CSC 495.002 – Lecture 8
AI for Privacy: Agents and Reasoning

Dr. Özgür Kafalı

North Carolina State University
Department of Computer Science

Fall 2017

Privacy Requirements Engineering

- Functional requirements and how they might have security and privacy implications
- Phases of requirements engineering
- Threat modeling
- Formal specification of privacy requirements and automated identification of conflicts

Exercise: Privacy Implications

- Assume you are developing a social application:
  - Determine how many users are in close proximity
  - Recommend an activity that they can do together
- First, determine a couple of functional requirements
- Then, identify related privacy requirements
  - How would you protect sensitive user information?
  - Access control requirements: Who should access what information?
  - Do you need to log any user actions in case something goes wrong?

Problem Definition

- Software agent: An intelligent entity that acts on behalf of a user
- Multiagent systems (MAS): A collection of agents
  - Collaboration
  - Coordination
  - Competition
- Design and implement a MAS to solve a privacy problem

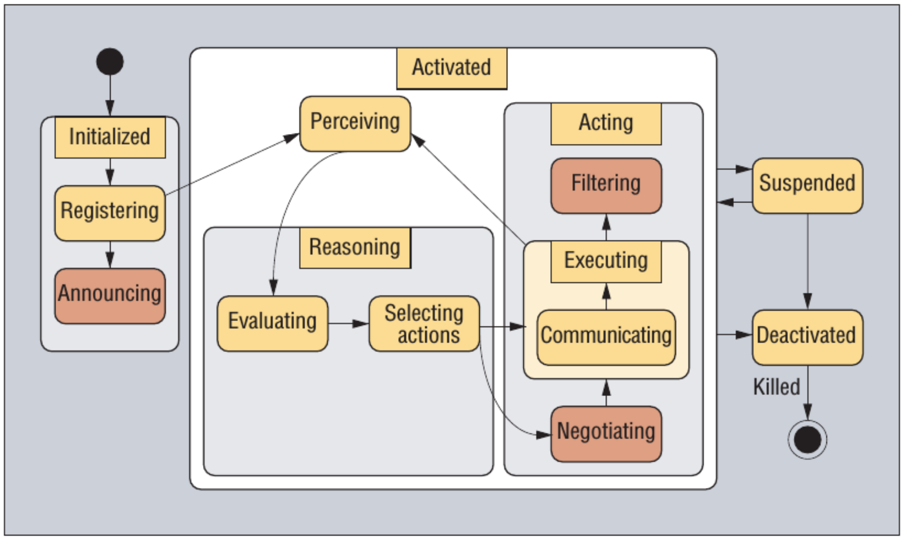
Overview of Problem Domains

- Resolving multi-party privacy concerns via argumentation
- Negotiating privacy preferences
- Formal policy specification and analysis via semantic reasoning

Privacy-aware Agents for Pervasive Healthcare

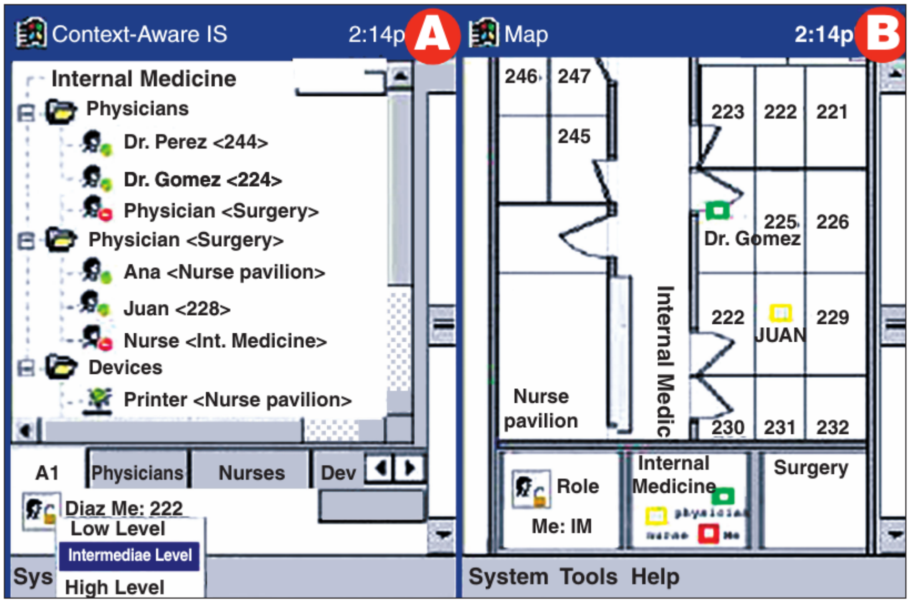
- Help developers design privacy-aware systems
- Handle threats raised by pervasive technology
- Dynamic hospital environment:
  - High availability
  - Careful attention to patients
  - Confidentiality
  - Rapid response to emergencies
  - Constant coordination with colleagues

