# **Original Paper**

# Privacy Perceptions and Concerns in Image-Based Dietary Assessment Systems: Questionnaire-Based Study

Aakash Sharma<sup>1</sup>, MSc; Katja P Czerwinska<sup>2</sup>, BA; Lars Brenna<sup>1</sup>, PhD; Dag Johansen<sup>1</sup>, PhD; Håvard D Johansen<sup>1</sup>, PhD

<sup>1</sup>Department of Computer Science, UiT - The Arctic University of Norway, Tromsø, Norway

<sup>2</sup>Faculty of Design, Computer Science, Media, RheinMain University of Applied Sciences, Wiesbaden, Germany

#### **Corresponding Author:**

Aakash Sharma, MSc Department of Computer Science UiT - The Arctic University of Norway Hansine Hansens veg 18 Tromsø, 9019 Norway Phone: 47 94428984 Email: <u>aakash.sharma@uit.no</u>

# Abstract

**Background:** Complying with individual privacy perceptions is essential when processing personal information for research. Our specific research area is performance development of elite athletes, wherein nutritional aspects are important. Before adopting new automated tools that capture such data, it is crucial to understand and address the privacy concerns of the research subjects that are to be studied. Privacy as contextual integrity emphasizes understanding contextual sensitivity in an information flow. In this study, we explore privacy perceptions in image-based dietary assessments. This research field lacks empirical evidence on what will be considered as privacy violations when exploring trends in long-running studies. Prior studies have only classified images as either private or public depending on their basic content. An assessment and analysis are thus needed to prevent unwanted consequences of privacy breach and other issues perceived as sensitive when designing systems for dietary assessment by using food images.

**Objective:** The aim of this study was to investigate common perceptions of computer systems using food images for dietary assessment. The study delves into perceived risks and data-sharing behaviors.

**Methods:** We investigated the privacy perceptions of 105 individuals by using a web-based survey. We analyzed these perceptions along with perceived risks in sharing dietary information with third parties.

**Results:** We found that understanding the motive behind the use of data increases its chances of sharing with a social group.

**Conclusions:** In this study, we highlight various privacy concerns that can be addressed during the design phase. A system design that is compliant with general data protection regulations will increase participants' and stakeholders' trust in an image-based dietary assessment system. Innovative solutions are needed to reduce the intrusiveness of a continuous assessment. Individuals show varying behaviors for sharing metadata, as knowing what the data is being used for, increases the chance of it being shared.

(JMIR Hum Factors 2020;7(4):e19085) doi: 10.2196/19085

# **KEYWORDS**

privacy perception; privacy; dietary assessment; mobile food records; image-based dietary assessment; data sharing; human factors, mobile phone

# Introduction

# Background

Food images are highly relevant for use in medical research and sport science. They can capture continuous and accurate

http://humanfactors.jmir.org/2020/4/e19085/

measurement of diets, and therefore are imperative in understanding the relationship between food intake and athletic development [1] or between food intake and health problems such as noncommunicable diseases [2]. The ubiquitous and increasingly capable smartphone is, in particular, becoming an essential asset that many studies now include the use of

smartphones to gather data, thereby enabling new findings and development of new research methodologies [3]. The use of smartphone cameras to document meals has already been suggested as a way of improving nutrition research data and generating new insights [4-16]. By importing food pictures into an image-based dietary assessment (IBDA) system, trained professionals can go through individual dietary habits and offer personalized recommendations. Capling et al [17] surveyed the issues in dietary assessment methods in athletes and highlighted the problems of bias, accuracy, and burden on the user. IBDA systems are designed to address those issues. Thompson and Subar [8] argue that IBDA methods have the potential for research as they require less effort compared to traditional dietary assessment techniques for reaching comparable accuracy.

Although research on human subjects is already strictly regulated by local, national, and international boards and procedures, the increased usage of personal information recorded automatically through new technology comes with new concerns for the security and privacy of the subjects. Little attention has been given to the specific individual privacy requirements related to the design of IBDA systems [7,18], and how privacy awareness in larger cohorts can change over time and with regulatory discussions and coverage of privacy controversies in media [19]. For research studies based on IBDA data such as large epidemiological studies and sports science studies [1], the lack of a proper privacy framework for food-related images makes it difficult to follow Privacy by Design [20] guidelines, which recommend incorporating privacy requirements early on from the design phase, and risk not being compliant with legal and ethical laws and regulations. Participation in voluntary studies relies heavily on trust [21], and any damage to reputation can have severe consequences to organizations that obtain data based on informed consent. Thus, it is crucial to understand the privacy perceptions and concerns before implementing solutions at a population-wide scale.

To improve our understanding of the privacy requirements in population-based research data, we conducted a web-based survey wherein subjects were asked about their perceptions of privacy with regard to capturing food images by using a smartphone camera. Taking food pictures is already a trend on many social networks, where people typically post images of their meals during vacations and special events [22], and therefore, this is something that many can relate to. Since our general field of study is performance development of elite athletes, we selected a cohort of young participants (<35 years of age). Further, to avoid selection bias by using just a local cohort of athletes, we selected motivated cohort members from throughout the world. It is important to notice that our ongoing interdisciplinary work has involved sports science and elite athletes of several national (soccer) teams spread throughout Europe. Hence, having similar characteristics such as age, despite not elite athletes yet, in this first study, resembles this distributed target cohort. We purposely did not use the elite athlete cohort in this inaugural study owing to previous experiences of introducing new technologies to them [23,24]; our experience is that the dropout rate of such distant cohort members is way too large after the first week or so. Instead, we

http://humanfactors.jmir.org/2020/4/e19085/

selected a distributed cohort that we knew would be motivated to be data contributors over a longer period.

Another important lesson from previous epidemiological-related work is that the data capturing should not be intrusive. We have previously attempted to use, for instance, 24-hour dietary recall–inspired methods with pictures taken during meals, but this showed to be too intrusive and too time-consuming for elite athletes. Moreover, the dropout rates of these elite athletes were very steep with these methods. Hence, we developed this survey by using alternative schemes for data assessment, wherein pictures of meals were captured similar to that captured in social media engagement.

