The Pennsylvania State University

The Graduate School

Computer Science and Engineering Department

AN INFORMATION-TECHNOLOGY-PEOPLE INVESTIGATION OF INFORMATION PRIVACY:

THE CASE OF CHILDREN'S ONLINE PRIVACY

A Thesis in

Computer Science and Engineering

by

Nazneen Noshir Irani

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 2008

The thesis of Nazneen Noshir Irani was reviewed and approved* by the following:

Heng Xu
Assistant Professor (College of Information Sciences and Technology)
Thesis Co-Advisor

Sencun Zhu
Assistant Professor (Department of Computer Science and Engineering & College of Information Sciences and Technology)
Thesis Co-Advisor

Bhuvan Urgaonkar
Assistant Professor (Department of Computer Science and Engineering)

Mahmut Kandemir
Associate Professor (Department of Computer Science and Engineering)
Director of Graduate Affairs of the Department of Computer Science and Engineering

*Signatures are on file in the Graduate School

ABSTRACT

The current literature on privacy addresses the need for privacy, user's perceptions of privacy, privacy-enhancing solutions and other related aspects. What lacks is an attempt to present the essence of privacy in a coherent, definable manner within the context of all the related factors that affect it. Therefore, we first propose an Information-Technology-People (I-T-P) framework to integrate the existing literature on privacy and to provide a conceptual analysis of multi-disciplinary works on privacy from within the perspective of the I-T-P framework. We believe that the integrative framework through the I-T-P analysis not only provides a comprehensive list of factors to assess and understand information privacy in, but it also suggests design principles for development of effective privacy aware solutions.

The second objective of this research is to address the acute privacy challenge of protecting children's online safety by creating a tool to empower parental control over their child's personal information disclosed online. This research employs the Value Sensitive Design method to create an innovative toolkit named COP – Children's Online Privacy protection tool. The COP tool utilizes the value sensitive design approach that adopts a tripartite methodology by systematically integrating and iterating on conceptual, technical and empirical investigations of online privacy. This study reported here is novel to the extent that existing research has not systematically examined the privacy issues from the value sensitive design perspective. We believe that, using the groundwork laid down in this study, future research along these directions could contribute significantly to minimizing parental concerns for children's online safety.

TABLE OF CONTENTS

LIST OF FIGURES vii

LIST OF TABLES viii

ACKNOWLEDGEMENTS ix

Chapter 1 Introduction and Motivation 1

1.1 Internet Era 1
1.2 Online Privacy Implications 2
1.3 Objective of Study 4
1.4 Preview of the Report 5

Chapter 2 Background and Related Work 6

2.1 Need for a Framework 6
2.2 The Main Players 7
2.3 Information People (I-P) Interaction 9
2.4 Information Technology (I-T) Interaction 10
2.5 Technology People (T-P) Interaction 12

Chapter 3 I-T-P Design Principles 15

3.1 Deriving Design Principles from I-P, I-T, T-P Interactions 15
3.1.1 Degree of Concern 15
3.1.2 Degree of Protection 16
3.1.3 Degree of Convenience 16
3.2 Summary of Design Principles 17

Chapter 4 The I-T-P Framework 18

4.1 Introduction to Framework 18
4.2 Application of Framework for Development of Privacy Enhancing Solutions 20
4.3 Summary of Framework 20

Chapter 5 Analysis of Child Online Privacy 22

5.1 Introduction to Child Online Privacy 22
5.2 Child Online Privacy Protection Act (COPPA) 25
5.3 State of Art 26
5.3.1 General Solutions 27
5.3.2 Solutions Targeted Towards Children 28

5.3.3 Parental Online Consent for Kid's Electronic Transactions (POCKET)........30

Chapter 6 Value Sensitive Design........33

6.1 Introduction to Value Sensitive Design........33
6.2 Privacy as a Design Value........34
6.2.1 Contextual Nature of Privacy........36
6.2.2 Privacy as Control........37
6.2.3 Legal Expectations on Protecting Children's Online Privacy........38

Chapter 7 Technical Investigation- COP Development........40

7.1 Introduction to COP........40
7.1.1 Categorization of Websites........41
7.1.2 Logging Facility........42
7.2 System Design Overview........42
7.3 Methods for Client-Side Data Perturbation........45

Chapter 8 Implementation of COP........50

8.1 COP- A Firefox Extension........50
8.2 User Interface Design for COP Configuration........51
8.2.1 Preference Modification........60
8.3 Technical Challenges........62

Chapter 9 COP User Evaluation Study........68

9.1 Data Collection........68
9.2 Data Analysis........69
9.2.1 COPPA and Existing Solutions........70
9.2.2 COP- System Design and Technical Implementation........72
9.2.3 Installation Phase and Usability........72
9.2.4 Working Phase........73
9.2.5 COP- An effective solution?........74

Chapter 10 Analysis of COP........77

10.1 Mapping COPPA Requirements to COP........77
10.2 Mapping ITP Framework Design Principles to COP Design........78
10.3 Mapping Value Sensitive Design Principles to COP Design........80

Chapter 11 Conclusion and Future Direction........81

Bibliography........84

Appendix A Children's Online Privacy Protection Act (COPPA)........94

Appendix B COP Evaluation Material 104

B.1 Tutorial Slides 104
B.2 Handout- Installation Steps 106
B.3 Interactive Session Questionnaire 109
B.4 Web-Based Post-Tutorial Questionnaire 110

Appendix C Mapping COP to Given Requirements 111

C.1 Mapping COPPA Requirements to COP 111
C.2 Mapping ITP Framework Design Principles to COP Design 118
C.3 Mapping Value Sensitive Design Principles to COP Design 120

LIST OF FIGURES

Figure **1**: The I-T-P Framework..19

Figure **2**: Value Sensitive Design Method..33

Figure **3**: System Design Overview..43

Figure **4**: Installation Step 1- Welcome Screen..53

Figure **5**: Installation Step 2- Parental Verification..54

Figure **6**: Installation Step 3- Set Password..55

Figure **7**: Installation Step 4- Preference Setup..56

Figure **8**: Privacy Level Help Popup Window..57

Figure **9**: Installation Step 5- Child Information Form...58

Figure **10**: Installation Step 6- Notification Settings..59

Figure **11**: Tool Menu...61

Figure **12**: About Menu Option..62

Figure **13**: Log Window..62

Figure 14: Tutorial Slides..106

LIST OF TABLES

Table **1**: Paradigms regarding the Concept of Privacy 35

Table **2**: Data perturbation Approaches for Different Data Types 47

Table **3**: User Response Statistics of COP survey 75

Table **4**: Summary of COPPA Law 94

Table **5**: Mapping COPPA Requirements to COP 111

Table **6**: Mapping ITP Framework Design Principles to COP Design 118

Table **7**: Mapping Value Sensitive Design Principles to COP Design 120

ACKNOWLEDGEMENTS

First, I would like to thank my co-advisors, Dr. Heng Xu and Dr. Sencun Zhu, for providing me with all the advice, guidance and encouragement that I could possibly require during the course of my research work and thesis writing period. I am extremely grateful for their patience and support, without which I could not have been successful.

I would also like to thank my student colleague, Wei Xu, with whom I collaborated during my research, for his advice and insight.

I would also like to thank Dr. Bhuvan Urgaonkar, my committee member, for supporting my thesis and my research.

I am thankful for all the support I received from both the Computer Science department and Information Sciences and Technology department, during the course of my degree. I am especially grateful to the IST department for giving me the opportunity to serve as a teaching assistant, something that I thoroughly enjoyed doing besides research.

Finally, I would like to thank my friends and family, in particular my mother Homai Irani, for supporting my decision to pursue a Masters degree.

Chapter 1

Introduction and Motivation

> “We don’t need to learn something completely new; we need to learn to be smarter, more skeptical, and more skilled about what we already know.” – Bruce Schneier

1.1 Internet Era

The Internet era is defined as a period in the information age in which communication and commerce via the Internet became a central focus for businesses, consumers, government, and the media [**1**]. The Internet era also marks the convergence of the computer and communications industries and their associated services and products.

The internet has revolutionized the way the world and everything around works. People today find it hard to imagine a world without electronic-mails and search engines where every piece of information is just a few clicks away. The internet is not a steady environment because each day there are new applications, new software being released which changes the purpose people use the internet for. To support this statement, we can see that the internet was born out of an attempt of the U.S. Air Force to do a study on how it could maintain its command and control over its missiles and bombers, after a nuclear attack. They needed a communication medium which could survive a nuclear attack. Today as we see, it is not the sole (or if I may say, not even the primary) use of the internet.

The internet has evolved into a means of communicating on a daily basis, exchanging information, carrying out trade and commerce, streaming media, social networking among many other uses. The Internet has made possible entirely new forms of social interaction, activities and organizing, thanks to its basic features such as widespread usability and access. In democratic societies, the Internet has achieved new relevance as a political tool. It is also a large market for companies which take advantage of the low-cost advertising and commerce through the Internet, known as e-commerce. It is the fastest way to spread information to a vast number of people simultaneously. During the 1990s, it was estimated that the Internet grew by 100% per year, with a brief period of explosive growth in 1996 and 1997 [**2**].

To leap thus across the centuries and to the more recent decades, we realize in a glimpse the incredible dynamic involved in the world of information and online technologies. Information from around the globe is available today in the air we breathe, some of it, essential to our survival.

The pervasive spread of computer networking has had numerous widespread effects. Like earlier periods witnessing the rise of radio, and then television, the birth of the Internet era has generated extensive speculation about the potential consequences of this development towards people's daily lives. The internet era has a lot to offer and the opportunities are endless but with this it brings the possibility of this very technology being the biggest threat to the safety of mankind.

1.2 Online Privacy Implications

> "Commerce, bandwidth, equity of access, copyright- these are some of the battles that are taking place in the internet environment, but the one that will affect all of us is the battle over privacy." - Steve Cister, Advanced Technology Group, Apple Computer

Privacy is the capacity to negotiate social relationships by controlling access to information about oneself [**3**]. Recent changes in the realm of technology and privacy have created a landscape that is both dangerous and encouraging.

New technology has always challenged the way people thought about privacy and the conceptualization of privacy. The Internet era brings an on-rush of changes, both revolutionary and subtle, to the field of privacy—changes to the very concept of privacy;

changes to the mindset of scholars who study privacy; and an effect on the effort put towards protecting people's privacy.

In recent decades there has been a steady growth in the use and manipulation of vast quantities of personal data supported by computing technology. Though advancement in technology is essential to the progress of the world today, it has opened up new avenues to misuse and exploit this information that people may consider as private data about themselves.

Growth of information technologies has enhanced the ability for organizations to store, process, and share personal data. Privacy is thus at the center of discussion and controversy among multiple stakeholders including IT professionals, business leaders, privacy activists, government regulators and the common man. Information privacy continues to be eroded as a result of technology innovations. Unanticipated release, and subsequent use or misuse of personal information is of particular concern. Privacy is now of utmost concern to everyone and the need to protect it is being felt more than ever.

Every week brings new headlines of yet another major identity breach or theft: 320,000 voter records stolen in Nashville, 260,000 social security numbers released in Wisconsin and 45.7 million credit and debit card records stolen from a national retail chain, the fact being that another identity is stolen every three seconds! The question is what exactly is making these thefts so simple and the answer is obvious. The capabilities of the internet are being exploited to gain easy access to personal information.

The five most popular search engines routinely archive a user's search terms, their computer's address, and the unique identifier for their Web browser for 13-18 months. Web 2.0, or the participatory web, has also raised new questions about the definition of "personal information" and the impacts it will have over privacy. More generally, consumers are now expressing a more consistent interest in control over personal information. A recent poll showed that 59% of adults have refused to provide information to a business or company because they thought it was not really necessary or was too personal [**4**]. People however continue to upload pictures to Flickr, personal profile information to Facebook, choosing to connect their online identities with pieces of their personal information.

