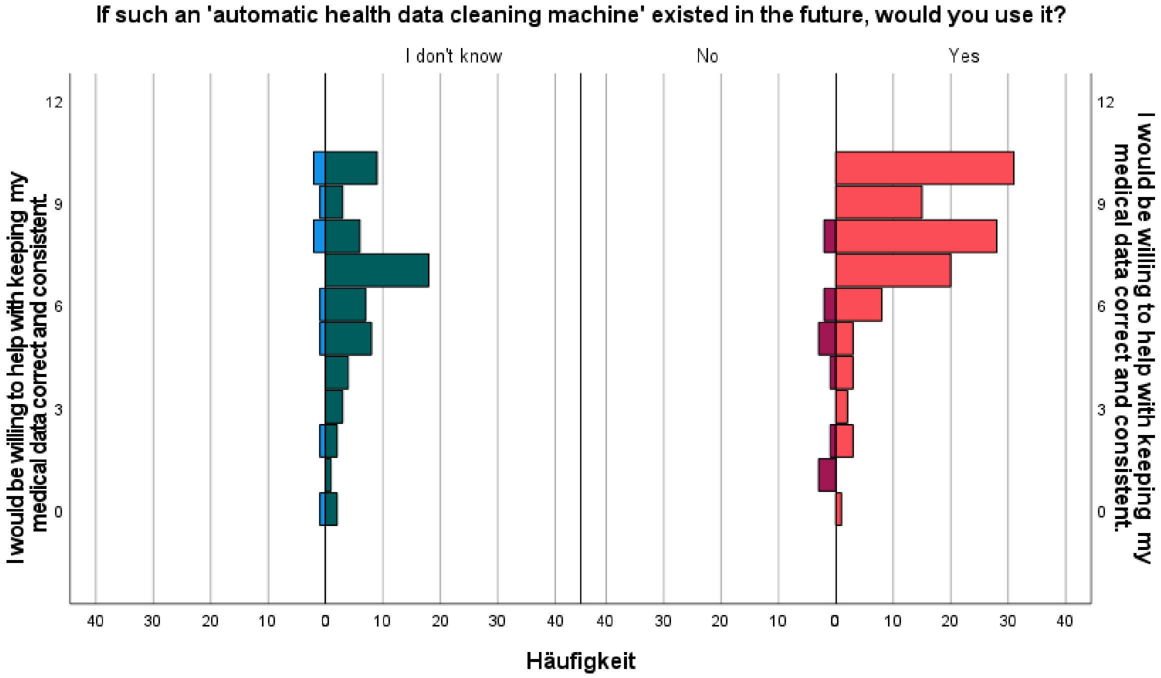

If such an 'automatic health data cleaning machine' existed in the future, would you use it?
66
115
13
Yes
No
I don't know

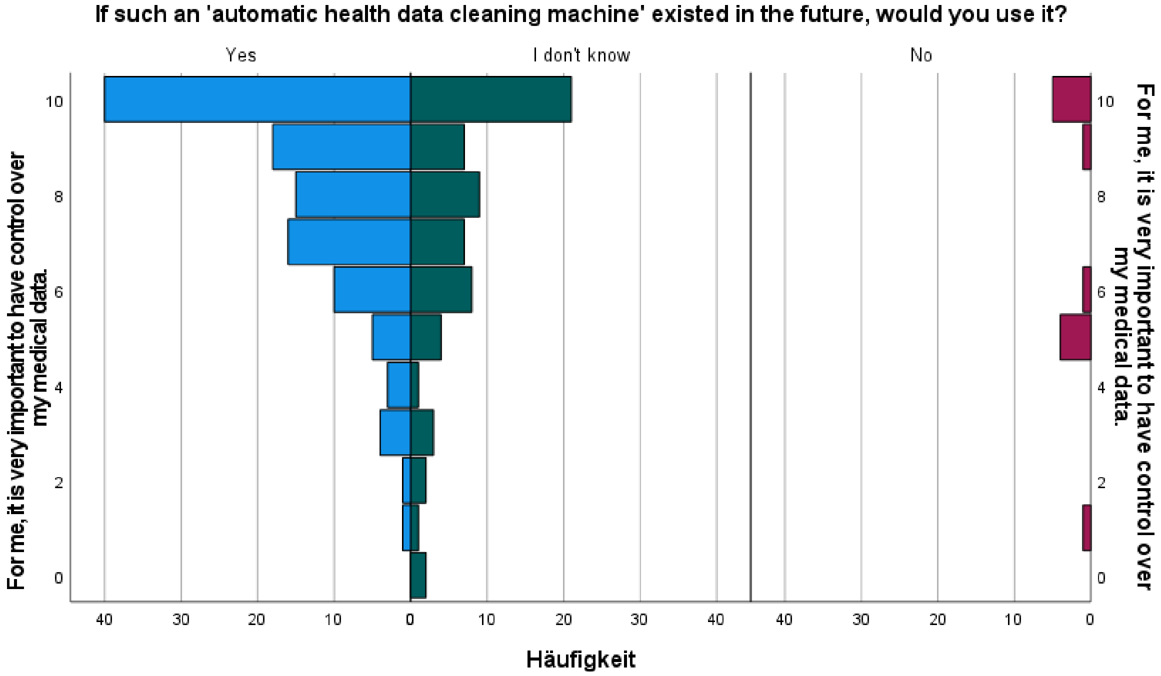


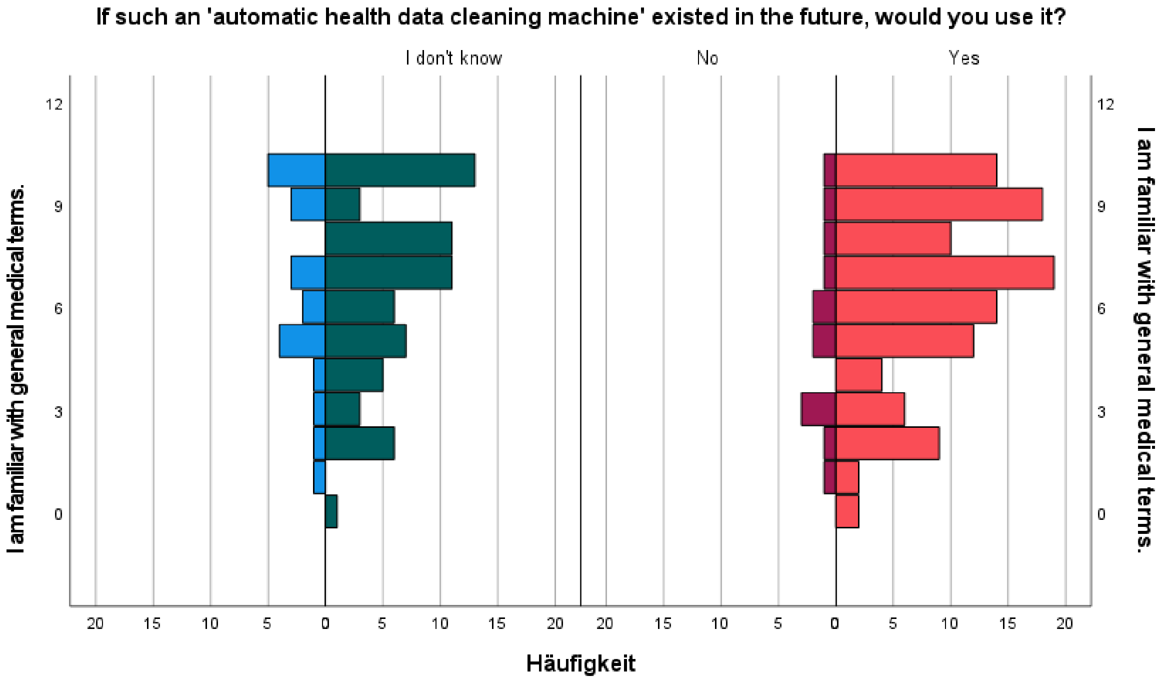
Why would you like to use such a tool? What benefits do you expect?

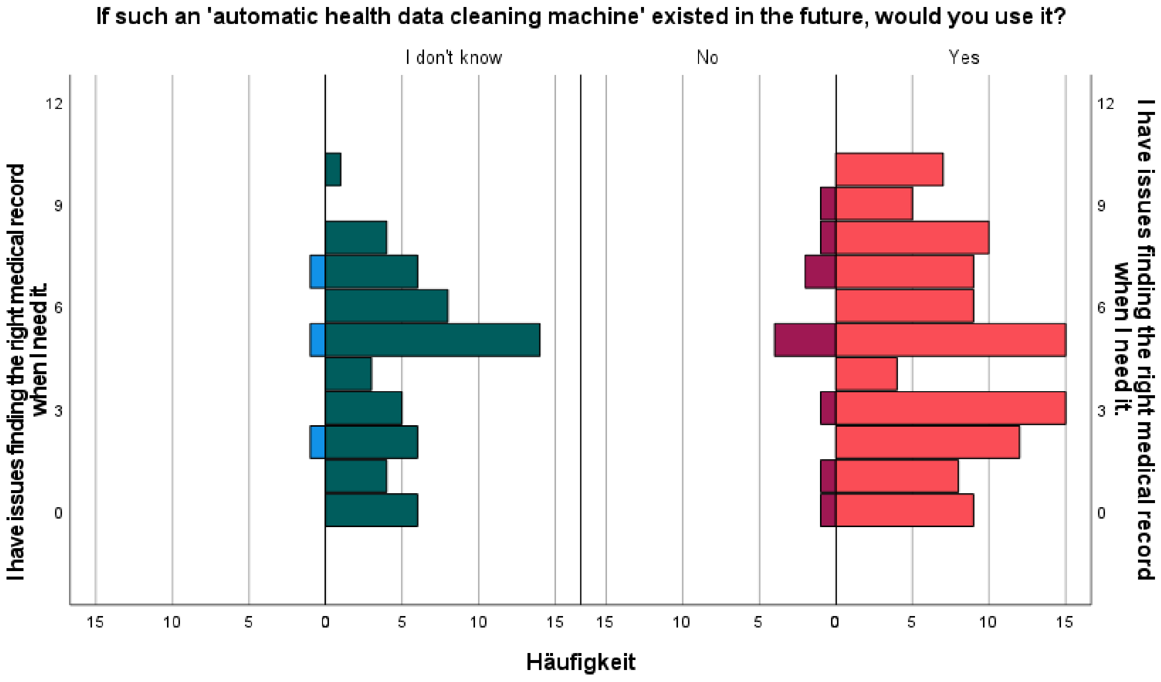


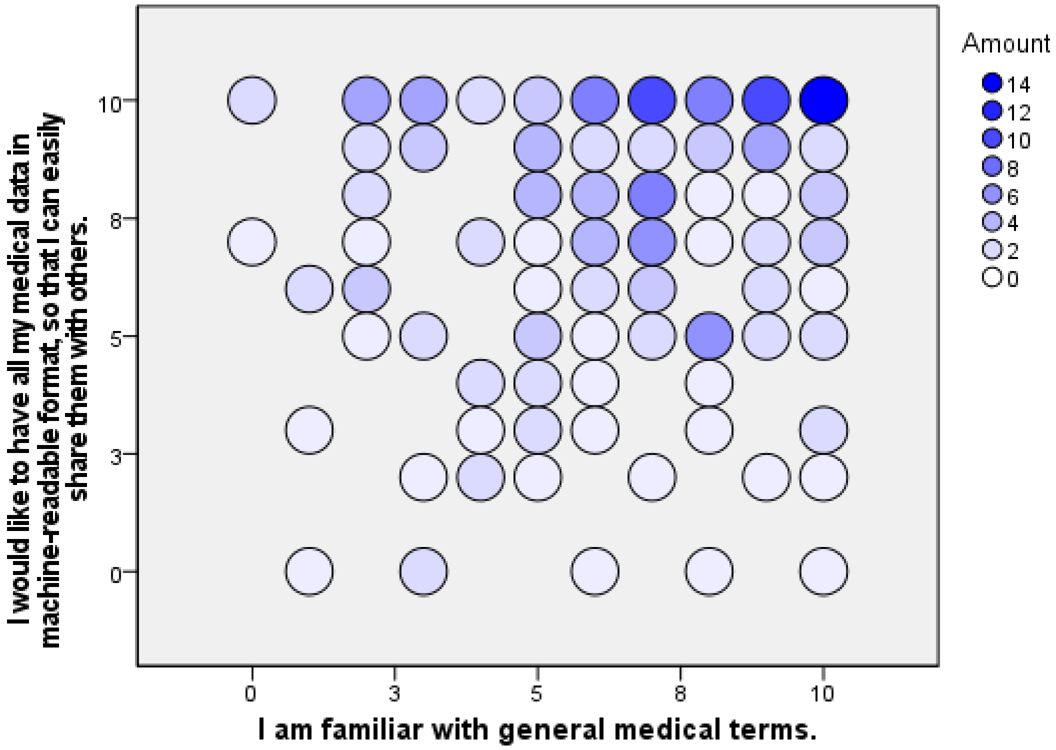
If such an 'automatic health data cleaning machine' existed in the future, would you use it?