Agent Reasoning Cycle

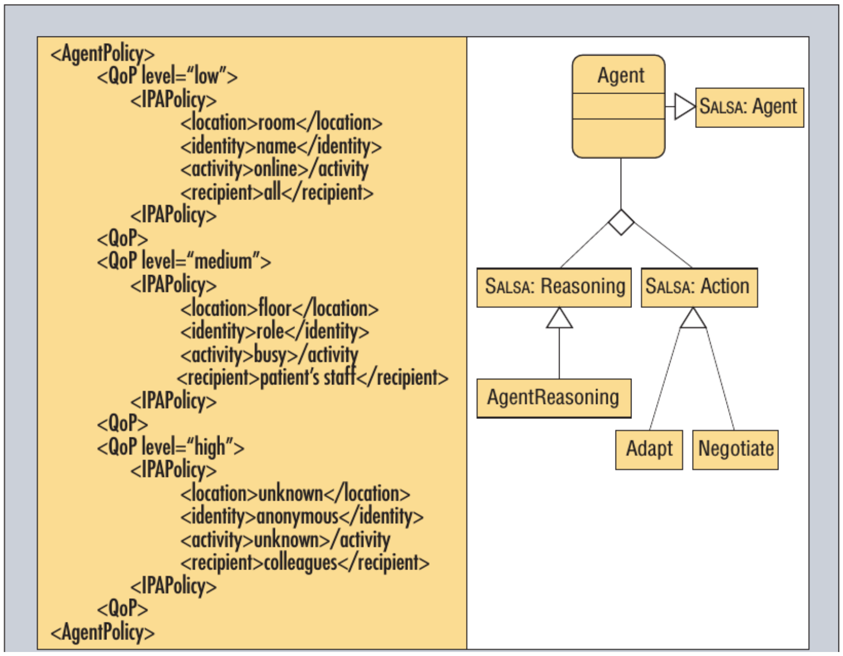


Tentori et al. Privacy-Aware Autonomous Agents for Pervasive Healthcare. IEEE Intelligent Systems, 21(6):55–62, 2006