The privacy implications of these actions of people who are unaware or choose to be ignorant can be numerous and in some cases extremely serious. Identity theft for example can cause a person to lose their money, social status and in the most extreme case their life. The possibilities are not bounded which makes protecting this information more important than ever. The privacy implications of having our information readily available online and the computing and communication power of the internet is something that needs immediate attention. Within the field of online privacy itself, there are a number of areas which remain unexplored. For example, the rising number of child abductions because of more and more children accessing the internet is a problem which is very real and requires attention.

In summary, the internet has improved the lives of most people and has had more positive impacts than negative. However, technology is growing at a fast pace and so is the value of personal information. Online privacy is an area which continues to remain unsafe. The internet has not been designed with issues like security and privacy in mind and thus effort is essential to study this problem and develop solutions before it is too late.

1.3 Objective of Study

The objective of this study is to analyze the current state of art of online privacy and specifically study the case of child online privacy. Developing a solution to the problem of child online privacy is the major objective. To realize this objective, a detailed study of the privacy literature was carried out and a conceptual analysis of privacy in the internet era was performed. The sub-objectives were defined as follows;

- To develop a framework that provides a holistic view of the literature on privacy
- Derive design principles based on the framework
- Study the state of art in protecting child online privacy and the related laws
- Implement a solution to protect child online privacy

- Provide an evaluation and analysis of the proposed solution

1.4 Preview of the Report

Chapter 1 is an introduction to the report and clarifies the objective of the study. The report is further divided into two main parts.

Part I introduces the ITP framework to the user and is divided into three chapters. Chapter 2 provides background and work related to the field of privacy. It explains the need for a framework and the interactions between information, technology and people. Chapter 3 is where design principles are derived based on the material discussed in Chapter 2. The ITP framework is introduced in Chapter 4 and the applications of the framework are explained in detail.

Part II is a detailed guide through the implementation of a tool for protecting child online privacy. For developing the tool, it is essential to analyze the current literature on child online privacy and Chapter 5 provides the background required to understand the same. Chapter 6 introduces the value sensitive design model used as guidance in the development of the tool. Chapters 7 and 8 are related to the implementation of COP as a Firefox extension and the detailed working of the tool along with all technical details and challenges faced during the development process. User evaluation study and the analysis of COP based on the models studied throughout the report are an important part of the study to prove the viability of COP as a tool for protecting children’s privacy online. Chapters 9 and 10 are dedicated to the evaluation studies and the analysis of the tool. Finally, Chapter 11 is a conclusion of Part II and highlights suggestions for future research.

Chapter 2

Background and Related Work

2.1 Need for a Framework

Vast amounts of information and services have caused the internet and its online resources to grow tremendously. Internet has been growing at a rapid rate since its conception, on a curve geometric and sometimes exponential [**5**]. In an environment where information flows among people in a seamless, sometimes unpredictable manner, the notion of privacy protection needs to be overhauled and understood clearly. In addition, the recent focus on national security and fighting terrorism has brought with it new concerns about governmental intrusions on personal privacy.

Privacy has been defined as 'the right to be left alone' [**6**]. With the growth of social networking sites, the growing popularity of blogs and increase in the number of people having an online identity, people are living a parallel online life which needs to be protected. The right to be left alone is no longer sufficient to describe privacy and the requirement to protect privacy has an added new dimension to it. It is necessary to assure people their privacy on the internet as the threat is no longer limited to online privacy but to the real identities of people. Growing information about people on the internet and the difficulty of keeping track of one's personal information is causing privacy breaches whose consequences extend beyond the threat to their online existence.

In short, lack of privacy in online environments is the major threat to people and finding a solution to this real threat is important today. The current literature addresses the need for privacy, user's perception of privacy, privacy-enhancing solutions and other related aspects. What lacks is an attempt to present the essence of privacy in a coherent, definable manner within the context of all the related factors that affect it. The Information-Technology-People (I-T-P) framework is a proposal to integrate the existing literature on privacy and to provide a conceptual analysis of multi-disciplinary works on privacy from within the perspective of the I-T-P framework. The integrative framework through the I-T-P analysis not only provides a comprehensive list of factors to assess and understand information privacy, but it also provides suggestions and directions for future research.

In order to address privacy issues, the aim is to: firstly integrate the existing literature on privacy and privacy concerns and introduce an integrative framework to study privacy concerns of people on the internet; secondly, define design principles derived from the study of this novel framework; and finally make an effort to develop a technical solution to protect children's online privacy using the principles set by the framework.

2.2 The Main Players

With privacy being clearly established as a concern, we need to approach the issue by first identifying and defining the factors that play the most important role in context of the internet. The main 'players' to understand the privacy issue are; (i) the people who use the technology; (ii) the information that is being captured and processed and; (iii) the technology itself.

After understanding these players, we try to analyze works that focus on the interactions between them. Finally we introduce and explain the Information-Technology-People (I-T-P) framework to understand the privacy issue and suggest solutions.

While analyzing the first player which is information, it can be broadly categorized into Personally Identifiable Information (PII) and non Personally Identifiable Information (non-PII) [**7**]. Data that uniquely identifies an individual is called as personally identifiable information (e.g., social security number, birth date, etc). On the other hand, information such as age, gender and personal interests are considered as non-PII as many people may share the same values for these and so the information cannot uniquely identify any particular individual.

Disclosure of information can raise huge concerns among people. The world has witnessed that every new technology that is introduced brings with it issues that may not be obvious upfront. Privacy of information was not seen as a major concern at the conception of the internet era. However today it is the area attracting a major portion of the attention.

Technology, the second player, is important for the analysis of information privacy since it is the technology that brings to the forefront the privacy concern in the first place. We classify technology based on four characteristics that can help analyze various situations and provide comprehensive dimensions to understand possible solutions to the privacy problem:

- Data capture: This includes the technologies which capture data from users.
- Data storage: This includes the technologies which not only collect but also store the information that is captured.
- Data analysis and integration: With advancing technology, it is now possible to derive results that are not directly captured by technology. This category includes technologies which analyze the gathered information and integrate results from various sources.
- Data dissemination: This refers to distribution of the information that is captured or stored.

All people do not have uniform privacy concerns. People are considered the third major player. To support the fact that people's privacy concerns vary, Westin and Harris [**8**] established categories of users with varying degrees of privacy concerns. They are:

- Privacy Fundamentalist: People who are extremists in their concern for privacy. These people would be extremely reluctant in giving away their personal information for fear that they may be affected adversely.
- Pragmatic Majority: People who are concerned about their privacy and provide personal information only under certain circumstances. Privacy pragmatists have a balanced opinion about privacy.

- Marginally Concerned: These people are marginally or almost unconcerned about their privacy and personal information and disclose information readily on demand. They display a perfunctory attitude towards disseminating their personal information.

The categorization of user's perceptions of privacy is very important, if not the most critical factor in accurately understanding the adequate requirement for a system built for privacy.

In order to study the aspect that relates each of these factors, it is the interactions between these players that need to be studied. In reviewing the literature, we find that many works have focused on the interactions between two or more of these players and formed conclusions based on their studies.

2.3 Information People (I-P) Interaction

To study the interaction between people and information, it is necessary to understand that privacy is a subjective concept and there is essentially a need to customize it based on user’s requirements. Privacy cannot be considered as ‘one size fits all” [**9**]. Concerns of people may vary based on the type of people as well as the sensitivity of the information in question. A wider variety of factors may influence a user's perception of privacy thereby making him or her, a privacy fundamentalist, unconcerned or a pragmatist user. Culture for example is such a factor [**10**].

Misplaced concerns amongst the users could cause an incorrect analysis of a particular privacy concern thereby leading to an inefficient design of the privacy enhanced system. In the context of their information, people usually have two kinds of privacy concerns [**11**]. People are first, concerned over the unauthorized access to their personal data through security breaches or lack of internal controls. Second, people are concerned about the secondary usage of their data referring to reuse of personal information for other undeclared purposes without consent of user. The act of collecting user’s data and their inability to correct errors is an additional concern that people might have [**12**].

Sensitivity of the information is an important factor that shapes people’s outlook towards privacy. Highly sensitive data like PII may raise higher concerns as compared to non-sensitive data like non-PII. The sensitivity of information should further be studied in the context of user’s familiarity to the party the information is being disclosed to. The dynamics of privacy can often change based on the entity concerned with it. To illustrate this, consider persona information that identifies the existence of an entity, and activity information that is information about the subject directly involving the person. Sensitivity of persona information has a direct relationship with the subject's unfamiliarity with the observer while unfamiliarity decreases the sensitivity of the subject's activity information because of anonymity [**9**]. A study showed that when presented with scenarios involving the provision of personal data to Web sites, respondents were much less willing to provide information when personally identifiable information was requested [**13**]. These findings suggest additional dimensions to be considered to conceptualize privacy, specifically in relation to information and people. Implications of privacy would vary depending on the comfort levels and trust between the subject and the observer.

The context in which the information is collected is another factor that plays a major part in determining the comfort level people have in giving out their information. It can be different in a workplace than in a more public sector. In some cases simple qualms or bad experiences prevent people from sharing their information [**12**]. Phelps et al. [**14**] found that people are more willing to disclose demographic or lifestyle information than financial or personal identification information. Another study states that for the design of personalized web-based systems, highly sensitive data should never be requested without the presence of mitigating factors [**7**].

2.4 Information Technology (I-T) Interaction

Technology today is able to capture, store, process, analyze and integrate, and disseminate information. It is important to understand that the term information cannot be generalized because privacy concerns vary depending on the sensitivity and context of usage of the data. The study of the interaction between information and technology can lead to useful conclusions in determining features that can be included in future applications to ensure privacy of all kinds of information in the most suitable manner.

Technical mechanisms for protecting privacy have been classified into four categories namely, encryption and security mechanisms, anonymizing mechanisms, infrastructures, and labeling protocols [**15**]. Privacy enhancing technologies (PETs) include anonymizing and de-identifying mechanisms and can be very useful for protecting privacy as opposed to having only encryption and security mechanisms which can be considered necessary but not sufficient for privacy protection.

Privacy enhancing technologies (PETs) refer to technical and organizational concepts that usually involve encryption in the form of (e.g.) digital signatures, line signatures or digital pseudonyms [**16**]. PET concepts are classified into four categories as stated below. These are based on the degree of interactions as actions between subjects relating to objects within a system.

- Subject oriented concept: It refers to dissociating user identity from a transaction or existing data. Proxies can be viewed as one such example where user identity is hidden by assigning individual identifiers that are untraceable with respect to a transaction.

- Object oriented concept: Many transactions involve exchange of the objects and many of these objects may carry user identity data. Ensuring that each object carries no traces of users and at the same time ensuring that the object is not eliminated is the central idea. Using cash instead of the credit card is the perfect example since it carries no traces of subjects.

- Transaction oriented concept: This refers to a trace that is left behind by the transaction procedure without being related to the object being exchange. Video-taping a cash over-the-counter dealing at a bank is a tracing process.

- System oriented concept: This creates to what is called as ‘zones of interaction’ where subject identities are hidden, objects do not bear information of subjects and transaction records are not created or maintained.

In addition, there are three trends which are relevant to future development of the concept of privacy [**16**]. These are:

- Information Balance: Designing PETs has to include individuals and their preferences more directly.
- Identity: PETs should be able to maintain the identity of users.
- Trust: These systems need to be reliable and safe within a given probability.

Providing adequate security is also an important design principle and certain encryption techniques and other such mechanisms should be put in place in order to enforce it.

2.5 Technology People (T-P) Interaction

Key to understanding privacy-aware technology is a fundamental grasp of how people use technology, how they understand it and how they make meaning out of it. The relationship between people and technology in reviewing privacy would provide profound insights on whether different interactions with technology can yield privacy enhancing approaches in designing systems.