From a more general perspective on privacy, food images offer an interesting case to study, as few food images might not carry much sensitive information. However, a large individual data set of images that is continuously recorded over long periods (>2 weeks) and linked to an individual's identity might disclose information that many find too sensitive to share. Such disclosure is a growing public concern [25,26] and therefore is an interesting use case for us to explore. Our hope is that the insight gained in our survey will be useful when designing data collection projects, thereby increasing trust and compliance, which are both necessary for public engagement that most cohort studies rely so heavily upon.

#### **Literature Review**

Systematic reviews of dietary assessment methods [17,27] have highlighted the issues of the burden on the user, accuracy of reported data, and bias in existing methods. Both studies argue about the potential of using IBDAs to address some of these concerns. Various studies [9-16] have validated the effectiveness of an IBDA system. However, they do not discuss any privacy concerns that might arise due to data collection in such tools. Furthermore, privacy risks are amplified by the increased willingness to self-disclose on one's smartphone [28]. In this regard, Christin et al [29] studied potential privacy violations in participatory studies that collect and process sensory data recorded by mobile devices. This work investigates violations such as revealing the location by a global positioning sensor in a mobile phone and provides strategies for safeguarding privacy. Their approach only attempts to identify privacy risks linked to sensory data. Avancha et al [30] studied privacy requirements for personal health care by using mobile technology. Their extensive work investigates privacy in a mobile health (mHealth) context. They elaborately defined a conceptual framework for privacy in mHealth from legal and technological aspects. Their work also provides properties for a privacy-aware mHealth system. Some of the privacy-relevant requirements discussed in this work are inspired by their work.

Zerr et al [31] explored classifying an image as private or public based on its contents. Their work highlights preliminary research focus in this domain. They built a machine learning model from photos marked by humans as private or public. Spyromitros-Xioufis et al [32] expanded on the work of Zerr et al by using classifiers based on the content of the image by using tags (eg, erotic, alcohol, drinking). A further layer of personalization was added by training the categories that a user wants to keep private. Squicciarini et al [33] further improved

XSL•FO RenderX

the classification of Zerr et al [31]. While these approaches lay important groundwork for privacy perception on images, they do not study contextual privacy implications for a specialized purpose such as for dietary assessment. In this work, we attempt to understand the privacy implications of recording dietary intake by using images, and to the best of our knowledge, this topic has not been covered earlier.

The work on IBDA methods by Boushey et al [4,11] and O'Loughlin et al [7] mainly discuss the identification of food from images. Their work does not discuss the privacy implications of recording diets over a long period. Thomaz et al [34] investigated privacy violations while recording eating behaviors with a wearable camera. Their work tries to understand the privacy implications of using a wearable camera. The wearable camera discussed in their work takes an image at periodic intervals, which might capture other images as well. They identify privacy implications such as capturing other people's faces or taking pictures inside a restricted location. The work further addresses these issues by offering novel solutions such as capturing images during certain hours instead of continuous captures. Similar to the work of Thomaz et al [34], Greiner and Yang [35] investigated issues with continuous recording by using a wearable camera for dietary assessments in obesity studies. They recommend postprocessing of the captured video in order to avoid any privacy violations. In our work, we do not target a wearable camera. Rather, we investigate the privacy implications of recording diets by taking images of food by individuals in a continuous study. We define a continuous study as taking pictures of food over a period of time instead of taking pictures periodically as in the study by Thomaz et al [34]. We build upon the fact that individuals are already taking selective pictures of food during vacations and sharing them on social networks [22]. Individuals are often not aware of their privacy being exposed by their data [19,25]. It appears that the awareness about data use, when shared with third parties, is often overlooked.

# Methods

# **Questionnaire and Ethical Approval**

To improve our understanding of the privacy perception related to the capture and use of food images, we conducted a study by using a web-based questionnaire (Multimedia Appendix 1) hosted by Nettskjema, a secure web-based survey tool hosted by the University of Oslo, Norway [36]. The questionnaire was developed using close-ended questions for their statistical analyses [37]. We simplified the questions, added a probability-severity matrix, and refined our goals through multiple pretesting/run-throughs in the laboratory. Some questions are repeated in the questionnaire to reduce biased context [37]. Responses were collected between February and June 2019. Based on our institution's research policy, we applied for ethics approval from the Norwegian Centre for Research Data. We did not collect any personally identifiable information. After a review, we obtained an exemption from the Norwegian Centre for Research Data. The full questionnaire is available in Multimedia Appendix 2 along with the collected data.

## Design

The survey was designed to record participants' perceptions in the following scenarios: (1) Scenario A, the participant having to record and share dietary data as an athlete; (2) Scenario B, the possibility and severity of privacy leak from one's dietary data; and (3) Scenario C, sharing dietary data and reports among different social groups.

The scenarios were designed to be familiar to our participants and to cover various angles on data sharing. For Scenario A, sharing is both internally and externally motivated and controlled by the subject; however, the subject is not in control of the processing. For Scenario B, the subject is not in control of the data processing but has concerns about the processing outcomes, and for C, the sharing is consensual but is based on external motivation from different social groups, for example, receiving feedback from a doctor or sharing with family/friends as part of social behavior/interaction.

We used these scenarios to record our participants' perceptions and attitudes toward sharing data. Note that we considered the perceptions on a scenario valid even for participants who never encountered that scenario in real life beforehand. One's perceptions can affect one's participation in a study if the concerns are not addressed at the beginning of a study.

Our questionnaire starts by familiarizing participants with food pictures on social networks (Scenario A). It then asks about the use of social networks and experience with taking food pictures. Additionally, participants are asked about their preference of IBDA methods over other similarly used techniques for dietary assessment. Attitudes toward privacy and personal control over data were collected on a 5-level *Likert* item, ranging from *strongly disagree* to *strongly agree*. These include perceptions toward responsibility for privacy, intrusiveness, and general attitude toward dietary practices [22].

Next, the participants are asked to consider Scenario A—an athlete who records his/her diet by taking pictures. We specify that every meal is recorded by taking a picture, including drinks, and even at events outside training by using a mobile app. We specify that the app allows the team owner, manager, coach, and doctor to monitor his/her diet and recommend diet plans. Continuing with questions from Scenario A, we further obtain responses toward the privacy and usability of metadata collected through such a system. Table 1 shows an excerpt of the questionnaire. For the complete questionnaire, refer to the Multimedia Appendix 2.