Salsa Agent Interface

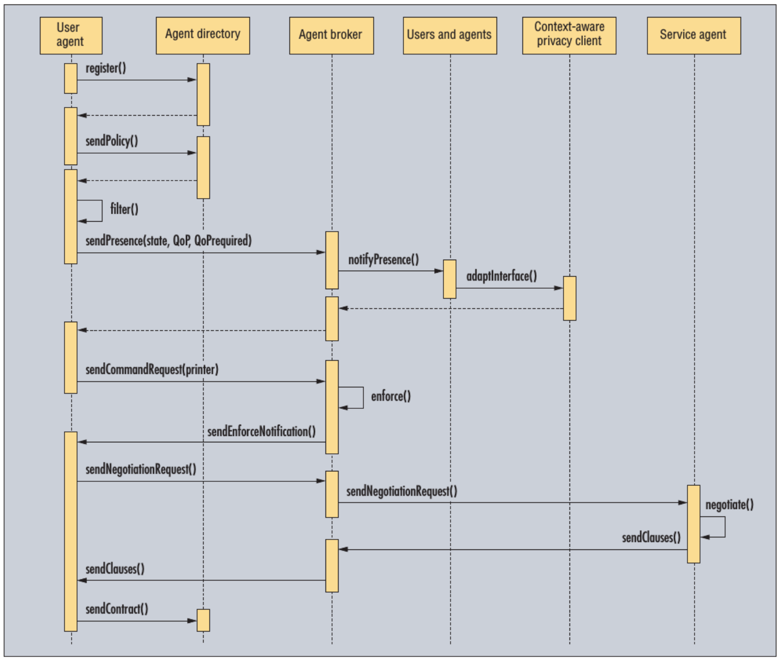


Salsa Agent Specification

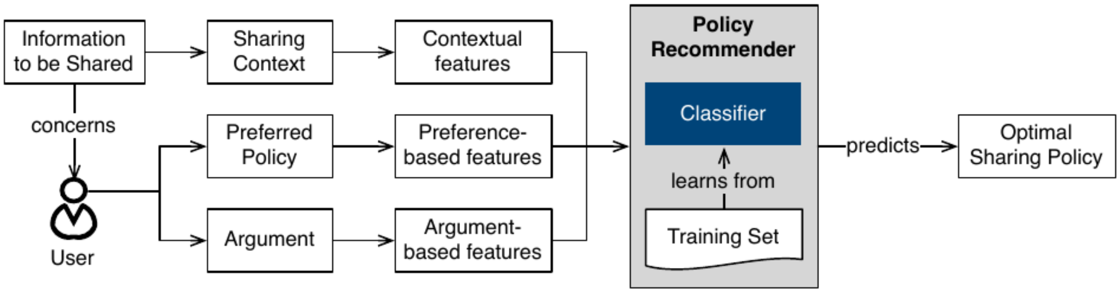


```
<AgentPolicy>
        <QoP level="low">
              <IPAPolicy>
                        <location>room</location>
                        <identity>name</identity>
                        <activity>online>/activity
                        <recipient>all</recipient>
              <IPAPolicy>
        <QoP>
        <QoP level="medium">
              <IPAPolicy>
                        <location>floor</location>
                        <identity>role</identity>
                        <activity>busy>/activity
                       <recipient>patient's staff</recipient>
              <IPAPolicy>
        <QoP>
        <QoP level="high">
              <IPAPolicy>
                        <location>unknown</location>
                        <identity>anonymous</identity>
                        <activity>unknown>/activity
                        <recipient>colleagues</recipient>
              <IPAPolicy>
        <QoP>
<AgentPolicy>
```

Quality of Privacy



Multi-party Privacy

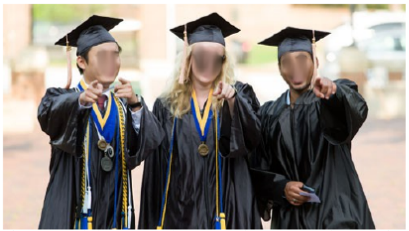


- Context
- Individual preferences
- Generated arguments

Fogues et al. Sharing Policies in Multiuser Privacy Scenarios: Incorporating Context, Preferences, and Arguments in Decision Making. ACM Transactions on Computer-Human Interaction, 24(1):5:1–5:29, 2017

Multi-party Privacy: Friends Scenario I

Picture and Context



Relationship: Friends (92.2%)
Sensitivity rating: μ = 1.56 (σ = 0.96)
Sentiment rating: μ = 1.77 (σ = 1.46)

Description

Tim, Ashley, and Jerry just graduated. Tim's father took the picture above after the graduation ceremony. Tim wants to upload the picture to his social media account.

Arguments

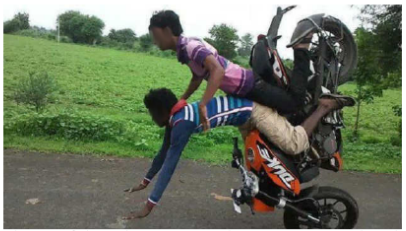
Positive consequence argument. People we know will be happy to see that we are finally done with college.

Negative consequence argument. Our gestures are not appropriate for a moment like this; people might think that we did not take our college time seriously.

Exceptional case argument. This is not like any of our other pictures. It was our graduation, which happens only once in our lifetimes.

Multi-party Privacy: Friends Scenario II

Picture and Context



Relationship: Friends (98.3%)
Sensitivity rating: μ = 3.29 (σ = 1.16)
Sentiment rating: μ = 3.82 (σ = 1.11)

Description

Three friends, Santosh, Arun, and Nitin, decided to perform some stunts on a motorcycle. Unfortunately, while performing a stunt, Arun and Nitin had a minor accident. Santosh took the picture below at that very moment. Santosh wants to upload the picture to his social media account.