Certain policies can be introduced for protecting people's privacy but there is an inherent tradeoff between the ease of use of technology and the detailing of policy terms [**17**]. To study people's usage of patterns in relation to different technologies, it is first essential to understand the people factor. This is primarily due to the fact that people wish to project themselves differently with respect to different groups, persona and institutions [**18**]. Thus once the people factor is studied and understood, technologies can be made flexible enough to deal with human nuances and ambiguities.

There are many factors that can be held responsible for a user's attitude towards a given application. As an example, five discrete contextual factors namely, society, government, industry, company and media influence a user's perception and are likely to be essential to an understanding of prospective adopter's attitudes and thus the actual behavior towards a given application [**19**].

Assessing precisely the nature of the application and its potential for abuse, designing for privacy mandates an understanding of the type of interaction involved during the use of the application itself [**20**]. This may involve considering the actor relations and the environment for use. For example, applications used in a closed environment and which are primarily single user centric may require straightforward and simplistic types of protection such as opt-in or spam filter protections to prevent abuse from other sources. Based on this, design considerations can be prioritized while designing privacy aware applications.

This could also reduce the overall development cost of developing such applications by avoiding un-necessary sophistication and complexity. Therefore it is essential to understand the task for which the system or technology is being developed. Technology that serves more than promised can sometimes be complicated to use from the user's perspective while technology that delivers less than required can be considered not useful. Junglas and Watson [**21**] have described this phenomenon by categorizing technologies as underfit and overfit technologies.

It is also essential that privacy aware system should minimize the asymmetry of information by decreasing the flow of information from data owners to data collectors and increasing the flow of information from data collectors back to data owners [**22**]. There are certain other design principles to guide the design of privacy aware systems which add value to the user [**23**]. First principle is to provide notice to the user or have a way to announce that data is being collected. Giving the user a choice is another design principle.

Access and recourse is the other feature that needs to be incorporated into the design of systems. It may require data collectors to collect only minimum required data, for well defined purposes, and keep it only for as long as required. Also, users should be given access to their data after it has been collected giving them enhanced control over their personal data.

It has been concluded that feedback and control are two design issues that can assist in gaining user confidence that their privacy is being protected[**9**]. If the user is kept aware of the manner in which his personal information is being handled even after it is submitted and if the user is given more control over his information, there is a sense of enhanced privacy.

The user can easily and less reluctantly accept technology that meets his needs and at the same time provides a guarantee of protection of his or her privacy. Thus studying the user's comfort level with respect to the technology can promote useful and successful implementations of technology to protect privacy on the internet.

Chapter 3

I-T-P Design Principles

3.1 Deriving Design Principles from I-P, I-T, T-P Interactions

The literature related to the interaction between the players; information, technology, people, has been highlighted and studied. It is now essential to analyze it further and make conclusions which can prove useful for designing technology.

The interactions provides a means of studying the important factors to be considered when designing privacy enhancing technology. These design principles have been categorized and discussed below.

3.1.1 Degree of Concern

Design principles derived from the interaction between information and people fall within this category. They highlight the varying degree of concern that different people may have regarding their information and suggest design principles that should be accounted for to mitigate this concern. The design principles can be stated as follows;

- Systems should be customizable to accommodate for different kinds of people.
- Users should be given control over their data. They should be allowed to change it, view it and unauthorized access to their data should be denied.
- Systems should be configurable for data of varying sensitivity levels.
- The context of disclosure and the party to which information is being disclosed are important to consider.

These conclusions can be viewed as the design principles governing the Degree of Concern.

3.1.2 Degree of Protection

These design principles are derived from the interaction between Information and Technology. They highlight the way that current technology handles information and is used to suggest design principles that should be accounted for mitigating the risks to information privacy. The design principles can be stated as follows;

- It is essential to understand the ideal category of PET.
- It is necessary to achieve information balance, maintain user's identity and achieve a degree of reliability and safety.
- Security is an important feature of systems for privacy protection.

These conclusions can be viewed as the design principles governing the Degree of Protection.

3.1.3 Degree of Convenience

These design principles are derived from the interaction between People and Technology. They highlight the way that people view current technology and the different concerns they might have with the usability and convenience of use of the technology. It is further used to suggest design principles that should be accounted to enhance the usability of the technology. These can be stated as follows;

- Consider the environment in which information is being collected and the actor relations.
- Try to avoid developing overfit and underfit technologies.
- Users should be provided notice on disclosure of information, sufficient choices for disclosure and allowed to access their information after disclosure.
- Feedback and control are necessary to be accounted for while developing a system.

These conclusions can be viewed as the design principles governing the Degree of Convenience.

3.2 Summary of Design Principles

In the above section, design principles have been derived and summarized based on the interactions studied throughout the previous chapter between the various players namely, information, technology and people. These three sets of design principles highlight the important factors to be considered in order to mitigate the degree of concern, and enhance the degree of protection and convenience.

In this research study, the design principles have been used in the context of protecting child online privacy and develop a corresponding tool. All design principles have been accommodated to arrive at a solution that fits best with the requirements of the I-T-P framework.

Chapter 4

The I-T-P Framework

4.1 Introduction to Framework

We have identified three main players, information, technology and people that need to be studied to understand privacy. Based on the interactions between these players, we have highlighted certain design principles which need to be accommodated for while designing a privacy aware system. The ITP framework relates these three players in a way that can assist in better understanding certain privacy enhancing features that need to be accommodated in the applications of the future. It relates the conclusions about the design concepts to finally list the design principles that add value to a system designed for privacy protection.

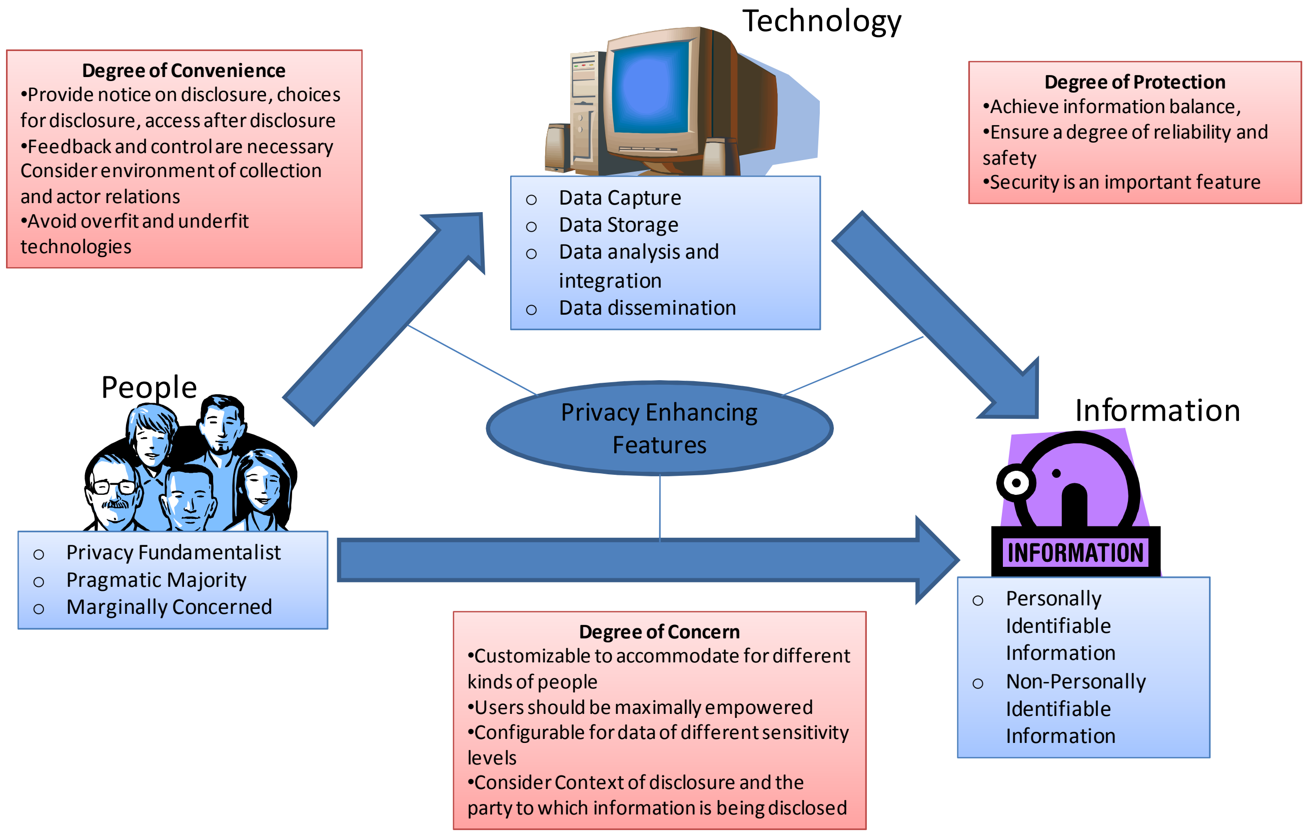
Figure **1** is a visual representation of the ITP framework. The main players, technology, people and information have been depicted with their categorizations as discussed earlier. After understanding these players individually, it is their interactions with each other that are important to analyze rather than independently studying them. These interactions can be explained as follows;

- Degree of Concern: People are very concerned about the privacy of their personal information and based on the kind and sensitivity of information, and also on the type of people in question, the degree of concern of people might vary. The design principles essential for achieving this balance have been highlighted.

- Degree of Protection: Technology that handles information is responsible for maintaining the privacy of that information. This is the information- technology interaction explaining how different technology should handle different kinds of information in certain ways. This interaction can help us deduce certain design principles as shown.

- Degree of Convenience: People use technology that serves some purpose for them and makes their tasks simpler. Therefore it is necessary that the technology being developed for users understands the needs of the users. The technology should be useful in a sense that it should serve the purpose and deliver results. Also, it should not be complex so that it would deter people from using it. The design principles essential for achieving this have been highlighted.

Figure **1**



Degree of Convenience
- Provide notice on disclosure, choices for disclosure, access after disclosure
- Feedback and control are necessary Consider environment of collection and actor relations
- Avoid overfit and underfit technologies

Degree of Protection
- Achieve information balance,
- Ensure a degree of reliability and safety
- Security is an important feature

○ Data Capture
○ Data Storage
○ Data analysis and integration
○ Data dissemination

○ Privacy Fundamentalist
○ Pragmatic Majority
○ Marginally Concerned

○ Personally Identifiable Information
○ Non-Personally Identifiable Information

Degree of Concern
- Customizable to accommodate for different kinds of people
- Users should be maximally empowered
- Configurable for data of different sensitivity levels
- Consider Context of disclosure and the party to which information is being disclosed

Figure **1**: The I-T-P Framework

From the framework, information privacy can be viewed as concerns that different people including privacy fundamentalists, privacy pragmatists and those that are marginally concerned have over the unauthorized capture, storage, analysis and integration and dissemination of personally identifiable as well as non-personally identifiable information. This framework provides guidelines to assist researchers and developers to build effective privacy enhancing technologies.

4.2 Application of Framework for Development of Privacy Enhancing Solutions

I-T-P framework provides a holistic approach for analyzing and reviewing works on privacy from within a framework, integrating three key factors, information, technology and people. Interactions among these factors when studied can provide a better understanding of the issues and can assist in formalizing solutions to the privacy problem. Furthermore, it aids in providing a larger picture in a field that often focuses on singular concepts. I-T-P framework is an integrative framework that can offer developers and researchers with a guideline to define a structure for incorporating privacy by design and thus enable them to create privacy-aware applications that reduce the concern of consumers. The objective is to provide a conceptual analysis of the diverse works on privacy from within the perspective of the I-T-P framework.