Table 1. Few questions from the questionnaire with their possible responses.

| Question category,<br>symbol         | Question text  | Response             |  |  |  |
|--------------------------------------|--|----------------------|--|--|--|
| Social media usage                   |  |                      |  |  |  |
| G1                                   | Do you use any social media platforms like Facebook, Twitter, Instagram, Snapchat, etc?    | Yes/No               |  |  |  |
| IBDA <sup>a</sup> -related questions |  |                      |  |  |  |
| E1                                   | Is taking a picture of food easier than writing down what you ate?                         | 5-point Likert scale |  |  |  |
| E2                                   | Is taking a picture of food easier than recording an audio describing what you ate?        | 5-point Likert scale |  |  |  |
| E3                                   | Is it intrusive to take pictures of food every time you eat?                               | 5-point Likert scale |  |  |  |
| E4                                   | Should any use of your data require an explanation in simple clear words?                  | 5-point Likert scale |  |  |  |
| E5                                   | Can your doctor share your data for research with his/her colleagues without your consent? | 5-point Likert scale |  |  |  |
| Demographic questions                |  |                      |  |  |  |
| D1                                   | What is your educational qualification?  | Level                |  |  |  |
| D2                                   | Do you follow a religious diet?  | Yes/No               |  |  |  |

<sup>a</sup>IBDA: image-based dietary assessment.

Regarding metadata collected through an IBDA system, we presented existing social network jargon that many are familiar with. For example, some users *tag* the location of a restaurant when posting a food picture. Further, we presented the hypothetical situation of a third party that gains control of a participant's diet data of 1 year. Based on that data, a few aspects of the individual might be inferred. We presented Scenario B and collected responses on what the participants thought can be inferred. The inferred information examples were hypothetical, and to our knowledge, no such work exists. It was designed to evaluate perceptions toward what is possible and how sensitive particular information is to the participant. The responses were collected on a 3-item likelihood and sensitivity *Likert* scale.

In addition to perceived threats with sharing food pictures and subsequently data set, we collected responses about which social groups an individual was voluntarily willing to share information about their diet with. The information as food images and attached metadata typically associated with an image was considered for sharing. In addition to food images, we added additional parameters such as medications and diet plans. The groups provided were Family, Friends, Doctor, Team, and Fans. Participants indicated their binary responses by checking corresponding blocks in the questionnaire. Additionally, we provided an option if they thought the information is sensitive and they do not wish to share with anyone. At the end of the questionnaire, the participants were asked a series of demographic questions, including some additional ones about their diet and allergy. The collected data from Nettskjema were downloaded and analyzed after the end of the study. The results are presented in the Results section.

#### Analysis

RenderX

We performed the Kruskal-Wallis test [38] to determine the differences among responses for E1-E5 based on participants' religious diets. Consistently, we obtained P>.05 supporting the hypotheses that the responses are uniform across the participants. The actual *P* values that we obtained were *P*=.14, *P*=.15, *P*=.56,

```
http://humanfactors.jmir.org/2020/4/e19085/
```

P=.18, and P=.78. We performed another set of Kruskal-Wallis tests to determine whether the responses among European and non-European participants had statistically significant differences. For E1-E5, the observed P values, that is, P=.20, P=.14, P=.28, P=.92, and P=.50 indicated that they were not different. Therefore, we proceeded with reporting ordinal variables in our results by using compound bar charts.

Additionally, we measured consensus among the reported ordinal values by using the Tastle and Wierman's consensus measure [39]. We report the consensus values for E1-E5 as Cns(E1)=0.59, Cns(E2)=0.64, Cns(E3)=0.56, Cns(E4)=0.61, and Cns(E5)=0.52. We observed a moderate amount of agreement in the reported data. For the concerns and the likelihood of them being inferred from one's dietary data, we performed the Pearson correlation analysis. The perceived likelihood did not indicate a strong correlation with the concern. The maximum correlation coefficient we observed was between the likelihood and severity of concern for "financial status" with r=0.41 (P<.001). However, it was still a low correlation. We report our observed ordinal values later in the Results section. We did not use a prediction model in our analysis as we did not find any predictor variables to be significantly related to the outcome in our analysis (P>.05).

#### Recruitment

Chung at al [22] showed that people are more likely to take food pictures when traveling. We leveraged this insight in combination with Goodman's *snowball sampling* methodology [40] to recruit a varied cohort to our study. Initially, we recruited 5 individuals in different regions of the world who were traveling and hence more likely to have reflected on the use of food pictures in our scenarios. We briefed our initial cohort about the goal of our study and provided them with information about the collected data. The initial cohort was then instructed to further recruit other individuals they met throughout their travels who had personal characteristics matching the selection criteria, in accordance with Goodman's methodology. We provided direct support for the participant that had questions

about the study. In total, 105 participants responded to our questionnaire.

Of the 105 participants, 99 (94.2%) indicated having an account on a social network. Of the 99 participants, 95 (96%) reported seeing food pictures at least once over social networks. Table 2 summarizes the participants' distribution across various factors such as age, gender, and region. Approximately 47.6% (50/105) of our participants identified themselves as male, 51.4% (54/105) as female, and 0.9% (1/105) as nonbinary. These participants were from 35 different countries around the world, spread over 6 continents. Of the 105 participants, 99 (94.3%) were younger than 35 years, which corresponds to the age up to which peak performances can be maintained by athletes [41]. We also collected information about the participants' education levels. Approximately 90.5% (95/105) of the participants indicated that they attained an education more than high school; 44 participants indicated having higher than a bachelor's degree (postgraduate degree, n=35; doctorate degree, n=9). Only 16.2% (17/105) of the participants were following a strict religious diet. More than one-third of the participants (39/105, 37.1%) indicated having food allergies.

