



# Automated privacy negotiations with preference uncertainty

Dorota Filipczuk<sup>1</sup> · Tim Baarslag<sup>2,3</sup> · Enrico H. Gerding<sup>1</sup> · m.c. schraefel<sup>1</sup>

Accepted: 27 July 2022 / Published online: 26 August 2022  
© The Author(s) 2022

## Abstract

Many service providers require permissions to access privacy-sensitive data that are not necessary for their core functionality. To support users' privacy management, we propose a novel agent-based negotiation framework to negotiate privacy permissions between users and service providers using a new multi-issue alternating-offer protocol based on exchanges of partial and complete offers. Additionally, we introduce a novel approach to learning users' preferences in negotiation and present two variants of this approach: one variant personalised to each individual user, and one personalised depending on the user's privacy type. To evaluate them, we perform a user study with participants, using an experimental tool installed on the participants' mobile devices. We compare the take-it-or-leave-it approach, in which users are required to accept all permissions requested by a service, to negotiation, which respects their preferences. Our results show that users share personal data 2.5 times more often when they are able to negotiate while maintaining the same level of decision regret. Moreover, negotiation can be less mentally demanding than the take-it-or-leave-it approach and it allows users to align their privacy choices with their preferences. Finally, our findings provide insight into users' data sharing strategies to guide the future of automated and negotiable privacy management mechanisms.

**Keywords** Multi-issue negotiation · Automation · Privacy management

---

✉ Dorota Filipczuk  
dorota@ecs.soton.ac.uk

Tim Baarslag  
t.baarslag@cwi.nl

Enrico H. Gerding  
eg@ecs.soton.ac.uk

m. c. schraefel  
mc@ecs.soton.ac.uk

<sup>1</sup> University of Southampton, Southampton, UK

<sup>2</sup> Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands

<sup>3</sup> Utrecht University, Utrecht, The Netherlands

## 1 Introduction

With the advancement of digital economy, our personal data has become the ‘new oil of the Internet’ [104]. From gaming apps to social media platforms, many popular online services are indirectly funded through the collection of personal data which use it as a currency: users share the data in return for access to the service. Service providers then monetise the information through personalised advertising, use it to improve the service or sometimes sell to third parties. Typically, those services rely on privacy mechanisms that allow the user to manually grant them access to the data. For instance, mobile operating systems such as iOS and Android allow service providers to ask for permission to access the user’s resources via a dialog box. Similarly, websites embed cookie banners which inform the user of online tracking and ask them to accept or reject such tracking. Other examples include authentication mechanisms such as Facebook Login<sup>1</sup> which allows apps to ask for permissions to access users’ Facebook data.

However, at the other end of this data exchange, there is an individual user who receives an unwieldy number of data sharing requests. As the number of services asking for personal data grows, there is an increasing risk that more and more users may start experiencing the so-called *consent transaction overload*, i.e. a situation when there are too many data sharing requests for an individual to consider [21, 103]. Consequently, the increasing difficulty on the user’s side in managing their personal data leads to them feeling a loss of control and a sense of ‘weariness’ towards privacy issues, in which users believe that there is no effective means of managing their personal data [27]. As a result, excessive data sharing requests lead to *consent fatigue* [103] which reflects users’ tendency to simply accept a request without reading it [27]. When users start accepting data sharing requests without thinking whether they aligns with their privacy preferences or not, they are no longer able to control their personal data manually. For this reason, we believe that there is a need for automation in the area of privacy management.

Furthermore, service providers have an unfair advantage over users. This is not only because users almost always grant access to their personal data when it is required by the service provider [22], but also because they may be unaware of the data collection. In fact, recent research shows that even European websites, where the most demanding privacy regulations apply [69], tend to collect visitor’s personal data without consent, collect the data even when the user has explicitly opted out, or nudge users towards sharing by pre-selecting options [74, 102]. This allows service providers to take advantage of users which does not result in fair agreements between the two parties. However, as users are becoming more aware of the ultimate value of their data and potential consequences of disclosure, a new trend is emerging to compensate users directly for the loss of privacy they may suffer as a result of this exchange [104]. More and more often, users can choose to pay to opt out of advertising on an app that is usually free, or to share their data in return for an incentive. For instance, startups such as Gener8 Ads<sup>2</sup> and BitsaboutMe<sup>3</sup> allow users to exchange their personal data for gifts, discounts, charity donations and monetary rewards. In fact, Perera et al. proposed the idea of personal data marketplaces empowered by agent negotiations where consent of data owners is exchanged in an efficient and effective manner [88, 89].

<sup>1</sup> Facebook Login, <https://developers.facebook.com/docs/facebook-login/>.

<sup>2</sup> Gener8 Ads, <https://gener8ads.com/>.

<sup>3</sup> BitsaboutMe, <https://bitsabout.me/en/>.

Yet, ensuring that users are compensated fairly for the privacy loss is a non-trivial problem, because the users' views of fairness are subjective within the context of privacy [20, 121]. To address this problem, Krol and Preibusch envisioned *effortless privacy negotiations* which provide the opportunity to reach individual agreements about personal data collection between users with heterogeneous privacy preferences and service providers [62]. In their article, they posed the question of how such negotiations could be implemented, highlighting that making such them 'effortless' is challenging and 'might not be achievable' at all [62]. Although privacy issues have received a lot of attention from the multi-agent systems community [108], to date the opportunity of providing an effective solution to enable privacy negotiations between the user and the service provider has not been well-explored. There have been very few studies which propose automated negotiation solutions in the privacy domain, and none of these have been evaluated with real users using their real data. A major challenge in the implementation of such negotiations is the uncertainty around the user's individual privacy preferences.

Therefore, we address this gap by proposing a novel agent-based approach for privacy negotiations and by testing this approach with human participants using their actual private data. In general, agent-based negotiation can offer a lot of benefits: win-win outcomes for all parties, reduction in negotiation duration and reduced users' effort [5]. In this paper, we argue that privacy negotiations can be automated effectively using autonomous agents that represent users and service providers. In particular, our goal is to automate this negotiation on behalf of a user to reduce the number of data sharing requests that the user needs to process manually. As such, we specifically focus on the problem of the agent's uncertainty about the privacy preferences of the user it represents.

Specifically, we introduce a novel framework for automated negotiation with preference uncertainty, i.e. where the agent needs to learn the user's preferences. As part of the framework, we propose a new alternating offers multi-issue negotiation protocol with costly quoting, in which one agent can propose partial offers by specifying values for some issues, and the other agent completes those offers. We argue that this framework: 1) is more suitable for this negotiation domain than the original alternating offers protocol, 2) allows for more collaborative exploration of the negotiation space to find mutually beneficial agreements, and 3) avoids distributive negotiation on single issues. Using this protocol, the agent employs its strategy to negotiate autonomously on the user's behalf. In a subsequent phase, the agent presents a potential agreement to the user, which the user can accept, or manually override and continue the negotiation in order to further improve the offer. Importantly, this framework generalises to other domains where the negotiation takes place between an individual and a service provider.

Additionally, to address the uncertainty about the user's preferences, we introduce a novel approach to *implicit preference elicitation* [58]. Specifically, the agent representing the interests of the user needs to learn the user's preferences in order to accurately represent them. Instead of asking directly, the agent can implicitly derive the user's preferences in the form of a utility function from feedback received from the user either by accepting or manually overriding offers. Once the user's utility function is established, in subsequent negotiations, it is used to generate optimal offers – that is, offers that maximise the user's utility and are likely to be accepted by the service provider. We propose two different variants of this approach. In the first variant (Agent 1), the utility function is personalised for each user based on their previous decisions (accepted or rejected offers). For the second variant (Agent 2), the agent classifies a user as being one of a three different privacy types, which is initially determined based on their Westin's Privacy Segmentation Index [40, 64] and later adjusted based on their sharing behaviour.

Finally, we perform the first empirical study of agent-based privacy negotiation<sup>4</sup>. We investigate how it impacts users' personal data sharing decisions (i.e. *do the users allow access to their private data more or less when an agent negotiates on their behalf?*) and on the regret of those decisions (i.e. *do the users regret granting the access more or less when an agent negotiates on their behalf?*), and how these decisions relate to the users' self-reported data sharing sensitivity scores (i.e. *are the permission-granting decisions consistent with users' privacy concerns and data sharing preferences?*). We also evaluate the approaches in terms of the accuracy of the offers and the users' effort. In general, the results from our lab study show that Agent 2 is more accurate than both Agent 1 and a condition with randomly-chosen default settings. However, over the time of eight negotiation rounds, we can observe a rising trend of the accuracy of Agent 1, which in the end, exceeds the accuracy of Agent 2. Moreover, we observe that users grant permissions significantly more often (on average, over 2.5 times more often) when they are able to negotiate their permission settings, while the percentage of the regretted permission-granting decisions remains the same. The results also suggest that agent-supported negotiation might be less mentally demanding compared to setting the permissions manually and, at the same time, it better enables users to align their privacy choices with their actual preferences. Our findings provide insight into data sharing strategies adopted by our participants to guide the future design of automated and negotiable privacy management mechanisms.

The remainder of the paper is structured as follows. We first review related work in Sect. 2. In Sect. 3, we propose the negotiation framework with a novel protocol for bilateral multi-issue negotiation. Then, Sect. 4 we demonstrate how the general framework can be implemented to develop an agent that represents the an individual. After that, in Sect. 5, we apply the framework and its proposed implementation to the privacy permission negotiation domain. Next, Sect. 6 describes the apparatus, methodology and results of the experimental evaluation. Furthermore, we discuss the implications of our results for future design choices for privacy management and automated negotiation in Sect. 7. Finally, Sect. 8 concludes with key contributions and a future direction for this work.

## 2 Related work

In this section, we expand the previous discussion by examining the existing literature related to automated privacy negotiations. Specifically, we start with a brief overview of the key papers in the vast area of privacy management. Then, we provide a brief overview of the literature in automated negotiation. Following this, we discuss more specific work that has looked at automated negotiation in the context of privacy. Lastly, we briefly summarize the recent developments in learning users' privacy preferences and preference elicitation methods used in automated negotiation.

### 2.1 Privacy management

Manifested in various legal frameworks around the world, data protection law provides online users with a set of rights to enable them to control their personal data. As advancements in computing made it much easier to collect, store, assemble, correlate, and use

---

<sup>4</sup> The initial study results were published in: [6].

information, there has been an emergence of initiatives and tools whose aim is to aid users in managing their privacy. Early efforts include the Platform for Privacy Preferences (P3P) project developed by the World Wide Web Consortium (W3C), which aimed to enable machine-readable privacy policies [29]. With P3P, such privacy policies could be automatically retrieved by Web browsers that could prompt users or take other appropriate actions. Some of the P3P user agents were also able to compare each policy against the user's privacy preferences and assist the user in deciding when to allow sharing their personal data with websites [29]. However, the binary nature of the choice offered to the users has been concerning. In practice, users could either accept service providers' unlimited use of their personal data or give up using the service [83, 103, 110]. This failure to offer the user any real choices has been described as the *take-it-or-leave-it* approach to privacy [12, 90].

To address this problem, many fine-grained solutions have been proposed for users to communicate their privacy preferences to the service provider (e.g. [2, 15, 16, 31, 42, 49, 75, 80, 85, 119]). In fact, there is evidence that when given a choice, users prefer to exercise fine-grained control and select the data they are happy to share, rather than refuse to share any data at all. Specifically, in one study, participants were able to choose the option of erasing their personal data from all optional form fields, which would save them a lot of time [63]. Regardless, most of them preferred to go from question to question and decide whether to share each data item, even though it required more effort on their side. While the reasons behind these choices varied (e.g. some participants simply enjoyed disclosing the information), some participants did so hoping for some benefit in return. Moreover, another study provides evidence that optional disclosure delivers a 'good data return' for service providers [93]. This result suggests that as the society becomes more privacy-knowledgeable, fine-grained data sharing options may not be just preferable by users, but also beneficial to service providers.

However, as data is becoming a valuable resource for service providers, experts suggest that users should be compensated for providing access to their personal data [104]. While users differ in what they consider fair use of their data, overstepping their individual boundaries risks violating their trust, which may have implications for the service provider's brand [104]. Thus, in this paper, we explore questions about the viability of an exchange between user and service that goes beyond communicating fine-grained privacy preferences: we investigate the use of automated negotiation as a sustainable, win-win solution that supports a meaningful dialog between the user and the service provider.

## 2.2 Automated negotiation

Negotiation is about a joint exploration of outcomes in search for mutual gains. In doing so, the ultimate goal of negotiation is to resolve the conflict of interest present among different parties. Since negotiation covers so many aspects of people's lives this has led to an increasing focus on the design of automated negotiators, i.e., autonomous agents capable of negotiating with other agents in a specific environment [48, 61].

There is a significant body of literature that deals with automated negotiation. This interest has been growing since the beginning of the 1980s with the work of early adopters such as Smith's Contract Net Protocol [105], Sycara's persuader [111, 112], Robinson's oz [95], as well as the work by Rosenschein [96] and Klein & Lu [57]. Broadly, there are two approaches: those using protocols based on the alternating offers approach where only offers are exchanged, and argumentation-based approaches where additional information is conveyed in an attempt to convince their counter part [52]. In the former the agent

preferences are typically modelled using utility theory and techniques such as game theory and decision theory are applied, whereas the latter is mostly based on logical inference. Our approach in this paper is based on alternating offers.

Specifically, the alternating offers protocol [86, 99] is the best-known and most widely studied model for bargaining [35]. Following this protocol, offers are exchanged between two agents over a sequence of rounds. When one agent (the proposer) submits an offer, the other one (the responder) can either accept it or reject it. If the offer is rejected, the responder proposes an alternative offer, which the proposer can accept or reject. The negotiation continues until one of the following conditions is satisfied: an offer is accepted, a deadline is reached or one of the parties terminates the negotiation. While there also exist other negotiation protocols such as the monotonic concession protocol [97], one of the main advantages of following this approach is that an abundance of agents have already been formulated for the alternating offers protocol, e.g. [4, 8, 25, 37, 39, 44, 53, 61, 122], which could be easily adapted to our model.

More specifically, the protocol we propose in this paper is based on the multi-issue bilateral alternating-offers protocol where two agents negotiate offers with regard to not just a single item or a single bundle of items, but on many issues. In general, there are two approaches to such negotiations: one way is to negotiate all the issues together and the other is to negotiate them one by one (issue-by-issue) [36]. As users' privacy preferences are influenced by the context [82, 83], in our work, we negotiate all the issues together.