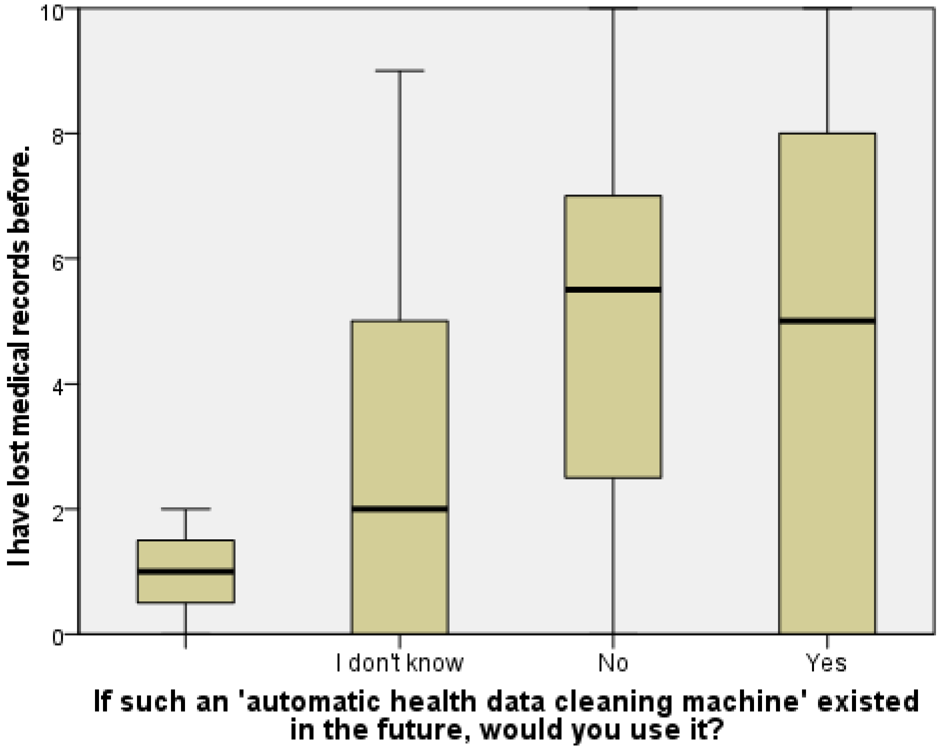

If such an 'automatic health data cleaning machine' existed in the future, would you use it?
I don't know
No
Yes
I would be willing to help with keeping my medical data correct and consistent.
12
9
6
3
0
40
30
20
10
0
10
20
30
40
Häufigkeit

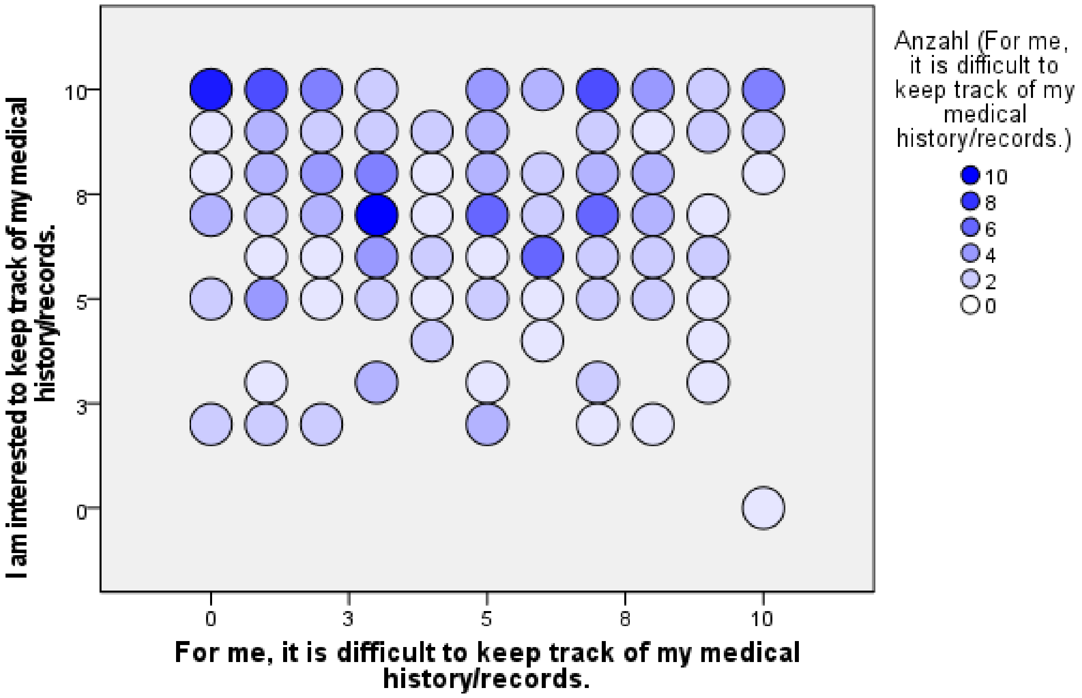

If such an 'automatic health data cleaning machine' existed in the future, would you use it?
Yes
I don't know
No
For me, it is very important to have control over my medical data.
10
8
6
4
2
0
40
30
20
10
0
10
20
30
40
Häufigkeit

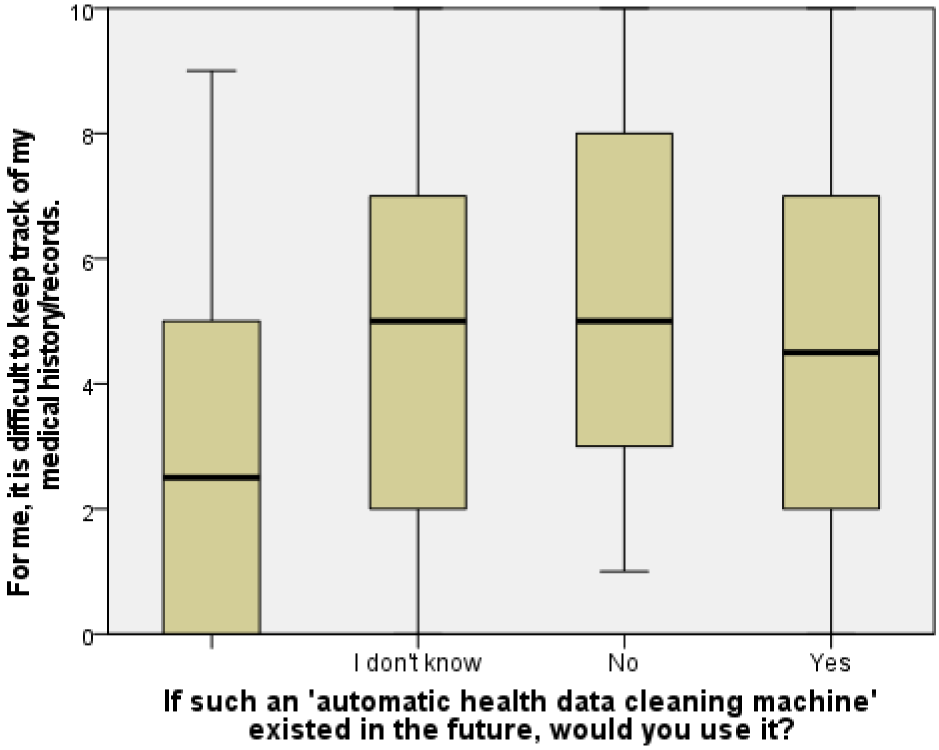

If such an 'automatic health data cleaning machine' existed in the future, would you use it?
I don't know
No
Yes
I am familiar with general medical terms.
12
9
6
3
0
20
15
10
5
0
Häufigkeit

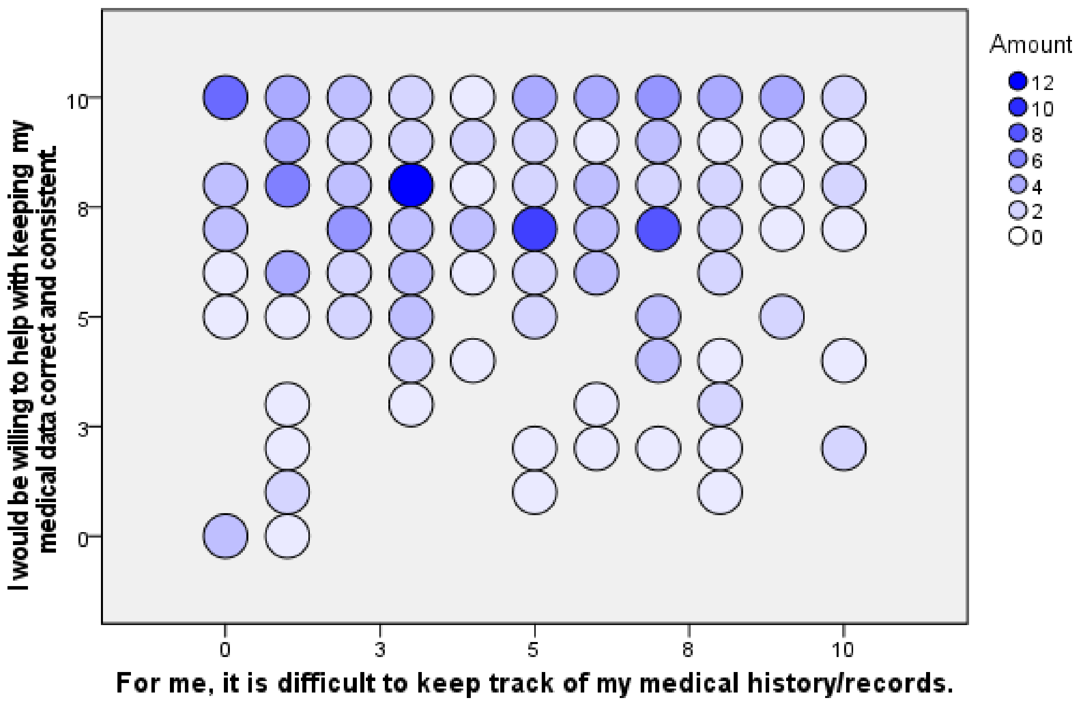

If such an 'automatic health data cleaning machine' existed in the future, would you use it?
I don't know
No
Yes
I have issues finding the right medical record when I need it.
12
9
6
3
0
15
10
5
0
Häufigkeit

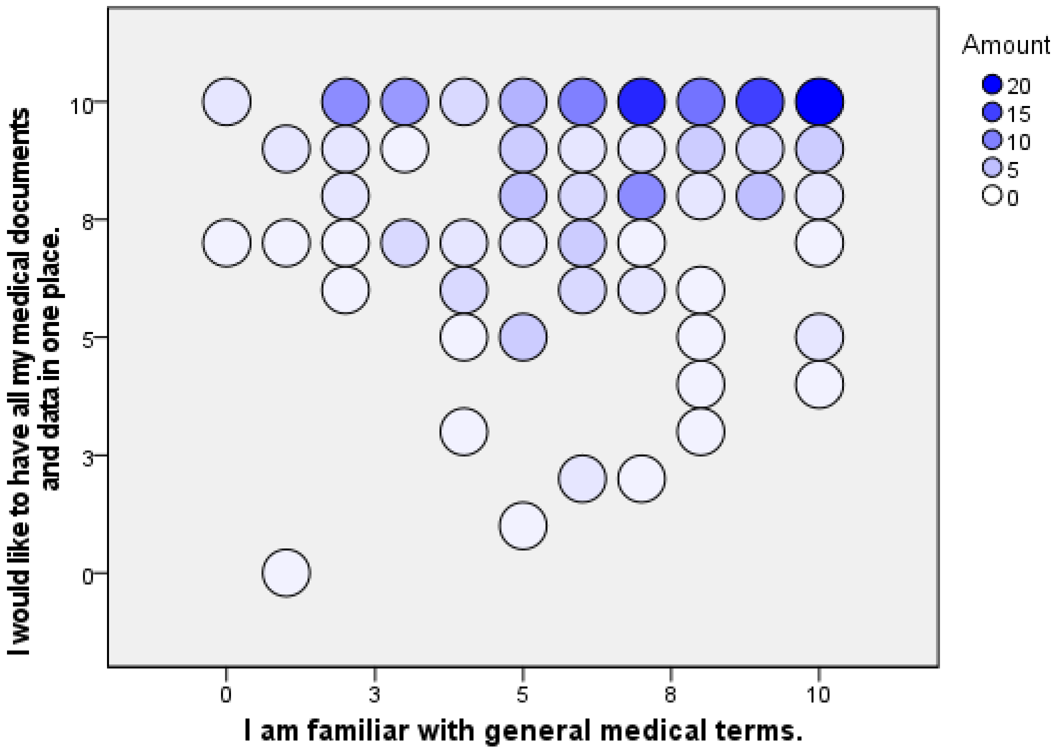

I would like to have all my medical data in machine-readable format, so that I can easily share them with others.
10
8
5
3
0
0
3
5
8
10
I am familiar with general medical terms.
Amount
14
12
10
8
6
4
2
0