Table 2. Demographic information of the participants (N=105).

| Demographic information | Total population, N=105, n (%) | Males, n=50, n (%) | Females, n=54, n (%) | Nonbinary, n=1, n (%) |  |
|-------------------------|--------------------------------|--------------------|----------------------|-----------------------|--|
| Region                  | ·                              |                    | -                    |                       |  |
| Africa                  | 4 (3.8)                        | 0 (0)              | 4 (100)              | 0 (0)                 |  |
| Asia                    | 28 (26.7)                      | 19 (68)            | 9 (32)               | 0 (0)                 |  |
| Australia               | 10 (9.5)                       | 5 (50)             | 5 (50)               | 0 (0)                 |  |
| Europe                  | 54 (51.4)                      | 22 (41)            | 31 (57)              | 1 (2)                 |  |
| North America           | 6 (5.7)                        | 4 (67)             | 2 (33)               | 0 (0)                 |  |
| South America           | 3 (2.9)                        | 0 (0)              | 3 (100)              | 0 (0)                 |  |
| Age group (years)       |                                |                    |                      |                       |  |
| 18-25                   | 36 (34.3)                      | 8 (22)             | 28 (78)              | 0 (0)                 |  |
| 25-35                   | 63 (60.0)                      | 37 (59)            | 25 (40)              | 1 (2)                 |  |
| 35-45                   | 4 (3.8)                        | 4 (100)            | 0 (0)                | 0 (0)                 |  |
| 45-55                   | 1 (0.9)                        | 0 (0)              | 1 (100)              | 0 (0)                 |  |
| 55-65                   | 1 (0.9)                        | 1 (100)            | 0 (0)                | 0 (0)                 |  |
| Religious diet          |                                |                    |                      |                       |  |
| No                      | 88 (83.8)                      | 40 (46)            | 47 (53)              | 1 (1)                 |  |
| Yes                     | 17 (16.2)                      | 10 (59)            | 7 (41)               | 0 (0)                 |  |
| Allergies               |                                |                    |                      |                       |  |
| No                      | 65 (61.9)                      | 38 (59)            | 27 (42)              | 0 (0)                 |  |
| Yes                     | 39 (37.1)                      | 11 (29)            | 27 (69)              | 1 (3)                 |  |
| Not indicated           | 1 (0.9)                        | 1 (100)            | 0 (0)                | 0 (0)                 |  |

# Results

#### Overview

RenderX

In this section, we discuss the perception of privacy and related attitudes based on the findings in our study. We divided our results into expectations and concerns. The expectations cover general perception toward privacy, IBDA methods, and data use. The concerns cover information that can be inferred from their dietary data. Additionally, we discuss the concerns toward exposing such information to a third party from their mobile Food Records (mFRs). Finally, we present our findings regarding the sharing of collected dietary information with different social groups.

## Expectations

We present the general expectations that participants have indicated toward IBDA methods. We also explored their attitudes toward data collection and use.

## Effort

Approximately 80.9% (85/105) of the participants agreed that capturing diet records by using a phone camera is easier than writing down their dietary intake (see E1, Figure 1). Individuals could also record audios describing their diets for accurately recording their diets. More than four-fifths of the participants (86/105, 81.9%) preferred capturing photos over recording their diets by voice (see E2, Figure 1). Only some participants were undecided about preferring image capture over writing down or recording audio (7/105, 6.7% and 8/105, 7.6%; respectively). About half of the participants (52/105, 49.6%) had previously posted food images on social networks. We considered these

participants as experienced because they are familiar with the required training for an IBDA-based study. Even those participants who lacked experience (53/105, 50.4%) in posting food images showed similar attitudes toward ease of recording their diet intakes by using photography. For a successful IBDA-based study, continuous recording of participants' diets

is required. Compared to the irregular posting of images on social networks, continuous recording requires extra effort from a participant. When we asked the intrusiveness of this requirement (see E3, Figure 1), about two-thirds of the participants (69/105, 65.7%) indicated that it would be intrusive.

Figure 1. Collected responses toward E1-E5 (see Table 1). The values in the graph indicate the percentage of responses.



#### Data Use

Individuals tend to have very little or no information about how their data are being used. Many individuals feel that they have no control [42]. In 2018, general data protection regulation (GDPR) [43,44] granted additional rights to individuals about their data in Europe. In our study, nearly half of our participants (54/105, 51.4%) were based in Europe (see Table 2). We recorded our participants' expectations and attitudes toward data use, some of which are enforced by GDPR. Overall, three-fourths of the participants (78/105, 74.2%) wanted to know about any use of their data (see E4, Figure 1). They preferred it to be explained in simpler terms over the complex "terms of use." While Europeans have a legal right to demand such explanations, we excluded data from them to see what the participants from outside Europe prefer. Even among non-Europeans (38/51, 75%), we observed a similar interest in the participants in knowing what their data are being used for.

#### Concerns

RenderX

Trust is important for participation in epidemiological research [21]. The early stages of newly developed methods rely heavily on voluntary participation from willing individuals. Building and maintaining trust in research is critical, especially while handling personal information. Data leaks can expose information about individuals that can be sold to third parties with potentially malicious intent. In our study, we presented a

http://humanfactors.jmir.org/2020/4/e19085/

scenario in which mFR data of an athlete were leaked to a third party. We are not aware of works inferring information about an individual from their mFRs. We investigated the participants' attitudes toward issues that might arise after their information is leaked. We discuss attitudes toward what is perceived to be exploitable and how much concerned the participants are. We collected the responses on a 3-point Likert scale.

# Allergies

With regard to food allergies, only a fraction of the participants (24/105, 22.8%) were very concerned about it being exposed to a third party (Figure 2). About one-third of the participants (30/105, 28.6%) were somewhat concerned. If we only consider only the participants with allergies (39/105, 37.1%), they were found to be relatively less worried. Only 13% (5/39) of the participants with allergies were particularly concerned, while 28% (11/39) indicated somewhat concerned. The majority (23/39, 59%) of the participants with allergies were not concerned about a third party learning about their allergies.

About the possibility of deriving allergies from their mFRs, more than two-thirds (75/105, 71.4%) thought that allergies can be inferred. A little more than a quarter (30/105, 28.6%) thought that it was not very likely to be derived from the mFR of an individual. This trend was very similar among participants with allergies. Approximately 30% (12/39) of the participants with allergies thought that it was not likely that it can be inferred from their mFRs.

