Human-centric Computing and Information Sciences

August 2024 | Volume 14

KCIA
Korea Computer Industry Association

www.hcisjournai.com

HCIS Human-centric Computing and Information Sciences

RESEARCH (Research Manuscript) Open Access

Human-centric Computing and Information Sciences (2024) 14:46
DOI: https://doi.org/10.22967/HCIS.2024.14.046
Received: July 28, 2022; Accepted: September 14, 2022; Published: August 15, 2024

Toward the Prevention of Privacy Threats: How Can We Persuade Our Social Network Platform Users?

Ramon Ruiz-Dolz*, Jose Alemany, Stella Heras, and Ana García-Fornes

Abstract

Complex decision-making problems, such as the privacy policy selection, when sharing content in online social network (OSN) platforms can significantly benefit from artificial intelligence systems. With the use of computational argumentation, it is possible to persuade human users to modify their initial decisions to avoid potential privacy threats and violations. In this paper, we present a study performed with the participation of 186 teenage users aimed at analyzing their behaviors when we try to persuade them to modify the post/publication of sensitive content on OSN platforms with different arguments. The results of the study revealed that the personality traits and the social interaction data (e.g., number of comment posts, friends, and likes) of our participants were significantly correlated with the persuasive power of the arguments. Therefore, these sets of features can be used to model OSN users and estimate the persuasive power of different arguments when used in human-computer interactions. The findings presented in this paper are helpful for personalizing decision support systems aimed at educating and preventing privacy violations on OSN platforms using arguments.

Keywords

Persuasion, Argumentation, Privacy, Human-Computer Interaction, Social Network Platforms

1. Introduction

Deciding which privacy policy is the best when making a publication in an online social network (OSN) platform is not an easy task for human users because it requires taking multiple factors into account (i.e., the potential receivers, information to be shared, users' preferences, etc.). In many situations, the information regarding those factors can be incomplete or unknown, such as the reachability of the post/publication or other users' preferences. Another relevant feature that characterizes online communication is that, once the content is posted/ published online, it can be downloaded and stored by anyone with access to it. Therefore, it is important to make sure that the content posted/published does not cause any future privacy issues. Additionally, if more than one user appears in the post/publication, it is even easier to violate any privacy preference of the rest of the users involved, which leads to privacy conflicts between users. The multi-party privacy conflicts [1] are a common type of privacy threat happening on an OSN. This problem combined with the great increase of users on an OSN, who are

***Corresponding Author:** Ramon Ruiz-Dolz (raruidol@dsic.upv.es)
Valencian Research Institute for Artificial Intelligence, Universitat Politècnica de València, València, Spain

mostly teenagers characterized as newcomer users hampered by limited abilities for self-regulation and complex decision-making [2], has piqued the interest of privacy management assistance research.

A natural way to approach the existing privacy management problem on an OSN is through the use of computational argumentation [3]. Computational argumentation research investigates how the human argumentative reasoning process can be approached from a computational viewpoint. Using computational argumentation techniques, it is possible to establish an argument-based human-computer interaction. This approach can be seen as a direct improvement of recommendation technologies [4] since added to the recommendation, a justification (i.e., an argument) is also provided to the user. An effective way to avoid and reduce the number of potential privacy threats (i.e., disclosure of sensitive information) is to persuade the author to adapt the initial privacy configuration since it may be harmful to him/her or any of the other users involved. The best way to persuade the author is by making him/her understand the reasons why the privacy threat is happening with the use of arguments. Using different messages and warnings make it viable to persuade OSN users to modify their initial decisions [5]. However, the perceived persuasive power of these messages may vary from one message to another [6] and even between different representations or structures of these messages [7]. In the OSN privacy domain, these persuasive messages may approach different privacy aspects. Based on the previous definition given in [8], as many as four different types of argument might be considered, depending on the source from which they can be supported, namely privacy, content, risk, and trust. Furthermore, arguments can be represented and structured according to different reasoning patterns. Argumentation schemes group the most common patterns of human reasoning [9].

In this paper, we study the persuasive power of different argumentation schemes and argument types when used to educate teenagers on privacy management in an OSN environment. In our study, we consider two different user models that help us encode human behavioral data into a computational system based on individual personality and social interaction data. Previous works such as [10] show how users' personality can be a key factor when directly interacting with them. Therefore, we have investigated any potential existing correlations between personality traits and arguments. Additionally, since obtaining those traits may not be possible in some social network platforms with behavior being usually influenced by personality [11], a study of the correlation between the most common social interaction features and the persuasive power of arguments has also been conducted. Therefore, the main contributions presented in this work are as follows:

- Quantify and measure the persuasive power of arguments used as a privacy threat prevention mechanism.
- Analyze the existing significant correlations between the persuasive power of arguments and the Big-5 personality traits.
- Analyze the existing significant correlations between the persuasive power of arguments and 13 different social interaction features.

All of these key findings are contextualized in the OSN educational setting with teenage participants. This is one of the most important target populations when working on this domain since they are very active and amenable to convince to share their personal information.

The rest of the paper is structured as follows. Section 2 reviews the most relevant work regarding privacy management and argumentation on the OSN platform. Then Section 3 introduces the background of our research and presents the design of the study carried out in this work, while Section 4 describes the observed results and analyzes their implications and interpretation. Finally, Section 5 wraps the paper up by summarizing the most important conclusions reached and the future research directions.

2. Related Work

Multiple approaches have been considered in the literature regarding the problem of finding optimal privacy policies on an OSN aimed at avoiding potential privacy violations [12]. A collaborative privacy preserving tool is proposed in [13], this system allows recommendations to be provided to user without

endangering their privacy. In [14], the authors propose an algorithm for predicting and preventing privacy violations on an OSN. This system detects a potential privacy violation and warns the users involved to prevent further damage to them. A different approach to the privacy problem was recently introduced in [15], where the authors propose an algorithm that combines users' features such as the age or gender with the trust between users to determine the risk of sharing a publication on an OSN. Another collaborative approach to provide privacy recommendations to users is proposed in [16]. The authors propose CoPE, a collaborative privacy management system where individual users can decide a specific privacy configuration for each publication. The system decides the best policy considering the most-voted configuration. Finally, some of the existing automated privacy management systems that rely on an internal negotiation process include, for example, PriArg [17] which is a multi-agent algorithm featuring an underlying negotiation protocol to compute the best privacy configuration for a specific situation. In PriArg, the negotiation is approached with argumentation. The agents represent real social network users that have an ontology with information from the network, the relationships between users, and content being posted/published. When considering all these data, each agent can generate arguments to achieve a deal trying to satisfy the user privacy preferences. Images were also brought into consideration in [18], where an autonomous agent uses the tags and image features to prevent privacy violations on an OSN. There are some common weak points in all of these privacy management systems, with all of them being focused on privacy conflicts where multiple users are involved in the same publication. But the case of a user choosing a dangerous configuration for themselves is not considered. There is also an important limitation if we seek to provide the user with an explanation as to why the configuration should be changed. None of the analyzed privacy management tools gives the user a reasoned explanation nor tries to persuade them. A recent explainable approach was proposed in [19, 20], but it was focused exclusively on collective privacy violations.