### 2.3 Negotiation in the context of privacy

To date, various techniques have been used to automate privacy negotiations, including rule-based reasoning [106], game theory [125, 126], question-answer-based profiling [43] and learning through historical negotiations [51]. Ontologies have also been commonly used to model privacy requirements [46, 47, 59, 106] and to measure the privacy sensitivity of different pieces of data based on relationships to each other [46].

In general, privacy negotiations can be divided into horizontal, where users negotiate among each other, and vertical, where the negotiation is conducted between a user and a service provider [62]. In the context of horizontal privacy, automated negotiations were proposed to resolve privacy conflicts among users. For example, to avoid privacy violations among users posting on social networks, automated negotiation frameworks and protocols were developed to help users manage the semantic rules of their privacy agreements with others [76]. Furthermore, a personal assistant was developed to help end-users with managing the privacy of their content by employing an auction-based mechanism [117]. Another work compared automated negotiation to users' behaviour in privacy conflict situations, proving that negotiation can help to reduce the amount of manual user interventions to achieve a satisfactory solution for social media users involved in multi-party privacy conflicts [107, 109]. In addition, an argumentation approach was proposed for user agents to argue with each other on the privacy rules by generating facts and assumptions from their ontologies [55, 59, 60].

For vertical privacy, several studies based their approach on P3P. Although P3P itself lacks a negotiation mechanism, researchers utilised P3P enhancements to enable privacy policy negotiations between the user and the service provider [14, 26, 51, 71, 72, 92]. For example [26] proposed a model for automatic privacy policy conflict detection and resolution, using the eXtensible Access Control Markup Language (XACML) as a policy

description language; [92] extended P3P with a negotiation process modelled as a Bayesian game where the service provider faces users of different privacy types.

Closest to our work, [126] proposed an intelligent agent-based system to quantify and measure privacy payoff through private data valuation and privacy risks quantification. Their system includes five different agents, including one responsible for negotiation with the service provider. However, the agent works on behalf of a list of users whose privacy preferences need to be exactly specified by the users themselves prior to the negotiation. However, expecting users to spend the time to fully configure their privacy preference settings has been shown to be unrealistic and difficult even for willing users [100]. Therefore, this approach is not suitable for our requirement of utilising automated negotiation to reduce the users' effort. Instead, we propose a preference elicitation method for the agent to learn the user's privacy preferences automatically.

## 2.4 Privacy preference elicitation

An important part of our work is modelling and learning the user's privacy preferences. A number of studies have looked at these issues. In particular, early work has focused on modelling users' location sharing preferences [13, 30, 94, 100]. For instance, personas and incremental suggestions were used to learn users' location privacy rules, resulting in users sharing significantly more without a substantial difference in comfort [56, 79, 123]. However, as users have different sensitivity to different types of personal data and their willingness to provide information varies depending on the perceived risk of sharing the particular information [77], location sharing preferences do not easily translate to other types of personal data. Thus, a more general approach is needed.

To understand users' privacy preferences for other types of data, researchers surveyed users' expectations and subjective feelings about different kinds of sensitive data, and identified key factors that influence them [67]. Then, clustering techniques were used to define a set of *privacy profiles* based on users' self-reported privacy preferences [68, 70]. However, privacy profiles do not take individual user differences into account. To address this issue, in our study, we compare the profile-based approach to one where the prediction outcomes are based on the behaviour of an individual user.

In general, while research has focused on opponent preference modelling, costly user preference elicitation in automated negotiation has received little attention [7, 18]. The work by Chajewska et al. [23, 24] provides one of the first starting points to this problem, but their solution method often deals with computationally intractable decision procedures. Costly preference elicitation has also been studied in the setting of auctions; notably by Conen and Sandholm [28] and Parkes [87]. These works are primarily aimed at designing *mechanisms* that can avoid unnecessary elicitation. Costly preference elicitation may alternatively be cast as a problem in which agents have to allocate costly resources to compute their valuation [65], but this type of work focuses on interactions *between* different strategies.

A large family of strategies proposed for user preference modelling are the UTA methods (UTilités Additives) which obtain a ranked set of outcomes as input and formulate a linear utility function through the use of linear programming [38, 45, 98, 118]. CP-nets, which provide a qualitative representation of preferences that reflects conditional dependence [19], were also studied in the automated negotiation context [7, 78]. Furthermore, Tsimpoukis et al. proposed a decision model that uses linear optimization to translate partial information into utility estimates based on a set of ranked outcomes [116]. However,



as we explain later in Sect. 4.2, when the input data stems from (possibly erroneous) information from human interaction, the problem quickly becomes over-constrained and linear constraint solvers can no longer be applied. Instead, in this paper, we propose a preference elicitation method based on approximation techniques studied by Popescu [91] to find the most preferable outcomes. Another related area is discrete choice modelling [73, 114]. Similar to our approach, the aim is to estimate a utility function based on limited pair-wise comparisons of alternatives. However, while we assume a deterministic utility function, the discrete choice approach assumes unobserved variables which are modelled as probabilistic noise, and the solution is found through logistic regression. In addition, it assumes decision makers are rational, i.e. the choices are consistent (e.g., if  $A$  is preferred to  $B$ , and  $B$  is preferred to  $C$  then  $C$  cannot be preferred to  $A$ ). In contrast, our approach also works with irrational choices. An empirical comparison of the approaches would be interesting but is beyond the scope of this paper and left for future work.

Specifically, we build upon previous work on a user agent for an automated negotiation of privacy permissions [6]. To this end, the agent was able to model the permission preferences of different user profiles and apply the right model when it represented a user assigned to one of these profiles. Conversely, in this paper, we propose a new user preference elicitation method, personalised to the needs of an individual user. To compare the two approaches, we present results of a new user study and re-analyse the data collected during the study published in [6].

### 3 Negotiation framework

In this paper, we propose a negotiation framework with a novel protocol that we refer to as the *partial-complete offer protocol*, which is a variant of the multi-issue alternating-offers protocol. In this section, we first motivate the development of this framework. Then, the negotiation setting is introduced. After that, we describe the partial-complete offer protocol, where two parties propose offers consisting of values of the negotiable issues. Finally, we explain the optimisation problem that needs to be solved in order for the negotiation to conclude.

#### 3.1 Motivation

In the classic multi-issue alternating offer protocol, offers are exchanged between two parties that specify values for each of the negotiable issues. As detailed in Sect. 2.2, the proposer submits a fully specified offer, which the responder can either accept or reject. If the offer is rejected, the responder proposes an alternative offer, which the proposer can accept or reject, and the negotiation continues until an offer is accepted, a deadline is reached or one of the parties terminates the negotiation.

However, in practice, parties typically have asymmetric roles such as buyer and seller. Often, a party can find it difficult to determine the value of some issues, which are more naturally determined by the counter party. For example, when negotiating a software development contract with negotiable issues such as the functionality requirements, maintenance level, hardware infrastructure and delivery timeframe, it is more natural for the customer to solely specify the functionality requirements, whereas the developer may choose how much infrastructure they can make available for the project given the requirements. In a similar way, during a negotiation of an insurance policy, it is usually the buyer who



specifies the conditions for the extent of the cover – the seller then completes the contract by proposing a price.

In particular, the ability of a party to leave some issues unspecified is important for privacy negotiations. In such negotiations, some essential issues need to be controlled by the user, e.g. what kind of data access the user is happy to consent to, while others must stay under control of the service provider, e.g. what reward or extra functionality can be offered in exchange for access to the data. Thus, the user may want to initiate the negotiation, specifying the requirements for a subset of important issues (e.g. *I am happy to share my contacts, but not my browsing history; what will I get for sharing this?*). The service provider is then able to submit a complete counter-offer, based on the proposed offer (e.g. *without your browsing history, I miss out on ad revenue; given that, I can offer you £2.99*). The user can either accept the proposal or submit a new enquiry (e.g. *What if I choose to share my contacts and location but not my browsing history?*).

To model this kind of automated negotiation, our premise is that, before the negotiation takes place, the proposer and the responder prescribe which issues they specify values for. This makes sure that, during the process, essential issues are under control of the party exchanging them, whereas others are left open. We do so by introducing *partial offers*, which we define as part of the negotiation setting.

### 3.2 Negotiation setting

In this paper, we focus on bilateral multi-issue negotiations. Specifically, a negotiation domain is specified by  $m$  issues, where  $m \in \mathbb{N}$  and  $m > 1$ . To reach an agreement, the agents must settle on an outcome that is accepted by both parties of the form  $\omega = (\omega_1, \dots, \omega_m)$ , where  $\omega_i$  denotes a value associated with the  $i$ -th negotiable issue for  $i \in \{1, \dots, m\}$ . Then,  $\Omega = \prod_{i=1}^m \Omega_i$ , where  $\Omega_i$  represents the set of possible values of  $\omega_i$ . For example, in the context of privacy permissions, if the issues  $\Omega_i$  correspond to *Shared data*, *Purpose of sharing*, *Retention policy*, and *Price discount*, then an example agreement in  $\omega \in \Omega$  is (*GPS location, Targeted ads, Shared with third parties only, £0.20*).

In this framework, we allow the value of an issue to be undefined. We denote so by a special character  $\perp \in \Omega_i$  for each  $i \in \{1, \dots, m\}$ . Moreover, we define a subset  $S \subseteq \Omega$  containing all offers where all values are defined. Formally, for each offer  $\omega \in S$ ,  $\omega_i \neq \perp$  for any  $i \in \{1, \dots, m\}$ . Because of this property, the elements of  $S$  are called the *complete offers*. On the contrary, the elements of the complement of  $S$ ,  $\bar{S} \subseteq \Omega$ , are called the *partial offers*.

While the domain is common knowledge, the preferences of each agent are its private information. Therefore, in addition to its own preferences, every agent also has an *opponent model*, which is an abstract description of the opponent's preferences. This allows the agent to employ a *negotiation strategy* to determine its optimal action in a given state of the negotiation.

Since we focus on bilateral negotiation, there are two negotiating agents involved. We call them the *proposer* and the *responder*, and present the protocol that dictates their moves.

### 3.3 Negotiation protocol

In the Partial-Complete Offer Protocol, the proposer submits a partial offer specifying the requirements for a subset of issues. Formally, the proposer offers  $\omega^p = (\omega_1^p, \dots, \omega_m^p)$

such that there exists (possibly more than one)  $i \in \{1, \dots, m\}$  for which the value remains unspecified, i.e.  $\omega_i^p = \perp$ . The responder is then able to complete the offer, taking into account the proposed partial offer. That is, the responder replies with a complete offer  $\omega^r = (\omega_1^r, \dots, \omega_m^r)$  such that  $\omega_i^r = \omega_i^p$  for all  $\omega_i^p \neq \perp$ . This ends the first *negotiation round*. From here, the proposer can either accept a complete offer and end the negotiation, reject it and submit a new partial one starting the next negotiation round, or break off the negotiation.

As the negotiation continues, we assume that previous offers remain valid. That is, the complete offers returned by the responder generate a growing set  $Q \subseteq S$  of possible outcomes that the proposer can agree to. Then, the negotiation ends when the proposer either accepts an offer  $\omega \in Q$  or actively ends the negotiation by signalling a break off.

**Protocol 1** Partial-Complete Offer Protocol

*Inputs.* The number of issues  $m$ ; for each  $i \in \{1, \dots, m\}$ , a set of values  $\Omega_i$  associated with the  $i$ th issue; the set of complete offers  $S$ ; the set of partial offers  $\bar{S}$ ; the cost per  $n$  rounds for the proposer  $c^p(n) \in \mathbb{R}$  and the responder  $c^r(n) \in \mathbb{R}$ .

*Goal.* The proposer and the responder agree on the outcome of the negotiation.

*Initialisation.* The number of completed rounds is  $n = 0$ .

*The protocol:*

1. The proposer submits a partial offer  $\omega^p = (\omega_1^p, \dots, \omega_m^p) \in \bar{S}$  such that there exists  $i \in \{1, \dots, m\}$  for which  $\omega_i^p = \perp$ .
2. The responder replies with a complete offer  $\omega^r = (\omega_1^r, \dots, \omega_m^r) \in S$ , where  $\omega_i^r = \omega_i^p$  for all  $\omega_i^p \neq \perp$ .
3. The number of completed rounds is incremented:  $n = n + 1$ .
4. The proposer performs one of the following actions:
  - (a) rejects  $\omega^r \rightarrow$  go to 1.
  - (b) accepts  $\omega^r \rightarrow$  the negotiation ends with outcome  $\omega^r$ , and with costs  $c^p(n)$  for the proposer and  $c^r(n)$  for the responder.
  - (c) breaks off the negotiation  $\rightarrow$  the negotiation ends with no outcome, and with costs  $c^p(n)$  for the proposer and  $c^r(n)$  for the responder.

In addition, negotiations often involve some form of time pressure to ensure that they finish in a timely manner. In order to avoid the proposer exploring all possible partial offers, the proposer incurs a bargaining cost  $c^p(n) \in \mathbb{R}$  where  $n$  is the total number of negotiation rounds. Similarly, to make sure that the responder’s offers are more likely to be accepted by the proposer, additional bargaining cost  $c^r(n) \in \mathbb{R}$  is levied on the responder.

To summarise, this the step-by-step process is presented in Protocol 1. Additionally, the protocol is illustrated in Fig. 1 as a sequence diagram. Next, we explain how the negotiation outcome and costs affect the utility of each of the negotiating agents.

**3.4 Utility**

When negotiation ends, the utility of each agent is updated. This depends on the cost the agent incurred and, if an offer is agreed, the agent’s valuation of the offer. Furthermore, each agent has a *reservation value*, which is the utility of a disagreement.