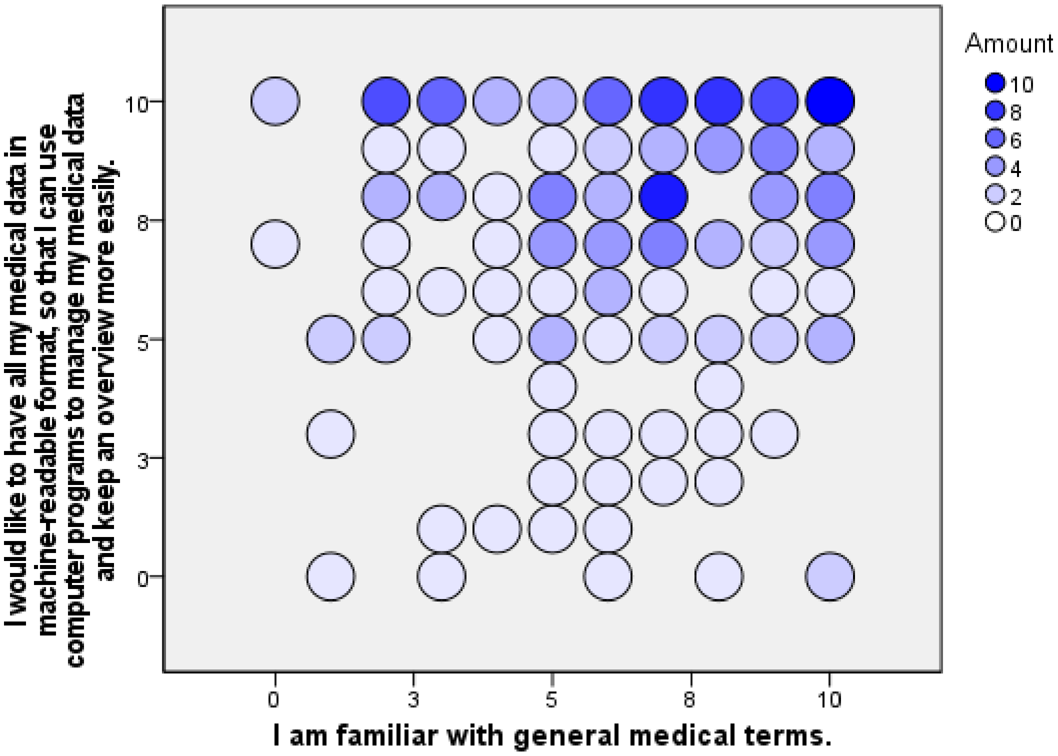

I have lost medical records before.
10
8
6
4
2
0
I don't know
No
Yes
If such an 'automatic health data cleaning machine' existed in the future, would you use it?

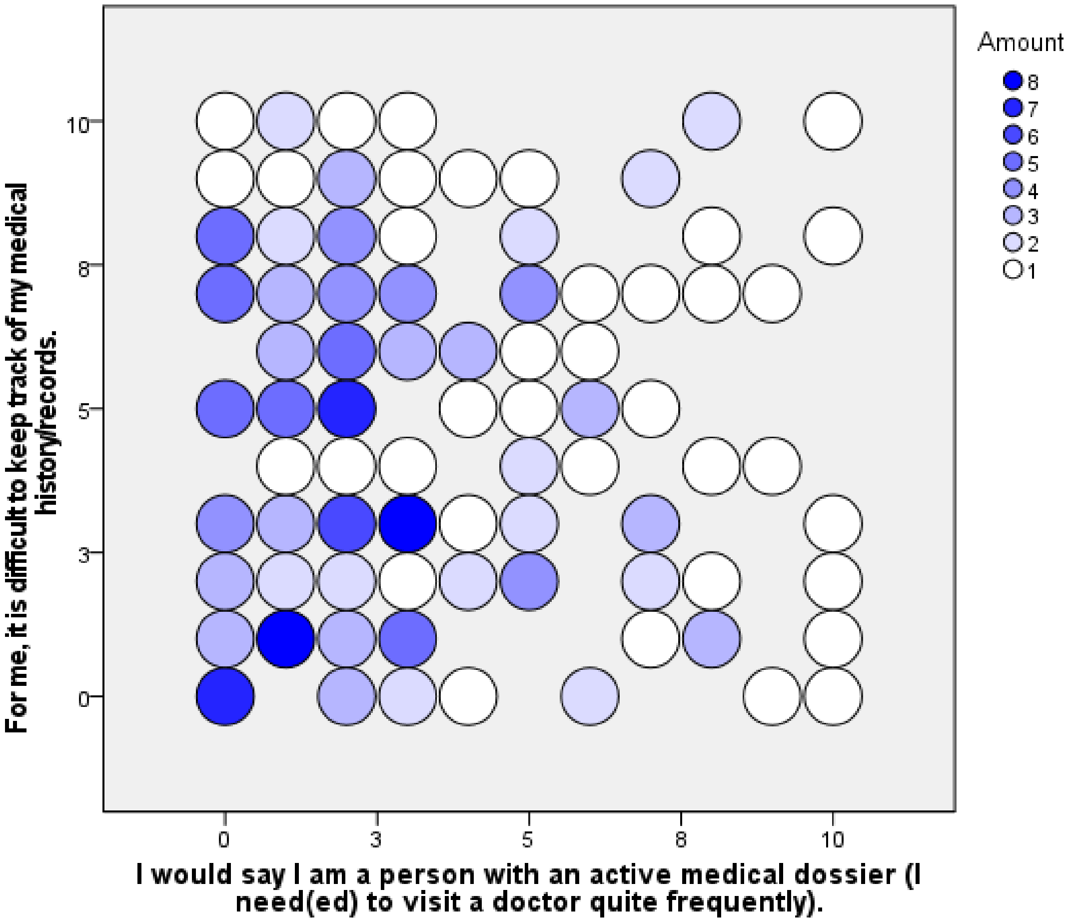

I am interested to keep track of my medical history/records.
10
8
5
3
0
0
3
5
8
10
For me, it is difficult to keep track of my medical history/records.
Anzahl (For me, it is difficult to keep track of my medical history/records.)
10
8
6
4
2
0

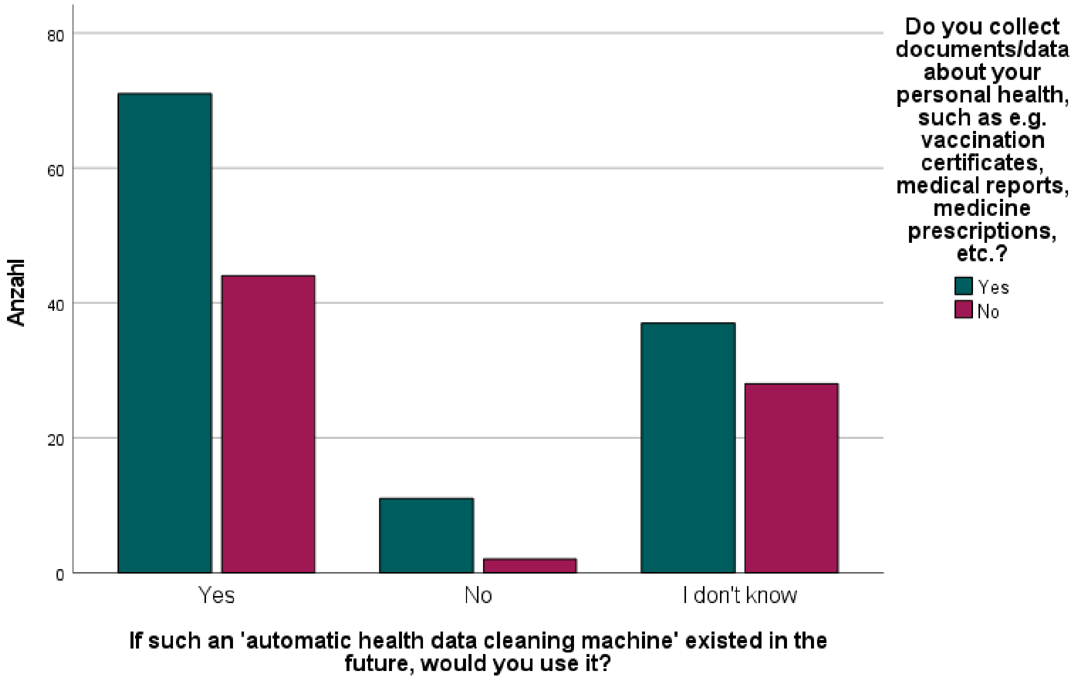

For me, it is difficult to keep track of my medical history/records.
10
8
6
4
2
0
I don't know
No
Yes
If such an 'automatic health data cleaning machine' existed in the future, would you use it?

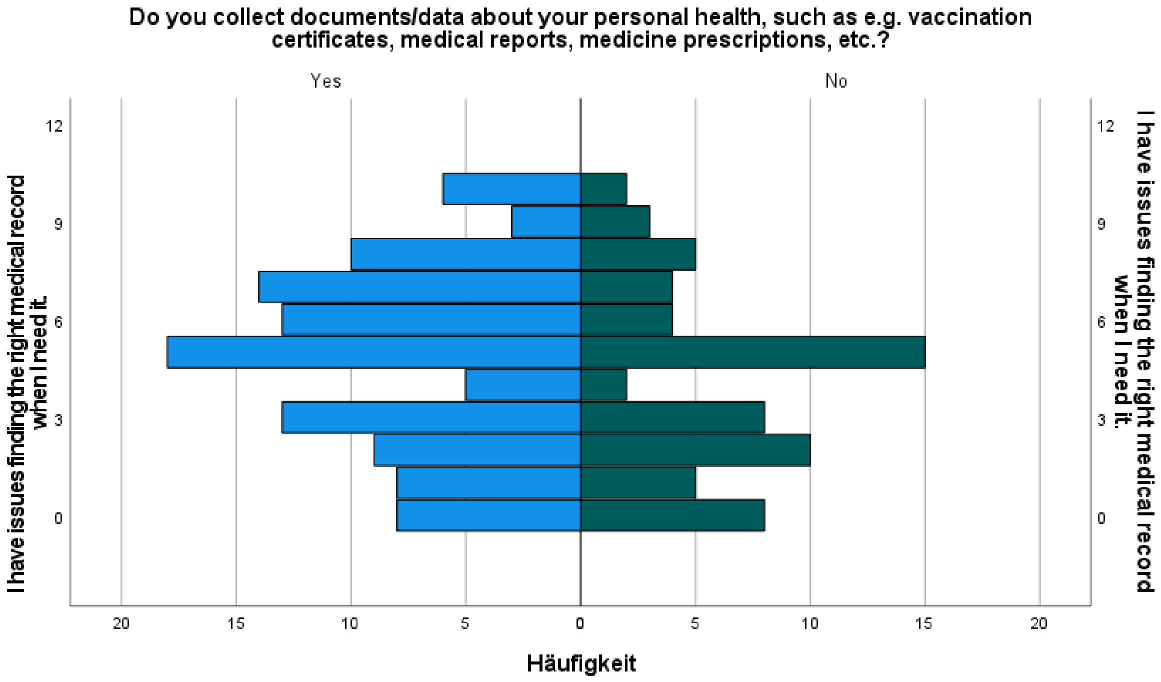

Amount
12
10
8
6
4
2
0
I would be willing to help with keeping my medical data correct and consistent.
10
8
5
3
0
0
3
5
8
10
For me, it is difficult to keep track of my medical history/records.