Though people, technology and information have been identified as the main players, their interaction is what guides the design principles and not their nature alone. The literature concentrating on the interactions have been utilized and studied to arrive at design principles which are essential while developing technology to guarantee privacy. These design principles provide the holistic view of the factors that are being dealt with when developing any technical solution. Essentially, they provide the view of things to be factored while arriving at good solutions. These design principles have been used to guide the research stated herein and develop a tool for protecting children’s privacy to demonstrate the effectiveness of this framework.

4.3 Summary of Framework

By analyzing the ITP framework, future research could concentrate on finding privacy enhancing features integrating the conclusions drawn from the different kinds of works discussed in this paper. The question that might need to be answered is whether certain privacy features and practices could be identified by viewing the works on privacy from within the ITP framework. Certainly a comprehensive study of the framework and its implications might be needed for this purpose.

In conclusion, the integrative ITP framework not only provides a comprehensive list of factors to assess and understand privacy concerns, but it also provides suggestions and directions for future research. The framework provides guidance in two ways; first, it identifies and elaborates on the privacy factors and the relevant players whose interactions need to be understood; second, the framework presents the multi-faceted nature of privacy the issues that need to be considered when suggesting privacy enhancing solutions for the technologies by suggesting design principles.

Chapter 5

Analysis of Child Online Privacy

5.1 Introduction to Child Online Privacy

Technology today can be viewed as a double edged sword because it provides children with a whole new world of information and great opportunities to learn, but at the same time is a major threat to their privacy and safety. Children face many risks online and growing cases of child abuse, online predators, growing child pornography industry and other such matters are causing great concern. Website operators too are interested in collecting children's personal information such as name, e-mail address, home address and telephone number in order to gather statistics for marketing and other related purposes. Recognizing that children of ages below 13 years are not well equipped to make intelligent choices while surfing the web or to make informed decisions before disclosing their private information to websites, the U.S. Congress enacted the Children's Online Privacy Protection Act (COPPA) [Refer to Appendix A]. COPPA governs the online collection of the personal information of children under age 13 and spells out what a Web site operator must include in a privacy policy, when and how to seek verifiable consent from a parent and the responsibilities an operator has to protect children. Though the law has been enacted, few technical solutions have been developed to implement this law making it difficult to achieve the ultimate goal of protecting children's online privacy.

The number of children accessing the internet is constantly on the rise and protecting their privacy is becoming a major challenge. By nature, children are more gullible than seniors and prove to be naïve in their decisions. Nearly half of teens (47%)

aren't worried about others using their personal information in ways they don't want.[1] Operators online exploit this factor by luring children to attractive prizes and offers in exchange of private information. Even registration procedures require children to submit personal information in most cases. Internet is even to be blamed for the rise in child porn as the offenders have the resources to remain anonymous online while children reveal their information.[2]

COPPA is a thorough law governing websites' practice of collecting personal information from children. It clearly states what a Web site operator must include in a privacy policy, when and how to seek verifiable consent from a parent and the responsibilities an operator has to protect children of ages below 13 years. Though the law has been enacted, the effect that this law has had so far in protecting children's privacy online is very limited due to various factors.

Enforcement is thus seen as the main reason that online child abuse and other threats to child privacy are still rampant. Children's organizations around the world are urging technology companies to make the internet safer for children.[3] Limited technical solutions have been researched to help enforce requirements of COPPA. This has caused websites to conveniently disregard the requirements of the rule while dealing with children.

The largest penalty that a company had to pay for allegedly violating the COPPA requirements was the social networking website, Xanga.com [**24**]. According to the FTC, Xanga.com collected, used, and disclosed personal information from children under the age of 13 without first notifying parents and obtaining their consent. Though the biggest case, it is certainly not the only case of violation of COPPA by a website and a number of such cases have been reported. A more recent case was filed against Imbee.com which was promoted as a free, secure, social networking and blogging destination specifically designed for children ages 8 to 14 [**25**]. According to the FTC, imbee.com collected and maintained personal information from children under the age of 13 without first notifying

[1] http://www.webwisekids.org/index.asp?page=statistics ((Teen Research Unlimited."Cox Communications Teen Internet safety Survey Wave II," March 2007)

[2] http://news.bbc.co.uk/1/hi/technology/3387377.stm (Net blamed for rise in child porn)

[3] http://news.bbc.co.uk/2/hi/uk_news/4454355.stm (Safety urged for child web users)

parents and obtaining their consent. Imbee enabled more than 10,500 children to create imbee accounts by submitting their first and last names, dates of birth, personal e-mail addresses, parents' e-mail addresses, gender, user-names and passwords prior to the site's providing notice to parents or obtaining their consent. The settlement included a $130,000 Civil Penalty and clearly shows that disregarding the law is possible.

Studies indicate that most parents are not aware of laws that can protect their children's privacy online [**26**], though all of them consider it a necessity to enact such laws. Parents in most cases do not take the extra effort of learning or researching such laws unless absolutely required or unless a problem occurs. Therefore, instead of expecting parents to understand and take advantage of these laws, there is a need for solutions that will take care of this responsibility for them. The task needs to be delegated to technology itself for handling responsibilities that parents are either not aware of or just do not find a reason enough to cooperate.

No effective solution, so far, has been introduced but instead, ample stress has been given to spreading awareness through education and to building child-parent relationship [**27**]. The belief that education can help reduce the threat related to children disclosing their personal information online is certainly a good step. However, it cannot serve as a comprehensive solution to the problem. For example, through education, parents can be informed of the existence and importance of following laws such as COPPA but with the lack of availability of techniques and tools to materialize this objective, the task seems improbable.

Children as discussed can be naive and expecting them to provide false information to websites again is a faulty expectation. Children below 13 years of age should not be encouraged to lie about their information but the responsibility should be materialized by a technical solution. In some cases, children may be tempted into indicating a false age greater than 13 years to gain access to certain objectionable material. This being the case, the website operator cannot be held responsible. A sound client side check should be implemented to account for this fact and not permit children to indicate a false age. Protecting children's innocence and at the same time protecting their privacy is a challenge.

Another challenge is obtaining a balance between accessibility and privacy. What this means is that, as discussed later, blocking information is not a solution as blocking information submitted during registration may cause the registration process to fail and prevent a child from gaining access to required online material.

In summary, enforcing COPPA and making sure that all websites abide by the rule is a difficult task which cannot be achieved by relying on website operators alone. Parents should be given complete control of their child's privacy preferences but at the same time should not be burdened with additional tasks which can make such laws unattractive. Relying on children again is not an option and asking them to lie about information is socially not a healthy practice. This shows that there is requirement for an automated and technical approach to protect children's online privacy while requiring minimal parental involvement. One important requirement of the tool would be that it should allow access and registration to all sites, not disclose true personal information and yet allow websites with data from which meaningful statistical models can be built.

This part of the study is focused towards the very goal of finding an automated and technically viable solution for enforcing COPPA.

5.2 Child Online Privacy Protection Act (COPPA)

COPPA was enacted in response to the growing danger to the online privacy of children. The stated goals of the Act are: (1) to enhance parental involvement in their children's online activities in order to protect children's privacy in the online environment; (2) to protect the safety of children at places in the online environment such as chat rooms, home pages, email accounts, and bulletin boards in which children may make public postings of identifying information; (3) to maintain the security of children's personal information collected online; and (4) to limit the collection of personal information from children without parental consent.[4]

[4] 144 Cong. Rec. S12741 (Oct 7, 1998) (Statement of Sen. Bryan)

COPPA specifically states that, "*Prior to collection, use, and/or disclosure of Personal information about a Child, an operator must obtain from a Parent of the Child verifiable Parental consent*". The method of collecting the verifiable parental consent has been an issue of debate for long. At the time the law was enacted, no inexpensive and viable technological means existed that simplified this process. Therefore, the FTC adopted a sliding scale approach for obtaining verifiable consent and this was declared to be a temporary approach assuming that more sophisticated methods would be introduced through advances in technology.[5]

This approach offers two different levels of parental consent. There is a more reliable method which suggests a print consent form, a credit card transaction or a toll free telephone number maintained by trained staff. The approach called Email Plus is the less reliable approach and is reserved for Web sites

that use children's personal information for internal uses only. This method requires obtaining the parent's consent through e-mail and then verifying that the person providing the consent is in fact the parent using an additional step. Recently the sliding scale approach has been adopted as the permanent approach for obtaining the consent from parents since technology has not served to provide any better options.

5.3 State of Art

Most of the technical solutions which exist are focused towards protecting online privacy of users, in general. As discussed, COPPA has very specific requirements for protecting children's privacy and no tools have been successfully implemented to address the rules. We will discuss tools in general for online protection of user's privacy and then efforts which have been targeted at protecting children's privacy.

[5] http://promomagazine.com/specialreports/COPPA-Sliding-Scale/ (New Study Shows Rise in Youth Exposure to Porn & Cyberbullying, Decrease in Online Solicitations, August 11, 2006)

5.3.1 General Solutions

Cookies, a unique identifier that a web server places on user’s computer that can be used to retrieve their records from the databases, authenticate, identify and track users, were seen as a major threat to user’s online privacy. Third party cookies could be linked to user’s collected browsing history which makes them a greater threat [**28**]. COPPA recognizes cookies to be a privacy threat and disallows operators from online collection of identifying code linked to a child, such as cookie. As a solution, most web browsers provide cookie control and blocking features to give users the option of protecting their privacy. Cookie blocking software is effective but addresses a very small part of the requirements of COPPA. This software does not contribute to the scenarios where websites explicitly collect personally identifiable information from children.

Anonymizer is another solution that protects user’s privacy by providing a way for anonymous web browsing. All communication is directed through an intermediary proxy server to hide the true origination of a message. Thus cookies cannot be placed on the user’s browser and the user’s true IP address cannot be tracked. Anonymizers serve as good solutions to protect a user but this may not be the best option for children. For example, anonymizers allow for anonymous browsing but the goal is contradictory since we need the website operator to recognize the client as a child and take additional precautionary steps to make sure their privacy is protected. Also, anonymous browsing can encourage children to access objectionable material once they are aware that their identity is not being disclosed.

Another initiative that has, to some extent, been effective and widely adopted is privacy policies. Privacy policies are documents detailing exactly how a company uses the information it collects from online registrations, purchases and other activity collecting user’s information and are an important way to protect user privacy. COPPA requires websites targeted towards children to post a privacy policy to inform of their activities and what they intend to do with the child’s data. However, privacy policies were not created specifically as a result of COPPA and many other specific laws also require websites to post privacy policies. Privacy preferences project (P3P) [**29**] is a project aimed at establishing a common machine-readable vocabulary that websites can use to publish their policies and privacy practices online. P3P again was not targeted towards enforcing COPPA and it does not provide for mechanisms for parents to specify privacy preferences for their children.

Privacy policies overall are not alone a solution to the privacy problem. One is the problem of websites not following the FTC regulations for their privacy policies while collecting children's information. The other major problem is that it is very difficult to understand what most privacy policies are trying to say as they are mostly written using official jargon and ambiguous language [**26**]. Such policies when presented to parents, who are not technically sound or are unaware of the laws and rules, fail to make informed decisions and in most cases may agree to the policies. This again is a danger to the child's privacy.

There are also several companies that provide trust seals in their various forms, including TRUSTe (http://www.truste.org/), BBBOnline reliability (http://www.bbbonline.org/), Hacker Safe (https://www.hackersafe.com/), among others. Trust seals are an image, certificate, or badge placed on a website demonstrating that a third party has verified that the company exists, has a privacy policy, and or is safe to deal with. This solution includes involvement from, and trust in, a third party. These third parties are simply viewed as a way of improving customer confidence and do not essentially implement any technical or security mechanisms to protect a user's privacy. With findings showing that only 25% of the consumers recognize trust seals, they are cannot be adopted as a solution but rather as an additional protection layer combined with other solutions. Sole reliance on trusting a website may not be good in some cases, especially where children are involved [**26**].