Specifically, the negotiation either ends with an outcome  $\omega \in Q$  or with no outcome. Then, if the proposer’s valuation is defined by a valuation function  $v^p : \Omega \rightarrow [0, 1]$

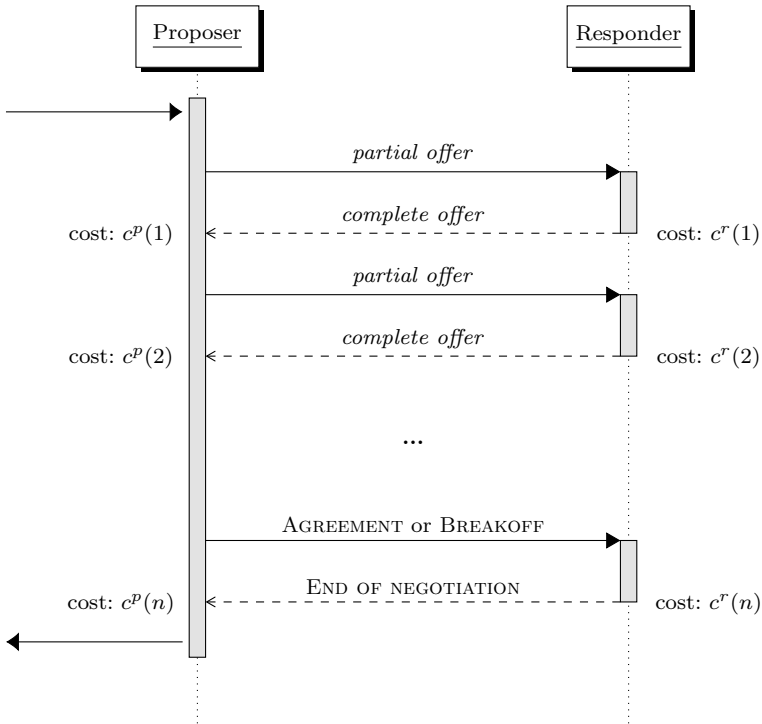


Fig. 1 Sequence diagram of a negotiation that ended after  $n$  rounds

and the proposer’s reservation value is  $r^p \in [0, 1]$ , the utility of the proposer when the negotiation stops is given by:

$$U^p(Q) = \begin{cases} v^p(\omega) - c^p(|Q|) & \text{if } \omega \in Q \text{ is the agreed outcome,} \\ r^p - c^p(|Q|) & \text{if no agreement is reached.} \end{cases} \tag{1}$$

Consequently, if the responder’s valuation is encoded by  $v^r : \Omega \rightarrow [0, 1]$  and the responder’s reservation value is  $r^r \in [0, 1]$ , then the utility of the responder is defined as:

$$U^r(Q) = \begin{cases} v^r(\omega) - c^r(|Q|) & \text{if } \omega \in Q \text{ is the agreed outcome,} \\ r^r - c^r(|Q|) & \text{if no agreement is reached.} \end{cases} \tag{2}$$

Given that the valuation function of one agent is not necessarily know to the other one, each of the agents aims to maximise their expectation over utility. Thus, the challenge the agent faces lies in employing an adequate negotiation strategy to determine the optimal sequence of actions. Furthermore, in some situations where human users are involved, the exact valuation function may not necessarily be defined for an agent due to the uncertainty over the user’s preferences. To address this, we show how such negotiations can be modelled using this framework.

## 4 Negotiation agent with preference uncertainty

In this section, we propose novel preference elicitation and negotiation strategies to demonstrate how the framework can be implemented to account for situations where a user (or a customer) is negotiating with a service provider. Specifically, we consider a setting, where the proposer is negotiating on behalf of a user who is unwilling or unable to fully specify the valuation function. This is particularly the case in many realistic scenarios, where collecting the necessary preference information to define the valuation function is time-consuming or costly, or communicating them is difficult for the user. To address the challenge that preference uncertainty brings into the negotiation, we first discuss how the agent builds the user and the opponent models. Second, we propose a new method of preference elicitation which allows the agent to learn the user's valuation. Finally, we propose a negotiation strategy which allows the agent to decide what offer, if any, to send to the opponent and when to terminate the negotiation.

### 4.1 Models of uncertainty

In order for an agent to faithfully represent a user, it is important that the agent's offers are aligned with the user's preferences. In our setup, we assume that the proposer agent first negotiates the offer with the responder (its opponent) on behalf of the user, and then proactively interacts with the user to establish whether the negotiated offer aligns with their preferences. That is, as opposed to asking the user directly beforehand, the proposer agent requests from the user feedback on the negotiation when it ends, and derives the user's preferences from the feedback. Specifically, in our setup, the user is presented with the complete offer negotiated by the agent and has an opportunity to communicate whether they approve the decision made on their behalf. This way, the agent is able to incrementally collect information on user's preferences for future negotiations at the most relevant time, while constantly keeping the user in the negotiation loop, as explained in detail in Sect. 5.3. Similarly, information on the opponent's valuation can be updated through the offers that are exchanged with the opponent.

To this end, we can identify two kinds of preference uncertainty in this negotiation:

- the user's preferences regarding the possible negotiation outcomes, and
- the opponent's reactions to the offers.

In order to model this uncertainty, the agent builds a *user model* and an *opponent model*. Specifically, the user model consists of the agent's beliefs about the user's preferences. This can be elicited from the agent's interactions with the user. Conversely, the opponent model, which reflects the agent's beliefs about the opponent's reactions, depends on the negotiation strategy used by the service provider. This model could be constructed from prior knowledge or previous interactions with the opponent, and can be based on the relative likelihood of the counter-proposal from the service provider (for an overview, see [9, 10]). This model of uncertainty has been illustrated in Fig. 2.

### 4.2 Preference elicitation

In a negotiation, a user has a specific set of preferences regarding the possible outcomes. The so-called *preference profile* is given by an ordinal ranking over the set of the outcomes: an

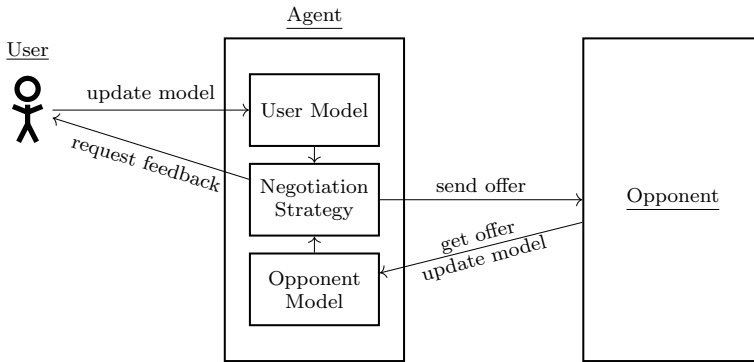


Fig. 2 The interaction model between the User, the Agent and the Opponent

outcome  $\omega$  is said to be weakly preferred over an outcome  $\omega'$  if  $\omega \succeq \omega'$  where  $\omega, \omega' \in \Omega$  or strictly preferred if  $\omega > \omega'$  [116]. Given the outcome ranking, the agent’s goal is to formulate its estimate of the valuation function  $v^p$  that approximates the real user’s valuation as much as possible so that the preferences are expressed in a cardinal way:

$$\omega \succeq \omega' \iff v^p(\omega) \geq v^p(\omega'). \tag{3}$$

Following the literature on multi-issue negotiation [54], we make the common assumption that the agent’s valuation function is linearly additive. That is, it has the following form:

$$v^p(\omega) = \omega_1 w_1 + \dots + \omega_m w_m, \tag{4}$$

where  $w_i$  is a weight indicating the importance of the  $i$ -th negotiable issue, and the weights are normalized such that:

$$\sum_{i=1}^m w_i = 1. \tag{5}$$

Therefore, deriving the valuation for the issues specified in the partial offer means deriving the weights  $w_1, \dots, w_m$ . This can be performed with feedback from the human user on the previous negotiation outcomes as explained in detail in Sect. 5. That is, if a user previously disliked a negotiated outcome  $\omega$  (which could include granting access to some personal data or sharing nothing) in favour of approving an outcome  $\omega'$ , we can assume that  $v^p(\omega') \geq v^p(\omega)$ . This can be written as:

$$(\omega'_1 - \omega_1)w_1 + \dots + (\omega'_m - \omega_m)w_m \leq v^p(\omega') - v^p(\omega). \tag{6}$$

If we do this for all previously approved and disapproved complete offers, we obtain a set of inequalities, from which we can deduce the most appropriate overall weights  $w_1, \dots, w_m$ . To this end, note that this procedure transposes the problem into a set of linear inequalities of the form:

$$Aw \leq b, \tag{7}$$

where the entries of  $A$  and  $b$  correspond, for each equation, to the values of  $(\omega'_i - \omega_i) \in \{-1, 0, 1\}$  and  $b = v^p(\omega') - v^p(\omega)$  respectively. We refer to these combinations of  $A$  and  $b$  as *constraints*.

However, as this data stems from human interaction, the problem quickly becomes over-constrained – that is, these inequalities are typically not consistent and any complete assignment violates some constraint. Therefore, we cannot simply use standard linear constraint solvers such as [116] for negotiation elicitation in practice. To address this, instead, we find a solution that best satisfies the constraints following the techniques described in [91]. Specifically, we determine the weights  $w^*$  that minimize the least squares norm:

$$w^* = \arg \min_w \|(Aw - b)_+\|^2, \tag{8}$$

where  $(Aw - b)_+$  is the vector whose  $i$ -th component equals  $\max\{(Aw - b)_i, 0\}$ .

The weights  $w^*$  can be then plugged to Eq. 4 to derive the valuation function, which the negotiation strategy relies on to select the most preferable offers.

### 4.3 Negotiation strategy

Using the valuation function, the negotiation strategy needs to determine which of the partial offers, if any, the agent should propose to the opponent. What makes this problem non-trivial is its sequential nature: whether or not to propose a partial offer depends on the offers proposed so far. Therefore, the goal is to find an optimal *sequence* of offers to propose and a corresponding strategy which specifies when to conclude the negotiation process. To find this optimal negotiation strategy, we propose a similar approach to the one used in related work for preference elicitation (cf. [7]).

To evaluate partial offers, we assume that the agent has a model of the likelihood of receiving a certain complete offer from the opponent. That is, the probability of a complete offer given a partial offer  $\omega_p$  is given by a stochastic variable  $X_{\omega_p}$ , with a cumulative distribution function  $G_{\omega_p}(x)$  known to the agent. From this, the expected value of a partial offer can be derived. Specifically, the valuation of a complete offer returned by the opponent is described by a stochastic variable  $Y_{\omega_p} = v^p(X_{\omega_p}) \sim [0, 1]$  with a corresponding cumulative distribution function  $H_{\omega_p}(y)$ . Additionally, we assume that the total cost of the negotiation for the agent after  $n$  rounds is defined as  $c^p(n) = C^p n$ .

With these assumptions, our goal is to formulate a negotiation policy  $\pi$  which, given a state  $\mathcal{E}$ , decides whether to continue or stop the negotiation. Since the costs of offer proposals are sunk, it is easily verified that the negotiation strategy depends on the offers proposed,  $P \subseteq \bar{S}$ , and the offers received,  $Q \subseteq S$  by the agent so far. Thus, the current negotiation state can be summarised by  $\mathcal{E} = \langle P, Q \rangle$ .

Given the state  $\mathcal{E}$ , the negotiation policy  $\pi$  should either stop the negotiation (with an outcome in  $Q$  or with no outcome) and obtain utility  $U^p(Q)$  (defined in Eq. 1), or continue the negotiation by proposing a new partial offer  $\omega^p \in \{\bar{S} \setminus P\}$  at cost  $C^p$ . If  $\omega^p$  is proposed and a new complete offer  $\omega^r \in S$  received, the negotiation enters a new state which can be described as  $\mathcal{E}' = \langle P \cup \{\omega^p\}, Q \cup \{\omega^r\} \rangle$ . Since at that point the future utility in state  $\mathcal{E}'$  cannot be observed, the expectation over utility can be represented by the expected value  $\mathbb{E}_{\omega^r|\omega^p} \{U(\pi, \mathcal{E}')\}$ .

Therefore, the utility of a policy  $\pi$  given the state  $\mathcal{E}$  can be computed as follows:

$$U(\pi, \mathcal{E}) = \begin{cases} U^p(Q) & \text{if the negotiation stops,} \\ \mathbb{E}_{\omega^r|\omega^p} \{U(\pi, \mathcal{E}')\} & \text{otherwise.} \end{cases} \tag{9}$$

Now, given the state  $\mathcal{E}$ , we are looking for the optimal negotiation policy  $\pi^* = \arg \max_{\pi} U(\pi, \mathcal{E})$ . Note that when all offers are observed,  $U(\pi^*, \langle \bar{S}, S \rangle) = U^p(S)$ ; otherwise, the agent may choose to propose one or more partial offers  $\omega^p \in \{\bar{S} \setminus P\}$ . Given this, the optimal negotiation policy  $\pi^*$  should consider the following: either stop the negotiation and obtain  $U^p(Q)$ , or propose a new partial offer  $\omega^p \in \{\bar{S} \setminus P\}$  which maximises the expected value.

More formally,  $U(\pi^*, \mathcal{E})$  must satisfy the following recursive relation:

$$U(\pi^*, \mathcal{E}) = \max \left\{ U^p(Q), \max_{\omega^p \in \bar{S} \setminus P} \left\{ \mathbb{E}_{\omega^r|\omega^p} \{U(\pi, \mathcal{E}')\} \right\} \right\}. \tag{10}$$

The relation for  $\pi^*$  given in Eq. 10 is essentially a Bellman equation which, in principle, could be solved by backward induction. However, even for a moderate-size negotiation space, this approach quickly becomes intractable. Instead, we use a simple index-based alternative method to decide which partial offers to propose and whether to break off the negotiation.

Specifically, the negotiation strategy of the agent can be mapped onto a variant of the Pandora’s Problem [120]: a search problem involving boxes that contain a stochastic reward. As such, each partial offer  $\omega_p \in \{\bar{S} \setminus P\}$  can be regarded as a *closed* box with stochastic reward  $Y_{\omega_p}$  that can be opened at cost  $C^p$ , while every partial offer  $\omega_p \in P$  can be represented by an *open* box with a known reward  $v(\omega_r)$ , where  $\omega_r$  is the complete offer observed after proposing  $\omega_p$ . As a consequence of Pandora’s Rule [120], we can assign an index  $z_{\omega_p}$  for every partial offer  $\omega_p \in \{\bar{S} \setminus P\}$ , satisfying:

$$\int_{y=z_{\omega_p}}^1 (y - z_{\omega_p}) dH_{\omega_p}(y) = C^p. \tag{11}$$

After identifying the offer  $\omega_p^* \in \{\bar{S} \setminus P\}$  with the highest index  $z_{\omega_p}^*$ , we apply the following negotiation strategy:

**SELECTION RULE:** *If an offer is proposed, it should be  $\omega_p^*$ .*

**STOPPING RULE:** *Terminate the negotiation whenever the reservation value  $r^p$  or the highest valuation  $\max_{\omega_r \in Q} (v^p(\omega_r))$  exceeds  $z_{\omega_p}^*$ . Choose the negotiation outcome as follows:*

- If  $\max_{\omega_r \in Q} (v^p(\omega_r)) > z_{\omega_p}^*$ , then  $\arg \max_{\omega_r \in Q} (v^p(\omega_r))$  is the outcome.
- If  $r^p > z_{\omega_p}^*$ , the negotiation ends with no outcome.