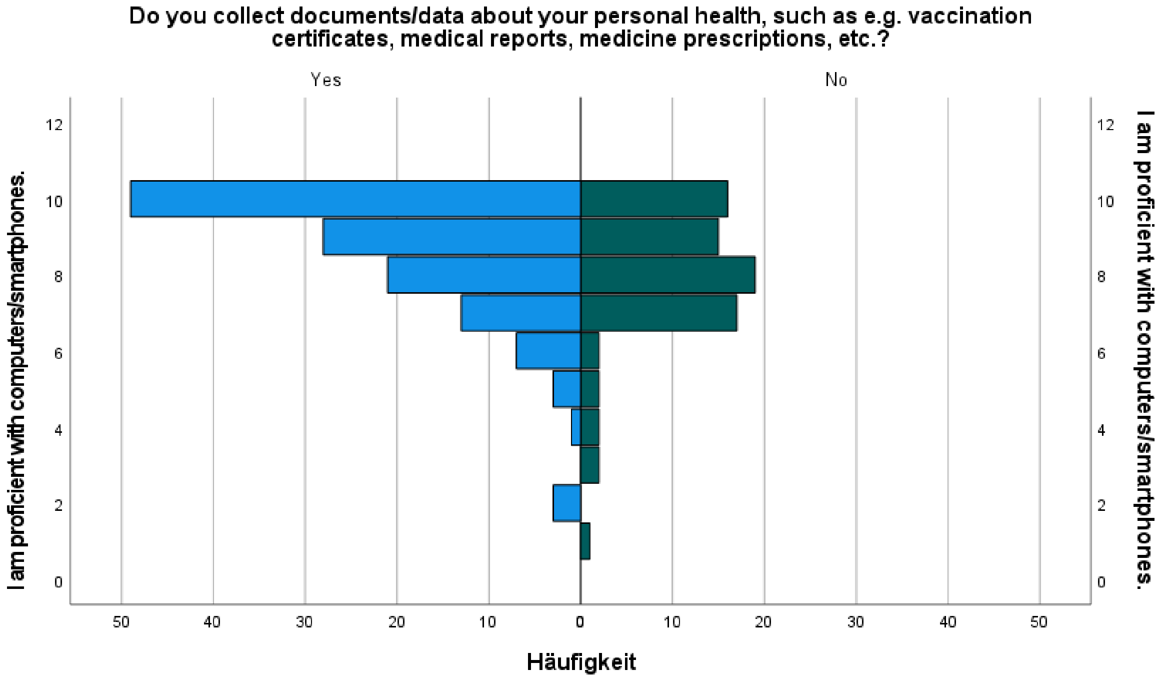

Amount
20
15
10
5
0
I would like to have all my medical documents and data in one place.
10
8
5
3
0
0
3
5
8
10
I am familiar with general medical terms.

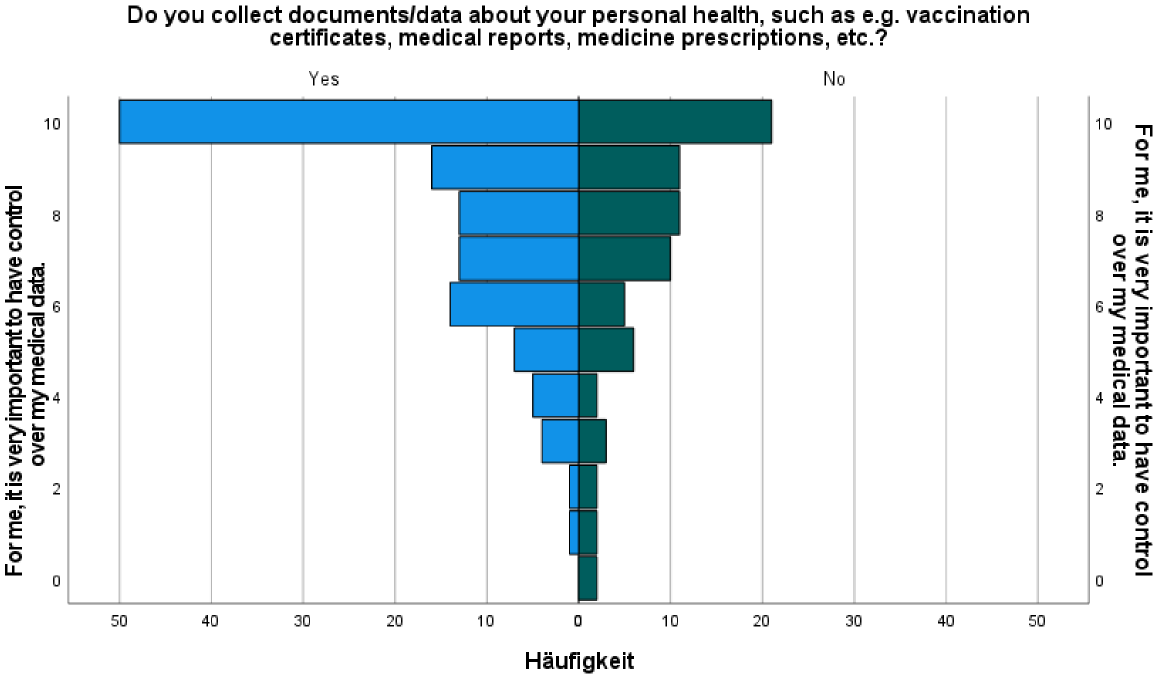

Amount
10
8
6
4
2
0
I would like to have all my medical data in machine-readable format, so that I can use computer programs to manage my medical data and keep an overview more easily.
10
8
5
3
0
0
3
5
8
10
I am familiar with general medical terms.

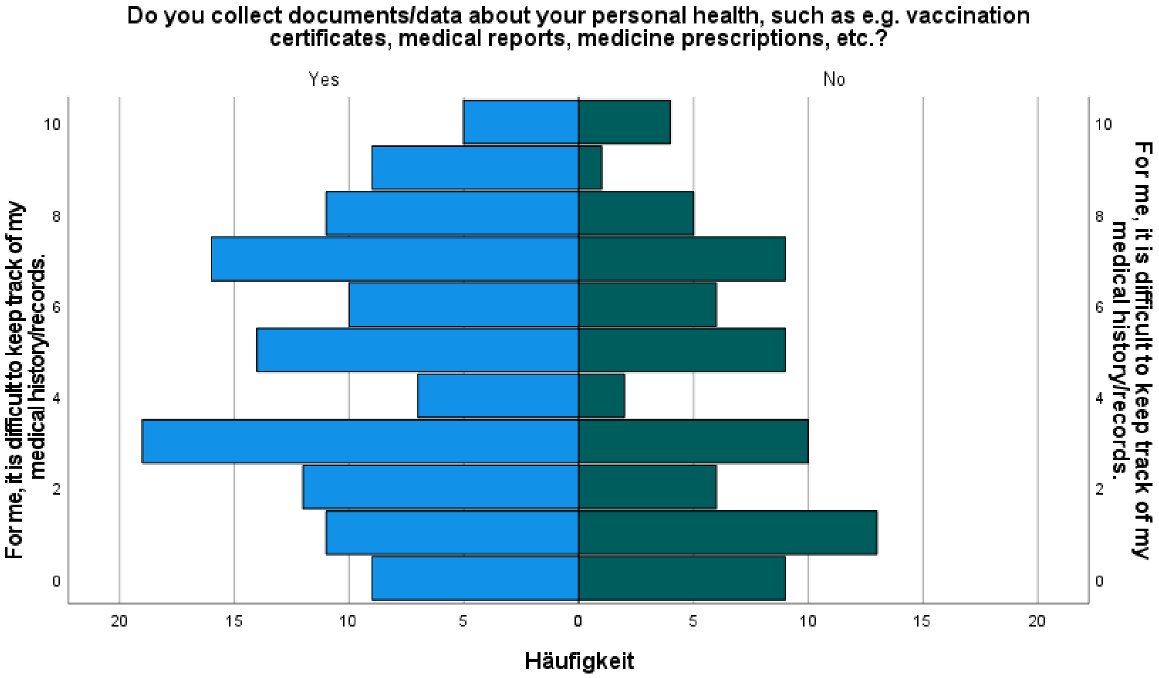

Amount
8
7
6
5
4
3
2
1
For me, it is difficult to keep track of my medical history/records.
0
3
5
8
10
I would say I am a person with an active medical dossier (I need(ed) to visit a doctor quite frequently).

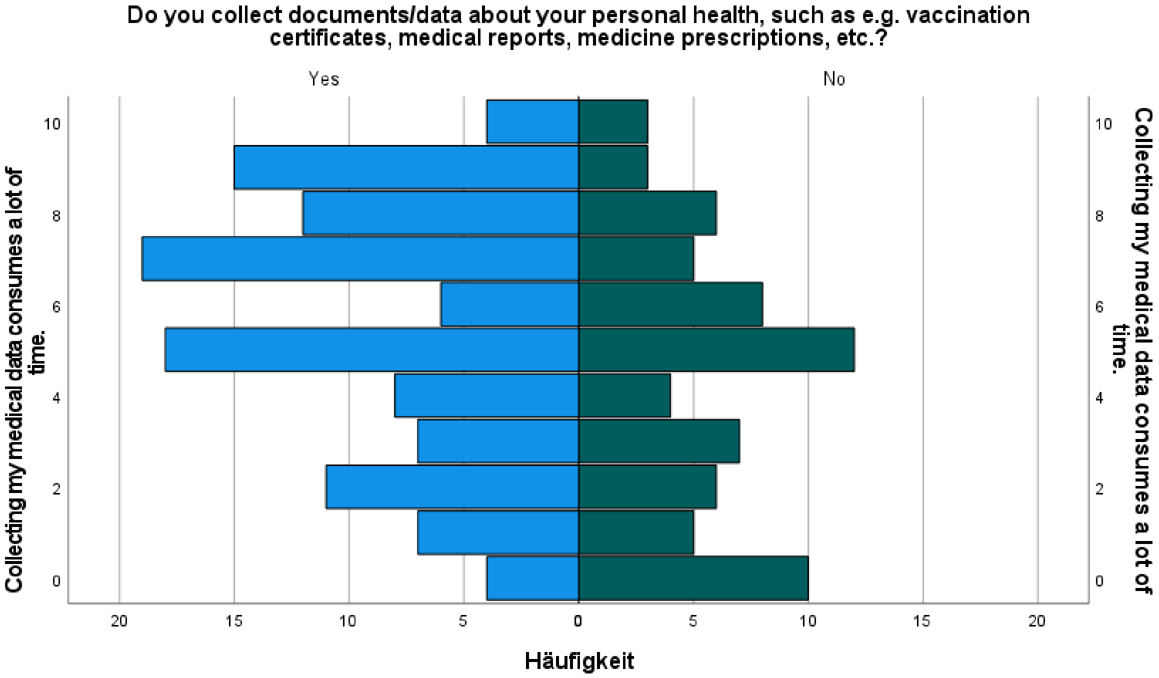

Anzahl
0
20
40
60
80
Yes
No
I don't know
If such an 'automatic health data cleaning machine' existed in the future, would you use it?
Do you collect documents/data about your personal health, such as e.g. vaccination certificates, medical reports, medicine prescriptions, etc.?
Yes
No