5.3.2 Solutions Targeted Towards Children

COPPA was responsible for raising awareness that children's online privacy is in fact a serious issue and needs legal consideration. This promoted research of tools and technology specifically targeted towards protecting children on the internet. Blocking software and filtering tools which block objectionable material and prevent children from accessing “bad” sites were developed. However, none of them have been focused towards children giving out their PII online.

Blocking software are not foolproof solutions because they rely on blocking known threats but may not effectively block a newly discovered threat. Besides, websites are constantly researching ways of overcoming blocking software. A study conducted for online victimization of youth showed that one third of respondents had been exposed to unwanted sexual material, up from one quarter of respondents in a similar study done five years earlier. The researchers attributed this increase to the growing aggressive tactics of porn marketers which shows that blocking software alone is not effective in this age.[6]

Many software controls for children too have been introduced to give parents the control of their child’s privacy. Microsoft has introduced parental controls built into Windows Vista, designed to help parents manage what their children can do on the computer. These controls help parents determine which games their children can play, which programs they can use, and which websites they can visit, and when. Parents can restrict computer use to specific times and trust that Windows Vista will enforce those restrictions, even when they're away from home.

Apart from OS features, there also exist dedicated software packages such as Net Nanny and browser extensions such as the Parental Control toolbar. Net nanny functions as blocking software by restricting access to certain software, blocking pre-configurable outbound information, among other features. Parental Control toolbar is similar as it disallows children from viewing objectionable online material. However, as in most of the tools, there is certain level of cooperation required from the website operator. For example, the toolbar blocks content based on the knowledge of labels assigned to web pages following a standard vocabulary.

Solutions targeted towards children are mostly variations of blocking software and content filters. These software do not in any way support the function of obtaining verifiable parental consent. Though some of them block outbound information from

[6] http://www.pbs.org/teachers/learning.now/2006/08/new_study_shows_rise_in_youth.html (New Study Shows Rise in Youth Exposure to Porn & Cyberbullying, Decrease in Online Solicitations)

client, preventing children from revealing sensitive information, they also prevent children from accessing certain sites and material online. Completely blocking information that a child submits to websites during registration may prevent a child from gaining access to a service. There needs to be another approach besides blocking that can strike a balance between giving children access and protecting their privacy. Both these goals are as important and need to be kept in mind while developing technology.

5.3.3 Parental Online Consent for Kid's Electronic Transactions (POCKET)

POCKET [**30**] is part of the category of solutions targeted towards children but is a unique technical solution and is hence being discussed independently here. POCKET is an artifact providing a way of obtaining verifiable parent consent as required by COPPA. Based on this artifact, there is a tool proposed which can be used to protect children's privacy online. It allows parents to configure privacy settings for their children and then automatically enforces the policies. POCKET is an extension of the P3P policy and it incorporates a TTP, extends the merchant policy to include data items specified by COPPA and automates the exchange of personal information between the child and server.

POCKET uses a 3-phase protocol: Registration, Setup and Transaction phase. During the registration phase, the parent registers with the TTP and obtains a unique identifier and the POCKET installer after verification. The merchant is also expected to register with the TTP to be POCKET compliant and receives a POCKET certificate. A merchant privacy preference file (MPPF) is created at this point which specifies the information that will be collected from a child accessing that particular merchant's website.

Next is the setup phase where the POCKET User Agent (UA), configured by the installer, is used by the parent to select the child's data that a merchant can collect. This preference is used to generate a user privacy preference file (UPPF). The POCKET UA also enables the browser helper object (BHO) that enforces the UPPF.

Finally in the transaction phase the BHO request the MPPF, POCKET certificate and the merchant's information collection practices. It then validates the MPPF against the UPPF and makes the decision of whether a particular web site's policies are comparable to the parent's privacy requirements.

The POCKET tool is a promising approach for protecting children's privacy online. It automates the function of obtaining parental consent and strongly abides by the clauses in COPPA. Our work is also an automated way of obtaining parental consent and for providing a viable technical solution for protecting children's privacy online, in accordance with COPPA. However our work is unique in many aspects and tries to overcome few weak links identified in the POCKET model.

A server-side solution for our model is not a possibility since out threat model involves the operator itself. We cannot expect the malicious party to implement or even cooperate towards solutions for protecting privacy of the children. We could either use client-side solution or involve a trusted third party (TTP). A solution with a TTP is possible but known to have major drawbacks. Compromise of these TTP's is another factor to be considered since all trust has been vested in the TTP's ability to protect the client and a successful attack on the TTP could bring down the entire trust model. This shows that the fewer the number of parties involved in the trust model, the better the solution is.

POCKET uses a third party (TTP) during its interaction. It requires merchants to abide by the rules and register with the TTP. Also, it is expected that once the MPPF is created, the merchant does not change their information collection practices. This may not be true if we are assuming that the merchant is not a trusted party.

Another inherent flaw in the solutions centered around TTP is that once the TTP has approved of the practices of a website and granted its seal of approval, or certification, few, if not no steps are taken to verify if the website operators conform to the policies agreed upon. This can lead to problems if a website receives the TTP's approval and then carries out malicious activities. Revocation of such policies and the procedures for that matter remain ambiguous. In case we assume that the website operators conform to their agreed upon practices, it can take time to build trust in such external third parties. We cannot expect people to trust their services instantly without any experience.

Our main goal here is to give maximum control to the client, i.e. the parents in our model, so that they feel more confident with their children's privacy rather than expecting them to trust an external third party and the website operator for that matter. Transferring trust to an external party may not lead to a good overall solution. Therefore, giving control to parents and moving the enforcement to client side serves as a better solution.

The POCKET tool has a BHO which compares the UPPF and the MPPF and grants access to the website in cases only where the merchant's policy requires no more information than that indicated by parent in UPPF. This implies that in all cases, the MPPF must be a subset of or at most equal to the UPPF. This might be a farfetched goal with the number of parents using the tool and the different variations of UPPF depending on each parent's preference. Our tool makes sure that the child gains access to sites it desires and at the same time the UPPF is enforced, irrespective of the merchant's data collection policy.

Chapter 6

Value Sensitive Design

6.1 Introduction to Value Sensitive Design

The goal of this research is to find an automated, technically viable and socially responsible solution for enforcing COPPA. This research will employ the Value Sensitive Design method (see Figure **2**) to create an innovative toolkit named COP – **C**hildren's **O**nline **P**rivacy protection tool.

Figure **2**

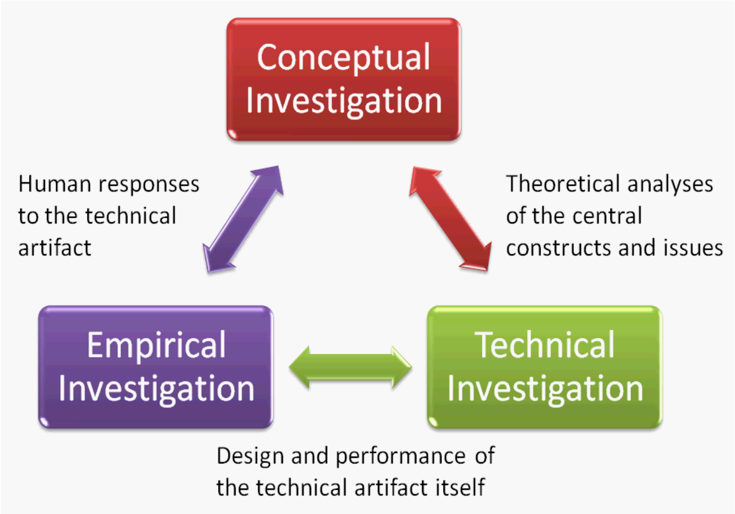

Conceptual Investigation
Human responses to the technical artifact
Theoretical analyses of the central constructs and issues
Empirical Investigation
Technical Investigation
Design and performance of the technical artifact itself

Figure **2**: Value Sensitive Design Method

Our proposed approach presents significant challenges that are theoretically intriguing as well as practically significant. Design for privacy is complicated by the fact that privacy is a socially constructed value that differs significantly across environments and individuals [**31**]. In particular, protecting children's online privacy is more complicated by as the tradeoff between children's rights of "being alone" and parental permission-based Internet access. Thus, any technical solution must be infused with a profound understanding of social implications, a design that balances children's privacy with their online safety, a technically sound approach to COPPA compliance, and an easy-to-use solution that bolster perceived trustworthiness. The COP toolkit utilizes the value sensitive design approach that adopts a tripartite methodology by systematically integrating and iterating on conceptual, technical and empirical investigations of privacy. It is socially responsive in that it provides a balance between allowing children to visit websites they desire to and at the same time making sure that their personal information, and in turn privacy, is protected. It is user-friendly in that it provides verifiable parental consent to websites without requiring intervention from parents. It further provides customization features for parents to configure privacy preferences and it enforces the set preferences automatically. Furthermore, respecting privacy as user control, we design COP as a solely client-side solution without dependence on the server or a trusted third party.

6.2 Privacy as a Design Value

Value sensitive design is an approach to the design of information and computer systems that accounts for human values in a principled and comprehensive manner throughout the design process [**32**][**33**]. Value sensitive design approach is particularly useful for our research because such method emphasizes values with moral import such as privacy and trust [**32**][**33**]. This design method embeds explicit values choices, documents those choices, and thus enables adoption and alteration of technologies to be informed choices for the appropriate social context [**31**].

As Camp et al. [**31**] pointed out, the sheer complexity of understanding a value as amorphous as privacy has been a serious difficulty in applying value-sensitive design. And, in fact, the difficulty in defining common ground of privacy will likely become more pronounced in the next few years. According to a 2007 study sponsored by the National Research Council [**34**], the relationship between information privacy and society is now under pressure due to several factors that are "changing and expanding in scale with unprecedented speed in terms of our ability to understand and contend with their implications to our world, in general, and our privacy, in particular." Factors related to technological change (e.g., data collection, communications), to societal trends (e.g., globalization, cross-border data flow, increases in social networking) are combining to force a reconsideration of basic privacy concepts and their implications. Thus rather than drawing on a monolithic concept of privacy from a single discipline, we try to build upon the previous literature from multiple disciplines to create a common understanding of Parents' Concerns for Children's Online Privacy (PCCOP). Clarification of values toward PCCOP is particularly important in discussing children's online privacy, as these are so often confused in technical design, websites' data collection practices and parents' perceptions.

Table **1**

Table **1**: Paradigms regarding the Concept of Privacy

Paradigms	**Theoretical Lenses**	**Driven Force**	**Consequences of Privacy Violation**
Contextual Nature of Privacy	Social	Individuals' own experiences and social expectations	Potential embarrassment or breakdown in relationship(s) etc.
Privacy as Control	Psychological	Autonomy, self-efficacy and trust	Concern/worry about data misuse and identity theft
Legal Protections of Privacy	Normative	National or supra-national legislative act	Civil and/or criminal penalties

Source: Adapted from Patil and Kobsa [**35**]

Value sensitive design adopts a tripartite methodology by systematically integrating and iterating on three types of investigations [**32**][**33**]: *conceptual investigations* comprise philosophically informed analyses of the central constructs and issues under investigation; *technical investigations* focus on the design and performance of the technology itself; *empirical investigations* focus on the human responses to the technical artifact. In this proposal, we offer our initial start at a conceptual investigation based on three main perspectives from which the notions of privacy are commonly described and analyzed (see Table **1**). Based on our initial conceptual investigations, we present a preliminary technical investigation of COP development, followed by our plan for empirical investigation of COP.