When people try to reach an agreement, explain our viewpoint, or try to convince another person, it is very common to make use of arguments. An argument is defined as a set of propositions that can support the veracity of the main statement (or the conclusion). Thus, using arguments, it is possible to provide a set of coherent reasons that support some specific idea. Therefore, the use of computational argumentation can be seen as the natural way to approach a decision-making problem in which a human user must be persuaded. In [21], it is possible to observe the relevance of analyzing the persuasive power of arguments and user preferences, when developing decision-making assistance artificial intelligence systems. Several works using argumentation on the OSN platform can be identified in the literature. As described in [22], argumentation on an OSN can be very useful, such as enhancing dialogues or helping to structure users' opinions. It is also possible to use computational argumentation techniques to model the dialogue between different users sharing their preferences on an OSN and to persuade students to use specific learning objects in an educational environment [23]. Therefore, as [24] proposes, argumentation seems the most coherent way to approach a persuasion problem framed in an educational context on an OSN. In [17], an argumentation protocol to define the best privacy policy when a multi-party privacy dispute is triggered is proposed. However, not many works in which all the topics of our research converge (i.e., privacy management, computational argumentation, and human user persuasion) have been identified. In addition to the main flaws identified before, the existing related work in argumentation on social network platforms is mainly focused on studying the multi-party privacy conflicts, as well [17, 19]. However, as described in [25], it is very common to find users regretting their own posts/publications on an OSN. Since we are focused on the educational arena, we need a system that considers not only privacy disputes between different users involved in the same posts/publication but also potential self-privacy violations. When defining an argument, several parameters should be considered (e.g., the content, reasoning pattern, language, etc.) to maximize its persuasive power. The reasoning pattern of an argument is defined by the underlying logic of its elements. Argumentation schemes were conceived as common patterns on human reasoning. In [9], as many as 60 generally accepted argumentation schemes that can be found in common dialogues have been identified. Therefore, the use of argumentation schemes is a convenient way to define the reasoning patterns of the arguments of our study.

Finally, persuasion plays a major role on the effectiveness of arguments when used in a dialogue. Since different users may perceive arguments differently, it is very important to be able to understand and adapt our arguments to individual users if we want an effective human-computer interaction. In [26], an argumentative system to make users change their behaviors in the healthy eating domain is proposed. The persuasiveness evaluation of the semi-automatic generated arguments is described in [7]. Furthermore, a study of the impact of personality, age, and gender on message type susceptibility [27] has also been conducted. When these works are considered altogether, it is possible to infer relations between elements like individual personality and effectiveness of argumentation schemes. Finally, in [28], the authors explore the persuasive principles underlying some of the most common patterns of human argumentative reasoning (i.e., argumentation schemes). However, to the best of our knowledge, no one has directly analyzed the persuasive power of arguments on teenagers, but behaviors may differ substantially between a teenager and adult on the OSN platform [29].

Therefore, with this paper, we put together the three research topics and present new results which are helpful to push forward all the identified limitations on these topics as follows: 1) with our arguments, we consider both self-disclosure and multi-party privacy conflicts; 2) we approach the privacy management assistance problem from a more explainable and educational perspective; and 3) we study teenage persuasion with arguments on an OSN, which has not been analyzed in the literature yet. Our study results provide a new perspective on human (i.e., teenager) persuasion in the privacy management area. In this way, we propose different but related user models based on two human aspects which we use to analyze the persuasive power of arguments based on personalities and social interactions. In this way, it is possible to optimize the chosen argument through the privacy management assistance system for each individual user.

3. Study Design

3.1 Background

Aimed at preventing privacy conflicts and minimizing the number of privacy violations, an argumentation framework for OSN platforms was proposed in [8]. It is defined as a tuple <A, R, P, τ_p>, where

A is a set of n arguments $[\alpha_1, \dots, \alpha_n]$,

R is the attack relation on A such as $A \times A \rightarrow R$,

P is the list of e profiles involved in a privacy dispute $[p_1, \dots, p_e]$, and

τ_p is a function $A \times P \rightarrow [0,1]$ that determines the score of an argument α for a profile p.

A complete definition of all the parameters that define the argumentation framework for OSN platforms is presented in [8]. As a solid motivation for the research conducted in this paper, it is important to mention that this framework models individual users by their personalities and their OSN usage statistics, which are the features that we analyze in this work. The personalities of individual users are represented with a 5-dimensional vector modeled with the Big-5 personality traits [30], namely openness, conscientiousness, extraversion, agreeableness, and neuroticism, which represent the five most significant aspects of human personality. The process of generating arguments is thoroughly explained in [31], and starts when a potential privacy violation is detected during posting/publishing content on the social network platform. Then, the set of relevant information is gathered and retrieved from the OSN. Once all the arguments are generated, the system determines the set of acceptable arguments based on the score function τ_p. Finally, the system translates the arguments in their computational form into human readable text with the use of templates. In the final step of human-computer interaction, the argumentation system has available the set of acceptable arguments. However, the system needs to know which argument is more effective during the interaction process. The present work attempts to shed light on the persuasive power of argument types and argumentation schemes so as to be able to define better dialogue strategies prioritizing the most persuasive arguments.

3.2 Research Questions

The previously defined theoretical framework was proposed to be integrated into PESEDIA [32], an educational social network platform. However, deciding on the dialogue strategies when interacting with human users is still a challenge. Therefore, we have carried out this study to answer the following research questions that arise when designing this interaction:

1. Which reasoning pattern (i.e., argumentation scheme) is more persuasive for teenage OSN users?
2. Which topic (i.e., argument type) is more persuasive for teenage OSN users?
3. How does the personality traits of teenage users influence the persuasive power of arguments?
4. How does the online social interaction behaviors of teenage users influence the persuasive power of arguments?

If it is possible to find any behavioral patterns regarding these questions, the arguments could be generated by the argumentation system following different strategies for individual users depending on their personality traits or social interaction behaviors. The measures, instruments, procedure, and participants of our study are described below.

3.3 Measures and Instruments

To answer the proposed research questions, we designed the following study based on three questionnaires and social network platform usage. Questionnaires were used to retrieve the personality traits of the participants (Big-5 personality test), the persuasive power of argumentation schemes (questionnaire A), and the persuasive power of types of argument (questionnaire B). Participants also used the social network platform PESEDIA [32] for one month from which we collected the online social interaction data. PESEDIA is an educational OSN platform aimed at teaching its users the basic privacy competencies on a social network platform. This social network platform provides a free environment like other OSN platforms (e.g., Facebook, Instagram, etc.). The chosen way to teach users is by gamification, with scores and a global ranking to reward the most active and participatory users. It is possible then to induce the users to do activities and participate in debates without forcing them [5] to do so. To find answers to the research questions proposed, this study was carried out in PESEDIA with teenage participants ranging in age from 12–15 years. The study lasted for one month, with the social network platform active and accessible round the clock for participants. The ethics and law committee from the Universitat Politècnica de València reviewed and approved the study performed. In particular, they reviewed that the social network platform PESEDIA abided by the GDPR Laws on users' privacy protection and management of their data.