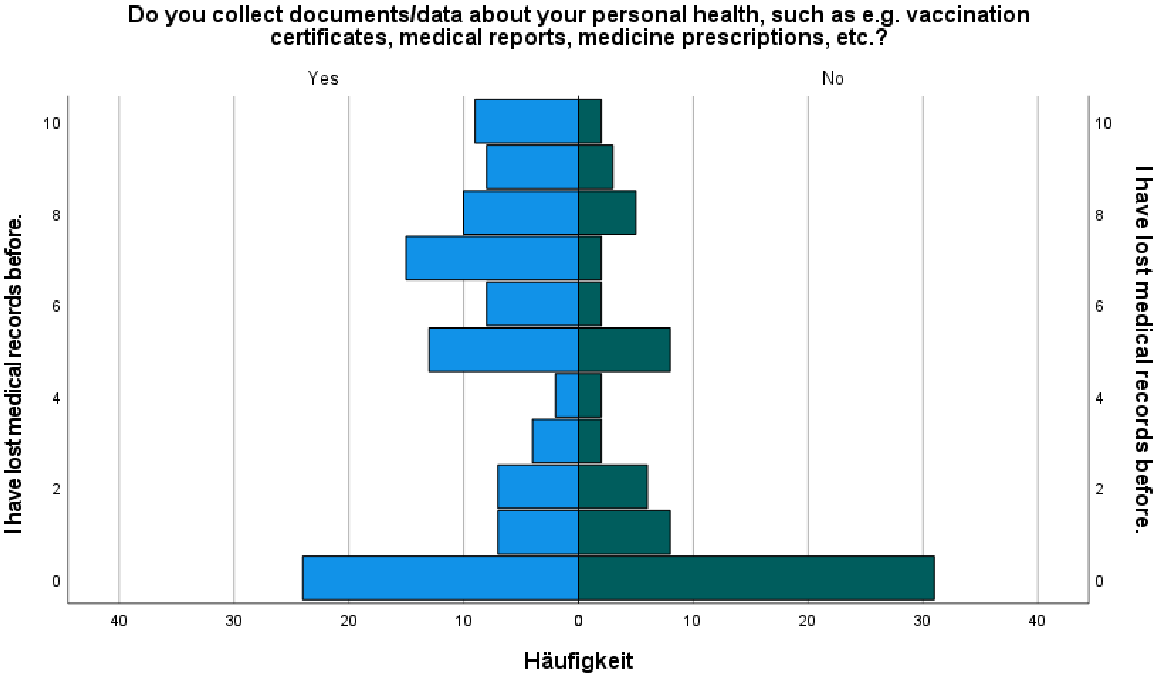

Do you collect documents/data about your personal health, such as e.g. vaccination certificates, medical reports, medicine prescriptions, etc.?
Yes
No
I have issues finding the right medical record when I need it.
12
9
6
3
0
20
15
10
5
0
5
10
15
20
Häufigkeit


Do you collect documents/data about your personal health, such as e.g. vaccination certificates, medical reports, medicine prescriptions, etc.?
Yes
No
I am proficient with computers/smartphones.
12
10
8
6
4
2
0
50
40
30
20
10
0
10
20
30
40
50
Häufigkeit

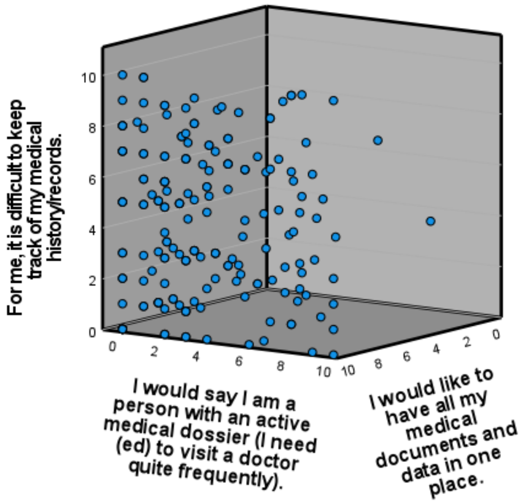

Do you collect documents/data about your personal health, such as e.g. vaccination certificates, medical reports, medicine prescriptions, etc.?
Yes
No
For me, it is very important to have control over my medical data.
10
8
6
4
2
0
50
40
30
20
10
0
10
20
30
40
50
Häufigkeit

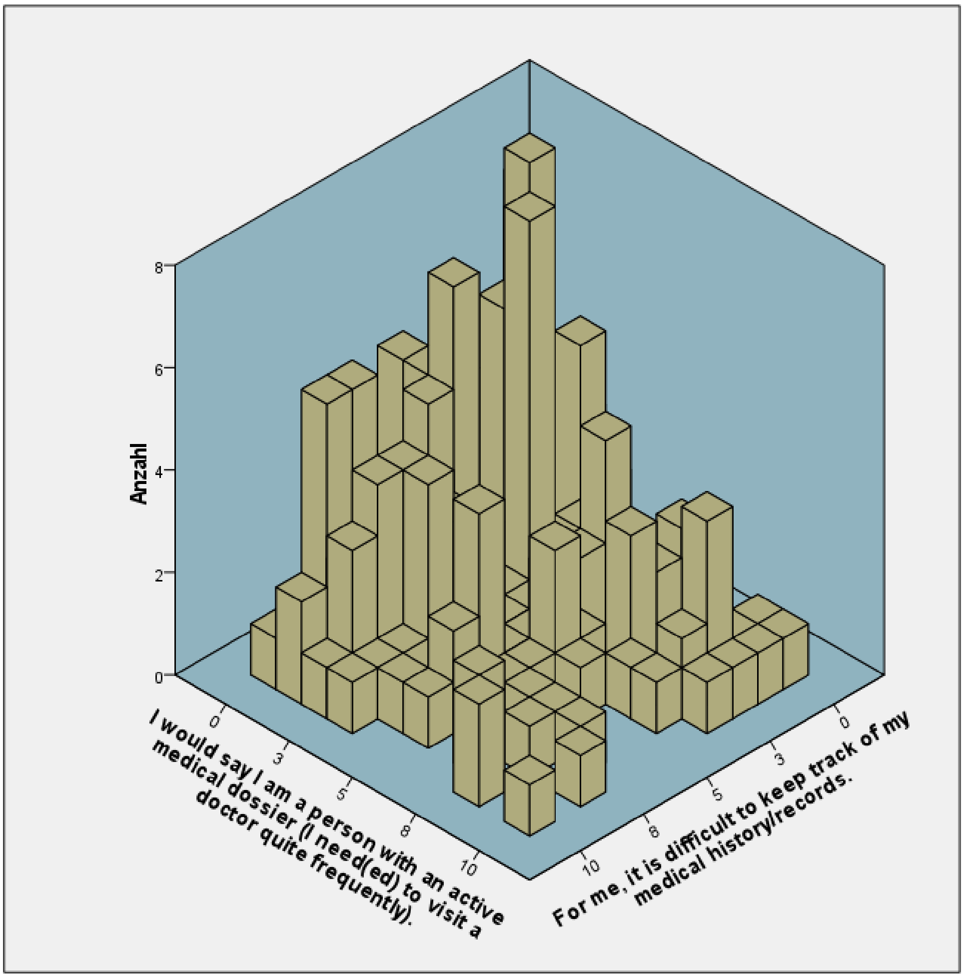

Do you collect documents/data about your personal health, such as e.g. vaccination certificates, medical reports, medicine prescriptions, etc.?
Yes
No
For me, it is difficult to keep track of my medical history/records.
10
8
6
4
2
0
20
15
10
5
0
5
10
15
20
Häufigkeit


Do you collect documents/data about your personal health, such as e.g. vaccination certificates, medical reports, medicine prescriptions, etc.?
Yes
No
Collecting my medical data consumes a lot of time.
10
8
6
4
2
0
20
15
10
5
0
5
10
15
20
Häufigkeit


Do you collect documents/data about your personal health, such as e.g. vaccination certificates, medical reports, medicine prescriptions, etc.?
Yes
No
I have lost medical records before.
10
8
6
4
2
0
40
30
20
10
0
10
20
30
40
Häufigkeit

Einfaches 3-D-Streudiagramm von For me, it is difficult to keep track of my medical history/records. Schritt: I would say I am a person with an active medical dossier (I need(ed) to visit a doctor quite frequently). Schritt: I would like to have all my medical documents and data in one place.





8.3 Interview guidelines

8.3.1 Patients Interviews

AIDAVA
AI powered Data Curation & Publishing Virtual Assistant.

Task 1.2
Guidelines for Patients’ Interview

Lead partners: MUG
Other partners: NEMC, b!loba, UM, ECPC, EHN, MID, DME

Coordinator: Michel Dumontier,
University of Maastricht,
The Netherland

Project website: www.aidava.eu

Grant Agreement: 101057062

Horizon Europe
THE NEXT EU RESEARCH & INNOVATION PROGRAMME (2021 – 2027)
European Commission

Table of Contents

General Instructions for the Interviewer

• Selection Criteria for Interviewees

1. Inclusion criteria:
    a. Age: at least 18 years
    b. Ability (and willingness) to use smartphone/computer applications
    c. History of a confirmed diagnosis of a chronic disease* resulting in frequent visits to doctors and hospitals
    (*not restricted to cardiovascular disease or cancer)
    d. Gender: All
2. Exclusion criteria:

. Child/minor
a. Vulnerable person
b. Does not use a smartphone/computer

IMPORTANT: Ensure that you are authorised to approach interview participants.

• Informed Consent

Before you can conduct such an interview, you need to obtain informed consent from your interview partner. Thus, it is essential to ask your interview partner to sign the Informed Consent Sheet. Thereby, the suggested procedure to obtain informed consent from your interview partner depends on whether you plan to conduct a face-to-face or remote interview:

- In case you conduct a *remote interview*:
  Send the Informed Consent sheet to your interview partner via email well before the

interview (e.g. when you agree with your interview partner on a date) and ask them to return you a signed copy.

- In case you conduct a *personal face-to-face interview*:
  Ask your interview partner to sign the Informed Consent sheet immediately before starting with the interview questions, after you have explained and “set the scene” (see section “1 Set the Scene” below). Please prepare two(!) printouts of the Informed Consent sheet to be signed → Give one to the interview participant and keep the second one for your records.

Aims of this Interview

We want to learn more about patients, who are potential future users of the “AIDAVA data cleaning assistant”. This will help us follow human-centred design principles and consider the users' needs, abilities, skills, constraints and preferences during the design and development of the AIDAVA system.

The results of this interview will together with the results of several other similar patient’s interviews form the basis for the development of archetypical users, so-called *personas. Personas* is a method well-known in the human-computer interaction (HCI) field. It was introduced to help designers and developers focus on the needs and goals of the target users throughout the product development process. In AIDAVA these personas will form the basis for elaborating user profiles.

Specifically, this interview shall help to understand better

- patients’ attitudes, goals, motivations, frustrations, challenges and pain points
- patients’ aptitudes, competencies, knowledge, skills and experience
- patients’ personal context and framework conditions
- patients’ interest and willingness to control and curate their health data
- patients’ vision of an “ideal” data curation tool

Important Notes

- This is an interview guideline. Do not strictly stick to these questions, but react on your interview partner:
  - Ask follow-up questions, whenever clarification of some aspects is needed;
  - Change the order of the questions or skip questions, whenever you deem this appropriate for the specific interview;
  - Adapt to your interview partner and use the appropriate language: you may need to modify the wording of some questions, or you may need to conduct the interview in your interview partner’s national language.
- Please, always keep in mind: Your goal is to learn from your interview partner! It is not the goal to confirm your opinion or any stereotypes. Therefore:
  - Ask open questions to motivate your interview partner to tell their story.
  - Avoid leading questions and do not introduce bias to the answers of your interview partner.
- Recommendations for conducting this interview:
  - The interview should be conducted by two persons, if possible: one asks the questions and interacts with the interviewee, and the other observes and takes notes.
  - Take enough time, avoid time pressure, and hurry: This interview will usually take about 40-60 minutes.
- After the interview:
  - Transcribe the answers of the interviewee, and translate them into English if the interview was conducted in another language.
  - Anonymise your interview report, so that it does not contain any identifying information: for example, replace names of persons with roles, and remove any explicit naming of workplace, organisations or city.

  - Download a copy of the provided template and use it to prepare a results' summary by grouping the results according to the topics: Personal Information, Digital/Data Literacy, Attitude and Approach towards new technologies, Medical/Health Literacy, Current Approaches regarding personal health data, Attitude towards "Automatic Health Data Cleaning"
  - Send this results' summary to the interviewee to check whether you've understood everything right and the interviewee is o.k. with it.
  - If the interviewee raises no objections, please upload the results summary to the respective folder in the AIDAVA google drive.

Interview Guideline

• 1 Set the Scene

1. Thank your interview partner for taking the time to participate in this interview.
2. Briefly introduce yourself, and the note-taker if applicable.
3. Briefly introduce the AIDAVA project.
   (You can use (parts of) the presentation from patients consultants onboarding - https://docs.google.com/presentation/d/16KA_tuL1QVzVK3RMsOetev7DOq66OVobbMckDaZPmb4/edit#slide=id.p1)
4. Briefly introduce the purpose of this interview:
   The results of this interview will be used by the AIDAVA consortium to develop descriptions of archetypical users, so-called "personas", which will support the user-centred design and development of the AIDAVA system.
5. Inform your interview partner that

a. Your interview partner may refuse to answer a question without stating a reason.

b. Your interview partner may stop the interview anytime.

c. You will handle all personal information from this interview confidentially.

d. Your summary of the results of this interview will contain no identifying information such as e.g., name, date of birth, contact details, workplace, etc.

e. You will send this summary of the interview results to your interview partner to check whether you've understood everything right and whether your interview partner is o.k. with it.

f. After receiving the "ok" from your interview partner, you will forward the summary of the results of this interview to the AIDAVA project partners. These interview results will then be merged with the results of several similar other patient interviews, and those aggregated results will form the basis for developing descriptions of archetypical AIDAVA users to support user-centric design and development of the AIDAVA "automatic data cleaning machine".