Figure 2. A. Perceived inference from image-based dietary assessment data set; B. Concerns toward a third party learning such attributes. The values in the graph indicate the percentage of responses.



# Religion

XSL•F() RenderX

Religious belief is often considered sensitive information in regional laws [43]. It can also affect the dietary choices of an individual. However, depending on the social and cultural aspects, individuals may share this information openly. In our study, about two-thirds (67/105, 63.8%) of the participants indicated that it is likely that their religion can be inferred from

their mFRs. As stated earlier in Table 2, 16% (17/105) of our participants followed a strict religious diet. They showed similar traits. Three-fourths (12/17, 76%) of the participants following religious diets indicated that it is likely that their religion can be derived from their mFRs.

With regard to a third party learning about their religion, a little more than half (56/105, 53.3%) of the participants were not concerned. Only 17.1% (18/105) were *very concerned* about a

third party learning about their religion. Participants following religious diets (n=17) were slightly more concerned about a third party learning about their religion. Approximately 65% (11/17) of the participants following religious diets indicated that they were concerned about a third party learning about their religion from the mFRs.

# Identity

As personalized dietary interventions are more effective [45], IBDA systems may require personally identifiable information. Exposing the identity of an individual is one of the prominent privacy concerns in the modern era [46,47]. With regard to inferring identity from mFRs, participants of both genders showed similar attitudes. The majority (64/105, 60.9%) thought that it was not likely to infer identity from their mFRs (Figure 2). Approximately 70% (38/54) of the females and 50% (25/50) of the males responded that way. While many found it unlikely to infer one's identity from mFRs, exposed identity was still a concern to many of the participants. In terms of gender, male participants showed a slightly higher (41/50, 82% vs 43/54, 80%) concern toward their identity being exposed to a third party when compared to females.

## **Information Sharing**

Information collected by an IBDA system about an athlete's diet provides insights into dietary habits and can guide toward

a proper diet. There might be additional metadata collection through an IBDA system. The information collected in the form of food images along with metadata can be mined for other purposes as well. For example, the time of dietary intake can be useful for maximizing performance on the field or predicting burnout. Similar to trends on social networks [22], an athlete might be interested in sharing this information with different social circles. We collected responses about sharing this information with different social groups as an athlete. The results showed that participants favored sharing information mostly with their doctors and family (see Figure 3).

About three-fourths (79/105, 75.2%) of the participants showed willingness to share food pictures continuously with family. In comparison to the social group *family*, the participants were more willing to share food images with their *doctors* (94/105, 89.5%). Only a quarter (26/105, 24.8%) of the participants showed willingness to share food pictures with their sports *team* while nearly half (47/105, 44.8%) showed willingness to share food pictures with their sports *team* while nearly half (47/105, 44.8%) showed willingness to share food pictures with *friends*. In terms of the metadata associated with dietary data, such as the time of the meal, the willingness to share further drops. Only 68.6% (72/105) of the participants agreed to share the *time of the meal* with their families in comparison to 82.9% (87/105) sharing the time of the meal with their doctors. Time of food intake is in fact an important consideration for elite athletes and coaches with respect to restitution and training planning.

Figure 3. Radar plot showing willingness to share information with different social groups.



RenderX

Individuals are more cautious about sharing their location. Only 70.5% (74/105) expressed willingness to share the location associated with diet records with *family* and also with *doctor*. Unsurprisingly, participants were not very keen on sharing location with their fans (10/105, 9.5%). About 16.2% (17/105) of the participants did not want to share their location of places they eat with anyone.

Figure 4 shows the correlation values between different metadata sharing behaviors within a group. We ignored the responses

from participants who indicated their unwillingness to share with anyone. All of the remaining participants agreed on sharing the food image, diet plan, and time of eating food with a *doctor*. There was a weak correlation between willingness to share location along with the food image with a *doctor*. Within the *family* group, there was a strong correlation in sharing diet plan and time. In terms of sharing metadata with one's team, we observed a strong correlation between diet plan and food image. The willingness to share time and location with one's team was also strongly correlated.

**Figure 4.** Pearson correlation values for sharing meta-information such as food image, diet plan, location and time between different social groups A. Family; B. Doctor; C. Friends; D. Team. The values indicate the correlation between the willingness, or lack thereof, to share meta-information.





# Discussion

In this paper, we present our findings on the perception of privacy for an IBDA system. Our findings provide a coarse view of privacy attitudes toward conducting dietary assessments with food images. Expanding upon prior works [7,18], these results explore contextual privacy violations for an IBDA system. Trust is crucial for voluntary participation in epidemiological studies [21]. When designing such systems, following Privacy by Design guidelines is beneficial for addressing privacy concerns early on and building trust. We found that participants indicated a strong dislike toward data use without consent. Explaining data collection and processing with easy-to-understand terms seems to be of interest to users. Public engagement in cohort studies is crucial for their success. We conjecture that incorporating these parameters can deliver a pleasant experience and increase users' trust in the system. Personalized dietary interventions are more effective than universal recommendations [45]. However, these interventions require additional information about the user, some of which may be considered sensitive from the legal or an individual's point of view. In this context, information such as food allergies can be acquired by third parties to improve recommendations. The additional information about a user can lead to identity leaks. Similar to previous works [46,47], identity is still a top concern for individuals interacting with web-based systems. It might be useful to prevent that in design by separating the authentication and data storage for mFRs. Accordingly, careful considerations should be taken to share inferred details about an individual, preferably only to individuals that the user has consented to share with. In this study, the participants were not worried about whether a third party could learn about their religion from their mFRs. However, regional privacy laws may restrict sharing such information [43]. These findings are not

RenderX

intended to replace legal requirements while building systems. They complement them to build a trustworthy system.

Our initial assumption from the study by Chung et al [22] was that individuals are willing to share food pictures over social networks. However, we experienced the opposite for long-running continuous studies. In our study, individuals preferred to share dietary data with groups who have a clear and stated use for it, such as their dietician. Another trusted group for information sharing is family. Individuals show different behaviors for sharing metadata. Thus, metadata require different sharing policies than images. For an IBDA system, this means that metadata need to be scrubbed from a food image and stored separately.