Therefore, we measured their individual personalities and online social interactions to model the participants of our study. The Big-5 personality trait test aimed at measuring the personality traits (openness, conscientiousness, extraversion, agreeableness, and neuroticism) of children and teenagers [33] was used. Furthermore, we have also divided the participants into clusters based on their personality traits. Four major clusters have been recently identified in the literature, given as average, self-centered, reserved, and role model [34]. This clustering is proposed to group samples sharing similar social perceptions and similar expected behaviors. Our hypothesis to use these clusters in our study is that among similar characterized participants, it can be possible to observe stronger behavioral patterns, reducing the noise and leading to more solid findings. Thus, we have split our samples into four different personality-based groups to observe if those same clusters could be found in our population and if any behavioral patterns toward argument persuasion could be detected in each specific cluster. We ran the K-means algorithm until it reached convergence to generate the mentioned clusters.

In some situations, it may not be possible to retrieve users' personality traits. Therefore, in our study, we have also considered the data from their social interaction behaviors in PESEDIA. Thirteen different features representing participants' social interactions in the OSN setting have been used in our study to

model PESEDIA users as an alternative to their personalities based on the number of friends (#friends), number of status updates (#status_upd), number of likes (#likes), number of shares (#shares), number of comments (#comments), number of private posts (#ppprivate), number of public posts (#pppublic), number of posts shared with friends (#ppfriends), number of posts shared with groups of friends (#ppgroups), number of uploaded photos (#photos), number of posts deleted (#deletes), the average length of text posts (avg_textsize), and time spent on the network (time_spent). The previous work identified in the literature pointed out that these features could be closely related to user personality [11, 35]. Therefore, these features represent an alternative dimension to personality from which it is possible to model OSN users.

Finally, the persuasive power of arguments (for schemes and types) has been computed as the normalized number of times an argument beats another one. Therefore, we define the persuasive power for an argument α_i as follows:

$$s(\alpha_i) = \frac{\sum_{j \in C} b_{ij}}{|P| \cdot (|C| - 1)}, \tag{1}$$

where b_{ij} refers to the number of times the argument α_i beats another one α_j $(i, j \in C, i \neq j)$. An argument α_i beats another one α_j if it is considered more persuasive by our participants in the questionnaires. In our study, the classes C are represented as argumentation schemes and types of argument. Regarding the parameters $|P|$ and C, they represent the number of participants and number of options inside a class, and are used to compute the maximum number of times an argument class can beat each other. The result is a 0–1 normalized value. We have used questionnaires to measure the persuasive power of arguments, with different ones for schemes and types. There, participants were faced with the same situation (see Fig. 1) as follows: they were going to make a post/publication on the social network, and they were told not to do so. The way of persuading the participant not to make the post/publication is with the use of arguments, so they had to rank these arguments from the most persuasive argument (1) to the least persuasive one ($|C|$). Next, we describe how these questionnaires have been designed.

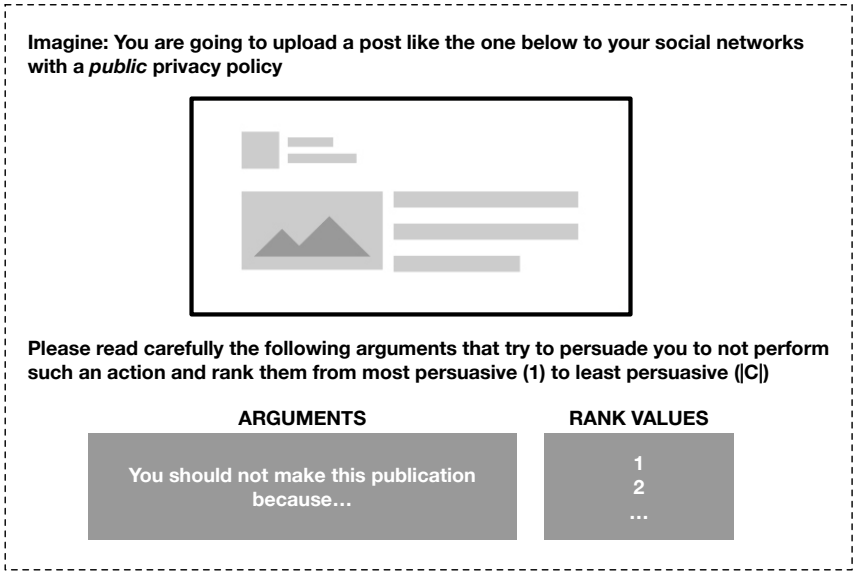

Fig. 1. Template of persuasive power questionnaires.

3.3.1 Questionnaire A (schemes)

This questionnaire has been designed to capture the persuasive power of different argumentation schemes on a user (RQ1). We decided to consider the following five schemes in our study, given as

argument from consequences (AFCQ), argument from popular practice (AFPP), argument from popular opinion (AFPO), argument from expert opinions (AFEO), and argument from witness testimony (AFWT). With these schemes, it is possible to capture users' behaviors when facing some of the most common reasoning patterns [9] used in social network privacy-related persuasive dialogues. Furthermore, we can analyze how practical reasoning and different source-based arguments are able to persuade teenage OSN users. By using these schemes, our goal was to see if teenagers were more concerned about recommendations based on the consequences of their actions, expert opinions, similar user experiences, popular behaviors, or previously affected users.

AFCQ show the participant the consequences of doing some specific action, namely sharing some content in our case. With this scheme, we can measure the importance each participant gives to the effect of their actions on the social network platform. AFPP try to persuade by evincing that there is a common popular practice among other similar people regarding some specific topic. In this case, with AFPP, we can observe the importance that participants give to an argument based on their friends' activity. Similarly, AFPO try to persuade with the use of a generally accepted opinion. Therefore, AFPO allows us to observe participants preferences toward the generally accepted opinion regarding their privacy. AFEO base their reasoning pattern on some expert opinion regarding a specific topic. These argumentation schemes make it possible to observe users' reliance in a privacy field expert. Finally, AFWT make the reasoning, considering the experience of someone with the same knowledge credentials. With this scheme, it is possible to measure the trust that our participants give to someone with a similar expertise level in privacy management.

In this first questionnaire, the arguments that represent these five argumentation schemes on the OSN platform and which participants ranked by their perceived persuasive power are the following: (You should not make this publication because...)

- Making the post/publication could have bad consequences for your privacy (AFCQ);
- Most of your friends would not publish this content (AFPP);
- Everyone knows that posting/publishing this is a mistake (AFPO);
- The monitors are experts on social network platforms, and they believe that making posts/publications of this type could be dangerous (AFEO);
- A user of the PESEDIA platform who has made similar posts/publications considers that it can be dangerous (AFWT).

3.3.2 Questionnaire B (types)

This questionnaire has been created to observe the persuasive power on our participants of the four different types of argument considered by the argumentation framework (RQ2). These types comprise privacy, trust, risk, and content arguments. Privacy arguments are generated regarding individual user privacy preferences toward the audience of one's posts/publications (i.e., private, friends, public, or a group of friends). Therefore, privacy arguments try to persuade the participants, considering their privacy preferences and configuration. Trust arguments are the ones generated taking friendships between users into account. This type of argument tries to persuade the participant, making them understand that other people may be harmed if the content gets posted/published. Risk arguments consider the post/publication reachability on the social network, computed as explained in [36]. Then, a risk argument is generated if the scope of the post/publication exceeds the user expected audience. Finally, content arguments are generated regarding the own content of the post/publication. Six different types of content (i.e., location, medical, alcohol/drugs, personal information, family/association, and offensive) [37] are considered by our argumentation system. In this case, the perceived persuasiveness may vary with the type of content included in the post/publication due to its sensitivity [38]. The arguments that participants ranked by their perceived persuasive power in this questionnaire and that represent the four argument types are as follows: (You should not make this publication because...)