6. Before proceeding with the interview, check with your interview partner if everything is clear. Answer any questions your interview partner may have regarding AIDAVA or regarding this interview.
7. If you want to record the interview: inform your interview partner why you want to record this interview session, explain to your interview partner what you are going to do with the recordings, and ask for permission.
8. If you do a face-to-face interview, provide two(!) paper versions of the Informed Consent sheet and ask your interview partner to sign these. (One copy is for your interview partner, and the other is for your records.)
   [In case you do a remote interview, you should have obtained the informed consent of your interview partner already before the interview via email.]
9. When all questions of the interviewee are answered and the Informed Consent has been signed by the interviewee, ask the interviewee, whether they are happy to proceed (and remind them that they can withdraw at any time).
   → If everything is o.k., start with the interview (see the following sections).

2 Collect some Personal Information about the Interviewee

1. Age (range): 18-40 years, 40-60 years, >60 years
2. How would you briefly describe/introduce yourself?

3 Ask about Digital Literacy and Data Literacy

1. Do you frequently work with a computer?
   a. *Follow up question, if yes:* What tasks/activities are you usually using the computer for?
2. Do you use a smartphone?
   . *If yes:* What tasks/activities are you usually using your smartphone for?
3. (*If not already known from previous answers of the interviewee:*)
   Do you have an educational background related to computer or data science?
4. What do you think about data rights and data privacy aspects?

4 Ask about Attitude and Approach towards new Technologies

1. Are you usually reading the manual when you use a new device or application for the first time?
2. When you think of getting a new smartphone, which feelings/thoughts come to your mind? - Would you be excited and eager to explore and try out all these new functionalities? Or would you rather be worried about giving up the old device you are familiar with and having to spend time learning and adapting to the new technology?

5 Ask about Medical/Health Literacy

1. How would you rate your level of specific medical knowledge regarding your (chronic) disease on a scale from 0 to 10, where 10 means “excellent knowledge” and 0 means “no knowledge”?
2. How would you rate your level of familiarity with general medical terms / your level of understanding of general medical terms on a scale from 0 to 10, where 10 means “very familiar with general medical terms”/”very good understanding of general medical terms” and 0 means “not at all familiar with medical terms”/”do not understand medical terms”
3. (*If not already known from previous answers of the interviewee:*)
   Do you have an educational background related to medicine and healthcare?

6 Ask about Current Approaches Regarding Personal Health Data

1. Are you collecting any data, documents or medical records related to your personal health?
   a. *Follow-up questions, if yes:*
      i. Which data or documents about your personal health do you collect?
      ii. How do you get these data/documents? (Describe the process/steps you have to take to retrieve these data/documents)
      iii. In which format do you get these data/documents? (electronic/paper?)
      iv. Where do you keep these data and documents and how do you organise them to keep the overview?
      v. Why (for what purpose) do you collect your personal health data and medical records?
      vi. What are the main issues you face when collecting your medical records?
   b. Follow-up questions, if no:

.Why don't you collect data or documents about your personal health?

i.Do you know how you could access your personal health records kept by your doctor or hospital?

2. Are you monitoring any health-related parameters at home (e.g. blood pressure, blood-glucose level…)?

. *Follow-up questions, if yes:*

.Do you share the results of these measurements with your doctor?

- How do you share these measurements with your doctor?

3. Are you using a fitness-tracker app or device?
4. What do you think about sharing your personal health data?

• 7 Ask about Attitude towards "Automatic Health Data Cleaning"

AIDAVA is developing a concept for a "health data cleaning machine" for automated consolidation and curation of a patient's health data from various sources. This is meant to support patients in creating and keeping a longitudinal record of their personal health data, and it should help them to share these data with their doctors or for research purposes.

1. If such an "automatic health data cleaning machine" would exist in the future, would you use it?

a. Follow-up question: Why? / Why not?

2. What would you expect from an **ideal** "automatic health data cleaning machine" in the future?

8.3.2 Data User Interviews

AIDAVA

AI powered Data Curation & Publishing Virtual Assistant.

AIDAVA

Task 1.2

Guidelines for Data Users' Interview

Lead partners: MUG
Other partners: NEMC, b!loba, UM, ECPC, EHN, MID, DME

Coordinator: Michel Dumontier,
University of Maastricht,
The Netherland

Project website: www.aidava.eu

Grant Agreement: 101057062

Horizon Europe
THE NEXT EU RESEARCH & INNOVATION PROGRAMME (2021 – 2027)
European Commission

Table of Contents

General Instructions for the Interviewer

• Selection Criteria for Interviewees

1. Inclusion criteria:
    a. Age: at least 18 years
    b. Job: the task of extracting (hospital) patients’ data from several databases and making data available for research as well as organising and managing research data
    c. Gender: all
2. Exclusion criteria:

. Child/minor
a. Vulnerable person
b. A person, who is not involved in retrieving (hospital) patients’ data for research purposes and in organising and managing research data

---

IMPORTANT: Ensure that you are authorised to approach interview participants.

• Informed Consent

Before you can conduct such an interview, you need to obtain informed consent from your interview partner. Thus, it is essential to ask your interview partner to sign the Informed Consent Sheet. Thereby, the suggested procedure to obtain informed consent from your interview partner depends on whether you plan to conduct a face-to-face or remote interview:

- In case you conduct a *remote interview*:
  Send the Informed Consent sheet to your interview partner via email well before the interview (e.g. when you agree with your interview partner on a date) and ask them to return you a signed copy.
- In case you conduct a *personal face-to-face interview*:
  Ask your interview partner to sign the Informed Consent sheet immediately before starting with the interview questions, after you have explained and "set the scene" (see section "1 Set the Scene" below). Please prepare two(!) printouts of the Informed Consent sheet to be signed → Give one to the interview participant and keep the second one for your records.

• Aims of this Interview

We want to learn more about people who make (hospital) patients' data accessible and usable for research (- in the AIDAVA context we call them "data users") since they are potential future users of the "AIDAVA data cleaning assistant". This will help us follow human-centred design principles and consider the users' needs, abilities, skills, constraints and preferences during the design and development of the AIDAVA system.
The results of this interview together with the results of several similar interviews with other "data users" will form the basis for the development of archetypical users, so-called *personas. Personas* is a method well-known in the human-computer interaction (HCI) field. It was introduced to help designers and developers focus on the needs and goals of the target users throughout the product development process. In AIDAVA these personas will form the basis for elaborating user profiles.
Specifically, this interview shall help to understand better

- data users' attitudes, goals, motivations, frustrations, challenges and pain points
- data users' aptitudes, competencies, knowledge, skills and experience
- data users' work context and framework conditions
- data users' motivations to make (hospital) patients' health data accessible and available for research
- data users' vision of an ideal "automatic data curation & publication" tool

• Important Notes

- This is an interview guideline. Do not strictly stick to these questions, but react on your interview partner:
  - Ask follow-up questions, whenever clarification of some aspects is needed;
  - Change the order of the questions or skip questions, whenever you deem this appropriate for the specific interview;
  - Adapt to your interview partner and use the appropriate language: you may need to modify the wording of some questions, or you may need to conduct the interview in your interview partner's national language.
- Please, always keep in mind: Your goal is to learn from your interview partner! It is not the goal to confirm your opinion or any stereotypes. Therefore:
  - Ask open questions to motivate your interview partner to tell their story.
  - Avoid leading questions and do not introduce bias to the answers of your interview partner.
- Recommendations for conducting this interview:
  - The interview should be conducted by two persons, if possible: one asks the questions and interacts with the interviewee, and the other observes and takes notes.
  - Take enough time, avoid time pressure and hurry: This interview will usually take about 40-60 minutes.
- After the interview:

  - Transcribe the answers of the interviewee, and translate them into English if the interview was conducted in another language.
  - Anonymise your interview report, so that it does not contain any identifying information: for example, replace names of persons with roles, and remove any explicit naming of workplace, organisations or city.
  - Download a copy of the provided template and use it to prepare a results summary by grouping the results according to the topics: Personal Information, Digital/Data Literacy, Attitude and Approach towards new technologies, Medical/Health Literacy, Job / Knowledge / Skills / Traits, Current Approaches regarding health data retrieval, Attitude towards "Automatic Health Data Curation and Publishing"
  - Send this results' summary to the interviewee to check whether you've understood everything right and the interviewee is o.k. with it.
  - After receiving the interviewee's "ok", please upload the results summary to the respective folder in the google drive.

Interview Guideline

1 Set the Scene

1. Thank your interview partner for taking the time to participate in this interview.
2. Briefly introduce yourself, and the note-taker if applicable.
3. Briefly introduce the AIDAVA project.
   → You can use parts of the presentation from patients consultants onboarding (e.g. slides 3, 15, 22)
   https://docs.google.com/presentation/d/16KA_tuL1QVzVK3RMsOetev7DOq66OVobbMckDaZPmb4/edit#slide=id.p1
4. Briefly introduce the purpose of this interview:
   The results of this interview will be used by the AIDAVA consortium to develop descriptions of archetypical users, so-called "personas", which will support the user-centred design and development of the AIDAVA system.
5. Inform your interview partner that

a. Your interview partner may refuse to answer a question without stating a reason.
b. Your interview partner may stop the interview at any time.
c. You will handle all personal information from this interview confidentially.
d. Your summary of the results of this interview will contain no identifying information such as e.g., name, date of birth, contact details, workplace, etc.
e. You will send this summary of the interview results to your interview partner to check whether you've understood everything right and whether your interview partner is o.k. with it.
f. After receiving the "ok" from your interview partner, you will forward the summary of the results of this interview to the AIDAVA project partners.

The results of this interview will then be merged with the results of several other similar interviews, and these aggregated results will be used to develop descriptions of archetypical users to support user-centric design and development of the AIDAVA "automatic data curation and publishing machine".

6. Before proceeding with the interview, check with your interview partner, if everything is clear, and answer any questions your interview partner may have regarding AIDAVA or regarding this interview.
7. If you want to record the interview: inform your interview partner why you want to record this interview session, explain to your interview partner what you are going to do with the recordings, and ask for permission.
8. If you do a face-to-face interview, provide two paper versions of the Informed Consent sheet and ask your interview partner to sign these. (One copy is for your interview partner, and the other one is for your records.)

[In case you do a remote interview, you should have obtained the informed consent of your interview partner already before the interview via email.]

9. When all questions of the interviewee are answered and the Informed Consent has been signed by the interviewee, ask the interviewee, whether they are happy to proceed (and remind them that they can withdraw at any time).
   → If everything is o.k., start with the interview (see the following sections).

2 Collect some Personal Information about the Interviewee

1. Age (range): 18-40 years, 40-60 years, >60 years
2. How would you briefly describe/introduce yourself?

3 Ask about Digital Literacy and Data Literacy

1. How would you rate your knowledge and skills as a computer user on a scale from 0 to 10, where 10 means "excellent computer-user skills" and 0 means "no computer-user skills"?
2. (*If not already known from previous answers of the interviewee:*)
   Do you have an educational background related to computer or data science?
3. What do you personally think about data rights and data privacy aspects?

4 Ask about Attitude and Approach towards new Technologies

1. Are you usually reading the manual when you use a new device or application for the first time?
2. When you think of getting a new smartphone, which feelings and thoughts come to your mind? - Would you be excited and eager to explore and try out all these new functionalities? Or would you rather be worried about giving up the old device you are familiar with and having to spend time learning and adapting to the new technology?

5 Ask about Medical/Health Literacy

1. How would you rate your level of familiarity with medical terms / your level of understanding of medical terms on a scale from 0 to 10, where 10 means "very familiar with medical terms"/"very good understanding of medical terms" and 0 means "not at all familiar with medical terms"/"do not understand medical terms"
2. (*If not already known from previous answers of the interviewee:*)
   Do you have an educational background related to medicine and healthcare?

6 Ask about Job / Knowledge / Skills / Traits

1. What is your profession?
2. What are your (main) tasks on a typical workday?
3. Which knowledge and skills are necessary for your job?
4. Which education do you have?
5. In your opinion, which personality traits are needed to do your job well?
6. What do you especially like at your job?
7. What are the main challenges in your job?
8. What annoys or frustrates you the most?
9. What are your motivating factors for doing your job well?