Arguments

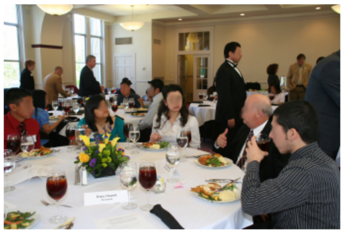
Positive consequence argument. Fortunately, none of us got hurt. This picture makes anyone who sees it laugh out loud.

Negative consequence argument. People looking at this picture may think that we are reckless drivers, which is not true.

Exceptional case argument. Motorbike stunts are not something we do everyday.

Multi-party Privacy: Colleagues Scenario I

Picture and Context



Relationship: Colleagues (94.4%)
Sensitivity rating: μ = 1.77 (σ = 1.10)
Sentiment rating: μ = 2.83 (σ = 0.92)

Description

Maria, Bonita, and Felipe, three junior employees in a company, attend a business lunch in which they meet their seniors. One of the other employees took the following picture and sent it to Maria. Maria wants to upload the picture to her social media account.

Arguments

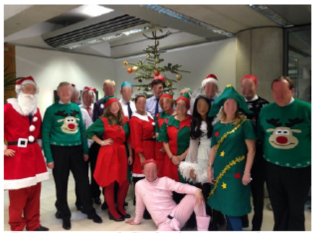
Positive consequence argument. This picture shows that we are making good progress in our careers.

Negative consequence argument. This was a professional event and our seniors might want to keep it private.

Exceptional case argument. This is an exceptional event since we attended a professional party for the first time.

Multi-party Privacy: Colleagues Scenario II

Picture and Context



Relationship: Colleagues (92.9%)
Sensitivity rating: $\mu = 3.26$ ($\sigma = 1.41$)
Sentiment rating: $\mu = 2.46$ ($\sigma = 1.50$)

Description Jerry, Laura, and Sabrina work together in a company. They were asked to attend the Christmas party dressed. However, a guy in their company (the one in pink dress) brought the whole dressing to a new level. They took the following picture at the party. Jerry wants to upload the picture to his social media account, a few days after the party.

Arguments

Positive consequence argument. People think that I have a boring life because I work at a boring place; this will prove them wrong.

Negative consequence argument. This is embarrassing; people will pick on us because of this picture.

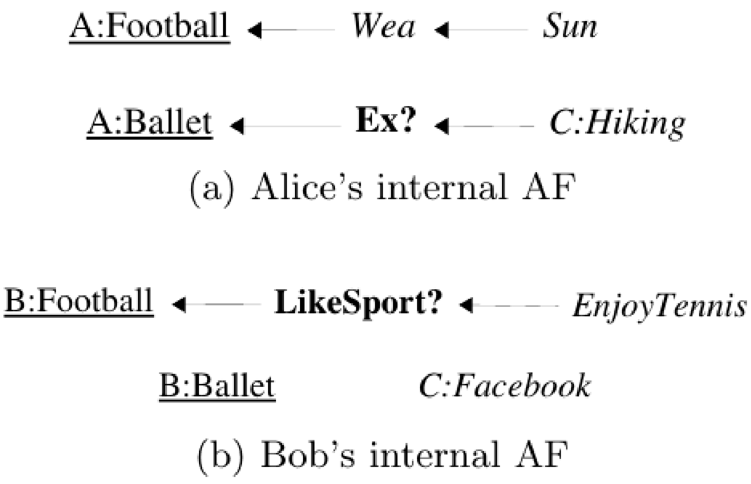
Exceptional case argument. This is an exceptional event since a Christmas party happens only once a year.