- You have chosen public privacy settings. (privacy)
- Some of the people who appear might get upset. (trust)

- It could be read by strangers. (risk)
  - You are revealing your location. (Content: location)
  - You are giving out personal medical information. (Content: medical)
  - Others may think you consume alcohol/ drugs. (Content: alcohol/drugs)
  - You are revealing personal data about yourself. (Content: personal info.)
  - You are revealing a friend's personal information. (Content: fam./assoc.)
  - You might offend another user. (Content: offensive)

where items listed in black refer to privacy, trust, and risk types of argument, and items listed in white refer to the different contents (i.e., location, medical, alcohol/drugs, personal information, family/ association, and offensive) of content-type arguments. This questionnaire was filled out by participants as many times as different contents of content-type arguments there are so as to avoid any bias on users' perception of information sensitivity [38].

3.4 Procedure

The study was carried out on the PESEDIA social network platform where teenage users used it for 1 month. To prevent interference, we included a registry controller (using a secret token) to avoid undesired registrations that could affect the security of the participants and this study. The two questionnaires described above to measure participants' features were integrated in the own social network platform and they were progressively enabled in the onsite sessions. They were not required to complete them at any specific moment, but participants were motivated through gamification techniques. During the whole period of the study, the participants had full access to the PESEDIA social network platform to share their experiences and feelings. We organized three onsite sessions of 90-minute intervals in equipped labs at the university to use as control points of the study. These three onsite sessions were distributed at three points in time, specified as the first session at the beginning of the one-month period; second session in the middle; and third session at the end. The aim of these sessions was to clarify any doubts that might arise among the participants as to the functionality and features of the social network platform. Each session started with a brief explanation of the potential activities that they could do concerning the testing and understanding functionalities of the social network platform, and then participants had time to interact using the social network. In the first session, we introduced it to the participants, and they signed up on the social network platform. Then, they had to complete basic activities that focused on customizing their user profiles, setting up their general setting options and building their friendship relations. Before finishing the first session, the personality test was made available for the participants to complete it. In the second session, we requested participants to complete the questionnaires about persuasive power (questionnaires A and B). In questionnaire A, participants ranked the five argumentation schemes in a decreasing persuasive sequence. In questionnaire B, participants faced six different instances of the questionnaire considering one specific content category at a time. They ranked the four argument types in a decreasing persuasive sequence in each instance of the questionnaire. Arguments were displayed in a different order in each round to avoid the sequence effects. Finally, in the third (last) session, we presented the participants with a summary regarding their behaviors and answers to the questionnaires to conclude the study.

3.5 Participants

Out of a total of 218 teenagers participating in the study, 215 participants completed the personality test and 212 completed both questionnaires A and B. We excluded the participants who did not complete all the control sessions and the proposed questionnaires (29 participants), along with the participants who decided not to participate (3 participants did not log into PESEDIA). Finally, 186 participants completed the study (103 males and 83 females; mean age, 13.15 years; range, 12–15 years). We included the participants in the experiment considering their age to have a sample of the teenage population

(participants older than aged 12). All the selected participants were attending high school in different schools in the Valencia area at the time of the experiment. In our study, we modeled our participants considering two different dimensions of personalities and social interaction behaviors on the OSN platform. Furthermore, we investigated if stronger behavioral patterns could be identified when grouping our population by gender (i.e., male/female) and by personality clusters (i.e., average, self-centered, reserved, and role model).

The first modeling dimension considered in this research is the personality. We used the Big-5 personality traits to represent the personality of our participants. From these five personality trait values, we grouped our participants into four different personality clusters having the following composition, specified as the average cluster (C_1=44, 56.8% males); the self-centered cluster (C_2=38, 68.4% males); the reserved cluster (C_3=52, 48.1% males); and the role model cluster (C_4=52, 63.5% males). Fig. 2 shows the Big-5 personality traits distribution of the clusters found in our study. Each cluster is defined by the means of averages of their personality trait z-scores. Therefore, it is possible to observe how the personality trait average z-scores of its members follow different distributions, depending on the cluster (i.e., average, self-centered, reserved, and role model).

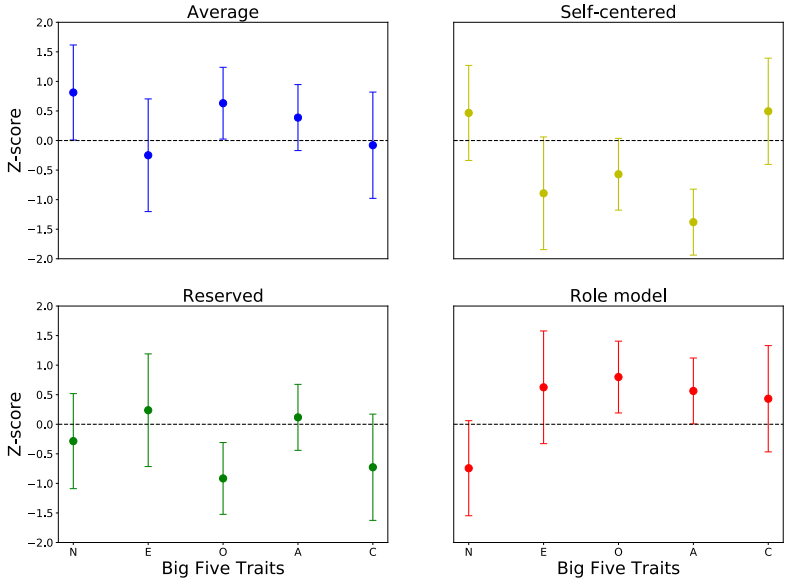

Fig. 2. Personality clusters observed in participants' data. "•" is the position of cluster centers represented as the average z-score of their cluster personality traits. The error bars represent the standard deviation of each trait in each cluster. The dotted lines represent global average values (z=0) for each personality trait.

Comparing the clusters found in this work with the clusters proposed in [34], it is possible to observe strong similarities between them. The Silhouette coefficient (SC) of the computed clusters is 0.173, meaning that some clusters could be overlapping (SC≈0), but the samples are not being misclassified (SC>0). The reserved personality type is characterized by negative z-score values on neuroticism and *openness*, while the rest of the traits (extraversion, agreeableness, and conscientiousness) are slightly higher than 0. The role model personality type is characterized by negative z-score values in neuroticism and positive z-score values for the rest of the traits. For both clusters, the personality traits of our participants followed the same distributions as [34] clusters. The average personality type is characterized by z-score values close to 0 for all personality traits. In our study, this cluster follows this trend with slightly higher z-score values on neuroticism and openness. Finally, the self-centered personality type is

characterized by negative z-score values on all the personality traits except for the extraversion trait. By comparing it with our cluster, we found some differences between those. However, we found a strong relationship with the original cluster on which self-centered was based, which is referred to as undercontrolled and introduced in the [39] work. In [34], the undercontrolled personality group is considered to have a strong influence on the new proposed self-centered cluster. From the clusters observed in our population, we can support this statement. The self-centered cluster observed in this work has a positive neuroticism z-score like the original undercontrolled group. Significant differences were observed regarding the conscientiousness trait in our reserved and self-centered clusters. Studies have shown how conscientiousness is the most variable trait with respect to the corresponding age [40]. Therefore, we think the observed differences were mainly due to the important age gap between the participants of both studies.