Sharing an athlete's data with fans can be an interesting opportunity to engage with followers, such as for crowd support using HeartLink [48]. However, our study shows that individuals might refrain from sharing their data. Participants show greater willingness to share information with groups they trust or when they know what it will be used for. Motivating users to share self-reported data can be challenging. Functionality, ease of use, and privacy are considered crucial for any self-reporting health app [49]. Even chronically ill patients are willing to share data if they receive personalized feedback [49]. Such attitudes can lead to exploring use-based privacy policies for their data [50].

Elite sport clubs, particularly in our elite soccer domain, have nutritional experts hired as part of their management and support team. Such experts are involved in providing detailed dietary plans for their athletes, and they know in detail about most of the common meals provided on-premise for the athletes. For instance, our main elite soccer clubs involved in our cooperation, as a rule, have breakfast and lunch together in their training facilities. Involving such experts using our proposed scheme means that they receive the needed data from their athletes when outside the training facilities to complete the picture. We investigated the privacy perceptions and concerns for conducting long-running studies using IBDA methods. For epidemiological studies, it is important for users to continuously record diets without any biases. In this study, individuals preferred recording diets using a digital camera over other methods. However, taking a picture of every meal is still perceived as intrusive for some users. For long-running studies, prediction models can be employed to reduce the labor of taking pictures [18]. In summary, this study provides initial insights into the privacy requirements for an IBDA system. Thus, our work provides the basis for discussion in the research community for building and deploying IBDA systems for population-wide studies.

Our study has the following limitation. A questionnaire-based study fails to identify the causation of behavior. For privacy reasons, we did not collect the contact information from the participants. Hence, any further study with the same set of participants is not possible.

In conclusion, we conducted a questionnaire-based study to understand the privacy perceptions and concerns for building IBDA systems. The privacy concerns can be addressed during the design phase to mitigate risks and strengthen participants' and stakeholders' trust in a system. We find a growing interest to know what the collected data are being used for. While IBDA methods are preferred for ease of use, continuous assessment is still seen as intrusive. GDPR compliance is an attractive feature for individuals worldwide. While uncertain about the inferences from mFRs, identity remains a top concern with regard to privacy for individuals. Knowing what the data is being used for, increases the chances of it being shared. Individuals are concerned about metadata sharing with third parties. We recently started a large interdisciplinary study involving computer scientists, sports scientists, psychologists, mathematicians, and medical experts (epidemiologists, nutritional scientists, physicians). Our select cohort includes over 400 female elite soccer athletes, from Norway to Portugal, and we intend to conduct our next study in this cohort.

#### Acknowledgments

This work is funded in part by Research Council of Norway project number 263248. The authors would like to thank Terje Holmlund for feedback during the design. The publication charges for this article have been funded by a grant from the publication fund of UiT - The Arctic University of Norway.

#### **Conflicts of Interest**

None declared.

#### **Multimedia Appendix 1**

The complete questionnaire from nettskjema. [PDF File (Adobe PDF File), 1706 KB-Multimedia Appendix 1]

#### Multimedia Appendix 2

The collected data from the study along with column descriptions for reproducing results. The readme and column descriptions are available in markdown format and PDF. The data file is available in TSV (tab-separated values) format. [ZIP File (Zip Archive), 72 KB-Multimedia Appendix 2]