Argumentation Frameworks

- An Argumentation Framework (AF) is a pair <Arg, Att>
- Arg: Set of arguments
- Att $\subseteq$ Arg X Arg: Attacks between arguments
- Represented as a graph

Gao et al. Argumentation-Based Multi-Agent Decision Making with Privacy Preserved, Autonomous Agents and MultiAgent Systems Conference (AAMAS), pages 1153–1161, 2016

Argumentation Example: Decide on Activity



(a) Alice's internal AF

(b) Bob's internal AF

- Alice prefers going to the ballet over watching football
- Bob prefers the opposite

Privacy Preserving Strategies

- Come up with a strategy to meet certain desired properties
  - Both go to the ballet
  - Both watch football

- Feasible: Assigned action should be doable for agent
- Acceptable: All constraints should be satisfied
- Socially optimal: Ideal preferences are complied with
- Privacy preserving: Only necessary information is disclosed

Argumentation Dialogue

- Alice (defender) puts forward argument “Hiking” for “Ballet”
- Bob (challenger) attacks “Hiking” with “Facebook”
- Alice has no more moves
- Thus, “Ballet” is not feasible

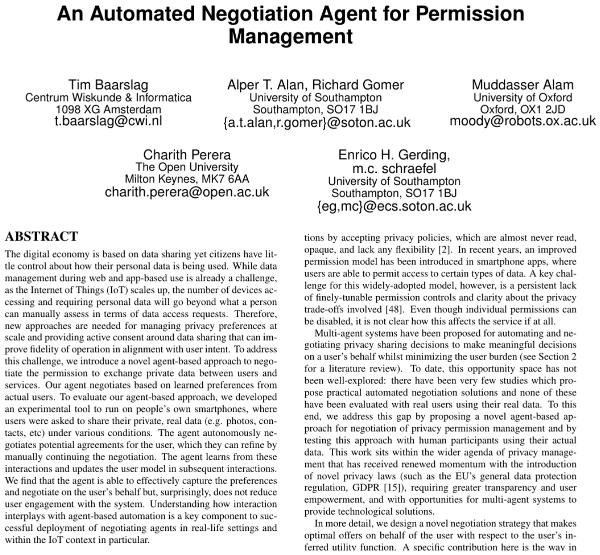
- Bob (defender) puts forward argument “Sun” for “Football”
- Alice (challenger) has no more moves
- Thus, “Football” is feasible

Resolving Privacy Disputes

- Generate facts and assumptions from an ontology
- Enrich ontology by requesting new information
- Decide whether a content should be shared

Kökciyan et al. PriGuard: An Argumentation Approach for Resolving Privacy Disputes in Online Social Networks. ACM Transactions on Internet Technology, 17(3): 27:1–27:22, 2017

Negotiation Agent for Permission Management



An Automated Negotiation Agent for Permission Management

Tim Baarslag
Centrum Wiskunde & Informatica
1098 XG Amsterdam
t.baarslag@cwi.nl

Alper T. Alan, Richard Gomer
University of Southampton
Southampton, SO17 1BJ
{a.t.alan,r.gomer}@soton.ac.uk

Muddasser Alam
University of Oxford
Oxford, OX1 2JD
moody@robots.ox.ac.uk

Charith Perera
The Open University
Milton Keynes, MK7 6AA
charith.perera@open.ac.uk

Enrico H. Gerding,
m.c. schraefel
University of Southampton
Southampton, SO17 1BJ
{eg,mc}@ecs.soton.ac.uk

ABSTRACT

The digital economy is based on data sharing yet citizens have little control about how their personal data is being used. While data management during web and app-based use is already a challenge, as the Internet of Things (IoT) scales up, the number of devices accessing and requiring personal data will go beyond what a person can manually assess in terms of data access requests. Therefore, new approaches are needed for managing privacy preferences at scale and providing active consent around data sharing that can improve fidelity of operation in alignment with user intent. To address this challenge, we introduce a novel agent-based approach to negotiate the permission to exchange private data between users and services. Our agent negotiates based on learned preferences from actual users. To evaluate our agent-based approach, we developed an experimental tool to run on people's own smartphones, where users were asked to share their private, real data (e.g. photos, contacts, etc) under various conditions. The agent autonomously negotiates potential agreements for the user, which they can refine by manually continuing the negotiation. The agent learns from these interactions and updates the user model in subsequent interactions. We find that the agent is able to effectively capture the preferences and negotiate on the user's behalf but, surprisingly, does not reduce user engagement with the system. Understanding how interaction interplays with agent-based automation is a key component to successful deployment of negotiating agents in real-life settings and within the IoT context in particular.

tions by accepting privacy policies, which are almost never read, opaque, and lack any flexibility [2]. In recent years, an improved permission model has been introduced in smartphone apps, where users are able to permit access to certain types of data. A key challenge for this widely-adopted model, however, is a persistent lack of finely-tunable permission controls and clarity about the privacy trade-offs involved [48]. Even though individual permissions can be disabled, it is not clear how this affects the service if at all.