This negotiation strategy completely characterises the optimal policy  $\pi^*$ , as it has been proved to be optimal in terms of maximizing expected utility [120]. In practice, the effectiveness of the optimal quoting strategy depends on the accuracy of the valuation function and the opponent model. However, with a faithful model, the agent’s strategy is optimal in a non-myopic sense: it will negotiate taking into account not only the costs, but also the incremental effect of any subsequent rounds.



Note that the complexity of the elicitation strategy is  $O(n \log n)$ , since the index values do not need updating after our selection rule has chosen to elicit a particular offer [7]; hence, to find the next offer with the highest index, it suffices to order the set of indexes once, in  $O(n \log n)$  time, before applying the stopping rule. Our algorithm has the further desirable property that if it is costless to acquire information about an offer then the algorithm will always elicit it, provided it is not dominated by an offer previously elicited [7].

## 5 Negotiation of privacy

In this section, the theoretical framework described in Sect. 3 is applied to privacy negotiations, where the user's interests are represented by the proposer, and the service provider's interests are represented by the responder. First, we formally define the negotiation domain. Second, we provide an overview of how the negotiation of privacy permissions is performed with human users in the loop. Finally, since the utility of both agents is similarly based on the valuation function and therefore, on the preference profile, as a first step, we focus specifically on modelling the user's agent, and present two variants of how this preference profile can be created.

### 5.1 Negotiation domain

In privacy permission management, the negotiation domain may consist of a set of permissions to access user's device resources such as contacts or text messages, as well as other negotiable issues such as access to certain functionalities of the online service owned by the service provider (e.g. a website, a mobile app or software running on a device). For the purpose of our proof of concept, consider a case where the permissions to access resources are granted in exchange for a monetary reward. This monetary reward is the total value a user can receive for granting permissions to a specific set of resources, and we refer to it as a *quote*. We assume that the user's agent is in control of the values of issues related to the set of permissions and the service provider's agent controls the value of the quote. For instance, if the permissions grant access to the user's *Contacts*, *Messages*, *Apps List*, *Photo Gallery* and *Browsing History*, an example partial offer proposed by the user's agent may be: (*permission granted*, *no permission*, *permission granted*, *no permission*, *permission granted*,  $\perp$ ). The service provider's agent may complete the offer and respond with: (*permission granted*, *no permission*, *permission granted*, *no permission*, *permission granted*,  $\pounds 0.55$ ).

Formally, following the notation from Sect. 3, we define the negotiation domain such that  $\omega_i \in \{0, 1\}$  for  $i \in \{1, \dots, m-1\}$  is a permission with binary values 1 and 0 indicating whether the permission is granted or not. In addition,  $\omega_m$  is an issue representing the quote, where the possible values are continuous and normalized such that  $\omega_m \in [0, 1]$ . Hence, the user's agent proposes a partial offer providing values for  $\omega_1, \dots, \omega_{m-1}$ , whereas the service provider's agent responds with a complete offer, including a value for  $\omega_m$ . This way, the example offer returned by the service provider's agent can be represented as: (1, 0, 1, 0, 1, 0.55). If the user's agent refuses to grant any permissions, the complete offer would be (0, 0, 0, 0, 0, 0).

## 5.2 Overview of the negotiation

To demonstrate the applicability of the framework, we consider a specific bargaining situation where the user's agent takes the role of the proposer by sending partial offers to the service provider's agent. The offers indicate the permissions the user is willing to grant access to. Next, the service provider's agent returns a complete offer, specifying the quote the service provider can offer for the permissions proposed. Consequently, the user's agent negotiates the best possible offer (given the current user and opponent models) with the service provider's agent on behalf of the human, and then proactively interacts with the human user to establish whether the negotiated offer met the user's expectations.

Although the agent performs the negotiation autonomously, this framework allows us to keep the human user in control over data sharing. As discussed in detail in Sect. 6.2.2, in our setup, it is the human user who has the final say on whether permissions are granted and which ones by selecting *Don't Share* for individual permissions. This way of negotiating privacy permissions provides two benefits: the user has full control over their privacy and the agent can use the feedback for constructing the user's preference profile. However, in order for the agent to accurately represent the user, establishing an accurate preference profile is crucial. This is because the weights in the valuation function depend on the data present in the preference profile (see Eq. 4). To this end, we propose two ways of using users' feedback to derive the user's valuation. Since we implement them as two different variants of the user's agent, we refer to them as Agent 1 and Agent 2.

## 5.3 Agent 1: Individual preferences

Agent 1 recomputes the weights after each negotiation based on an individual user's decisions in the previous negotiations. At the beginning, to deal with the 'cold-start' problem (i.e. where no information is available about the past negotiations), the agent applies the weight derivation method on data collected from other users' negotiations. Specifically, all data from control treatments without automated negotiation<sup>5</sup> is aggregated to generate a set of constraints according to Inequality 6 which are, in turn, used to derive an initial set of weights through Eq. 8. Thus, when no information about an individual user is known, the initial weights  $w_i$  are set the same for all users.

Then, whenever a user rejects an offer  $\omega$  in favour of accepting  $\omega'$ , a new inequality of the form of Inequality 6 is added to the set of already existing inequalities. However, if there is a conflict between constraints, the user's own constraints are always prioritised over the ones from the manual negotiation trials. This means that the individual user's choice replaces the conflicting constraint. Then, the new weights are again derived and updated using Equation 8 for use in the next negotiation.

## 5.4 Agent 2: Type-based preferences

Agent 2 assigns users into categories before the start of their negotiations. We define these categories as follows:

- *Fundamentalists* – users generally granting less than 33% permissions.

<sup>5</sup> For details on the control treatments, see Sect. 6.3.1.

- *Pragmatists* – users generally granting between 33% and 66% permissions.
- *Unconcerned* – users generally to granting more than 66% permissions.

Initially, the preferences of a newly added user are based on the negotiation data from users previously classified into the same category. Specifically, the agent applies the weight derivation method on negotiation data from the control treatments without negotiation where participants were categorised in the same way (see Sect. 6.3.1) to determine the weights offline.

After each negotiation, the cluster classification is updated according to the percentage of permissions that were actually granted. Specifically, if it is less than 33%, they are re-classified as *Fundamentalists*, between 33% and 66% – *Pragmatists*, and *Unconcerned* otherwise. This way, even though the weights are learnt offline, the classification of the user is performed online.

## 6 Experimental evaluation

To evaluate the suitability of the agent-based negotiation framework for privacy permission management, as well as compare the performance of Agents 1 and 2, we developed an experimental platform in the form of a mobile application, which allows users to negotiate privacy permissions on their own smartphones. Following the approval of the study by the University of Southampton's Ethics Committee (ref.: ERGO/FPSE/18082), we recruited 132 participants, who used the tool to negotiate access permissions to their actual personal data in exchange for points converted to a monetary value at the end of the experiment. In this section, we provide a detailed description of the assumptions made about the service provider's agent, the apparatus, the methodology used and the findings of our experimental evaluation. In particular, we compare the negotiation approach to the take-it-or-leave-it approach often reflected in cookie banners, which we use as a benchmark for our experiments.

### 6.1 Experimental setup

In order to conduct the experimental evaluation of the user's agent, we abstract away from modelling the service provider's agent and its actual strategy. Instead, we design a stochastic way of completing the partial offers, making assumptions about the service provider. Specifically, we first assume that the order in which partial offers are made does not impact the complete offers, i.e. the complete offers are not different depending on the negotiation round. However, we assume that the complete offers depend on the number of granted permissions  $N = \sum_{i=1}^{m-1} \omega_i$ . That is, the agent completes every partial offer with a uniformly random quote  $\omega_m \in [\max(0, N - 1/m), N/m)$ . Using the previous example, if the partial offer received from the user's agent is: (1, 0, 1, 0, 1,  $\perp$ ), then some of the possible complete offers are: (1, 0, 1, 0, 1, 0.4), (1, 0, 1, 0, 1, 0.45) and (1, 0, 1, 0, 1, 0.55). This approach to completing the offers ensures that granting more permissions results in a higher quote, while different combinations of permissions result in a different (not necessarily linearly additive) quotes. This reflects situations where some data types could complement each other (i.e. when a specific combination allows the service provider to derive more relevant information) or substitute each other (in which case the added data provides little additional benefit to the service provider).

Furthermore, as in this paper we study the strategy of the user's agent, we set the cost of negotiation to  $c^p(|Q|) = 10 * |Q|$ . For the service provider's agent, we set the cost  $c^r = 0$ . In addition, for the purpose of this experiment, we set the reservation value for the user's agent and the service provider's agent to  $r^p = r^r = 0$ . Lastly, we assume that the service provider's valuation is equal to the quote of the accepted offer, i.e. for an outcome  $\omega = (\omega_1, \dots, \omega_m)$ , the valuation  $v^r(\omega) = \omega_m$ . Thus, the service provider's utility after negotiation is  $U^r(G) = \omega_m$  if  $\omega$  is the agreed outcome or 0 otherwise.

## 6.2 Apparatus

We developed a tool that allows participants to negotiate combinations of selected permissions on their smartphones in exchange for points, which map directly to a monetary reward. In the setting screen, the user can see which permissions are going to be granted, the number of points they would receive as a reward, and the total number of points collected. The more data a user shares, the more points they receive. In this section, we describe the tool and the ways users interact with it.

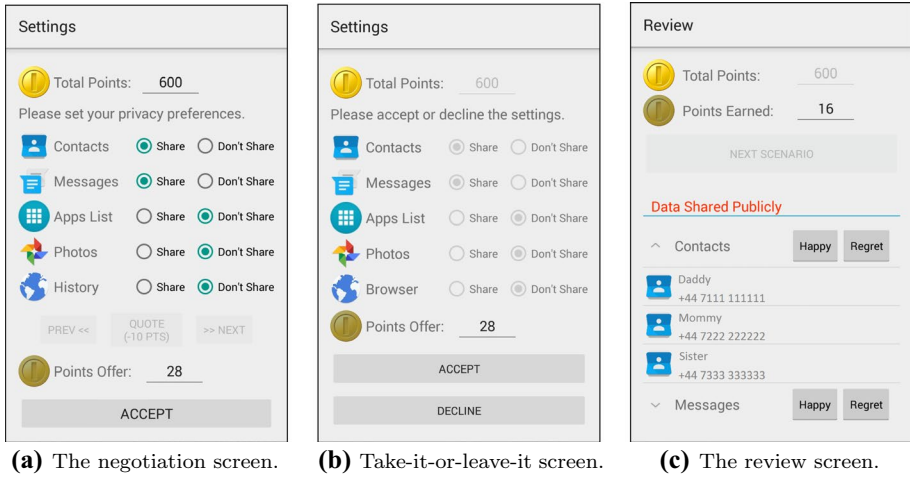
### 6.2.1 Permissions

In order to avoid any personal preference towards the service and to focus on users' permission preferences, the tool we developed has no intrinsic functionality other than capturing and displaying selected data points, and the testbed for negotiation. The advantage of using a mobile app as our experimental environment is that it allows us to use participants' own data that they personally care about in a real, privacy-sensitive situation. This level of realism is particularly important in privacy-related experiments, because past studies [11, 113] have consistently shown a discrepancy between a person's stated preferences from surveys and actual disclosure decisions (i.e. the so-called privacy paradox).

The app requests access permissions to the following data:

- the list of *contacts*,
- the text *messages*,
- the list of installed *apps*,
- the gallery of *photos*, and
- the browsing *history*.

The reason why we selected the above permissions is that they are among the most often used, as ranked by [66], and can be acquired, mined from users' smartphones and quantified. It is important to note that, although we decided to use the mobile app as our experimental environment, the aim is to test the negotiation framework in practice and to compare different privacy preference learning approaches. Hence, for simplicity of the interface, we selected only five permission types. We present two designs of this interface: the negotiation screen allows the user to request a new negotiation of the permissions and the "Take It Or Leave It" screen that represents the *take-it-or-leave-it* approach that is still used by many services requiring users to share data.



**Fig. 3** The interface of the experimental tool, which displayed the result of the negotiation or take-it-or-leave-it interaction during the user study, and allowed for a review of the user's decision

### 6.2.2 Setting screen: negotiation

On the negotiation screen, the user is presented with the outcome of the negotiation. When the negotiation between the user's agent and the service provider's agent ends, the agreed permission settings (*Share* or *Don't Share*) as well as the agreed quote (a number of points received in exchange) are displayed. If the user is happy with the offer, they can press the *Accept* button. In that case, the offered number of points is added to their total points. Otherwise, the user can communicate their permission preferences by selecting *Share* or *Don't Share* to grant or refuse granting a permission. An example configuration is presented in Fig. 3a, where a user is offered 28 points for access to their contacts and messages only, but can change these settings. By pressing the *Quote* button, the user can then request a new quote. To prevent the user from constantly doing so, 10 points are subtracted from their accumulated points every time they request the offer to be renegotiated. The user can also freely switch between the received offers using the *Prev* and *Next* buttons. Once they are happy with one of the offers, they can press *Accept* to approve it.

### 6.2.3 Setting screen: "Take It Or Leave It"

The *take-it-or-leave-it* approach represents the situation where the data sharing terms lack any kind of tunable control over the privacy trade-offs. For example, prior to Android 6.0 users were required to accept all data access permissions requested by mobile apps in order to proceed with the installation on their smartphones. Moreover, the take-it-or-leave-it approach is typical in many exchanges and not limited to apps. This approach is reflected in the differences between the negotiation screen and the "take it or leave it" one. Specifically, the "take it or leave it" screen does not allow the user to modify the permissions. The user can only accept or decline an offer by pressing the *Accept* or *Decline* buttons accordingly. The *Decline* option is equivalent to selecting *Don't Share* for all data types. For example,

**Table 1** A summary of the treatments used in the experimental evaluation

Treatment	Summary
TIOLI	<i>Take-it-or-leave-it</i> (no agent)
MN	Manual negotiation (no agent)
A1	Agent 1 negotiating
A2	Agent 2 negotiating

in Fig. 3b the user is able to accept or decline access to contacts and messages in return for 28 points.