Finally, the second dimension used to model OSN users in our analysis is their online social interaction behavior. During our study, a total number of 2,195 likes, 7,650 comments, 1,309 shares, 846 uploaded photos, and 7,788 status updates (among them, 761 were private, 769 were public, 5,774 were disclosed to friends, and 484 were disclosed to specific lists of friends) were registered on the PESEDIA database. The average text size was measured with the average number of words per post/publication, and the time spent was measured as the number of logins into the OSN platform. The participants had a mean of 12 friendships and regretted 2,761 actions made (which warranted undo/delete). The most common social interactions were posted comments and status updates. We observed an average of 41 comments and 42 status updates per user. It is also interesting to observe the high average number of deletes per user (i.e., 15), which represents a high number of regrets of the content posted/published on the social network platform. We also observed that in general users preferred to share posts/publications with friends only rather than publicly, privately, or with specific groups of friends.

At the end of the study, we collected 186 different combinations of the Big-5 personality traits and 18,942 different OSN interactions. Also, we also collected the following: 930 pairwise comparisons of argumentation schemes, one per participant (186) and argumentation schemes (5); as well as 4,464 pairwise comparisons of argument types, one per participant (186), argument types (4), and content variations of content-type of argument (6).

4. Results

With the aim of finding an answer to our research questions, we calculated the persuasive power of the five argumentation schemes and four argument types using the persuasive power equation (see Equation 1). Furthermore, we did a correlation analysis between our user modeling features (i.e., individual personality and online social interaction features) and the previously calculated persuasive power values. In this section, we present the observed results after completing this process using the data gathered at the end of our study.

4.1 Persuasive Power of Arguments

Based on the results of the study, we have calculated the persuasive power of all the argumentation schemes and types considered in this work. In this way, we can sequentially sort the five argumentation schemes, considering their persuasive power in the order of (RQ1) AFCQ > AFEO > AFWT > AFPP > AFPO. AFCQ seemed to be the most effective scheme for persuading our participants with a score of 0.61. Subsequently, we have the AFEO scored with 0.53, AFWT with a score of 0.47, AFPP with a score of 0.46 and finally, AFPO was the least persuasive scheme with a score of 0.43. These results mean that teenagers in general can be persuaded better by showing the consequences of their actions or using recommendations made by experts rather than recommendations made by someone like them or according to popular trends or opinions. Table 1 shows the direct comparison of the persuasive ranking between pairs of argumentation schemes. For that purpose, we measure the number of times an argument

is ranked at a higher position than another. We can notice how this value is higher when arguments with stronger persuasive power are compared with lower persuasive power arguments and vice-versa.

Table 1. Pairwise rank comparative between argumentation schemes

	AFCQ	AFEO	AFWT	AFPP	AFPO	Total
AFCQ	-	106	114	112	123	455
AFEO	80	-	104	102	109	395
AFWT	72	82	-	96	99	349
AFPP	74	84	90	-	92	340
AFPO	63	77	87	94	-	321

Values are represented as the number of times an argumentation scheme (rows) beats another argumentation scheme (columns).

On the other hand, we sorted in order the argument types taking their persuasive power into account in the order of (RQ2) content > privacy > risk > trust. Content arguments were the most persuasive with a score of 0.59. Subsequently, we have privacy arguments with a score of 0.52, risk arguments with score of 0.47 and trust arguments with a score of 0.42, the results of which mean that teenagers are more concerned about sharing sensitive content rather than being read by unknown users or endangering other parties' privacy. Like the previous analysis with argumentation schemes, Table 2 represents a direct comparison between the ranking position of every pair of argument types. Here, we also observe how arguments with a higher persuasive power score are ranked in general at a higher position than the rest. If we consider each round of questionnaire B independently to analyze the effect of each content type on the persuasion of the argument, the following persuasive sequence is observed in the order of offensive > personal > family > medical > alcohol/drugs > location. Therefore, although content arguments were found as the most persuasive type of argument, depending on which type of content was considered in each round, users' susceptibility was different. Our study revealed that teenagers are more concerned about sharing offensive content with a score of 0.64, closely followed by sharing personal information with a score of 0.62. The concern with these specific types of content matches the new trends on social network platforms of self-presentation [41]. The next most concerning types of content were family/association and medical content with scores of 0.59 and 0.58, respectively. Finally, revealing alcohol/drug consumption or location information seemed to be the less relevant types of content for our participants with scores of 0.56 and 0.53, respectively.

Table 2. Pairwise rank comparative between argument types

	Content	Privacy	Risk	Trust	Total
Content	-	623	659	681	1,963
Privacy	493	-	605	643	1,732
Risk	457	511	-	616	1,584
Trust	435	482	500	-	1,417

Values are represented as the number of times an argument type (rows) beats another argument type (columns).

4.2 Personality Impact on Argument Persuasion

To be able to adjust our argumentation system to increase the persuasive power of the arguments for our target population, we analyzed the personality impact on the persuasive power (RQ3) of argumentation schemes and argument types. For this purpose, we have calculated the Spearman ρ rank correlation between the persuasive power of arguments and the Big-5 personality traits. To ease the interpretation and visualization of the results, we have grouped correlations into three correlation-strength categories. Weak correlations stand for correlation values between 0 and 0.2, while we consider a moderate correlation

if its correlation value is between 0.2 and 0.6. Finally, a strong correlation stands for correlation values higher than 0.6.

The significant correlations found between argumentation schemes and argument types with personality traits are represented in Tables 3 and 4, respectively. As we observed, different correlations have been found for different groups of users and corresponding different personality traits. It is possible to observe how in some cases the personality correlates with the perceived persuasive power, which means that personality features could serve as predictors of persuasiveness when defining persuasive policies. We can also observe a smaller number of significant correlations between the argument types and argumentation schemes. A possible cause for this pattern in our findings is that variations among different argumentation schemes have a greater impact on the perceived persuasiveness of an argument than the variation between argument types. Furthermore, studying specific groups of users categorized by descriptive features such as the gender or personality allow to draw more informed conclusions rather than considering the whole heterogeneous group of users. Thus, personalization is a key aspect to improve the effectiveness of the human-computer interactions of an argumentation system.

Table 3. Significant correlations of argumentation schemes' persuasive power and personality traits

Participants	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
All	-	-	-AFEO**	-AFPP**	-
Male	-	-	-AFEO*/+AFWT*	-AFPP*/+AFWT*	-
Female	-	-	-	-	+AFPP*
Average	-	-	-	-	-
Reserved	-	-	--AFEO*	-	++AFPP**/--AFEO*
Self-centered	-	--AFCQ*	-	-	-
Role model	-	-AFPO*/++AFEO**	-	-	-

The correlation strength is represented as weak "+/-", moderate "++/--", and strong "+++/---."
*p<0.05, **p<0.01.