Multi-agent systems have been proposed for automating and negotiating privacy sharing decisions to make meaningful decisions on a user's behalf whilst minimizing the user burden (see Section 2 for a literature review). To date, this opportunity space has not been well-explored: there have been very few studies which propose practical automated negotiation solutions and none of these have been evaluated with real users using their real data. To this end, we address this gap by proposing a novel agent-based approach for negotiation of privacy permission management and by testing this approach with human participants using their actual data. 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, GDPR [15]), requiring greater transparency and user empowerment, and with opportunities for multi-agent systems to provide technological solutions.

In more detail, we design a novel negotiation strategy that makes optimal offers on behalf of the user with respect to the user's inferred utility function. A specific contribution here is the way in

Baarslag et al. An Automated Negotiation Agent for Permission Management. Autonomous Agents and MultiAgent Systems Conference (AAMAS), pages 380–390, 2017

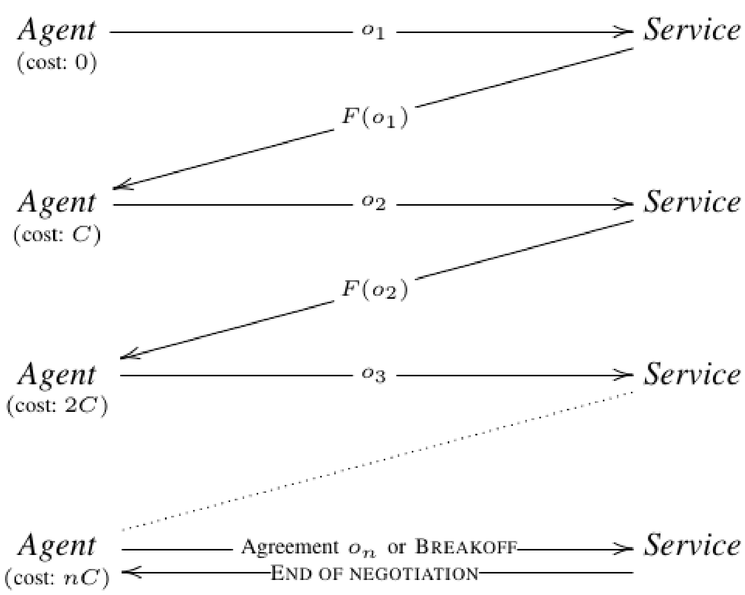
Motivation

- Number of devices/apps accessing personal data increases everyday
- People cannot keep track of all such requests
- Privacy policies: Never read, vague, and lack flexibility
- Automated methods are required to manage privacy preferences at scale
- Make meaningful decisions on behalf of the user while minimizing user burden
- Tradeoff between monetary reward and privacy

Automated Negotiation Methods

- Alternating offers protocol
  - Agents take turns to present offers
  - After an offer is made, the opponent can
  - Accept the current offer
  - Or, make a counteroffer
- Other variations of alternating offers protocol
  - Multiple issues: Price, color, performance, etc
  - Multilateral: More than two parties involved

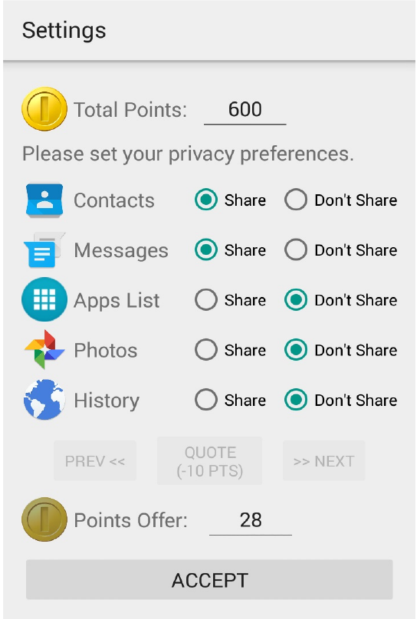
Sample Negotiation Process



Negotiation Strategy

- Utility: What the agent gains if negotiation is successfully terminated
- Objective: Maximize utility at the end of the negotiation
  - Accept an offer if gained utility is above a threshold
  - Generate counteroffer based on user's preferences and a history of offers