6.2.1 Contextual Nature of Privacy

One very important perspective considers the contextual nature of privacy [**36**]. In more recent privacy literature, such contextual paradigm of privacy recognizes that privacy *both* influences and can be influenced by various situational and societal forces. Individuals' desire for privacy is innately dynamic [**37**], and influenced by various situational forces, such as pressures from others, societal norms, and processes of surveillance used to enforce them [**36**]. Altman [**38**] conceptualized privacy decision-making as a dialectic and dynamic boundary regulation process. As a *dialectic* process, privacy is "conditioned by individuals' own experiences and social expectations, and by those of others with whom they interact" [**39**]. As a *dynamic* process, privacy is "understood to be under continuous negotiation and management, with the *boundary* that distinguishes privacy and publicity defined according to circumstance [**39**].

Protecting children's privacy is complicated by the fact that children's privacy is a socially constructed value that reflects the child-parent relationship as well as the balance between children's rights of "being alone" and parental permission-based Internet access. Unfortunately, parental permission-based approaches are more socially complicated for children's online behavior. Because websites are far away from the parents, how is the site operator going to ensure that the person vouching for the child's age is really the parent or even an adult? According to a recent FTC report, it is concluded that age verification technologies have not kept pace with other developments [**40**]. One of the social complexities associated with children's privacy is that, children quickly learned that if they say they are below thirteen they will be prohibited from using many sites. As a result, children regularly lie about their age everywhere online. Seeing such social complexity in the context of protecting children's privacy, we propose following design principle for addressing Parents' Concerns for Children's Online Privacy (PCCOP):

Design Principle #1: The technical systems for addressing PCCOP should make a balance between protecting children's personal information online and preserving their ability to access content.

6.2.2 Privacy as Control

A second major paradigm considers privacy in terms of psychological control of personal information. This perspective is found in various prior works [e.g. [**6**][**41**][**42**][**43**][**44**]] which have contributed to and stimulated the paradigm of privacy as a control related concept. A number of privacy theorists have put emphases on the concept of *control* when defining privacy [e.g. [**6**][**45**][**46**][**47**][**48**]]. For example, Wolfe and Laufer [**49**] suggested that "the need and ability to exert control over self, objects, spaces, information and behavior is a critical element in any concept of privacy". Empirical evidence revealed that control is one of the key factors which provide the greatest degree of explanation for privacy concern [**50**][**51**][**52**][**53**][**54**]. Individuals perceive less privacy concerns when, among other things, they believe that they will be able to control the use of the information and that the information will be used to draw accurate inferences about them [**55**].

Drawing on control agency theory [**56**], our earlier work identified two types of control in the privacy context: 1) *personal control* in which the self acts as the control agent, 2) *proxy control* in which external entities act as the control agent [**57**]. End-user privacy-protecting tools such as cookie managers allow users to protect their information privacy by directly controlling the flow of their own personal information to others [**40**]. As is evident, with end-user privacy protecting tools, the agent of control is the *self*; and the effects of this mechanism arise due to the opportunity for direct personal control. With regard to proxy control, trusted third party (TTP) is a commonly used approach that mainly consists of an entity facilitating interactions between users and websites who both trust the third party to secure their interactions. TTP solution to privacy is one example of proxy control that is created to provide third-party assurances to users based on a voluntary contractual relationship between websites and the third party. On behalf of users, the TTP acts as the control agent for users to exercise proxy control over the flow of personal information.

This paradigm of privacy as control brings rise to the debate among scholars and practitioners on the effectiveness of these two (and other) mechanisms for privacy control: Whose responsibility of protecting children's privacy – users themselves or websites or TTPs? Which approach will be more effective and trustworthy? Drawing on recent privacy literature on comparing the relative effectiveness of personal vs. proxy privacy control approaches [**58**][**59**], we propose that neither the server-side nor TTP is completely trustworthy due to at least two reasons: First, the threat model may involve the operator itself. Second, a successful attack on the TTP could destroy the entire trust model since all trust has been vested in the TTP's ability to protect the client. Therefore, we propose following design principle for addressing Parents' Concerns for Children's Online Privacy (PCCOP):

Design Principle #2: The technical systems for addressing PCCOP should maximally empower parental control in which user data (including both parents and children) and their preference files are located at the client side rather than the server side or any third party side.

6.2.3 Legal Expectations on Protecting Children's Online Privacy

COPPA was enacted in response to the growing danger to the online privacy of children. The stated goals of the Act are: (1) to enhance parental involvement in their children's online activities in order to protect children's privacy in the online environment; (2) to protect the safety of children at places in the online environment such as chat rooms, home pages, email accounts, and bulletin boards in which children may make public postings of identifying information; (3) to maintain the security of children's personal information collected online; and (4) to limit the collection of personal information from children without parental consent.

We performed a content analysis on COPPA in terms of scope of the law, definition of personal information, jurisdiction/territoriality, parental consent requirements, notice requirements, proceeding rules, rights of individuals, enforcement and sanctions. We identified COPPA's core *personal information collection*, *parental consent and access*. *Collection* is broadly defined and applies to the online gathering of any Personal Information form a Child under the age of 13. *Parental consent* requires that prior to collection, use, and/or disclosure of personal information about a child, an operator must obtain from a parent of the child for verifiable parental consent. The method of collecting the verifiable parental consent has been an issue of debate for long. At the time the law was enacted, no inexpensive and viable technological means existed that simplified this process. Therefore, the FTC adopted a sliding scale approach for obtaining verifiable consent and this was declared to be a temporary approach assuming that more sophisticated methods would be introduced through advances in technology. *Access* requires that the Parent of any Child who has provided Personal Information to an Operator has the right to request access to such information. Based on the content analysis of COPPA, we propose following design principle for addressing Parents' Concerns for Children's Online Privacy (PCCOP):

Design Principle #3: The technical systems for addressing PCCOP should comply with COPPA requirements on personal information collection, parental consent and access.

Chapter 7

Technical Investigation- COP Development

Following the philosophy of Value Sensitive Design, the above conceptual investigations can now be employed to help structure the first iteration of a technical investigation. A tool named as COP (Children's Online Privacy protection tool) is proposed to provide technical mechanisms for protecting children's online privacy.

7.1 Introduction to COP

The proposed tool is built to overcome the drawbacks of existing solutions to protecting children's privacy while being a technically viable solution for enforcing COPPA. COP is a reliable, easy-to-use, technical solution to give parents complete control of their child's personal information which the child can disclose to websites. It tries to provide a balance between allowing children to visit websites they desire to and at the same time making sure that their personal information, and in turn privacy, is protected. COP allows parents to configure privacy preferences for their children and control data that is disclosed to websites. Being aware that parents trust certain websites more than others, COP provides categorization of websites into trusted and not-trusted giving parents the option of enhanced personalization. Tool provides other customization features as well to ensure that parents feel in complete control of their child's private information. COP enforces the set preferences automatically and is a way of providing verifiable parental consent to websites without requiring intervention from parents. What is unique about COP is the fact that it is solely a client-side solution and requires no dependence on either the server or a trusted third party.

COP enables children to enjoy the experience of online browsing and exploit the vast resources of online knowledge. It does not expect children to submit false data or lie under any circumstances and at the same time requires minimal intervention on part of the parent. TTP is not involved and it is a client side solution to give complete control of a child's privacy to the parent.

Information submitted by children through online forms is not blocked but instead perturbed to a degree before it is submitted to websites, to allow children to register for sites but still be protected. Notification is used as a major aspect of this tool and parents are constantly kept updated of their child's activities, when and what information they submit to websites and what sites they visit. COP provides for an automated way of obtaining verifiable parental consent to prevent inconvenience caused to parents otherwise.

The parties involved in our framework include the child, the child's parent and the operator. The parent is the one who sets preferences for their child and receives notifications about their child's online activities. The child is the entity under consideration whose online privacy needs to be protected. The operator is the website that the child is visiting or wishes to access.

7.1.1 Categorization of Websites

In our tool, we allow parents to select data items which can be submitted to the website by the child without the fear of their privacy being compromised. Assuming here that the parents trust all sites uniformly would not be true. There are certain sites which parents trust more than others and could allow more information to be disclosed to those particular sites.

For this purpose, we have categorized web sites as "TRUSTED" and "NOT-TRUSTED". Trusted websites are sites which are known to be safe for children to access. All sites by default are not-trusted. By categorizing websites into two groups, greater amount of true information could be disclosed to trusted sites than that submitted to not-trusted ones. The tool also provides a suggested list of trusted sites for parents, to help them identify trusted sites.

7.1.2 Logging Facility

Our tool provides a logging facility that maintains a record of the websites visited by the child and also, the data submitted to each website, after perturbation. Logs are an important part of our proposed model. Consider the case where the child forgets their password or login information for a website. In most cases, before revealing the password to the user, the websites prompt users to verify their information submitted during registration.

In our model, the information submitted to a website is randomly perturbed for each website. From the child's point of view, they have entered the correct information. However, if the child submits this information and the tool perturbs the data using a different value, it may cause inconvenience to the child and can also cause him to lose his account access. For this purpose, the tool needs to record the values submitted to a website by the child to prevent any conflicts in such cases.

Another use of these logs is if the parent requests to be notified at a particular time of the child's online activities (method discussed later). The logged entries over the last 24hours are needed in that case to send parents the notification.

The logs can also be used as a proof of access and to prove exactly what information was submitted to a website by the child. This sort of proof could settle disputes in future if any.

7.2 System Design Overview

In our system, the parties involved include the child, the child's parent and the operator. The parent is the one who sets preferences for their child and receives notifications about their child's online activities. The child is the entity under consideration whose online privacy needs to be protected. The operator is the website that the child is visiting or wishes to access. To give an overview of our system design for the COP toolkit, we first show the process when a child registers with a website to access some service. This process involves a sequence of message flows, which are explained as follows.

Three design principles derived from our conceptual investigations have been applied to structure the first iterations of the COP design.

0) Assume that the COP toolkit has been installed, configured, and activated by the parent who has the system administrator privilege and COP monitors all outgoing traffic.

Figure **3**

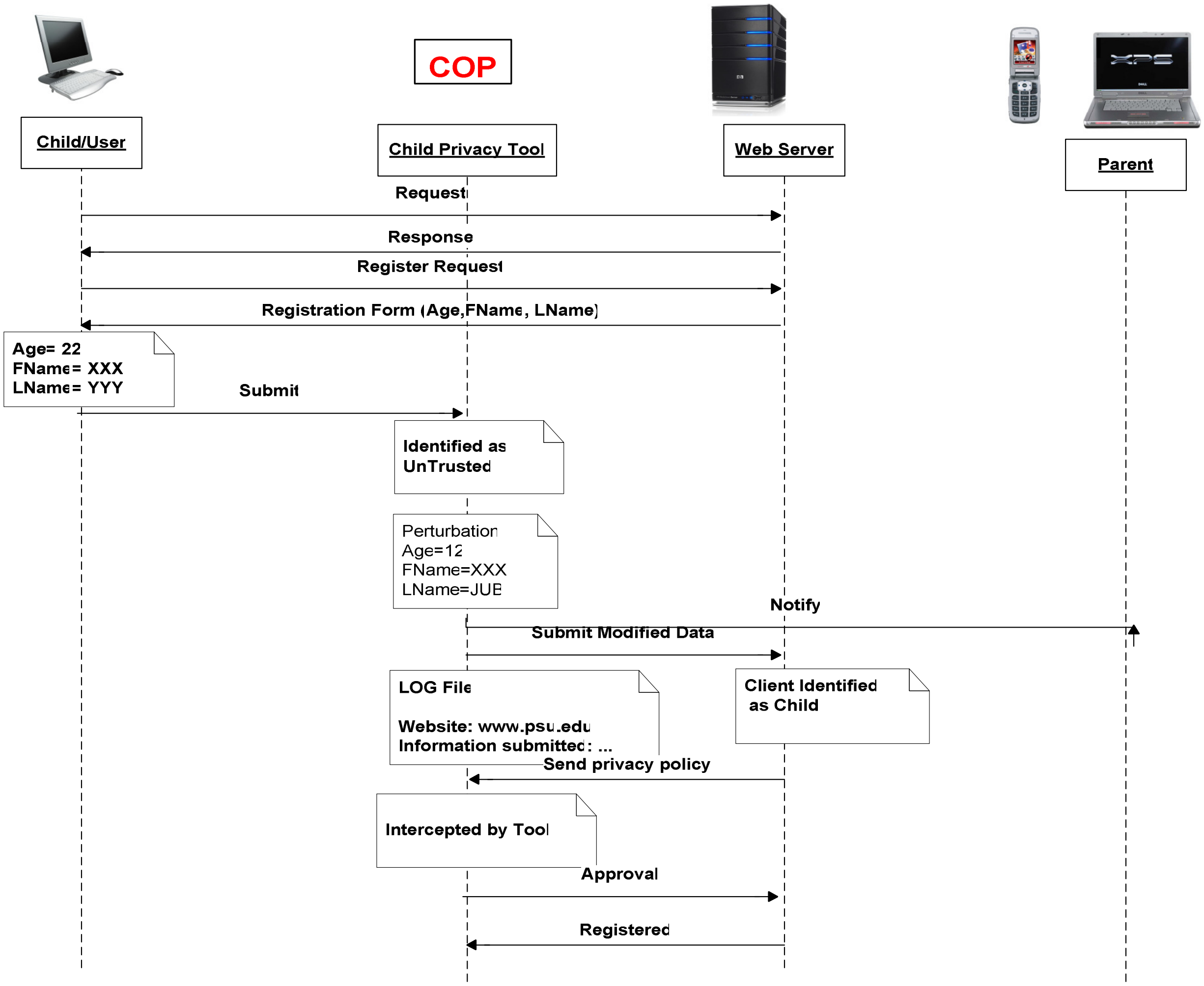


Figure **3**: System Design Overview

1) The child's browser sends an HTTP request to register with the website.