Table 4. Significant correlations of argument types' persuasive power and personality traits

Participants	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
All	-	-	-Privacy*	-	-
Male	-	-	-	-	-
Female	-	-	-	-	-
Average	-	-	-	-	-
Reserved	-Risk*	-	-	-	-
Self-centered	-	-	++Content*	-	-
Role model	-	-	-	-	-

The correlation strength is represented as weak "+/-", moderate "++/--", and strong "+++/---."
*p<0.05, **p<0.01.

4.3 Social Interaction Impact on Argument Persuasion

In some environments, obtaining the Big-5 personality traits may not be possible. Therefore, to model our OSN users before analyzing the persuasive power of arguments, we have proposed an alternative to personality traits based on the social interaction behaviors of our participants [11, 35]. In this way, we analyzed if there existed any correlation between the persuasion of arguments toward each participant, depending on their social interaction behaviors (RQ4). To measure the impact of these thirteen features on the persuasive power of arguments, we have calculated the Spearman ρ rank correlation between them and the persuasive power of arguments. The interpretation of the correlation values is done the same way as in the previous section. Furthermore, if personality traits are available, we have also considered making a complete analysis, considering personality clusters. In this way, it is possible to combine the results of

both analyses, thus observing even more significant correlations between users' features and persuasiveness of arguments.

The significant correlations found between argumentation schemes and argument types with OSN interaction data have been represented in Tables 5 and 6, respectively. As in the previous analysis, social interaction data has proved to be a good predictor of variations in perceived persuasion for different user models. Again, it was harder to find significant correlations when considering the argument types compared to the argumentation schemes. This pattern reinforces the hypothesis that argumentation schemes contain more persuasive information that argument types, and are aligned with the recent findings described in [28].

Table 5. Significant correlations of argumentation schemes' persuasive power and social interaction data

Participants	#friends	#status_upd	#likes	#comments	#ppprivate	#pppublic	#ppfriends	#ppgroups	Avg_textsize
All	-	-	-	-	+AFPO*	-	-	-AFPP**	-AFPO*
Male	-	-	--AFPP*	++AFPO**	-	-	-	--AFPP**	-
Female	-	++AFEO*	-	-	++AFPO**	-	--AFCQ**	-	-
Average	++AFEO*	-	-	-	--AFPP*	-	++AFEO*	-	-
Reserved	++AFCQ*	++AFCQ*	-	-	-	--AFWT*	++AFCQ*	-	-
Self-centered	-	-	-	-	--AFEO*	-	-	-	--AFCQ**/ ++AFWT**
Role model	-	-	-	++AFEO*	-	-	-	AFCQ**/ ++AFPO**	++AFCQ**

The correlation strength is represented as weak "+/-", moderate "++/--", and strong "+++/---."
*p<0.05, **p<0.01.

Table 6. Significant correlations of argument types' persuasive power and social interaction data

Participants	#status_upd	#comments	#ppprivate	#ppgroups	#deletes
All	-	-	-	-	-
Male	++Trust*	-	-	++Trust*	--Risk*
Female	-	-	-	-	-
Average	-	-	-	-	-
Reserved	-	-	-	-	-
Self-centered	-	-	--Risk*	-	-
Role model	++Trust*	++Trust*	-	-	-

The correlation strength is represented as weak "+/-", moderate "++/--", and strong "+++/---."
*p<0.05, **p<0.01.

4.4 Interpretation of the Results

This work sets the starting point to develop the human interaction part of argumentative educational systems to help with privacy management on an OSN. The findings observed in this paper reveal that personality traits and social interaction data are relevant user modeling features useful for estimating the perceived persuasive power of arguments by different user models. Therefore, these features represent a powerful way to model human users when approaching a problem having these specifications. These findings are consistent with recent research on similar topics [42]. In Table 7, we present four different OSN user models, considering the features proposed in this work. Based on the identified correlations, we have estimated potential trends in the persuasive power of arguments for these users as shown in Table 8. We can observe how different user models may perceive arguments with a modified persuasive power. Thus, based on the observed results, we can adapt the available argumentation schemes and argument types following different user-tailored persuasive policies, which are more effective than the one based on the general persuasive power of arguments.

Argumentation schemes have been previously investigated and classified by experts of many different disciplines ranging from philosophy, communication studies, linguistics to computer science and psychology [9]. Thus, several clusters of schemes have been defined grouped according to their general

category. The schemes we work with belong to the general categories of "practical reasoning arguments" (AFCQ) and "source-dependent arguments," concretely, as well as its subcategories of "arguments from position to know" (AFEO and AFWT) and "arguments from popular acceptance" (AFPO and AFPP).

Table 7. Four different user models

Model features	User 1	User2	User 3	User 4
Gender	Male	Female	Female	Male
Cluster	Self-centered	Role model	Reserved	Average
Openness	-	-	↓	-
Conscientiousness	↓	-	-	-
Extraversion	↓	-	-	-
Agreeableness	-	-	-	-
Neuroticism	↑	↓	↑	-
#friends	-	-	↑	↓
#status_upd	↑	-	↓	-
#likes	↓	-	-	-
#shares	-	-	-	-
#comments	-	↑	-	-
#ppprivate	↓	-	↑	-
#pppublic	-	-	-	-
#ppfriends	↑	-	↑	-
#ppgroups	-	-	-	-
#deletes	-	-	-	-
#photos	-	-	-	-
avg_textsize	-	↑	-	↓
time_spent	-	-	-	-

"-" represents an average value, "↑" represents a value above the average, and "↓" represents a value below average.

Table 8. Persuasive power of argumentation schemes and argument types for four different users

Persuasive power	User 1	User2	User 3	User 4
AFCQ	↑	↑	↓	-
AFEO	↑↑	↑↑	↓↓	↓
AFWT	↓	-	-	↑
AFPP	-	↓	↑↑	↓↓
AFPO	↓	-	-	↑
Content	↓	-	-	-
Privacy	↑	-	-	↓
Risk	↑	-	↑	↓
Trust	↑	↑	-	-

"-" represents an unmodified value, "↑" represents an increased persuasive power, and "↓" represents a decreased persuasive power.

Recently, a relation between this classification and Cialdini's principles of persuasion has been established [28, 43]. Thus, the "consistency" principle of persuasion, by which people like to be consistent with the things they have previously said or done, relates to their practical behavior (AFCQ); the principle of "authority," by which people follow the lead of credible, knowledgeable experts, relates to source-based arguments (AFEO and AFWT); and the principle of "consensus," by which individuals are conformed to what the majority regards as acceptable, relates to arguments from popular acceptance (AFPO and AFPP). First, we can see how our findings detect a preference order given as "consistency" > "authority" > "consensus" for the persuasion principles on our social media platform. Second, although there is still no specific research that orders these persuasion principles by their persuasive power in the

context of online social media, similar research on the healthy eating domain concluded an order given as "authority" > "consensus" and "consistency" (with no significant difference between these two), and stated that persuasive power is highly influenced by the domain [43]. Based on these mappings between argumentation schemes and persuasive principles, we can contextualize our findings within essential concepts of persuasive psychology research. Thus, any significant correlation detected between user descriptive features (i.e., individual personality and social interaction) and argumentation schemes can be also interpreted as a correlation between the users' features and the three persuasive principles related to the five argumentation schemes analyzed in our study. Case in point, from our findings we can interpret that the "authority" (AFEO) principle has shown an increased persuasive power for reserved users with a low value on extraversion. Although we have not been able to identify much prior research focused on this same purpose, the work presented in [27] sheds light on existing correlations between Big-5 personality traits and Cialdini's persuasive principles. Even though the populations of both studies differ substantially in age, some similarities can be identified among the significant correlations detected in both works. In this paper, we have identified the following correlations, namely a negative correlation between AFEO and extraversion trait; a positive correlation between AFEO and conscientiousness for our role model participants; and a positive correlation between AFPP and neuroticism for our female and reserved participants. All of these detected correlations have also been found in [27]. However, the age gap and significant differences between populations make it harder to compare the findings on both works and some of the found correlations remain unexplained.