#### References

```
XSL•FO
RenderX
```

- Vogt M, Puntschart A, Howald H, Mueller B, Mannhart C, Gfeller-tuescher L, et al. Effects of Dietary Fat on Muscle Substrates, Metabolism, and Performance in Athletes. Medicine & Science in Sports & Exercise 2003;35(6):952-960. [doi: 10.1249/01.mss.0000069336.30649.bd]
- Afshin A, Sur P, Fay K, Cornaby L, Ferrara G, Salama J, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 2019 May;393(10184):1958-1972. [doi: 10.1016/S0140-6736(19)30041-8]
- Holmlund TB, Foltz PW, Cohen AS, Johansen HD, Sigurdsen R, Fugelli P, et al. Moving psychological assessment out of the controlled laboratory setting: Practical challenges. Psychol Assess 2019 Mar;31(3):292-303. [doi: 10.1037/pas0000647] [Medline: 30802115]
- 4. Boushey CJ, Spoden M, Zhu FM, Delp EJ, Kerr DA. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. Proc Nutr Soc 2016 Dec 12;76(3):283-294. [doi: 10.1017/s0029665116002913]
- 5. Brenna L, Johansen HD, Johansen D. A Survey of Automatic Methods for Nutritional Assessment.: arXiv:1907.07245; 2019. URL: <u>https://arxiv.org/abs/1907.07245</u> [accessed 2020-01-01]
- 6. Howes E, Boushey C, Kerr D, Tomayko E, Cluskey M. Image-Based Dietary Assessment Ability of Dietetics Students and Interns. Nutrients 2017 Feb 07;9(2) [FREE Full text] [doi: 10.3390/nu9020114] [Medline: 28178196]
- O'Loughlin G, Cullen SJ, McGoldrick A, O'Connor S, Blain R, O'Malley S, et al. Using a wearable camera to increase the accuracy of dietary analysis. Am J Prev Med 2013 Mar;44(3):297-301. [doi: <u>10.1016/j.amepre.2012.11.007</u>] [Medline: <u>23415128</u>]
- 8. Thompson F, Subar AF. Dietary Assessment Methodology. In: Nutrition in the Prevention and Treatment of Disease. San Diego CA USA: Academic Press; 2017:5-48.
- Myers A, Johnston N, Rathod V, Korattikara A, Gorban AN, Silberman N, et al. Im2Calories: Towards an Automated Mobile Vision Food Diary. 2015 Dec Presented at: ICCV '15: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV); Dec 7-13, 2015; Santiago Chile p. 1233-1241. [doi: 10.1109/iccv.2015.146]
- Salvador A, Drozdzal M, Giro-i-Nieto X, Romero A. Inverse Cooking: Recipe Generation From Food Images. 2019 Presented at: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 15-20 June 2019; Long Beach CA USA p. 10453-10462. [doi: 10.1109/CVPR.2019.01070]
- 11. Boushey CJ, Kerr DA, Wright J, Lutes KD, Ebert DS, Delp EJ. Use of technology in children's dietary assessment. Eur J Clin Nutr 2009 Feb;63 Suppl 1:S50-S57 [FREE Full text] [doi: 10.1038/ejcn.2008.65] [Medline: 19190645]
- 12. de Vries JHM, de Groot LCPGM, van Staveren WA. Dietary assessment in elderly people: experiences gained from studies in the Netherlands. Eur J Clin Nutr 2009 Feb;63 Suppl 1:S69-S74. [doi: <u>10.1038/ejcn.2008.68</u>] [Medline: <u>19190649</u>]
- Lazarte CE, Encinas ME, Alegre C, Granfeldt Y. Validation of digital photographs, as a tool in 24-h recall, for the improvement of dietary assessment among rural populations in developing countries. Nutr J 2012 Aug 29;11:61 [FREE Full text] [doi: 10.1186/1475-2891-11-61] [Medline: 22931128]
- Beltran A, Dadabhoy H, Ryan C, Dholakia R, Jia W, Baranowski J, et al. Dietary Assessment with a Wearable Camera among Children: Feasibility and Intercoder Reliability. J Acad Nutr Diet 2018 Nov;118(11):2144-2153. [doi: <u>10.1016/j.jand.2018.05.013</u>] [Medline: <u>30115556</u>]
- 15. Ashman A, Collins C, Brown L, Rae K, Rollo M. Validation of a Smartphone Image-Based Dietary Assessment Method for Pregnant Women. Nutrients 2017 Jan 18;9(1) [FREE Full text] [doi: 10.3390/nu9010073] [Medline: 28106758]
- 16. Ambrosini GL, Hurworth M, Giglia R, Trapp G, Strauss P. Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutr J 2018 Jan 09;17(1):5 [FREE Full text] [doi: 10.1186/s12937-018-0315-4] [Medline: 29316930]
- 17. Capling L, Beck K, Gifford J, Slater G, Flood V, O'Connor H. Validity of Dietary Assessment in Athletes: A Systematic Review. Nutrients 2017 Dec 02;9(12) [FREE Full text] [doi: 10.3390/nu9121313] [Medline: 29207495]
- Wang Y, He Y, Boushey CJ, Zhu F, Delp EJ. Context Based Image Analysis With Application in Dietary Assessment and Evaluation. Multimed Tools Appl 2018 Aug;77(15):19769-19794 [FREE Full text] [doi: 10.1007/s11042-017-5346-x] [Medline: 30202237]
- Fiesler C, Hallinan B. "We Are the Product": Public Reactions to Online Data Sharing and Privacy Controversies in the Media. 2018 Apr Presented at: CHI '18: CHI Conference on Human Factors in Computing Systems; 2018; Montreal QC Canada p. 1-13. [doi: 10.1145/3173574.3173627]
- Langheinrich M. Privacy by Design Principles of Privacy-Aware Ubiquitous Systems. 2001 Oct 16 Presented at: Ubicomp 2001: Ubiquitous Computing; September 30 October 2, 2001; Atlanta Georgia USA p. 273-291. [doi: 10.1007/3-540-45427-6\_23]
- Slegers C, Zion D, Glass D, Kelsall H, Fritschi L, Brown N, et al. Why do people participate in epidemiological research? J Bioeth Inq 2015 Jun;12(2):227-237. [doi: 10.1007/s11673-015-9611-2] [Medline: 25672617]
- 22. Chung C, Agapie E, Schroeder J, Mishra S, Fogarty J, Munson SA. When Personal Tracking Becomes Social: Examining the Use of Instagram for Healthy Eating. In: Proc SIGCHI Conf Hum Factor Comput Syst. 2017 May 02 Presented at: CHI Conference on Human Factors in Computing Systems; 2017; Denver CO USA p. 1674-1687 URL: <u>http://europepmc.org/abstract/MED/28516174</u> [doi: 10.1145/3025453.3025747]