2) Web server responds by sending an HTTP response which is a registration form required to be filled in by the child. (Note that at this point the web server is unaware that the client's age is under 13 years).

3) The child completes the fields of the form with true information such as age, address, first and last name, etc.

4) Once the child clicks the submit button, the COP toolkit intercepts the communication. It first checks if the website is "trusted" or "not-trusted" and accordingly compares submitted information with the privacy preferences set for that category of websites by the parent.

5) It perturbs sensitive data according to some perturbation rules and sends this modified or perturbed information in an HTTP request to the server (Details discussed in Section 4.1.2).

6) COP makes an entry in its log of the website to which the data was submitted and the values that were eventually, after perturbation, sent out.

7) COP notifies the parent (either immediately or at a pre-selected time depending on the parent's preferences).

8) Web server on the other hand recognizes the client as a child. This is enforced by the tool because even if the child indicated a false higher age, the age would be perturbed to a value lesser than 13 years of age.

9) The web server (assuming the website conforms to COPPA) sends its data collection, usage and distribution policy across to the client side to be approved by the parent.

10) COP intercepts this communication and provides approval, irrespective of the website's policy. (Note that COP will enforce the parent's preferences in any case and hence providing the approval may not need extra effort).

11) The registration process continues and the child is registered with the website.

7.3 Methods for Client-Side Data Perturbation

An important aspect of our tool which distinguishes it from exiting solutions is that our tool does not block the information submitted by a child to websites. Children usually give out personal information online when asked for. These practices are common during registration, subscription processes, where children are presented with a form that needs to be filled in and submitted to gain access to the service. Blocking the outgoing data in such cases may limit the child from accessing the online services. Therefore, instead of blocking, we perturb the data to an extent before it is sent out in a way that the child's privacy is protected.

Several challenges arise in our research. First, we need to process many different types of data, including Numerical, String, Enumeration and others. Clearly, there is no single algorithm for perturbing all these kind of data. Second, there exist constraints on the perturbed data. What this implies is that a perturbed credit card number should seem valid in accordance with a card type; a perturbed zip code should correspond to a geographical location. Third, improving children's privacy by concealing their identity online may not be a good option if they are required to identify themselves in order to receive certain services. Hence, we need to have scope to allow a certain degree of personalization. Fourth, under the condition that COPPA has not been violated, we might allow the website to study the collected data for various purposes. As such, instead of perturbing the data rendering it totally useless, we would perturb the data such that its statistical characteristics are preserved. For example, although the age values submitted by children are perturbed, the average value would not be very different from the true value.

To address these challenges, we will utilize some of the existing data perturbation techniques from the area of privacy preserving data mining and data publishing[7]. There are three existing data perturbation methods for numerical data namely, additive perturbation [**60**], multiplicative perturbation [**61**][**62**] and probability distortion [**63**]. In additive perturbation, noise is added to the data in order to mask the attribute values of

[7] http://www.csee.umbc.edu/~kunliu1/research/privacy_review.html

records. The noise added is sufficiently large so that the individual record values cannot be recovered from the perturbed data. Note that although there are known drawbacks of additive perturbation including that additive noise may be easily filtered out through correlation of the data points from within a large data set, it will not be an issue for our data because a website only has one data point from each child. For multiplicative perturbation [**61**][**62**], there are two basic methods of performing it. The first method is based on generating random numbers that have a truncated Gaussian distribution with mean equal to one and a small variance, and then multiplying each element of the original data by this noise. The second method is to take a logarithmic transformation of the data (for positive data only), compute the covariance, and generate random noise following a multivariate Gaussian distribution with mean zero and variance equal to a constant times the covariance computed in the previous step. We then add this noise to each element of the transformed data, and take the antilog of the noise-added data. The latter method assures higher security while the former one only requires minor changes to the original data. Unlike other approaches, probability distortion perturbs the value of each data element (point distortion) and replaces it with another sample from the same (estimated) distribution. The merit of this approach is the difficulty of data being compromised using repeated queries.

We will employ the idea of k-anonymity for protecting data such as zip code, address and birth date. Traditionally, k-anonymity is defined as "each release of the data must be such that every combination of values of quasi-identifiers can be indistinguishably matched to at least k respondents" [**64**].It is known for being resilient against indirect identification of personal records from a public database. Two common k-anonymity approaches are generalization and suppression. In the method of generalization, the values in a database table are true, sensitive values and are hence generalized to ranges each covering k original values. Thus each record is indistinguishable from the other k-1 one records in the same group. In our context, generalization is performed over an individual value where its value domain is k. For example, to reduce the risk of identification, the date of birth could be generalized to either year, or month and year of birth. Note that although the former type of k-anonymity is known to be susceptible to homogeneity attack [**65**], this attack would not work in our scenario because the data provided by each user is individually perturbed and it has nothing to do with the data provided by other users. In the method of suppression, the value of the attribute is removed completely.

Table **2** shows different types of data items that might be requested from a child during registration. For each data item, we list its data type, potential perturbation methods, range/formats and special notes. Data Types under Consideration

The types and items of the data in our work are listed below. For numerical data, they can be further classified as formatted or unformatted data where formatted refers to data items with an identifiable trait or pre-set format making them easier to identify and validate, and unformatted refers to data that has no pre-set identifiable trait. The additional category of "other" is data items which do not classify under any of these three categories.

Table **2**

Table **2**: Data perturbation Approaches for Different Data Types

Data Type	**Data Item**	**Possible Perturbation Methods**	**Range/Format**	**Notes**
Numerical Value	Age	1: Follow certain predefined distribution 2: ε, normal distribution with mean μ=1 and σ=0.5; a=(Age× ε)mod 13, if >6, R(X)=a; otherwise R(X)=a+6	6~12	1. Assume we know the distribution of the ages from 6-12. 2. This is a demonstration of the multiplicative perturbation approach
Numerical Value	Phone number	Reserve the area code; generate random 7 digits (uniformly) for the rest	XXX-XXX-XXXX	Certain rules for phone number
	SSN	Randomly Perturb	XXX-XX-	Certain rules for SSN

	number		XXXX	number
	Date of birth	1:Year must be in accordance with age; 2: For month and day, following the same perturbation procedure as for age.	XX-XX-(>1996)	Consistent with age
	ZIP code	Reserve the first 3 digits, randomly select the last two digits or use 00	168XX or 16800	Consistent with address
	Credit Card	Follow the CCN rules, use random number or select from pre-defined dataset	16 or 15 digits (depending on card type)	CCN validation check
String	Name	Choose from certain data set (e.g. Cartoon characters)	Mickey Mouse	-
	Username	Do not perturb unless real name used	N/A	Kids may like a certain name
	Address	Keep state info. Change door number 1234, Use four identical letters (e.g. AAAA) as the road name, Change city name to all "X"s	1234, AAAA Street, XXX (City), PA	If city name is provided as a drop down menu option then the index number of the option will be perturbed to another valid index
	E-mail	Changed to parent's email address	N/A	-
	Home Page	Use common sites / Use www.username.com	N/A	-

	(URL)			
Enumeration	Gender	a% unchanged, (1-a)% change	N/A	-
	Hobbies/ Habits/ Interests	Perturb index, use binominal distribution with π=a	N/A	Hobby selection is normally implemented as a checkbox, each choice has an index
Other	Picture/ Photo	Pre-defined dataset	N/A	-

Chapter 8

Implementation of COP

8.1 COP- A Firefox Extension

We have limited implementation options as we need to perturb the data before it is sent to the server and this may be problematic if the website has an SSL-encrypted communication with the browser. Therefore, there are mainly three feasible implementation methods for our tool and both of these are briefly discussed here.

1. Proxy Plug-in: The idea here is to insert a HTTP proxy between the browser and the web site. Once the proxy is active, any requests that the browser makes to the server will be routed through this proxy. WebScarab[8] is a good example of a similar implementation, since it can interpret the SSL-encrypted HTTPS requests from the web browser. It is similar to the man-in-the-middle attack with the modification that the "attacker" here is WebScarab itself.

    - WebScarab:

      It is a framework to analyze applications that communicate using HTTP and HTTPS protocols. It uses a proxy that observes traffic between the browser and the web server. The WebScarab proxy is able to observe both HTTP and encrypted HTTPS traffic, by negotiating an SSL connection between WebScarab and the browser instead of directly connecting the browser to the server and allowing an encrypted stream to pass through it. Various proxy plugins have also been developed to allow the operator to control the requests and responses that pass through the proxy. However, for a proxy plug-in however, the parents need to understand how to change the proxy settings of their browser.

[8] WebScarab is a project in Open Web Application Security Project (OWASP) http://www.owasp.org/index.php/Category:OWASP_WebScarab_Project

2. Development of add-ons to Firefox: All of the following are a kind of add-on to Firefox, and they can be described as;
    - Extension:

      Extensions add new functionality to Mozilla applications such as Firefox and Thunderbird. They can add anything from a toolbar button to a completely new feature. They allow the application to be customized to fit the personal needs of each user if they need additional features.
    - Plug-in:

      In context of Mozilla-based applications, they are binary components that, when registered with an application, can display content that the application itself can't display natively. For example, Adobe Reader plugin lets the user open PDF files directly inside the browser, and the QuickTime and RealPlayer plugins are used to play special format videos in a web page.

Though the Proxy implementation is better than a Firefox add-on, the final decision was to implement COP as a Firefox extension for initial tests. The tool has two parts to it; (i) Installation Phase, and, (ii) Working Phase. The previous section on system design overview was the working phase of the tool. The next section on user interface design, explains in detail, the installation phase.

8.2 User Interface Design for COP Configuration

COP can be downloaded and installed by the parent. The installation step is where the parent's involvement is most required. Once the installation of the tool is complete, the responsibility of protecting the child's privacy is assigned to the tool. The tool is a representation of the parent's preferences for protecting the child's privacy and the parent's consent as well. Next we discuss several major steps in the user interface design.