• 7 Ask about Current Approaches Regarding Health Data Retrieval

1. Please tell me more about your tasks and responsibilities with respect to retrieving patients' health data and making these data accessible and usable for research.

a. At which occasions / for what purpose are you retrieving patients' health data?

b. Which kind of data? From which data sources do you have to collect the data? In which format do you get the data? How many data records do you usually have to retrieve in one batch?...

c. Describe in detail how you retrieve patients' health data and make these data accessible and usable for research (i.e. describe the process(es) and steps you follow).

2. What are the main challenges and issues/pains you are facing in these tasks?

• 8 Ask about Attitude towards "Automatic Health Data Curation & Publishing"

AIDAVA is developing a concept for an "automatic health data curation and publishing machine" for automated consolidation and curation of patients' health data from various sources to make it machine-readable and reusable, e.g. for research purposes.

1. If such an "automatic health data curation & publishing machine" would exist in the future, what pre-conditions/requirements must it fulfil so that you would use it?
2. What would you expect from an **ideal** "automatic health data curation and publishing machine" in the future?

8.3.3 Data Curator Interviews

AIDAVA
AI powered Data Curation & Publishing Virtual Assistant.

Task 1.2
Guidelines for Data Curators' Interview

Lead partners: MUG
Other partners: NEMC, b!loba, UM, ECPC, EHN, MID, DME

Coordinator: Michel Dumontier,
University of Maastricht,
The Netherland

Project website: www.aidava.eu

Grant Agreement: 101057062

Horizon Europe
THE NEXT EU RESEARCH & INNOVATION PROGRAMME (2021 – 2027)
European Commission

Table of Contents

General Instructions for the Interviewer

• Selection Criteria for Interviewees

1. Inclusion criteria:
    a. Age: at least 18 years
    b. Job: the task of collecting/curating (hospital) patients’ data and entering data into the (hospital’s) information system or compiling/editing research documentation for clinical studies is part of the interviewee’s work
    c. Gender: all
2. Exclusion criteria:

. Child/minor

a. Vulnerable person

b. A person, who is neither involved in collecting/curating (hospital) patients’ data and entering data into the (hospital’s) information system nor in compiling or editing research documentation for clinical studies

---

IMPORTANT: Ensure that you are authorised to approach interview participants.

• Informed Consent

Before you can conduct such an interview, you need to obtain informed consent from your interview partner. Thus, it is essential to ask your interview partner to sign the Informed Consent Sheet. Thereby, the suggested procedure to obtain informed consent from your interview partner depends on whether you plan to conduct a face-to-face or remote interview:

- In case you conduct a *remote interview*:
  Send the Informed Consent sheet to your interview partner via email well before the interview (e.g. when you agree with your interview partner on a date) and ask them to return you a signed copy.
- In case you conduct a *personal face-to-face interview*:
  Ask your interview partner to sign the Informed Consent sheet immediately before starting with the interview questions, after you have explained and "set the scene" (see section "1 Set the Scene" below). Please prepare two(!) printouts of the Informed Consent sheet to be signed → Give one to the interview participant and keep the second one for your records.

• Aims of this Interview

We want to learn more about data curators (i.e. people who curate (hospital) patients' data), who are potential future users of the "AIDAVA data cleaning assistant". This will help us follow human-centred design principles and consider the users' needs, abilities, skills, constraints and preferences during the design and development of the AIDAVA system.
The results of these interviews will form the basis for the development of archetypical users, so-called *personas. Personas* is a method well-known in the human-computer interaction (HCI) field. It was introduced to help designers and developers focus on the needs and goals of the target users throughout the product development process. In AIDAVA these personas will form the basis for elaborating user profiles.
Specifically, this interview shall help to understand better

- data curators' attitudes, goals, motivations, frustrations, challenges and pain points
- data curators' aptitudes, competencies, knowledge, skills and experience
- data curators' work context and framework conditions
- data curators' motivations to control and curate (hospital) patients' health data
- data curators' vision of an "ideal" data curation tool

• Important Notes

- This is an interview guideline. Do not strictly stick to these questions, but react on your interview partner:
  - Ask follow-up questions, whenever clarification of some aspects is needed;
  - Change the order of the questions or skip questions, whenever you deem this appropriate for the specific interview;
  - Adapt to your interview partner and use the appropriate language: you may need to modify the wording of some questions, or you may need to conduct the interview in your interview partner's national language.
- Please, always keep in mind: Your goal is to learn from your interview partner! It is not the goal to confirm your opinion or any stereotypes. Therefore:
  - Ask open questions to motivate your interview partner to tell their story.
  - Avoid leading questions and do not introduce bias to the answers of your interview partner.
- Recommendations for conducting this interview:
  - The interview should be conducted by two persons, if possible: one asks the questions and interacts with the interviewee, and the other observes and takes notes.
  - Take enough time, avoid time pressure and hurry: This interview will usually take about 40-60 minutes.
- After the interview:
  - Transcribe the answers of the interviewee, and translate them into English if the interview was conducted in another language.

- Anonymise your interview report, so that it does not contain any identifying information: for example, replace names of persons with roles, and remove any explicit naming of workplace, organisations or city.
- Download a copy of the provided template and use it to prepare a results summary by grouping the results according to the topics: Personal Information, Digital/Data Literacy, Attitude and Approach towards new technologies, Medical/Health Literacy, Job / Knowledge / Skills / Traits, Current Approaches regarding health data curation, Attitude towards "Automatic Health Data Cleaning"
- Send this results' summary to the interviewee to check whether you've understood everything right and whether the interviewee is o.k. with it.
- After receiving the interviewee's "ok", please upload the results summary to the respective folder in the google drive.

Interview Guideline

1 Set the Scene

1. Thank your interview partner for taking the time to participate in this interview.
2. Briefly introduce yourself, and the note-taker if applicable.
3. Briefly introduce the AIDAVA project.
→ You can use parts of the presentation from patients consultants onboarding (e.g. slides 3, 15, 21, 22)
https://docs.google.com/presentation/d/16KA_tuL1QVzVK3RMsOetev7DOq66OVobbMckDaZPmb4/edit#slide=id.p1
4. Briefly introduce the purpose of this interview:
The results of this interview will be used by the AIDAVA consortium to develop descriptions of archetypical users, so-called "personas", which will support the user-centred design and development of the AIDAVA system.
5. Inform your interview partner that

a. Your interview partner may refuse to answer a question without stating a reason.
b. Your interview partner may stop the interview at any time.
c. You will handle all personal information from this interview confidentially.
d. Your summary of the results of this interview will contain no identifying information such as e.g., name, date of birth, contact details, workplace, etc.
e. You will send this summary of the interview results to your interview partner to check whether you've understood everything right and whether your interview partner is o.k. with it.
f. After receiving the "ok" from your interview partner, you will forward the summary of the results of this interview to the AIDAVA project partners.

The results of this interview will then be merged with the results of several other similar interviews, and these aggregated results will be used to develop descriptions of archetypical users to support user-centric design and development of the AIDAVA "automatic health data cleaning machine".

6. Before proceeding with the interview, check with your interview partner, if everything is clear, and answer any questions the interviewee may have regarding AIDAVA or this interview.
7. If you want to record the interview: inform your interview partner why you want to record this interview session, explain to your interview partner what you are going to do with the recordings, and ask for permission.
8. If you do a face-to-face interview, provide two paper versions of the Informed Consent sheet and ask your interview partner to sign these. (One copy is for your interview partner, and the other one is for your records.)
[In case you do a remote interview, you should have obtained the informed consent of your interview partner already before the interview via email.]

9. When all questions of the interviewee are answered and the Informed Consent has been signed by the interviewee, ask the interviewee, whether they are happy to proceed (and remind them that they can withdraw at any time).
   → If everything is o.k., start with the interview (see the following sections).

2 Collect some Personal Information about the Interviewee

1. Age (range): 18-40 years, 40-60 years, >60 years
2. How would you briefly describe/introduce yourself?

3 Ask about Digital Literacy and Data Literacy

1. How would you rate your knowledge and skills as a computer user on a scale from 0 to 10, where 10 means “excellent computer-user skills” and 0 means “no computer-user skills”?
2. (*If not already known from previous answers of the interviewee:*)
   Do you have an educational background related to computer or data science?
3. What do you personally think about data rights and data privacy aspects?

4 Ask about Attitude and Approach towards new Technologies

1. Are you usually reading the manual when you use a new device or application for the first time?
2. When you think of getting a new smartphone, which feelings/thoughts come to your mind? - Would you be excited and eager to explore and try out all these new functionalities? Or would you rather be worried about giving up the old device you are familiar with and having to spend time learning and adapting to the new technology?

5 Ask about Medical/Health Literacy

1. How would you rate your level of familiarity with medical terms / your level of understanding of medical terms on a scale from 0 to 10, where 10 means “very familiar with medical terms”/”very good understanding of medical terms” and 0 means “not at all familiar with medical terms”/”do not understand medical terms”
2. (*If not already known from previous answers of the interviewee:*)
   Do you have an educational background related to medicine and healthcare?

6 Ask about Job / Knowledge / Skills / Traits

1. What is your profession?
2. What are your (main) tasks on a typical workday?
3. Which knowledge and skills are necessary for your job?
4. Which education do you have?
5. In your opinion, which personality traits are needed to do your job well?
6. What do you especially like at your job?
7. What are the main challenges in your job?
8. What annoys or frustrates you the most?
9. What are your motivating factors for doing your job well?

• 7 Ask about Current Approaches Regarding Health Data Curation

1. Please tell me more about your tasks and responsibilities with respect to handling patients' health data.

a. At which occasions / for what purpose are you handling patients' health data?
b. Which kind of data? From which data sources do you have to collect the data? In which format do you get the data?
c. Describe the process(es) and steps you follow / how you do these tasks

2. What are the main challenges and issues you are facing in these tasks?

• 8 Ask about Attitude towards "Automatic Health Data Curation"

AIDAVA is developing a concept for a "health data cleaning machine" for automated consolidation and curation of a patient's health data from various sources to make it machine-readable and re-usable.

1. If such an "automatic health data cleaning machine" would exist in the future, what pre-conditions/requirements must it fulfil that you would use it?
2. What would you expect from an **ideal** "automatic health data cleaning machine" in the future?

8.3.4 3rd Party App Developers Interviews

AIDAVA
AI powered Data Curation & Publishing Virtual Assistant.

AIDAVA

Task 1.2
Guidelines for App Developers' Interview

Lead partners: MUG
Other partners: NEMC, b!loba, UM, ECPC, EHN, MID, DME

Coordinator: Michel Dumontier,
University of Maastricht,
The Netherland

Project website: www.aidava.eu

Grant Agreement: 101057062

Horizon Europe
THE NEXT EU RESEARCH & INNOVATION PROGRAMME (2021 – 2027)
European Commission

Table of Contents

General Instructions for the Interviewer

• Selection Criteria for Interviewees

1. Inclusion criteria:
    a. Age: at least 18 years
    b. Job: Application developer, ideally someone who has got experience with providing/retrieving health-related data
    c. Gender: all
2. Exclusion criteria:

. Child/minor
a. Vulnerable person

IMPORTANT: Ensure that you are authorised to approach interview participants.

• Informed Consent

Before you can conduct such an interview, you need to obtain informed consent from your interview partner. Thus, it is essential to ask your interview partner to sign the Informed Consent Sheet. Thereby, the suggested procedure to obtain informed consent from your interview partner depends on whether you plan to conduct a face-to-face or remote interview:

- In case you conduct a *remote interview*:
  Send the Informed Consent sheet to your interview partner via email well before the interview (e.g. when you agree with your interview partner on a date) and ask them to return you a signed copy.

- In case you conduct a *personal face-to-face interview*:
  Ask your interview partner to sign the Informed Consent sheet immediately before starting with the interview questions, after you have explained and "set the scene" (see section "1 Set the Scene" below). Please prepare two(!) printouts of the Informed Consent sheet to be signed → Give one to the interview participant and keep the second one for your records.

Aims of this Interview

We want to learn more about people who may potentially in the future develop applications for smartphone/computer utilizing AIDAVAs APIs to either transfer medical data to the "AIDAVA data cleaning assistant" or retrieve medical data from the "AIDAVA data cleaning assistant". This will help us follow human-centred design principles and consider the users' needs, abilities, skills, constraints and preferences during the design and development of the AIDAVA system.

The results of this interview together with the results of several similar interviews with other app developers will form the basis for the development of archetypical users, so-called *personas. Personas* is a method well-known in the human-computer interaction (HCI) field. It was introduced to help designers and developers focus on the needs and goals of the target users throughout the product development process.