RenderX

- 23. Pettersen SA, Johansen HD, Baptista IAM, Halvorsen P, Johansen D. Quantified Soccer Using Positional Data: A Case Study. Front Physiol 2018;9:866 [FREE Full text] [doi: 10.3389/fphys.2018.00866] [Medline: 30034347]
- Thambawita V, Hicks S, Borgli H, Stensland H, Jha D, Svensen MK, et al. PMData: a sports logging dataset. 2020 May Presented at: MMSys '20: 11th ACM Multimedia Systems Conference; June 2020; Istanbul Turkey p. 231-236. [doi: 10.1145/3339825.3394926]
- 25. Schroeder R. Big Data and the brave new world of social media research. Big Data & Society 2014 Dec 17;1(2):205395171456319. [doi: 10.1177/2053951714563194]
- 26. Graeff T, Harmon S. Collecting and using personal data: consumers' awareness and concerns. Journal of Consumer Marketing 2002;19(4):302-318. [doi: 10.1108/07363760210433627]
- 27. Naska A, Lagiou A, Lagiou P. Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Res 2017;6:926 [FREE Full text] [doi: 10.12688/f1000research.10703.1] [Medline: 28690835]
- 28. Melumad S, Meyer R. Full Disclosure: How Smartphones Enhance Consumer Self-Disclosure. Journal of Marketing 2020 Mar 17;84(3):28-45. [doi: 10.1177/0022242920912732]
- 29. Christin D, Reinhardt A, Kanhere SS, Hollick M. A survey on privacy in mobile participatory sensing applications. Journal of Systems and Software 2011 Nov;84(11):1928-1946. [doi: 10.1016/j.jss.2011.06.073]
- 30. Avancha S, Baxi A, Kotz D. Privacy in mobile technology for personal healthcare. ACM Comput Surv 2012 Nov;45(1):1-54. [doi: 10.1145/2379776.2379779]
- Zerr S, Sierdorfer S, Hare JS, Demidova E. Privacy-aware image classification and search. 2012 Aug Presented at: SIGIR '12: The 35th International ACM SIGIR conference on research and development in Information Retrieval; 2012; Portland OR USA p. 35-44. [doi: 10.1145/2348283.2348292]
- Spyromitros-Xioufis E, Papadopoulos S, Popescu A, Kompatsiaris I. Personalized Privacy-aware Image Classification. 2016 Jun Presented at: ICMR'16: International Conference on Multimedia Retrieval; 2016; New York p. 71-78. [doi: 10.1145/2911996.2912018]
- Squicciarini A, Caragea C, Balakavi R. Toward Automated Online Photo Privacy. ACM Trans Web 2017 Apr 10;11(1):1-29. [doi: <u>10.1145/2983644</u>]
- 34. Thomaz E, Parnami A, Bidwell J, Essa I, Abows GD. Technological approaches for addressing privacy concerns when recognizing eating behaviors with wearable cameras. 2013 Presented at: UbiComp '13: The 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2013; Zurich Switzerland p. 739-748. [doi: 10.1145/2493432.2493509]
- 35. Greiner S. Privacy Protection in an Electronic Chronicle System. 2008 Presented at: 34th Annual Northeast Bioengineering Conference; 4-6 April 2008; Providence, RI URL: <u>http://isl.anthropomatik.kit.edu/downloads/850\_NBC08\_Simon.pdf</u>
- 36. Nettskjema. URL: https://nettskjema.no [accessed 2019-01-01]
- 37. Glasow P. Fundamentals of survey research methodology. Bedford, MA, USA: The MITRE Corporation; 2005. URL: https://www.mitre.org/sites/default/files/pdf/05\_0638.pdf [accessed 2020-01-01]
- 38. Kruskal WH, Wallis WA. Use of Ranks in One-Criterion Variance Analysis. Journal of the American Statistical Association 1952 Dec;47(260):583-621. [doi: 10.1080/01621459.1952.10483441]
- 39. Tastle W, Wierman M. Consensus and dissention: A measure of ordinal dispersion. International Journal of Approximate Reasoning 2007 Aug;45(3):531-545. [doi: 10.1016/j.ijar.2006.06.024]
- 40. Goodman LA. Snowball Sampling. Ann Math Statist 1961 Mar;32(1):148-170. [doi: 10.1214/aoms/1177705148]
- 41. Allen SV, Hopkins WG. Age of Peak Competitive Performance of Elite Athletes: A Systematic Review. Sports Med 2015 Oct;45(10):1431-1441. [doi: 10.1007/s40279-015-0354-3] [Medline: 26088954]
- 42. Shklovski I, Mainwaring SD, Skúladóttir HH, Borgthorsson H. Leakiness and creepiness in app space: perceptions of privacy and mobile app use. 2014 Presented at: CHI '14: CHI Conference on Human Factors in Computing Systems; April 2014; Toronto Ontario Canada p. 2347-2356. [doi: 10.1145/2556288.2557421]
- 43. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). 2016. URL: <u>https://eur-lex.europa.eu/legal-content/EN/ALL/</u>?uri=celex%3A32016R0679 [accessed 2019-01-01]
- 44. Kuner C. The European Commission's proposed data protection regulation: A copernican revolution in European data protection law. Bloomberg BNA Privacy and Security Law Report. 2012. URL: <u>https://ssrn.com/abstract=2162781</u> [accessed 2020-01-01]
- 45. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized Nutrition by Prediction of Glycemic Responses. Cell 2015 Nov 19;163(5):1079-1094 [FREE Full text] [doi: 10.1016/j.cell.2015.11.001] [Medline: 26590418]
- 46. Wang Y, Norice G, Cranor LF. Who Is Concerned about What? A Study of American, Chinese and Indian Users' Privacy Concerns on Social Network Sites. 2011 Presented at: Trust and Trustworthy Computing; 2011; Pittsburgh PA USA p. 146-153. [doi: 10.1007/978-3-642-21599-5\_11]
- 47. Krasnova H, Günther O, Spiekermann S, Koroleva K. Privacy concerns and identity in online social networks. IDIS 2009 Oct 1;2(1):39-63. [doi: 10.1007/s12394-009-0019-1]

```
XSL•FO
RenderX
```

- 48. Curmi F, Ferrario MA, Whittle J. BioShare: a research tool for analyzing social networks effects when sharing biometric data. 2014 Presented at: DIS '14: Designing Interactive Systems Conference; June 2014; Vancouver BC Canada p. 101-104 URL: <u>https://dl.acm.org/doi/10.1145/2598784.2602793</u> [doi: 10.1145/2598784.2602793]
- Woldaregay AZ, Henriksen A, Issom DZ, Pfuhl G, Sato K, Richard A, et al. User Expectations and Willingness to Share Self-Collected Health Data. Stud Health Technol Inform 2020 Jun 16;270:894-898. [doi: <u>10.3233/SHTI200290</u>] [Medline: <u>32570511</u>]
- 50. Cate FH. Principles of Internet Privacy. Articles by Maurer Faculty. 2002. URL: <u>https://www.repository.law.indiana.edu/cgi/viewcontent.cgi?article=1245&context=facpub</u> [accessed 2020-01-01]

# Abbreviations

**GDPR:** general data protection regulation **IBDA:** image-based dietary assessment **mFR:** mobile food record **mHealth:** mobile health

Edited by B Price; submitted 03.04.20; peer-reviewed by M Abdelhamid, F Prasser, T Yasaka; comments to author 12.05.20; revised version received 06.07.20; accepted 03.09.20; published 15.10.20 <u>Please cite as:</u> Sharma A, Czerwinska KP, Brenna L, Johansen D, Johansen HD Privacy Perceptions and Concerns in Image-Based Dietary Assessment Systems: Questionnaire-Based Study JMIR Hum Factors 2020;7(4):e19085 URL: http://humanfactors.jmir.org/2020/4/e19085/ doi: 10.2196/19085 PMID:

©Aakash Sharma, Katja P Czerwinska, Lars Brenna, Dag Johansen, Håvard D Johansen. Originally published in JMIR Human Factors (http://humanfactors.jmir.org), 15.10.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Human Factors, is properly cited. The complete bibliographic information, a link to the original publication on http://humanfactors.jmir.org, as well as this copyright and license information must be included.