### 6.2.4 Review screen

Following each sharing choice (through either negotiation or *take-it-or-leave-it*), the app collects three random data points of each shared permission type from the user's device. The user is then presented with a review screen showing those data points (which the users are made to believe is made public, as explained specifically in the experimental design, see Sect. 6.3.1), and asked to retrospectively express whether they are *Happy* about or *Regret* granting the permission. Fig. 3c presents an example review screen displaying three sample contacts and messages of the user.

Note that participants cannot revoke their decision at this point. The purpose of this screen to assess whether retrospectively the agent has made the right decision and how this compares to the *take-it-or-leave-it* approach. As we explain in detail in Sect. 6.4.3, aligning the agent's choices with the user's preferences is one of the success measures we use.

## 6.3 Methodology

A lab study is an experiment conducted in a controlled setting of a laboratory, as opposed to a real-world, natural setting. In order to allow for more control of the conditions, our experiments were conducted through lab studies. Participants were asked to use their own mobile phones, so that they truly cared about the data being shared. In this section, we report on the experimental design, the procedure of our lab studies and the participant recruitment.

### 6.3.1 Experimental design

Between-subject experimental design is a type of experimental design where participants of an experiment are assigned to different treatments, with each participant experiencing only one of the experimental conditions. In our pilot studies, we noticed a change in participants' behaviour over time, as they became more aware of the consequences of their sharing decisions through the review of the data. Thus, in order to avoid bias caused by the learning effect, we decided to employ between-subject experimental design with four treatments summarised in Table 1.

Firstly, our goal was to observe how a human negotiates without the help of an agent. For this reason, in the first treatment, Treatment Take-It-Or-Leave-It (TIOLI), we used the *take-it-or-leave-it* screen as a benchmark for our experiments. Secondly, we wanted to compare the *take-it-or-leave-it* approach to manual negotiation. We did so by introducing

**Table 2** A summary of the user study procedure as experienced by a participant

Part of experiment	Summary
Initial survey	Demographics (age, gender, nationality, university course) Westin's Privacy Segmentation Index
Main experiment	"Take it or leave it" or negotiation screen + review screen repeated eight times
Post-study questionnaire	Data sharing sensitivity survey NASA-TLX

a second treatment, Treatment Manual Negotiation (MN), which instead, used the negotiation screen. Finally, we aimed to test if the proposed designs of the agent-based negotiation can facilitate data sharing better than manual negotiation and to evaluate the performance of Agents 1 and 2, comparing them to the manual negotiation and to each other. Therefore, we developed Treatment Agent 1 (A1) and Treatment Agent 2 (A2).

### 6.3.2 Procedure

The lab study procedure involved four parts: the initial survey, the main experiment, a post-study questionnaire and a debrief. Table 2 summarises the user study procedure as it was experienced by the participant.

At the beginning, participants were asked to download the mobile app from the Play Store. They entered their demographics (age, gender, nationality, university course) into the app, and completed Westin's Privacy Segmentation Index survey [64]. Westin's Privacy Segmentation Index is a widely used tool for measuring privacy attitudes [124] and categorising individuals into three privacy types: Fundamentalists, Pragmatists, and Unconcerned. As part of the Index, they indicated on a 4-point Likert scale (1 – strongly disagree, 4 – strongly agree) the extent to which they agreed with three statements regarding control over personal information, personal information handling and legal privacy protection. The complete question set as defined by Westin's Privacy Segmentation Index reads as follows:

1. *Consumers have lost all control over how personal information is collected and used by companies.*
2. *Most businesses handle the personal information they collect about consumers in a proper and confidential way.*
3. *Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today.*

They were informed that their monetary reward will be based on the total number of points earned and (to elicit a genuine response) that any data shared during the experiment will be made available on a public website. They were not informed about the agent negotiation taking place in this experiment.

After reading the app manual, participants were asked to interact with either the negotiation or take-it-or-leave-it screen, depending on the treatment. In order to control the conditions, the offers for treatments with no agent negotiation (TIOLI and MN) were pre-defined using random sampling prior to the experiment. Once a participant accepted or



declined an offer, they proceeded to the review screen. As the experiment continued, these interactions were repeated such that each participant engaged in eight negotiation scenarios in total. To explore varying reward levels, those interactions differed in the maximum possible number of points to be gained by a participant: 25, 50 or 100. To cancel out possible interaction effects, this maximum reward was set to 50 in the first and last interaction for all participants, and a balanced Latin square design<sup>6</sup> was used to determine the order of maximum rewards in all others. Consequently, the number of points a participant gained during a single negotiation or a take-it-or-leave-it interaction was the maximum reward level multiplied by the negotiated quote  $\omega_m$  in the given scenario. In the end, the total number of points collected by a participant was the sum of points collected in all scenarios.

After that, they completed a questionnaire about their data sharing sensitivity. Specifically, they were asked to rate the following statements on a 7-point Likert scale (1 – strongly disagree, 7 – strongly agree):

1. *I am sensitive about sharing the contacts stored on my phone.*
2. *I am sensitive about sharing the text messages stored on my phone.*
3. *I am sensitive about sharing the apps stored on my phone.*
4. *I am sensitive about sharing the photos stored on my phone.*
5. *I am sensitive about sharing the browsing history stored on my phone.*

Additionally, participants were asked to complete NASA Task Load Index (NASA-TLX), which is a widely used, multidimensional assessment tool that rates perceived workload [41]. As part of this survey, they were asked to rate the effort they had to put into this experiment on a 20-point Likert scale.

Finally, the participants were debriefed about the purpose of the study specifically and informed that their data was never made publicly available on any website, despite the initial claim. All participants received a cash payment of between £5 and £10, directly depending on the number of points accumulated during the experiment, regardless of treatment allocation, e.g. if they collected 658 points, they received £6.60; if they collected less than 500 points, they received £5.

### 6.3.3 Participants

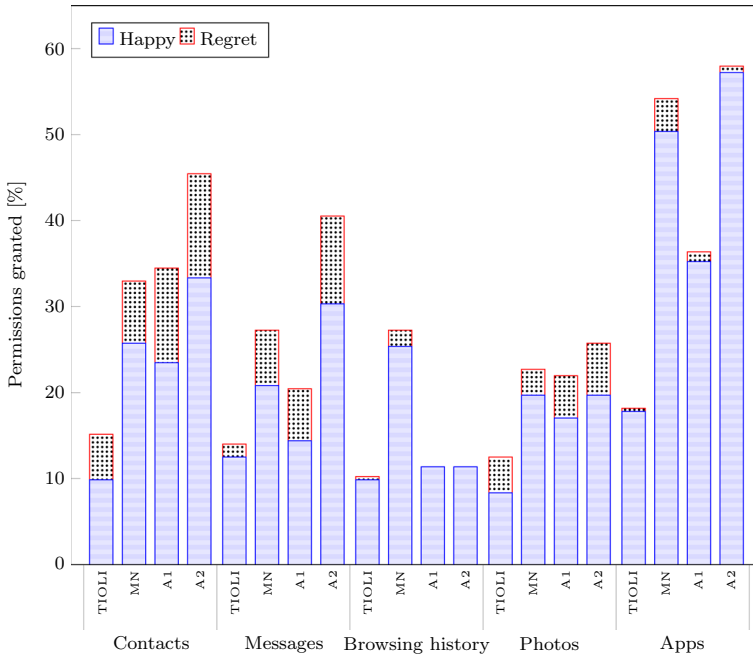
We recruited 132 participants from the University of Southampton. At the recruitment stage, they were informed that, during the experiment, they will be asked to download an Android application and make privacy-related decisions to earn between £5 and £10.

The participants were undergraduate, Master's or Ph.D. students from a variety of disciplines (e.g. Engineering, Medicine, Law). Since university students typically have a good level of digital literacy and a variety of attitudes towards privacy, this sample was suitable for the purpose of evaluating agent-based negotiation. 37.12% of them identified as women and 62.88% identified as men. 45.45% of the sample was British; others were nationals of 32 different countries such as Romania (7.58%), Malaysia (6.06%) and India (5.3%). Their age ranged from 18 to 43 (mean: 21.69, median: 21, st. dev.: 3.78). The participant poll consisted of 18.18% Fundamentalists, 75% Pragmatists and 6.82% Unconcerned (as

<sup>6</sup> Balanced Latin square design is a type of study design in which multiple conditions are administered to the same participant over time, and the order of the conditions varies among the participants such that the effects of extraneous sources of variation are balanced.

**Table 3** The number of participants of each privacy type per treatment

Privacy Type / Treatment	TIOLI	MN	A1	A2	Total
Fundamentalists	5	5	7	7	18.18%
Pragmatists	26	26	23	24	75%
Unconcerned	2	2	3	2	6.82%



**Fig. 4** The percentage of times the participants granted permissions to each of the resources

defined by Westin’s Privacy Segmentation Index) which is broadly consistent with the overall American population [64] – however attitudes to privacy may differ between the overall American population and the participants in our study.

The participants were randomly allocated into 4 treatments, i.e. there were 33 participants in each treatment. The allocations were non-significant such that any differences in privacy attitudes between the treatments were non-significant. To illustrate this, Table 3 presents the number of participants of each privacy type per treatment.

### 6.4 Results

In this section, we present the results of the user study. In particular, we report on the impact of automated negotiation on data sharing and user’s post-sharing regret, aligning users’ decisions with their self-reported data sharing sensitivity, the effort required from the users and the accuracy of the proposed agents.

### 6.4.1 Impact on data sharing

One aim of our research was to investigate how agent negotiation may influence users' data sharing behavior. The results, based on participants' own private data show that, on average, participants allowed access to data of the five data types over 2.5 times more often when they were able to negotiate. Figure 4 shows the percentages of how many times the participants allowed access to each of the data types in all scenarios. In particular, participants in treatment MN decided to share their list of installed applications 3.5 times more often than those in treatment TIOLI, and participants in treatment A2 shared their list of installed applications nearly four times more often than those in treatment TIOLI. The messages stored on the participants' mobile phones in treatment MN were shared almost twice more often than in treatment TIOLI. In treatment A2, messages were shared nearly three times more often than in treatment TIOLI. These findings suggest that negotiation leads to a win-win situation, for both the user who receives higher payoffs from sharing more data, and for our hypothetical service provider who receives more data from them.

### 6.4.2 Impact on regret

Based on the results, we can see that people are happily sharing certain kinds of data, and regret having shared others. However, the regret rate does not change when the negotiation agent is introduced. On average, the participants expressed regret after allowing access to their data in 15.96% cases. This is consistent with findings from related work [17], where in 10% of decisions, users were granting permissions reluctantly.

Figure 4 illustrates the percentage of times the participants were happy having shared their data of each type and how many times they regretted their decisions. The most shared resource was contacts (27.78%), the least shared one was the list of installed applications (3.55%). Nonetheless, there were no significant differences between the regret rates in the treatments. We consider this a positive outcome for the potential of automated negotiation in this area. That is, even though, as discussed in Sect. 6.4.1, users grant access to their data more often when a negotiation agent is involved, there is no significant increase in regret.

### 6.4.3 Aligning choices with privacy preferences

Our results show that, when users are allowed to negotiate, not only does their sharing behavior change radically, but their choices also better reflect their privacy preferences. For each treatment, Figure 5 shows the mean average of the users' self-reported data sharing sensitivity for granting each permission on 7-point Likert scale (see Sect. 6.3.2 for details). As expected, there are no statistical differences found between the treatments.

When we compare the reported sensitivity of each data type presented in Fig. 5 with users' sharing decisions in Fig. 4, we observe a more marked correspondence for the negotiation settings. Although the average scores hide a one-to-one correspondence between sensitivity scores and sharing actions, we can observe how often-shared data in these three treatments correspond with permissions of markedly lower sensitivity, such as apps, while access to photos and messages, which are both highly sensitive, is permitted far less often

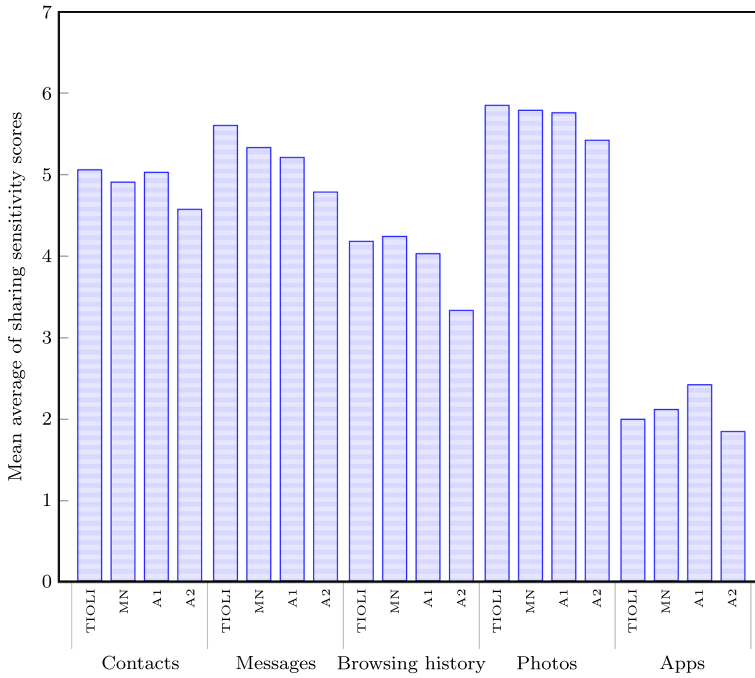


Fig. 5 Mean average of self-reported data sensitivity scores on a 7-point Likert scale

than other permissions. The only outlier in this order seems to be the browsing history, which we believe is due to the fact that a number of participants did not have any browsing history available<sup>7</sup>, independent of the sensitivity.

### 6.4.4 Accuracy of the negotiation agents

The last aim of our research was to examine the accuracy of the proposed negotiation agents. To do so, we calculated the accuracy by comparing the number of changes that the users made to the negotiated outcome.