Specifically, this interview shall help to understand better

- App developers' attitudes, goals, motivations, frustrations, challenges and pain points
- App developers' aptitudes, competencies, knowledge, skills and experience
- App developers' work context and framework conditions
- App developers' vision of an ideal "automatic data curation & publication" tool

Important Notes

- This is an interview guideline. Do not strictly stick to these questions, but react on your interview partner:
  - Ask follow-up questions, whenever clarification of some aspects is needed;
  - Change the order of the questions or skip questions, whenever you deem this appropriate for the specific interview;
  - Adapt to your interview partner and use the appropriate language: you may need to modify the wording of some questions, or you may need to conduct the interview in your interview partner's national language.
- Please, always keep in mind: Your goal is to learn from your interview partner! It is not the goal to confirm your opinion or any stereotypes. Therefore:
  - Ask open questions to motivate your interview partner to tell their story.
  - Avoid leading questions and do not introduce bias to the answers of your interview partner.
- Recommendations for conducting this interview:
  - The interview should be conducted by two persons, if possible: one asks the questions and interacts with the interviewee, and the other observes and takes notes.
  - Take enough time, avoid time pressure and hurry: This interview will usually take about 40-60 minutes.
- After the interview:
  - Transcribe the answers of the interviewee, and translate them into English if the interview was conducted in another language.
  - Anonymise your interview report, so that it does not contain any identifying information: for example, replace names of persons with roles, and remove any explicit naming of workplace, organisations or city.
  - Download a copy of the provided template and use it to prepare a results summary by grouping the results according to the topics: Personal

Information, Digital/Data Literacy, Attitude and Approach towards new technologies, Medical/Health Literacy, Job / Knowledge / Skills / Traits, Current Approaches regarding health data retrieval, Attitude towards "Automatic Health Data Curation and Publishing"

- Send this results' summary to the interviewee to check whether you've understood everything right and the interviewee is o.k. with it.
- After receiving the interviewee's "ok", please upload the results summary to the respective folder in the google drive.

Interview Guideline

1 Set the Scene

1. Thank your interview partner for taking the time to participate in this interview.
2. Briefly introduce yourself, and the note-taker if applicable.
3. Briefly introduce the AIDAVA project.
   → You can use parts of the presentation from patients consultants onboarding (e.g. slides 3, 15, 22)
   https://docs.google.com/presentation/d/16KA_tuL1QVzVK3RMsOetev7DOq66OVobbMckDaZPmb4/edit#slide=id.p1
4. Briefly introduce the purpose of this interview:
   The results of this interview will be used by the AIDAVA consortium to develop descriptions of archetypical users, so-called "personas", which will support the user-centred design and development of the AIDAVA system.
5. Inform your interview partner that

a. Your interview partner may refuse to answer a question without stating a reason.

b. Your interview partner may stop the interview at any time.

c. You will handle all personal information from this interview confidentially.

d. Your summary of the results of this interview will contain no identifying information such as e.g., name, contact details, workplace, etc.

e. You will send this summary of the interview results to your interview partner to check whether you've understood everything right and whether your interview partner is o.k. with it.

f. After receiving the "ok" from your interview partner, you will forward the summary of the results of this interview to the AIDAVA project partners.

The results of this interview will then be merged with the results of several other similar interviews, and these aggregated results will be used to develop descriptions of archetypical users to support user-centric design and development of the AIDAVA "automatic data curation and publishing machine".

6. Before proceeding with the interview, check with your interview partner, if everything is clear, and answer any questions your interview partner may have regarding AIDAVA or regarding this interview.
7. If you want to record the interview: inform your interview partner why you want to record this interview session, explain to your interview partner what you are going to do with the recordings, and ask for permission.
8. If you do a face-to-face interview, provide two paper versions of the Informed Consent sheet and ask your interview partner to sign these. (One copy is for your interview partner, and the other one is for your records.)
   [In case you do a remote interview, you should have obtained the informed consent of your interview partner already before the interview via email.]
9. When all questions of the interviewee are answered and the Informed Consent has been signed by the interviewee, ask the interviewee, whether they are happy to proceed (and remind them that they can withdraw at any time).
   → If everything is o.k., start with the interview (see the following sections).

2 Collect some Personal Information about the Interviewee

1. Age (range): 18-40 years, 40-60 years, >60 years
2. How would you briefly describe/introduce yourself?

3 Ask about Data Literacy and Approach towards new Technologies

1. Are you usually reading the manual when you use a new device or application for the first time?
2. Are you eager to try out new techniques or rather prefer to use technologies you are familiar with?
3. What do you personally think about data rights and data privacy aspects?
4. In your opinion, are there any peculiarities to take into account when handling medical/health data?

4 Ask about Medical/Health Literacy

1. How would you rate your level of familiarity with medical terms / your level of understanding of medical terms on a scale from 0 to 10, where 10 means “very familiar with medical terms”/”very good understanding of medical terms” and 0 means “not at all familiar with medical terms”/”do not understand medical terms”
2. (*If not already known from previous answers of the interviewee:*)
   Do you have an educational background related to medicine and healthcare?

5 Ask about Job / Knowledge / Skills / Traits

1. What are your (main) tasks on a typical workday?
2. Which knowledge and skills are necessary for your job?
3. In your opinion, which personality traits are needed to do your job well?
4. What do you especially like at your job?
5. What annoys or frustrates you the most?

6 Ask about Current Approaches towards health data and third-party APIs

1. Do you have experience with developing health data related apps?
2. How is your (company’s) work situated in the health data ecosystem?
3. What is your (company’s) approach to storing health data? (How do you store health data? (structures, ontologies))
4. What APIs are you using?
5. When you are utilizing APIs of third party applications, how do you usually approach this task?
6. What are the main challenges and hurdles you are facing in this task?

7 Discuss the AIDAVA concept from the perspective of the use cases their company has

AIDAVA is developing a concept for an “automatic health data curation and publishing machine” for automated consolidation and curation of patients' health data from various sources to make it machine-readable and reusable.

1. If such an “automatic health data curation & publishing machine” would exist in the future, what pre-conditions/requirements must it fulfil so that you as an app developer (and your company) would use it?
2. What would you as an app developer (and your company) expect from an **ideal** “automatic health data curation and publishing machine” in the future?

8.4 Informed consent

AIDAVA

AI powered Data Curation & Publishing Virtual Assistant.

Task 1.2

Informed Consent to Interviews

Lead partners: MUG
Other partners: NEMC, b!loba, UM, ECPC, EHN, MID, DME

Coordinator: Michel Dumontier,
University of Maastricht,
The Netherland

Project website: www.aidava.eu

Grant Agreement: 101057062

Horizon Europe
THE NEXT EU RESEARCH & INNOVATION PROGRAMME (2021 – 2027)
European Commission

Note:
This is a suggestion for an informed consent form for the T1.2 interviews in AIDAVA. Please, adjust it according to your organisation's usual procedures.

Informed Consent for Participation in a User Interview of the AIDAVA Project

Dear Participant,

We cordially invite you to participate in a user interview within the project AIDAVA. For that, we need your written consent. Please read the following text carefully and do not hesitate to ask, if anything is unclear to you.

1 What is the AIDAVA project?

AIDAVA ("AI-powered Data Curation & Publishing Virtual Assistant") is a cooperative research project funded by the European Union, including 14 organisations from 9 European countries. AIDAVA addresses a current problem: patient's health data are scattered across different clinics, physicians, hospitals, healthcare providers and, increasingly, across medical devices and personal health apps. There is inconsistent content across these data sources, and, in addition, much of the data is still in paper and narrative form. Consequently, patients do not have easy access to and full control of their data. Moreover, costly and recurring manual data cleaning work is needed to make patients' health data useful for clinical care and clinical research.

To tackle these problems, AIDAVA aims to develop a prototype of an "automatic health data cleaning machine" that should support patients and clinical data stewards (assistants) to integrate health data from different sources and increase the quality of these data. Availability of high-quality health data is an essential basis for further improvements in preventive medicine, quality of care and clinical research.

2 What is the purpose of this interview?

The AIDAVA consortium wants to learn more about patients, who are the potential future users of the "AIDAVA data cleaning assistant". Specifically, this interview shall help to better understand your motivations, challenges, aptitudes and personal context with regard to health data, as well as your vision of an "ideal" automatic health data cleaning tool.

The results of this and other patients' interviews will be used internally by the AIDAVA consortium to define the description of archetypical users. These will form the basis for elaborating user-profiles and help to consider the users' needs, aptitudes, constraints and preferences throughout the design and development of the AIDAVA system. W

3 How will this interview be conducted?

I. A team of two people will conduct this interview: One person (the interviewer) will guide you through the interview and ask you a series of questions. The second person (the note-taker) will carefully listen to your answers and take notes.
If you agree, we would like to record this interview to facilitate the work of the note-taker. After the interview, that recording will together with the notes be used by the

interview team to write a summary of the interview results and will be deleted afterwards.

II. The interview questions will address your professional/educational background, your computer and smartphone experience, your approach towards new technologies, your experience with personal health data, and your visions regarding automatic health data cleaning.

III. You may refuse to answer any of the questions without stating a reason.

IV. Please do not hesitate to ask anytime, if anything is unclear to you or if you need additional information.

V. The interview will take approximately one hour, but you may stop the interview at any time.

4 What happens with the results of the interview?

- All personal information from this interview will be handled confidentially by the interview team and will not be disclosed to other people.
- The written summary of the interview results will contain no identifying information such as e.g., name, date of birth, contact details, workplace, etc.
- You will receive the written summary of the interview results so that you can check and ensure that everything was understood right by the interviewer.
- If you raise no objections, the anonymised written summary of the results of this interview will be forwarded to the AIDAVA project partners. The AIDAVA partners will merge it with the summaries of the results of other similar interviews. These aggregated results from several patient interviews will then form the basis for developing descriptions of archetypical users to support user-centric development of the AIDAVA system.
- You may demand to stop the usage of the unaggregated results of this interview at any time.

5 Further information

AIDAVA project website: https://www.aidava.eu/

AIDAVA project coordination:

Remzi Celebi	*Isabelle de Zegher*
Universiteit Maastricht, The Netherlands	b!loba, Belgium
remzi.celebi@maastrichtuniversity.nl	isabelle@dezegher.com

6 Data protection

This interview involves the collection of personal data that can identify you. These personal data include contact information, personal experiences and opinions.

The personal data you provide will be used for the following purposes:

- To determine eligibility criteria for interview participation
- To contact you after the interview and let you check the written summary of the interview results
- To comply with legal and regulatory requirements, including requirements to share data with agencies overseeing the research
- To confirm proper conduct and integrity of the research
- To fulfil the purpose of this interview, as stated in point 2 of this Informed Consent Sheet

All personal data that can identify you will be retained by the interview team as long as necessary to fulfil the objectives of the interview and to ensure the integrity of the research. Your personal data will be deleted when it is no longer needed for the research project or if you withdraw your consent.

The personal data that can identify you will be kept solely by the interview team. As described in point 4 of this document, no personal data that can identify you will be given to third parties, and the results of this interview will be provided to the AIDAVA project partners only in anonymised form.

The European Union's General Data Protection Regulation (GDPR) affords you certain rights with respect to your personal data, including the right to:

- Access, correct, withdraw, or delete your personal data.
- Restrict the types of activities the research team can do with your personal data.
- Withdraw your consent to use your personal data for the purposes outlined in this document.

If you want to make a request relating to the rights listed above or if you have any concerns about how your personal data is being handled, please contact:

Affiliation of the interview team: Please insert here the name of the organisation
Name of the interviewer: Please insert here the name of the interviewer
Contact details of the interviewer: Please insert here EMAIL, PHONE of the interviewer

You may also contact the Data Protection Officer of Please enter here the name of your organisation: Please insert here the EMAIL of the Data Protection Officer

7 Informed Consent Declaration

Name of the participant: …………

I have been thoroughly informed about the AIDAVA project, the purpose of this interview and my voluntary participation. I have read and understood the text of this informed consent sheet, which comprises 4 pages in total. My questions have been answered completely and comprehensively.

☐ I confirm that I am over 18 years old.

☐ I consent to participate in the user interview for the AIDAVA project. My participation is fully voluntary, and I know that I can withdraw anytime without negative consequences.

☐ I consent to the processing of my information as described in point 4 of this informed consent sheet.

☐ I consent to the recording of this interview as described in point 3-I of this informed consent sheet.

I have got a copy of this informed consent sheet. The original of this informed consent sheet is kept in the records of the interview team.

…………………………………………………………………………………………

Date and Signature of the Participant

…………………………………………………………………………………………

Date, Name and Signature of the Interviewer