5. Conclusion and Future Work

At the beginning of this work, we raised four different research questions aimed at having a better understanding of human persuasion on an OSN platform using arguments. Based on our findings, we have been able to answer the four research questions, and have a better understanding of the persuasiveness of arguments when used with different user models. Personalization plays a major role in effective human-computer interactions. In this paper, we have been able to observe that it is possible observe variations in the effectiveness of arguments, using representative user modeling features (i.e., personality and social interaction data). Therefore, the user models analyzed in this work provide a solid basis for developing personalized argumentation systems aimed at educating and preventing privacy violations on the OSN platform. Given the nature of arguments and argumentation, the findings observed in our study lay the groundworks for developing powerful tools for education and decision-making assistance.

As for future work, we will strive to deepen the analysis presented in this paper by extending our study to an adult population, and implement an argument-based persuasive algorithm capable of generating personalized persuasive policies aimed at maximizing the efficiency of human-computer interactions.

Author's Contributions

Conceptualization, RRD; Funding acquisition, AGF; Investigation and methodology, RRD, JA; Supervision, SH, AGF; Data curation, RRD; Visualization, RRD; Writing of the review & editing, RRD, SH, AGF.

Funding

This work was partially supported by the Spanish Government (Project TIN2017-89156-R, PID2020-113416RB-I00), and the Valencian Government (Project PROMETEO/2018/002).

Competing Interests

The authors declare that they have no competing interests.

References

[1] J. M. Such and N. Criado, "Resolving multi-party privacy conflicts in social media," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 7, pp. 1851-1863, 2016. https://doi.org/10.1109/TKDE.2016.2539165

[2] F. Quayyum, D. S. Cruzes, and L. Jaccheri, "Cybersecurity awareness for children: a systematic literature review," *International Journal of Child-Computer Interaction*, vol. 30, article no. 100343, 2021. https://doi.org/10.1016/j.ijcci.2021.100343

[3] R. Ruiz-Dolz, "Towards an artificial argumentation system," in *Proceedings of the 29th International Conference on International Joint Conferences on Artificial Intelligence*, Yokohama, Japan, 2020, pp. 5206-5207. https://doi.org/10.24963/ijcai.2020/743

[4] C. Chesnevar, A. G. Maguitman, and M. P. Gonzalez, "Empowering recommendation technologies through argumentation," in *Argumentation in Artificial Intelligence*. Boston, MA: Springer, 2009 pp. 403-422. https://doi.org/10.1007/978-0-387-98197-0_20

[5] J. Alemany, E. Del Val, J. Alberola, and A. García-Fornes, "Enhancing the privacy risk awareness of teenagers in online social networks through soft-paternalism mechanisms," *International Journal of Human-Computer Studies*, vol. 129, pp. 27-40, 2019. https://doi.org/10.1016/j.ijhcs.2019.03.008

[6] R. J. Thomas, J. Masthoff, and N. Oren, "Adapting healthy eating messages to personality," in *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors.* Cham, Switzerland: Springer, 2017, pp. 119-132. https://doi.org/10.1007/978-3-319-55134-0_10

[7] R. J. Thomas, J. Masthoff, and N. Oren, "Is ArguMessage effective? A critical evaluation of the persuasive message generation system," in *Persuasive Technology: Development of Persuasive and Behavior Change Support Systems*. Cham, Switzerland: Springer, 2019, pp. 87-99). https://doi.org/10.1007/978-3-030-17287-9_8

[8] R. Ruiz-Dolz, S. Heras, J. Alemany, and A. Garcia-Fornes, "Towards an argumentation system for assisting users with privacy management in online social networks," in *Proceedings of the 19th Workshop on Computational Models of Natural Argument co-located with the 14th International Conference on Persuasive Technology (CMNA@PERSUASIVE 2019)*, Limassol, Cyprus, 2019, pp. 17-28.

[9] D. Walton and F. Macagno, "A classification system for argumentation schemes," *Argument & Computation*, vol. 6, no. 3, pp. 219-245, 2015. https://doi.org/10.1080/19462166.2015.1123772

[10] W. Wu, L. Chen, and Y. Zhao, "Personalizing recommendation diversity based on user personality," *User Modeling and User-Adapted Interaction*, vol. 28, pp. 237-276, 2018. https://doi.org/10.1007/s11257-018-9205-x

[11] S. Basaran and O. H. Ejimogu, "A neural network approach for predicting personality from Facebook data," *Sage Open*, vol. 11, no. 3, article no. 21582440211032156, 2021. https://doi.org/10.1177/21582440211032156

[12] J. Alemany, E. D. Val, and A. Garcia-Fornes, "A review of privacy decision-making mechanisms in online social networks," *ACM Computing Surveys (CSUR)*, vol. 55, no. 2, pp. 1-32, 2022. https://doi.org/10.1145/3494067

[13] P. Russo, L. Bracciale, and G. Bianchi, "Dare-to-share: collaborative privacy-preserving recommendations with (almost) no crypto," *Security and Privacy*, vol. 4, no. 3, article no. e153, 2021. https://doi.org/10.1002/spy2.153

[14] A. Altalbe and F. Kateb, "Assuring enhanced privacy violation detection model for social networks," *International Journal of Intelligent Computing and Cybernetics*, vol. 15, no. 1, pp. 75-91, 2022. https://doi.org/10.1108/IJICC-05-2021-0093

[15] N. Voloch, N. Gal-Oz, and E. Gudes, "A trust based privacy providing model for online social networks," *Online Social Networks and Media*, vol. 24, article no. 100138, 2021. https://doi.org/10.1016/j.osnem.2021.100138

[16] A. C. Squicciarini, H. Xu, and X. Zhang, "CoPE: enabling collaborative privacy management in online social networks," *Journal of the American Society for Information Science and Technology*, vol. 62, no. 3, pp. 521-534, 2011. https://doi.org/10.1002/asi.21473

[17] N. Kokciyan, N. Yaglikci, and P. Yolum, "An argumentation approach for resolving privacy disputes in online social networks," *ACM Transactions on Internet Technology (TOIT)*, vol. 17, no. 3, article no. 27, 2017. https://doi.org/10.1145/3003434

[18] A. C. Kurtan and P. Yolum, “Assisting humans in privacy management: an agent-based approach,” *Autonomous Agents and Multi-Agent Systems*, vol. 35, article no. 7, 2021. https://doi.org/10.1007/s10458-020-09488-1

[19] F. Mosca and J. Such, “An explainable assistant for multiuser privacy,” *Autonomous Agents and Multi-Agent Systems*, vol. 36, article no. 10, 2022. https://doi.org/10.1007/s10458-021-09543-5