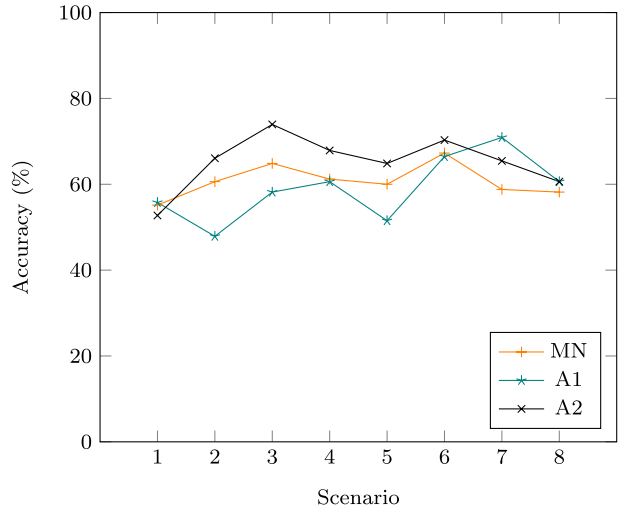
Specifically, we define the accuracy as the difference between the accepted offer and the offer negotiated by the agent. Building on the notation from Sect. 5.1, if the initial offer is  $\omega^1 = (\omega_1^1, \dots, \omega_{m-1}^1, \omega_m^1)$  and the final offer is  $\omega^k = (\omega_1^k, \dots, \omega_{m-1}^k, \omega_m^k)$ , the difference between them,  $\delta$ , is calculated as:

$$\delta = \sum_{i=1}^{m-1} |\omega_i^k - \omega_i^1|$$

where the first  $m - 1$  issues are the permissions (i.e., contacts, messages, browsing history, photos and list of applications installed) and the  $m^{th}$  issue is the number of points received

<sup>7</sup> This was verified by counting the number of items in the user’s browsing history.

**Fig. 6** The accuracy of Agent 1 (A1), Agent 2 (A2) and the manual negotiation (MN) in each scenario



**Table 4** Mean, median and standard deviation of the perceived effort of negotiation in each treatment, collected via the post-study NASA-TLX questionnaire

Effort / Treatment	TIOLI	MN	A1	A2
Mean	4.19	5.33	5.88	3.55
St. dev.	4.09	3.54	3.82	2.88
Median	3	5	5	3

in exchange for them, which is ignored here since this is set by the service provider and not the user. Hence, the *accuracy* of the negotiated outcome is calculated as:

$$1 - \frac{\delta}{m - 1}$$

Figure 6 presents the accuracy of Agent 1, Agent 2 and the manual negotiation in each scenario. Results show that, in all scenarios except the first one (when the agent is still relying on Westin’s Privacy Segmentation Index categorisation), the users made the least changes to the default settings when Agent 2 was negotiating on their behalf. On average, offers proposed by Agent 2 were the most accurate (65.23%). After the first scenario, the choices of Agent 2 for the default settings were more accurate at accommodating the users’ privacy preferences than the manual negotiation.

Although, on average, the accuracy of Agent 1 (58.86%) is lower than the accuracy of manual negotiation (60.76%), overall, we can observe a rising trend of the accuracy of Agent 1. In particular, in the penultimate scenario, it exceeds the accuracy of both Agent 2 and the default settings. This suggests that, with more learning, the individual approach could eventually outperform an agent which bases the preferences on a limited number of profiles.

Our experimental setup, in which participants were not aware of the agent, allows us to be confident that this accuracy is the result of correctly predicting preference, rather than a tendency of participants to “go along with” suggestions that they know are made by an

agent. Although we detect some bias resulting from the defaults, this is apparent in all three conditions – since defaults in the first scenario were set randomly, we expected them to be aligned with user preferences 50% of the time; in fact, the slightly higher percentage in all treatments (MN: 55.15%; A1: 55.76%; A2: 52.73%) show that the defaults exert some influence and act as a means to promote exploration of the different options.

### 6.4.5 Negotiation effort

Lastly, we measured the perceived effort of negotiation via the NASA-TLX questionnaire. Table 4 presents the mean, median and standard deviation of the results in each treatment. We can see that the effort required from the user supported by Agent 2 is not only less than in the *take-it-or-leave-it* approach but also during manual negotiation. This finding shows a potential for automated negotiation to be less demanding than the manual negotiation.

## 7 Discussion and future work

In this paper, we propose a novel multi-issue negotiation framework, where two agents exchange partial and complete offers, bargaining over a number of issues in a bundle. An advantage of this approach is that it prevents competitive, zero-sum negotiations on isolated issues and, instead, promotes mutually beneficial deals. Moreover, the protocol allows users to focus on issues that are important to them and leave out issues for which users find it difficult to determine a precise value and are more naturally determined by the counter party. This is especially important in negotiating privacy permissions, because the benefits of privacy protection are often uncertain and intangible [1] and, as a result, users find it difficult to express the exact willingness to pay for revealing certain information. It is easier for consumers to decide, through relative comparisons, which of the complete offers they prefer in order to assess the value of protecting their privacy [115]. In this way, users can easily explore the set of possible agreements, while the service provider, provided with information about monetizing the data (e.g. through advertising), has the ability to exercise the final say. Such negotiations often occur in practice in a number of settings not necessarily limited to permissions management. For example, when negotiating insurance policies, buyers often specify certain conditions for the extent of cover, for which the seller completes the possible contract by proposing a price. Other examples include negotiating mortgages and broadband packages. For this reason, we believe that the partial-complete offer protocol generalises to other negotiation domains with similar individual-vendor relationships.

Furthermore, we demonstrate the applicability of the framework in a specific bargaining situation, which is the negotiation of permissions between a user and a service provider. In doing so, we assume that it is the user's agent who starts the negotiation, specifying the requirements for a subset of important issues. The service provider's agent is then able to submit a complete counter-offer based on the proposed offer. The user's agent can either accept the complete offer or submit a new partial one. It can also break off the negotiation. Although this ensures that the service provider's agent can always "price out" any undesired partial offers, an equally valid use of this framework could be where the negotiation is initiated by the service provider's agent. Then, the user's agent could specify the conditions, under which they would agree to the service provider's conditions or "price out" any undesired proposals. Future work on this topic should investigate how this setting impacts

user's data sharing, regret, preferences and effort, comparing to the setting we used in this paper.

In addition, we show that the framework can be used in a practical context through a user study with human participants and their private data. To this end, we developed an experimental tool to run on the participants' own smartphones, which allowed them to negotiate various combinations of app permissions in exchange for monetary rewards. We compared negotiation to the *take-it-or-leave-it* approach through experimental evaluation. The results of the user study suggest that users can be incentivised to share much more data when they are able to negotiate, with no increase in regret about their decisions. We also show that negotiation enables users to align their privacy choices more closely with their preferences. In particular, we found that the deals negotiated by the agents are more accurate than the baseline in that the resulting agreements are better aligned with the user's actual preferences. These outcomes suggest that negotiation is a powerful interaction mechanism for achieving mutually beneficial data sharing agreements.

Moreover, we propose two variants of the user's agent, which differ in the way the user's preferences are elicited. Comparing the two agent implicit learning variants, results indicate that, with limited data, a profile-based variant might be better where users can be categorised into a limited number of types, but with more interactions, this can be further personalised. This suggests that the profile-based variant could be effectively used at the beginning when the number of previous negotiations is very small. However, as the number of negotiations increases, the agent can be more effective personalising at offers based on a single user's negotiation data. Alternatively, the negotiation could benefit from a variant that incorporates the two approaches at the same time. Future work should further explore different options of implicit preference elicitation.

As the main limitation of our study we consider the small number of participants in each treatment and, therefore, a limited amount of data. Although this sample size allowed us to prove the concept and make conclusions from our observations, it is hard to generalise findings based on a study of 132 participants. In addition, all the participants are university students, attending in-person in the same geographic location. While this experimental design allowed us to observe the application of our framework in a controlled setting, future work should explore negotiation in a real-life scenario, with a larger sample size that allows for more statistically significant results.

When we discuss this kind of negotiation, it is important to consider the broader ethical implications of it. Currently, many popular online services are indirectly funded through the collection of personal data which users share in return for access to the service and service providers monetise. Whereas in our study participants were told that their data will be posted on a public website, in a real-life scenario, personal information shared with the service provider might have been monetized to generate profit, e.g. by selling the data. In this paper, instead of modelling and using an actual service provider, we decided to explore the user's agent side in a controlled setting without additional variables which introducing an actual service provider adds, and assumed that the service provider compensates the user with monetary rewards. This approach, however, raises the question of whether the proposed solution might promote this view, including settings where monetary incentives could lead users to act against their own best interest in terms of other rights and values. Despite the fact that the human is always in the loop, if a negotiation agent is provided by the service provider, there is a risk that persuasive technologies may be used to influence the behaviour of the user, e.g. based on information mined from other users, or in terms of presenting information to the user that makes clever use of cognitive biases ('nudging'). One way of preventing this problem could be an appropriate legal framework regulating



the use of data negotiation agents, e.g. as part of the EU Artificial Intelligence Act [33], to prevent manipulation and exploiting users. Additionally, having a range of independently-developed and open-source agent providers could allow users to choose an agent they consider trustworthy and would like to be represent by when they exchange their personal data with online services.

In our experiment, the monetary rewards to participants created a certain element of ‘gamification’, where reward was linked to earning more points during the negotiations. As a result of using negotiation, we observed a significant increase in the amount of data shared by the users, which may have only been to maximise their reward within our experiment. On the one hand, if this approach was applied in practice, there is a risk of creating a ‘race to the bottom’ culture, where service providers could optimise the payoffs for privacy-reducing behaviour. This is particularly the case as negotiation is not used in a user-to-user but user-to-platform context, where there is clearly a power imbalance between both sides. To mitigate this risk, the data valuation could be computed by an independent, non-profit organisation, e.g. a data trust [84], that would act as intermediary facilitating data sharing. Future work should take into account these issues when choosing how the negotiation mechanism is applied.

On the other hand, one aspect might also be gamification as users might tend to see the process of granting access to their data more like a game to maximize their profit. This, in turn, might distract users from the actual issue that private data is being made accessible for service providers. This poses a question of responsibility of the platform owners facilitating the data selection process for a potential loss of reluctance when sharing data. In another study, when participants were asked about delegating their consent decisions to intelligent agents, they view those agents as potentially unfair and biased [81]. In order to become a valid delegation option, participants wanted reassurance that the development of the agent was controlled and regulated to reduce the risks of unwanted outcomes. Both regulation such as the Artificial Intelligence Act [33] and transparency through data trusts [84] could help autonomous agents engage with users whose trust is to be earned. Future work should consider the social responsibility of the agent’s actions.

From a user’s perspective, the willingness to share data also might strongly depend on the service that is requesting the data. One criterion might be where the service is located (e.g. US or EU) due to the data protection acts that apply. Also, some services might be perceived more trustworthy than others. To help users understand the level of risk, the user interactions could be designed such that in cases of concerning transactions, the agent would warn the user about potential negative consequences. To this end, future work should consider more realistic scenarios, where participants are asked to negotiate with a selection of specific services so that they can better understand the consequences of their action. Additionally, future work should explore more transparent interfaces, testing various wordings of the warning messages and modelling user’s trust as part of the preference profile.

When it comes to the user interface design, another factor that might have influenced the results of our study was the interface through which the user was kept in the loop. In particular, when asked for approval, the user was able to see the ‘Share’ and ‘Don’t Share’ radio buttons, on the Take-It-Or-Leave-It screen, even though they could not use them. Although it is unclear whether this influenced their choices at all, it is possible that having these buttons provoked them to make certain approval or declining decisions that a different interface (e.g. simply listing the data types without radio buttons) would not. Future work should explore how the interface affects users’ decisions and their trust to the offers negotiated by the agent. For example, providing options to view the data to be shared could mean that the user would

spend more time making the decision manually (providing the agent with feedback for future decisions on behalf of the user), but it could also help providing more transparency into the data sharing.

It is also important to note that some of the data types we selected for the experiment are more sensitive than others. Photos or messages, for instance, might be highly personal, whereas the app list does not provide equally deep insights into the user's private life. In fact, the agent could act autonomously for the data types that are not perceived sensitive (which the user should decide) but for the sensitive data, the user could be asked to make the decision manually. However, the formula that is used for generating the service provider's random quote in our study does not take this difference between the data types sensitivity into account. Instead, only the quantity of allowed permissions is used for proposing a quote. In reality, the offers companies make for getting access to pictures and messages might be significantly higher compared to the app list. Future work should also consider this when proposing a quote, to allow for investigating different permissions. In particular, when more information is shared, a more detailed analysis of the user can be made. Accordingly, it can be assumed the quotes not only increase linearly when the number of granted permissions increases. Future work should explore more appropriate valuation models for the negotiated data types and consider how those models impact the user's reactions to the offers.

Another limitation is that while the proposed variants of the user's agent are both quite general, our experimental setting is limited to reasoning about data types. Other issues that need to be considered during negotiation are the factors beyond data types, for example, the recipient, retention period, purpose, quality, and privacy risks. In addition, we noticed that, when expressing regret, users often did so for specific data points (e.g. specific contacts or photos). Therefore, it is clear that a model based on permissions alone is too coarse to accurately capture the privacy preferences. Combining a semi-autonomous agent with a more meaningful classification of data (perhaps using signals such as location, time of the day and relations to other users) is another avenue that warrants further exploration. Additionally, the data valuation could be dependent on the purpose of data processing (e.g. the service provider specifying the purpose as part of the negotiation). We also plan to study various designs of privacy agents that can learn to negotiate on users' behalf, and engage users directly only when in doubt.

Moreover, in some of the settings in our study, requesting a new offer manually meant additional cost for the user. While this choice allowed us to study the user behaviour in those settings and compare them to the take-it-or-leave-it setting, we would like to acknowledge that punishing users for requesting a new offer affects the strategy that users need to pursue to maximize their total points. The consequences of this choice should be considered when designing a negotiation mechanism for a real-life scenario.

Finally, in our study, the agent's user model was based on types derived from results of a short survey before the negotiations started. A more personalised model, derived from e.g. apps installed on the phone, data sensitivity and other factors, could increase the accuracy of the deal negotiated by the agent even further. Future work should aim to improve the accuracy of the agent.