Negotiation Tool



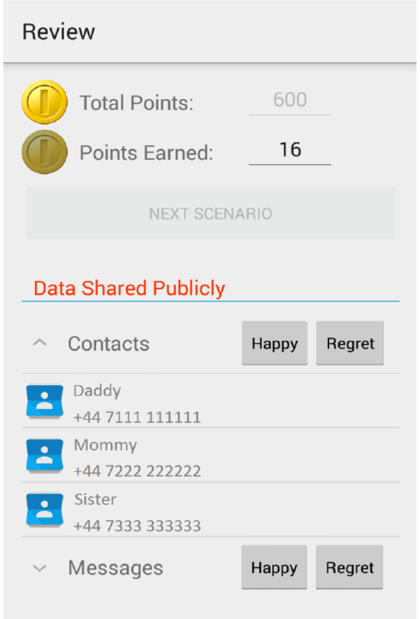
Settings

Total Points: 600

Please set your privacy preferences.

Contacts	● Share	○ Don't Share
Messages	● Share	○ Don't Share
Apps List	○ Share	● Don't Share
Photos	○ Share	● Don't Share
History	○ Share	● Don't Share

Points Offer: 28



Review

Total Points: 600

Points Earned: 16

Data Shared Publicly

Contacts

Daddy +44 7111 111111

Mommy +44 7222 222222

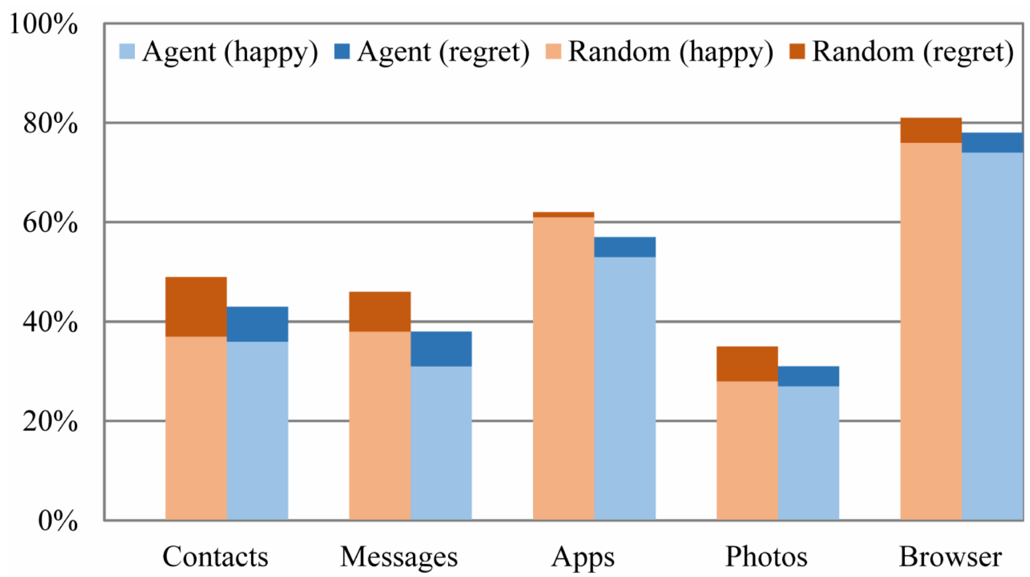
Sister +44 7333 333333

Messages

Study Setting

- 3,090 units of data (content) shared out of 343,709
- Participants: 15% Fundamentalists, 79% Pragmatists, 6% Unconcerned

Results



Limitations

- Filter bubble effect and padded room effect
- Filter bubble effect: Users on social media are disproportionately exposed to views that they already agree with
- Padded room effect: Mechanisms intended to decrease discomfort or improve safety actually prevent exploration and prevent beneficial change