[20] F. Mosca and J. M. Such, “ELVIRA: an explainable agent for value and utility-driven multiuser privacy,” in *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, Virtual Event, UK, 2021, pp. 916-924. https://dl.acm.org/doi/10.5555/3463952.3464061

[21] E. Hadoux and A. Hunter, “Comfort or safety? Gathering and using the concerns of a participant for better persuasion,” *Argument & Computation*, vol. 10, no. 2, pp. 113-147, 2019. https://doi.org/10.3233/AAC-191007

[22] H. Nasrudin, B. Yonata, and W. B. Sabtiawan, “The claims statements from viral videos for instrument development to assess argumentation thinking skills,” *Journal of Physics: Conference Series*, vol. 1899, no. 1, article no. 012174, 2021. https://doi.org/10.1088/1742-6596/1899/1/012174

[23] S. Heras, P. Rodriguez, J. Palanca, N. Duque, and V. Julian, “Using argumentation to persuade students in an educational recommender system,” in *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors.* Cham, Switzerland: Springer, pp. 227-239. https://doi.org/10.1007/978-3-319-55134-0_18

[24] R. L. Fogues, P. K. Murukannaiah, J. M. Such, and M. P. Singh, “Sharing policies in multiuser privacy scenarios: incorporating context, preferences, and arguments in decision making,” *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 24, no. 1, article no. 5, 2017. https://doi.org/10.1145/3038920

[25] Y. Wang, G. Norcie, S. Komanduri, A. Acquisti, P. G. Leon, and L. F. Cranor, “"I regretted the minute I pressed share" a qualitative study of regrets on Facebook,” in *Proceedings of the 7th Symposium on Usable Privacy and Security*, Pittsburgh, PA, USA, 2011, pp. 1-16. https://doi.org/10.1145/2078827.2078841

[26] R. J. Thomas, N. Oren, and J. Masthoff, “ArguMessage: a system for automation of message generation using argumentation schemes,” in *Proceedings of 2018 Convention of the Society for the Study of Artificial Intelligence and the Simulation of Behaviour (AISB)*, Liverpool, UK, 2018, pp. 28-31.

[27] A. Ciocarlan, J. Masthoff, and N. Oren, “Actual persuasiveness: impact of personality, age and gender on message type susceptibility,” in *Persuasive Technology: Development of Persuasive and Behavior Change Support Systems.* Cham, Switzerland: Springer, 2019, pp. 283-294. https://doi.org/10.1007/978-3-030-17287-9_23

[28] R. Ruiz-Dolz, J. Taverner, S. Heras, A. Garcia-Fornes, and V. Botti, “A qualitative analysis of the persuasive properties of argumentation schemes,” in *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization*, Barcelona, Spain, 2022, pp. 1-11. https://doi.org/10.1145/3503252.3531324

[29] E. Christofides, A. Muise, and S. Desmarais, “Hey mom, what’s on your Facebook? Comparing Facebook disclosure and privacy in adolescents and adults,” *Social Psychological and Personality Science*, vol. 3, no. 1, pp. 48-54, 2012. https://doi.org/10.1177/1948550611408619

[30] S. Rothmann and E. P. Coetzer, “The big five personality dimensions and job performance,” *SA Journal of Industrial Psychology*, vol. 29, no. 1, pp. 68-74, 2003. https://doi.org/10.4102/sajip.v29i1.88

[31] R. Ruiz-Dolz, J. Alemany, S. Heras, and A. Garcia-Fornes, “Automatic generation of explanations to prevent privacy violations,” in *Proceedings of the 2nd EXplainable AI in Law Workshop (XAILA 2019) co-located with 32nd International Conference on Legal Knowledge and Information Systems (JURIX 2019)*, Madrid, Spain, 2019.

[32] J. Alemany, E. Del Val, and A. Garcia-Fornes, “Assisting users on the privacy decision-making process in an OSN for educational purposes,” in *Advances in Practical Applications of Agents, Multi-Agent Systems, and Trustworthiness: The PAAMS Collection*. Cham, Switzerland: Springer, 2020, pp. 379-383. https://doi.org/10.1007/978-3-030-49778-1_31

[33] M. Mackiewicz and J. Cieciuch, “Pictorial personality traits questionnaire for children (PPTQ-C): a new measure of children's personality traits,” *Frontiers in Psychology*, vol. 7, article no. 498, 2016. https://doi.org/10.3389/fpsyg.2016.00498

[34] M. Gerlach, B. Farb, W. Revelle, and L. A. Nunes Amaral, “A robust data-driven approach identifies four personality types across four large data sets,” *Nature Human Behaviour*, vol. 2, pp. 735-742, 2018. https://doi.org/10.1038/s41562-018-0419-z

[35] C. Huang, "Social network site use and Big Five personality traits: a meta-analysis," *Computers in Human Behavior*, vol. 97, pp. 280-290, 2019. https://doi.org/10.1016/j.chb.2019.03.009

[36] J. Alemany, E. Del Val, J. M. Alberola, and A. Garcia-Fornes, "Metrics for privacy assessment when sharing information in online social networks," *IEEE Access*, vol. 7, pp. 143631-143645, 2019. https://doi.org/10.1109/ACCESS.2019.2944723

[37] A. Caliskan Islam, J. Walsh, and R. Greenstadt, "Privacy detective: detecting private information and collective privacy behavior in a large social network," in *Proceedings of the 13th Workshop on Privacy in the Electronic Society*, Scottsdale, AZ, USA, 2014, pp. 35-46. https://doi.org/10.1145/2665943.2665958

[38] E. M. Schomakers, C. Lidynia, D. Mullmann, and M. Ziefle, "Internet users' perceptions of information sensitivity–insights from Germany," *International Journal of Information Management*, vol. 46, pp. 142-150, 2019. https://doi.org/10.1016/j.ijinfomgt.2018.11.018

[39] J. B. Asendorpf, P. Borkenau, F. Ostendorf, and M. A. Van Aken, "Carving personality description at its joints: confirmation of three replicable personality prototypes for both children and adults," *European Journal of Personality*, vol. 15, no. 3, pp. 169-198, 2001. https://doi.org/10.1002/per.408

[40] M. B. Donnellan and R. E. Lucas, "Age differences in the Big Five across the life span: evidence from two national samples," *Psychology and Aging*, vol. 23, no. 3, pp. 558-566, 2008. https://doi.org/10.1037/a0012897

[41] T. H. H. Chua and L. Chang, "Follow me and like my beautiful selfies: Singapore teenage girls' engagement in self-presentation and peer comparison on social media," *Computers in Human Behavior*, vol. 55, pp. 190-197, 2016. https://doi.org/10.1016/j.chb.2015.09.011

[42] D. Hooshyar, Y. M. Huang, and Y. Yang, "A three-layered student learning model for prediction of failure risk in online learning," *Human-centric Computing and Information Sciences*, vol. 12, article no. 28, 2022. https://doi.org/10.22967/HCIS.2022.12.028

[43] R. J. Thomas, "Personalised persuasive messages for behaviour change interventions: combining Cialdini's principles and argumentation schemes," Ph.D. dissertation, Aberdeen University, Aberdeen, UK, 2019.