## 8 Conclusions

To conclude, this paper presents a novel multi-issue negotiation framework with a new variant of the alternating-offers protocol, based on exchanges of partial and complete counter-offers. Moreover, we demonstrate how this framework can be used in bilateral negotiations of privacy permissions between users and service providers. In order for an agent to faithfully represent a user, we investigate two approaches to implicit preference elicitation: one approach personalised to each individual user and one personalised depending on the user's privacy profile classification. Furthermore, we compare them to the common *take-it-or-leave-it* approach, in which users are required to accept all permissions requested by a service. We do so through a user study with human participants who negotiate access permissions to private data stored on their own smartphones through the experimental tool we developed.

The results of this experimental evaluation provide evidence that negotiating agents are able to automatically negotiate on behalf of users, while optimally balancing between utility and the sum of incurred costs. Specifically, we find that users grant permissions 2.5 times more often when they are able to negotiate while maintaining the same level of decision regret. Moreover, we observe that negotiation can be less mentally demanding than the *take-it-or-leave-it* approach and that it enables people to better align their privacy choices with their actual preferences. We discuss our findings, which point to several avenues of future work on automated and negotiable privacy management mechanisms.

This work sits within the wider agenda of privacy management that has received renewed momentum with the introduction of novel privacy laws such as the EU's General Data Protection Regulation [32], requiring greater transparency and user empowerment, and with opportunities for multi-agent systems to provide technological solutions. Our ultimate aim is to enable automated negotiation of the terms and conditions of personal data use. Building on approaches from the field of multi-agent systems, where automated negotiation is widely researched [34], a software agent may eventually negotiate on behalf of the user, based on preferences that are elicited at a time that is convenient to the user. The work described in this article is an essential step towards addressing this broader vision.

In addition, while this paper focuses on combining negotiation and preference elicitation techniques in the context of privacy negotiation, many of these techniques can be generalised to other settings. In particular, an interesting direction for future work is to apply the proposed user preference elicitation to other domains, and to compare it to other approaches, e.g. those developed as part of the Automated Negotiating Agents Competition (ANAC) [3, 50]. However, unlike our approach, the ANAC competition so far has assumed no noise in the user's own preferences, i.e. the partial preference order is consistent. Nevertheless, another aspect of the challenge of this competition is opponent modelling (where the aim is to learn the utility function of the opponent), which uses similar techniques and where noise is more likely. Hence, it would be interesting to compare our approach to recent advances in opponent modelling, including machine learning techniques (e.g. [101]). There is also a wide area of research which considers eliciting user preferences through surveys, often applied in areas such as transportation, marketing and social sciences. In particular, discrete choice models [73] assume a probabilistic model and have shown to be effective in capturing preferences from partial orders. It would be interesting to compare our approach to using profiles and minimising constraint violations to existing approaches in these related areas.

**Acknowledgements** This work is supported by a grant from the EPSRC for Meaningful Consent in the Digital Economy project (EP/K039989/1) and is part of the Vidi project COMBINE (VI.Vidi.203.044) which is (partly) financed by the Dutch Research Council (NWO). Dorota Filipczuk is funded by the EPSRC, award ref.: 1807923. All data supporting this study are openly available from the University of Southampton repository at <https://doi.org/10.5258/SOTON/D1323>. Study approved by the University of Southampton's Ethics Committee (ref.: ERGO/FPSE/18082). The authors wish to thank all of the participants and our anonymous reviewers for their comments.

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

## References

1. Acquisti, A., & Grossklags, J. (2005). Uncertainty, ambiguity and privacy. *In Fourth Workshop on the Economics of Information Security (WEIS)*, pp. 2–3.
2. Amadeo, R. (2013). App ops: Android 4.3's hidden app permission manager, control permissions for individual apps! <http://www.androidpolice.com>
3. Aydođan, R., Baarslag, T., Fujita, K., Mell, J., Gratch, J., Jonge, D.d., Mohammad, Y., Nakadai, S., Morinaga, S., & Osawa, H., et al. (2020). Challenges and main results of the automated negotiating agents competition (ANAC) 2019. *In Multi-agent systems and agreement technologies*, pp. 366–381. Springer
4. Aydođan, R., Festen, D., Hindriks, K.V., & Jonker, C.M. (2017). Alternating offers protocols for multilateral negotiation. *In Modern approaches to agent-based complex automated negotiation*, pp. 153–167. Springer.
5. Baarslag, T. (2016). Exploring the strategy space of negotiating agents: A framework for bidding, learning and accepting in automated negotiation. Springer.
6. Baarslag, T., Alan, A.T., Gomer, R., Alam, M., Perera, C., & Gerding, E.H., et al. (2017). An automated negotiation agent for permission management. *In Proceedings of the 16th Conference on Autonomous Agents and Multi-Agent Systems*, pp. 380–390. International Foundation for Autonomous Agents and Multiagent Systems.
7. Baarslag, T., & Gerding, E.H. (2015). Optimal incremental preference elicitation during negotiation. *In Proceedings of the 24th International Joint Conference on Artificial Intelligence*, pp. 3–9.
8. Baarslag, T., Gerding, E.H., Aydođan, R., & schraefel, m.c. (2015). Optimal negotiation decision functions in time-sensitive domains. *In Proceedings of the 2015 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, vol. 2, pp. 190–197. IEEE.
9. Baarslag, T., Hendriks, M. J., Hindriks, K. V., & Jonker, C. M. (2012). Measuring the performance of online opponent models in automated bilateral negotiation. *Lecture Notes in Computer Science* In M. Thielscher & D. Zhang (Eds.), *AI 2012: Advances in Artificial Intelligence* (Vol. 7691, pp. 1–14). Springer.
10. Baarslag, T., Hendriks, M.J., Hindriks, K.V., Jonker, C.M.: Learning about the opponent in automated bilateral negotiation: a comprehensive survey of opponent modeling techniques. *Autonomous Agents and Multi-Agent Systems* 30(5), 849–898 (2016).
11. Barth, S., & de Jong, M.D.: The privacy paradox—investigating discrepancies between expressed privacy concerns and actual online behavior—a systematic literature review. *Telematics and Informatics* 34(7), 1038–1058 (2017).
12. Bender, S.: Privacy in the cloud frontier: Abandoning the take it or leave it approach. *Drexel Law Review* 4, 487 (2011).
13. Benisch, M., Kelley, P.G., Sadeh, N., Cranor, L.F.: Capturing location-privacy preferences: Quantifying accuracy and user-burden tradeoffs. *Personal and Ubiquitous Computing* 15(7), 679–694 (2011).

14. Bennis, M., & Langendörfer, P. (2003). Towards automatic negotiation of privacy contracts for internet services. In *The 11th IEEE International Conference on Networks*, pp. 319–324. IEEE.
15. Beresford, A.R., Rice, A., Skehin, N., & Sohan, R. (2011). MockDroid: trading privacy for application functionality on smartphones. In *Proceedings of the 12th Workshop on Mobile Computing Systems and Applications*, pp. 49–54. ACM.
16. Bokhurst, M. (2015). Xprivacy. <https://github.com/M66B/XPrivacy>.
17. Bonné, B., Peddinti, S.T., Bilogrevic, I., & Taft, N. (2017). Exploring decision making with android's runtime permission dialogs using in-context surveys. In *Proceedings of the 13th Symposium on Usable Privacy and Security*. USENIX Association.
18. Boutilier, C. (2002). A POMDP formulation of preference elicitation problems. In *Eighteenth National Conference on Artificial Intelligence*, pp. 239–246. American Association for Artificial Intelligence, Menlo Park, CA, USA. <http://dl.acm.org/citation.cfm?id=777092.777132>.
19. Boutilier, C., Brafman, R.L., Domshlak, C., Hoos, H.H., Poole, D.: Cp-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. *Journal of Artificial Intelligence Research* 21, 135–191 (2004).
20. Campbell, A.J.: Relationship marketing in consumer markets: A comparison of managerial and consumer attitudes about information privacy. *Journal of Direct Marketing* 11(3), 44–57 (1997).
21. Cate, F.H. (2006). The failure of fair information practice principles. In J.K. Winn (ed.) *Consumer Protection in the Age of the Information Economy*, chap. 13, pp. 343–379. Ashgate.
22. Cate, F.H.: The limits of notice and choice. *IEEE Security & Privacy* 8(2), 59–62 (2010).
23. Chajewska, U., Getoor, L., Norman, J., & Shahar, Y. (1998). Utility elicitation as a classification problem. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, UAI'98, pp. 79–88. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. <http://dl.acm.org/citation.cfm?id=2074094.2074104>
24. Chajewska, U., Koller, D., & Parr, R. (2000). Making rational decisions using adaptive utility elicitation. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence*, pp. 363–369.
25. Chen, S., Ammar, H.B., Tuyls, K., & Weiss, G. (2013). Optimizing complex automated negotiation using sparse pseudo-input gaussian processes. In *Proceedings of the 12th International Conference on Autonomous Agents and Multi-agent Systems*, AAMAS '13, pp. 707–714. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC.
26. Cheng, V.S., Hung, P.C., & Chiu, D.K. (2007). Enabling web services policy negotiation with privacy preserved using XACML. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, pp. 33–33. IEEE.
27. Choi, H., Park, J., Jung, Y.: The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior* 81, 42–51 (2018).
28. Conen, W., & Sandholm, T. (2001). Minimal preference elicitation in combinatorial auctions. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, Workshop on Economic Agents, Models, and Mechanisms, pp. 71–80.
29. Cranor, L.F. (2002). Web privacy with P3P. O'Reilly Media, Inc.
30. Cranshaw, J., Mogan, J., & Sadeh, N. (2011). User-controllable learning of location privacy policies with gaussian mixture models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 25(1).
31. Do, Q., Martini, B., & Choo, K.K. (2014). Enhancing user privacy on android mobile devices via permissions removal. In *Proceedings of the 47th Hawaii International Conference on System Sciences*, pp. 5070–5079.
32. European Parliament and the Council: 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). *Official Journal of the European Union* pp. 1–88 (2016).
33. European Parliament and the Council: Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. *Official Journal of the European Union* pp. 1–107 (2021).
34. Fatima, S., Kraus, S., & Wooldridge, M. (2014). Principles of automated negotiation. Cambridge University Press.
35. Fatima, S., & Wooldridge, M. (2014). The negotiation game. *Intelligent Systems*, 29, 57–61.
36. Fatima, S., Wooldridge, M., & Jennings, N.R. (2003). Optimal Agendas for Multi-Issue Negotiation. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 129–136. ACM.

37. Fatima, S.S., Wooldridge, M.J., & Jennings, N.R. (2002). Multi-issue negotiation under time constraints. In *AAMAS '02: Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pp. 143–150. ACM, New York, NY, USA.
38. Greco, S., Kadziński, M., Mousseau, V., Slowiński, R.: Robust ordinal regression for multiple criteria group decision: Utagms-group and utadisgms-group. *Decision Support Systems* 52(3), 549–561 (2012).
39. Hao, J., & Leung, H. (2014). CUHK agent: an adaptive negotiation strategy for bilateral negotiations over multiple items. *Studies in Computational Intelligence* *Novel Insights in Agent-based Complex Automated Negotiation* (Vol. 535, pp. 171–179). Japan: Springer.
40. Harris, L., Westin, A.F.: E-commerce and privacy: What net users want. *Privacy and American Business*, (1998).
41. Hart, S.G., & Staveland, L.E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: *Advances in psychology*, vol. 52, pp. 139–183. Elsevier.
42. Hornyack, P., Han, S., Jung, J., Schechter, S., & Wetherall, D. (2011). These aren't the droids you're looking for: Retrofitting android to protect data from imperious applications. In *Proceedings of the 18th ACM Conference on Computer and Communications Security*, pp. 639–652. ACM.
43. Hutton, L., & Henderson, T. (2015). “I didn’t sign up for this!”: Informed consent in social network research. In *International AAAI Conference on Web and Social Media*.
44. Ilany, L., & Gal, Y. K. (2014). The simple-meta agent. *Studies in Computational Intelligence* *Novel Insights in Agent-based Complex Automated Negotiation* (Vol. 535, pp. 197–200). Japan: Springer.
45. Jacquet-Lagrez, E., Siskos, J.: Assessing a set of additive utility functions for multicriteria decision-making, the uta method. *European Journal of Operational Research* 10(2), 151–164 (1982).
46. Jang, I., & Yoo, H.S. (2009). Personal information classification for privacy negotiation. In *2009 Fourth International Conference on Computer Sciences and Convergence Information Technology*, pp. 1117–1122. IEEE.
47. Jang, I.J., Shi, W., & Yoo, H.S. (2008). Policy negotiation system architecture for privacy protection. In *2008 Fourth International Conference on Networked Computing and Advanced Information Management*, vol. 2, pp. 592–597. IEEE.
48. Jennings, N.R., Faratin, P., Lomuscio, A.R., Parsons, S., Wooldridge, M.J., Sierra, C.: Automated negotiation: Prospects, methods and challenges. *Group Decision and Negotiation* 10(2), 199–215 (2001).
49. Jeon, J., Micinski, K.K., Vaughan, J.A., Fogel, A., Reddy, N., Foster, J.S., & Millstein, T. (2012). Dr. Android and Mr. Hide: Fine-grained permissions in android applications. In *Proceedings of the Second ACM Workshop on Security and Privacy in Smartphones and Mobile Devices*, SPSM '12, pp. 3–14. ACM, New York, NY, USA.
50. Jonker, C., Aydogan, R., Baarslag, T., Fujita, K., Ito, T., & Hindriks, K. (2017). Automated negotiating agents competition (ANAC). In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31.
51. Kalyani, Y., & Adams, C. (2006). Privacy negotiation using a mobile agent. In *Electrical and Computer Engineering, 2006. CCECE'06. Canadian Conference on*,
52. Karunatilake, N.C. (2006). Argumentation-based negotiation in a social context. Ph.D. thesis, University of Southampton.
53. Kawaguchi, S., Fujita, K., & Ito, T. (2012). Compromising strategy based on estimated maximum utility for automated negotiating agents. *Series of Studies in Computational Intelligence* *New Trends in Agent-based Complex Automated Negotiations* (pp. 137–144). Berlin, Heidelberg: Springer-Verlag.
54. Keeney, R.L., & Raiffa, H. (1976). *Decisions with Multiple Objectives*. Cambridge University Press.
55. Kekulluoglu, D., Kokciyan, N., Yolum, P.: Preserving privacy as social responsibility in online social networks. *ACM Transactions on Internet Technology* 18(4), 1–22 (2018).
56. Kelley, P.G., Hanks Drielsma, P., Sadeh, N., & Cranor, L.F. (2008). User-controllable learning of security and privacy policies. In *Proceedings of the 1st ACM workshop on Workshop on AISec*, pp. 11–18. ACM.
57. Klein, M., Lu, S.C.Y.: Conflict resolution in cooperative design. *Artificial Intelligence in Engineering* 4(4), 168–180 (1989).
58. Knijnenburg, B.P., Reijmer, N.J., & Willemsen, M.C. (2011). Each to his own: how different users call for different interaction methods in recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*, pp. 141–148. ACM.
59. Kökciyan, N., Yaglikci, N., Yolum, P.: An argumentation approach for resolving privacy disputes in online social networks. *ACM Transactions on Internet Technology (TOIT)* 17(3), 1–22 (2017).
60. Kökciyan, N., Yolum, P.: Priguard: A semantic approach to detect privacy violations in online social networks. *IEEE Transactions on Knowledge and Data Engineering* 28(10), 2724–2737 (2016).

61. Kraus, S. (2001). *Strategic Negotiation in Multiagent Environments*. MIT press.
62. Krol, K., Preibusch, S.: Effortless privacy negotiations. *IEEE Security & Privacy* 13(3), 88–91 (2015).
63. Krol, K., & Preibusch, S. (2016). Control versus effort in privacy warnings for webforms. In *Proceedings of the 2016 ACM Workshop on Privacy in the Electronic Society*, pp. 13–23. ACM.
64. Kumaraguru, P., Cranor, L.F.: Privacy indexes: a survey of Westin's studies. Carnegie Mellon University, School of Computer Science, Institute for Software Research International (2005).
65. Larson, K., & Sandholm, T. (2001). Costly valuation computation in auctions. In *Proceedings of the 8th Conference on Theoretical Aspects of Rationality and Knowledge*, TARK '01, pp. 169–182. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA. <http://dl.acm.org/citation.cfm?id=1028128.1028148>.
66. Liccardi, I., Pato, J., Weitzner, D.J.: Improving user choice through better mobile apps transparency and permissions analysis. *Journal of Privacy and Confidentiality* 5(2), 1 (2014).
67. Lin, J., Amini, S., Hong, J.I., Sadeh, N., Lindqvist, J., & Zhang, J. (2012). Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 501–510. ACM.
68. Lin, J., Liu, B., Sadeh, N., & Hong, J.I. (2014). Modeling users' mobile app privacy preferences: Restoring usability in a sea of permission settings. In *Symposium On Usable Privacy and Security*, pp. 199–212.
69. Linden, T., Khandelwal, R., Harkous, H., Fawaz, K.: The privacy policy landscape after the GDPR. *Proceedings on Privacy Enhancing Technologies* 2020(1), 47–64 (2020).
70. Liu, B., Lin, J., & Sadeh, N. (2014). Reconciling mobile app privacy and usability on smartphones: Could user privacy profiles help? In *Proceedings of the 23rd international conference on World wide web*, pp. 201–212. ACM.
71. Maaser, M., & Langendoerfer, P. (2005). Automated negotiation of privacy contracts. In *Computer Software and Applications Conference, 2005. COMPSAC 2005. 29th Annual International*, vol. 1, pp. 505–510. IEEE.
72. Maaser, M., Ortmann, S., & Langendörfer, P. (2006). NEPP: Negotiation enhancements for privacy policies. In *W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement*.
73. Mariel, P., Hoyos, D., Meyerhoff, J., Czajkowski, M., Dekker, T., Glenk, K., Jacobsen, J.B., Liebe, U., Olsen, S.B., & Sagebiel, J., et al. (2021). Environmental valuation with discrete choice experiments: *Guidance on design, implementation and data analysis*. Springer Nature.
74. Matte, C., Bielova, N., & Santos, C. (2020). Do cookie banners respect my choice?: Measuring legal compliance of banners from iab europe's transparency and consent framework. In *Proceedings of the 41st IEEE Symposium on Security and Privacy*, pp. 791–809. IEEE.
75. d. Melo, L.L., & Zorzo, S.D. (2012) PUPDroid - personalized user privacy mechanism for Android. In *International Conference on Systems, Man, and Cybernetics*, pp. 1479–1484. IEEE, Seoul, South Korea.
76. Mester, Y., Kökciyan, N., & Yolum, P. (2015). Negotiating privacy constraints in online social networks. In *Advances in Social Computing and Multiagent Systems*, pp. 112–129. Springer.
77. Milne, G.R., Pettinico, G., Hajjat, F.M., Markos, E.: Information sensitivity typology: Mapping the degree and type of risk consumers perceive in personal data sharing. *Journal of Consumer Affairs* 51(1), 133–161 (2017).
78. Mohammad, Y., & Nakadai, S. (2018). Fastvoi: Efficient utility elicitation during negotiations. In *International Conference on Principles and Practice of Multi-Agent Systems*, pp. 560–567. Springer.
79. Mugan, J., Sharma, T., & Sadeh, N. (2011). Understandable learning of privacy preferences through default personas and suggestions. Institute for Software Research Technical Report CMU-ISR-11-112. Carnegie Mellon University, Pittsburgh, PA.
80. Nauman, M., Khan, S., & Zhang, X. (2010). Apex: extending android permission model and enforcement with user-defined runtime constraints. In *Proceedings of the 5th ACM symposium on Information, Computer and Communications Security*, pp. 328–332. ACM.
81. Nissen, B., Neumann, V., Mikusz, M., Gianni, R., Clinch, S., Speed, C., & Davies, N. (2019). Should i agree? delegating consent decisions beyond the individual. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13.
82. Nissenbaum, H.: *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press, (2009).
83. Nissenbaum, H.: A contextual approach to privacy online. *Daedalus* 140(4), 32–48 (2011).
84. O'Hara, K. (2019). *Data Trusts: Ethics, Architecture and Governance for Trustworthy Data Stewardship* (WSI White Papers, 1) Southampton. University of Southampton.



85. Ongtang, M., McLaughlin, S., Enck, W., & McDaniel, P. (2009). Semantically rich application-centric security in android. In *Annual Computer Security Applications Conference*, 2009, pp. 340–349. IEEE, Honolulu, HI, US.
86. Osborne, M.J., & Rubinstein, A. (1994). *A Course in Game Theory*, 1st edn. The MIT Press.
87. Parkes, D.C.: Auction design with costly preference elicitation. *Annals of Mathematics and Artificial Intelligence* 44(3), 269–302 (2005). DOI: 10.1007/s10472-005-4692-y.
88. Perera, C. (2017). Sensing as a service for internet of things: A roadmap. Lulu.com.
89. Perera, C., Ranjan, R., Wang, L.: End-to-end privacy for open big data markets. *IEEE Cloud Computing* 2(4), 44–53 (2015).
90. Polykalas, S.E. (2017). Assessing general data protection regulation for personal data privacy: is the end of “take it or leave it” approach for downloading apps? In *Proceedings of the Seventh International Conference on Social Media Technologies, Communication, and Informatics*.
91. Popescu, E.: On the approximation of inconsistent inequality systems. *Analele Științifice ale Universității Ovidius* 11(2), 109–118 (2003).
92. Preibusch, S. (2006). Implementing privacy negotiations in e-commerce. In *Asia-Pacific Web Conference*, pp. 604–615. Springer.
93. Preibusch, S., Krol, K., & Beresford, A.R. (2013). The privacy economics of voluntary over-disclosure in web forms. In *The Economics of Information Security and Privacy*, pp. 183–209. Springer.
94. Ravichandran, R., Benisch, M., Kelley, P.G., & Sadeh, N.M. (2009). Capturing social networking privacy preferences. In *Goldberg I., Atallah M.J. (eds) Privacy Enhancing Technologies. PETS 2009. Lecture Notes in Computer Science*, vol. 5672, pp. 1–18. Springer, Berlin, Heidelberg.
95. Robinson, W.N. (1990). Negotiation behavior during requirement specification. In *Proceedings of the 12th International Conference on Software Engineering*, pp. 268–276. IEEE.
96. Rosenschein, J.S. (1986). Rational interaction: cooperation among intelligent agents. Ph.D. thesis, Stanford University, Stanford, CA, United States.
97. Rosenschein, J.S., & Zlotkin, G. (1994). Rules of encounter: designing conventions for automated negotiation among computers. MIT Press.
98. Roszkowska, E., et al.: The application of uta method for support evaluation negotiation offers. *Optimum Economic Studies* 80(2), 144–162 (2016).
99. Rubinstein, A.: Perfect equilibrium in a bargaining model. *Econometrica* 50(1), 97–109 (1982).
100. Sadeh, N., Hong, J., Cranor, L., Fette, I., Kelley, P., Prabaker, M., Rao, J.: Understanding and capturing people’s privacy policies in a mobile social networking application. *Personal and Ubiquitous Computing* 13(6), 401–412 (2009).
101. Salmanian, F., Jazayeriy, H., & Kazemitabar, J. (2021). User preferences elicitation in bilateral automated negotiation using recursive least square estimation. In *2021 12th International Conference on Information and Knowledge Technology (IKT)*, pp. 1–5. IEEE.
102. Sanchez-Rola, I., Dell’Amico, M., Kotzias, P., Balzarotti, D., Bilge, L., Vervier, P.A., & Santos, I. (2019). Can I opt out yet? GDPR and the global illusion of cookie control. In *Proceedings of the 2019 ACM Asia Conference on Computer and Communications Security*, pp. 340–351.
103. Schermer, B.W., Custers, B., van der Hof, S.: The crisis of consent: How stronger legal protection may lead to weaker consent in data protection. *Ethics and Information Technology* 16(2), 171–182 (2014).
104. Schwab, K., Marcus, A., Oyola, J., Hoffman, W., & Luzi, M. (2011). Personal data: The emergence of a new asset class. In *An Initiative of the World Economic Forum*.
105. Smith, R.G.: The contract net protocol: High-level communication and control in a distributed problem solver. *IEEE Transactions on Computers* 29(12), 1104–1113 (1980).
106. Squicciarini, A.C., Bertino, E., Ferrari, E., Ray, I.: Achieving privacy in trust negotiations with an ontology-based approach. *IEEE Transactions on Dependable and Secure Computing* 3(1), 13–30 (2006).
107. Such, J.M., Criado, N.: Resolving multi-party privacy conflicts in social media. *IEEE Transactions on Knowledge and Data Engineering* 28(7), 1851–1863 (2016).
108. Such, J.M., Espinosa, A., García-Fornes, A.: A survey of privacy in multi-agent systems. *The Knowledge Engineering Review* 29(03), 314–344 (2014).
109. Such, J.M., Rovatsos, M.: Privacy policy negotiation in social media. *ACM Transactions on Autonomous and Adaptive Systems (TAAS)* 11(1), 4 (2016).
110. Susser, D.: Notice after notice-and-consent: Why privacy disclosures are valuable even if consent frameworks aren’t. *Journal of Information Policy* 9, 148–173 (2019).
111. Sycara, K. (1988). Resolving goal conflicts via negotiation. In *Proceedings of the Seventh AAAI National Conference on Artificial Intelligence*, vol. 88, pp. 245–250.



112. Sycara-Cyranski, K. (1985). Arguments of persuasion in labour mediation. In *Proceedings of the 9th international joint conference on Artificial intelligence*, pp. 294–296. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
113. Taddicken, M.: The ‘privacy paradox’ in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication* 19(2), 248–273 (2014).
114. Train, K.E. (2009). *Discrete choice methods with simulation*. Cambridge university press.
115. Tsai, J.Y., Egelman, S., Cranor, L., Acquisti, A.: The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research* 22(2), 254–268 (2011).
116. Tsimppoukis, D., Baarslag, T., Kaisers, M., & Paterakis, N.G. (2018). Automated negotiations under user preference uncertainty: A linear programming approach. In *International Conference on Agreement Technologies*, pp. 115–129. Springer.
117. Ulusoy, O., Yolum, P.: Panola: A personal assistant for supporting users in preserving privacy. *ACM Transactions on Internet Technology (TOIT)* 22(1), 1–32 (2021).
118. Van Nguyen, D.: Global maximization of uta functions in multi-objective optimization. *European Journal of Operational Research* 228(2), 397–404 (2013).
119. Vollmer, R. (2015). Xposed framework. <http://repo.xposed.info/>.
120. Weitzman, M.L.: Optimal search for the best alternative. *Econometrica* 47(3), 641–654 (1979).
121. Wilkinson, D., Namara, M., Badillo-Urquiola, K., Wisniewski, P.J., Knijnenburg, B.P., Page, X., & Toch, E., Romano-Bergstrom, J. (2018). Moving beyond a “one-size fits all” exploring individual differences in privacy. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–8. ACM.
122. Williams, C. R., Robu, V., Gerding, E. H., & Jennings, N. R. (2012). Iamhaggler: A negotiation agent for complex environments. *Studies in Computational Intelligence New Trends in Agent-based Complex Automated Negotiations* (pp. 151–158). Berlin, Heidelberg: Springer-Verlag.
123. Wilson, S., Cranshaw, J., Sadeh, N., Acquisti, A., Cranor, L.F., Springfield, J., Jeong, S.Y., & Balasubramanian, A. (2013). Privacy manipulation and acclimation in a location sharing application. In *Proceedings of the 2013 International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 549–558. ACM.
124. Woodruff, A., Pihur, V., Consolvo, S., Brandimarte, L., & Acquisti, A. (2014). Would a Privacy Fundamentalist Sell Their DNA for \$1000... If Nothing Bad Happened as a Result? The Westin Categories, Behavioral Intentions, and Consequences. In *Proceedings of the 10th Symposium On Usable Privacy and Security*, pp. 1–18.
125. Yassine, A., & Shirmohammadi, S. (2009). An intelligent agent-based framework for privacy payoff negotiation in virtual environments. In *Computational Intelligence in Virtual Environments, 2009. CIVE'09. IEEE Workshop on*, pp. 20–25. IEEE.
126. Yassine, A., & Shirmohammadi, S. (2009). Measuring users’ privacy payoff using intelligent agents. In *Computational Intelligence for Measurement Systems and Applications, 2009. CIMSA'09. IEEE International Conference on*, pp. 169–174. IEEE.

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