In our tool, we allow parents to select data items which can be submitted to the website by the child without the fear of their privacy being compromised. Recognizing that parents do not trust all sites uniformly, we categorize web sites as ‘TRUSTED’ and ‘NOT-TRUSTED’. Trusted websites are sites which are known to be safe for children to access. All sites by default are not-trusted. The parents specifically have to indicate which sites they consider to be trusted. A study indicated that parents would prefer if they were provided with a list of trustworthy websites as they are not experienced in most cases to judge for themselves [**26**]. Based on this finding, we provide an information icon which when clicked, lists a number of trusted sites recommended for each age group. After the parent has added the trusted sites to the list, the next step is to select the data items, which the child should not be allowed to reveal to these websites using checkboxes. The unchecked data items will not be perturbed when sent to the trusted websites. The checked items will be perturbed in case the child has disclosed any.

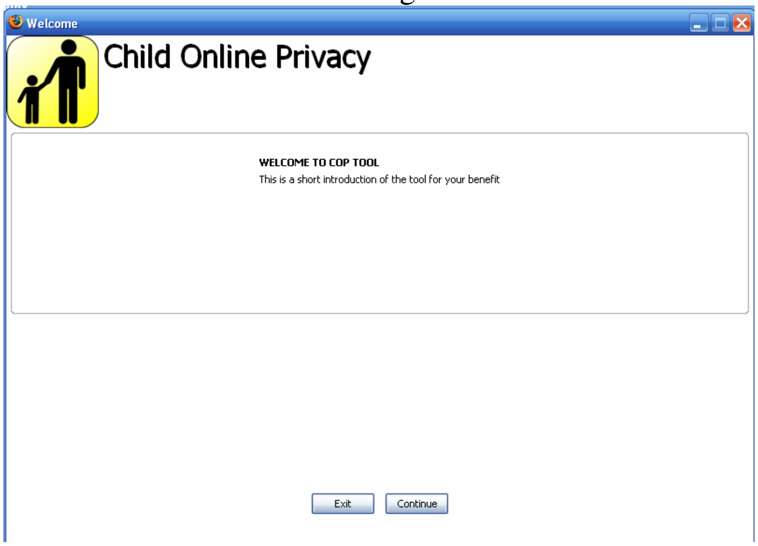
Not all parents have the same privacy preferences. For example, some parents may consider certain information as very private while other parents may not. We thus include an additional option for parents to indicate data items that they would specifically like to block. This is an important customization feature which can assist parents who have concerns over particular data items.

Step 1: Welcome screen, Privacy policy, License Agreement

This is the first screen displayed to the parent during the installation phase. It is a welcome screen which introduces the parent to the tool, explains its purpose, the extent of control and other features that it offers. This screen is to make the parents, who are not very familiar with technology, comfortable with the process. COP does not require the parent to have any background technical knowledge and is extremely simple to use, even for novice users.

Next, the parents are shown a license agreement which they need to agree for them to continue with the installation. It then presents the privacy policy which is written in simple language and is very easy to understand. Once the parent is aware of what they should expect, they can agree to continue with the installation. There is also an option of disagreeing with the policy which informs the parent using a pop-up window that the installation process will be aborted.

Figure **4**



Welcome

Child Online Privacy

WELCOME TO COP TOOL

This is a short introduction of the tool for your benefit

Figure **4**: Installation Step 1- Welcome Screen

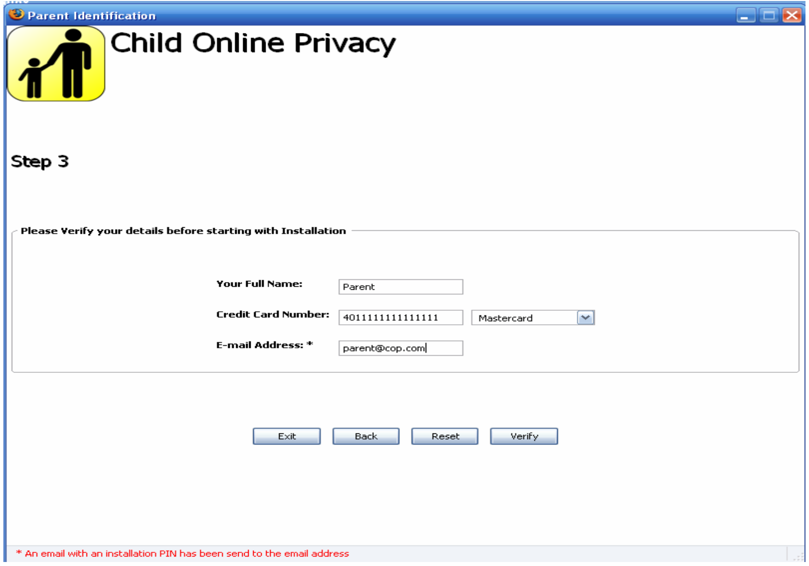
Step 2: Parental verification screen

As discussed earlier, COPPA requires that before a website can collect information from a child below 13 years of age, they should obtain verifiable parental consent. Obtaining this consent every time that information is requested from a child may get cumbersome both, for the child as well as the parent. Therefore we need a way of delegating this responsibility to the tool for approving website policies. Note here that since our tool will be enforcing the privacy preferences set by the parent, irrespective of the operator's policy, the tool can safely approve of any website policy.

This brings the issue of the tool obtaining the parent's consent and verifying their identity. For this purpose, COPPA suggests few methods including an electronic consent form, a toll free number or credit card verification. We have adopted the credit card verification process. Assumption here is that children below thirteen years of age do not have a credit card and hence a successful credit card transaction would suggest that it is the parent providing the verification. The transaction is for verification and the card is not charged. We would further like to mention that the children have no motivation what-so-ever to install this tool. Therefore in all cases it is the parent who will download and install the tool. Credit card verification is an added layer of security however, to comply with COPPA's requirements.

Once the parents have entered their name, credit card number, and email address, the card is verified and an email is sent to the email account indicated. The email address provided here will also be stored for all future communication with the parent.

Figure **5**



Parent Identification

Child Online Privacy

Step 3

Please Verify your details before starting with Installation

Your Full Name: Parent

Credit Card Number: 4011111111111111 Mastercard

E-mail Address: * parent@cop.com

* An email with an installation PIN has been send to the email address

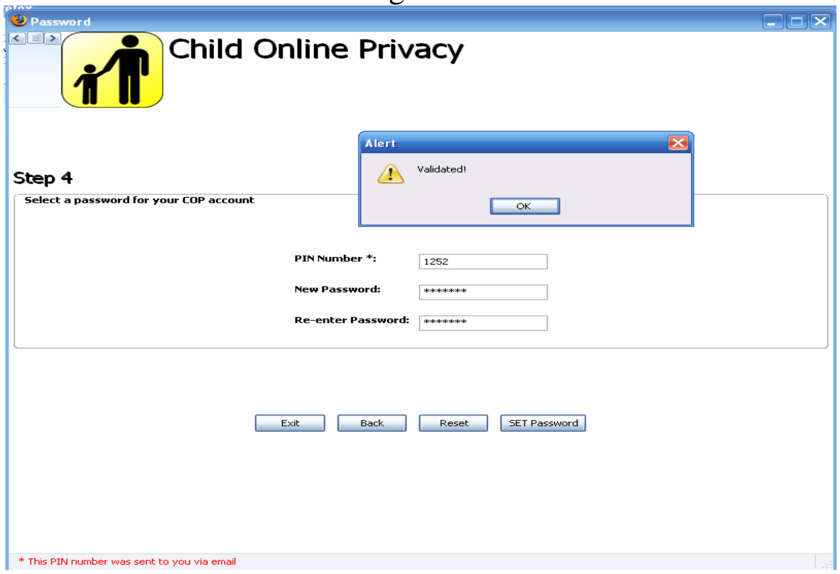
Figure **5**: Installation Step 2- Parental Verification

Step 3: Set password

The email sent to the parent in the previous step contains a one-time PIN number for further verification purposes. Since the installation process is a one-time process, the parents should take effort to set up the tool in the most secure way and the PIN number is an security feature.

The PIN number is verified and a password is set for the tool by the parent. This password will be used for any future access to the tool such as changing preferences, uninstalling the tool, etc.

Figure **6**



Child Online Privacy

Step 4

Select a password for your COP account

Validated!

PIN Number *: 1252

New Password: ********

Re-enter Password: ********

* This PIN number was sent to you via email

Figure **6**: Installation Step 3- Set Password

Step 4: Preferences Setup

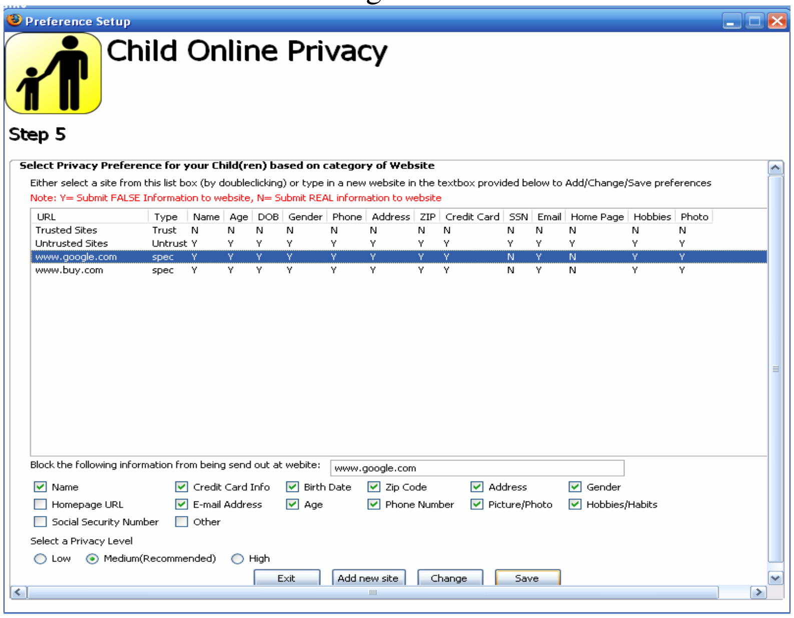
This screen allows parents to set privacy preferences for both, trusted and not-trusted sites. The parents specifically have to indicate for each site that they add to the trusted list, the data items that they wish their child to submit to the sites. Therefore, in addition to adding sites to the list of trusted sites, they need to further indicate the data items they think would be acceptable for the child to disclose to that site.

A study indicated that parents would prefer if they were provided with a list of trustworthy websites as they are not experienced in most cases to judge for themselves [**26**]. The information icon provides a list of trusted sites recommended by the tool for each age group.

After adding the trusted sites to a list, the next step is to select the data items which the child should not be allowed to reveal to these websites by checking the boxes provided. The unchecked data items will not be perturbed when sent to any trusted website. The checked items will be perturbed in case the child has disclosed any.

Not all parents have the same privacy preferences and not to mention some parents may consider certain information as very private which other parents find trivial. We thus include an added option for parents to indicate data items that they would specifically like to block. This is an important customization feature which can assist parents who have particular concerns.

Figure **7**



Child Online Privacy

Step 5

Select Privacy Preference for your Child(ren) based on category of Website

Either select a site from this list box (by doubleclicking) or type in a new website in the textbox provided below to Add/Change/Save preferences

Note: Y= Submit FALSE Information to website, N= Submit REAL information to website

URL	Type	Name	Age	DOB	Gender	Phone	Address	ZIP	Credit Card	SSN	Email	Home Page	Hobbies	Photo
Trusted Sites	Trust	N	N	N	N	N	N	N	N	N	N	N	N	N
Untrusted Sites	Untrust	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
www.google.com	spec	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y
www.buy.com	spec	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	Y	Y

Block the following information from being send out at webite: www.google.com

☑ Name ☑ Credit Card Info ☑ Birth Date ☑ Zip Code ☑ Address ☑ Gender
□ Homepage URL ☑ E-mail Address ☑ Age ☑ Phone Number ☑ Picture/Photo ☑ Hobbies/Habits
□ Social Security Number □ Other

Select a Privacy Level

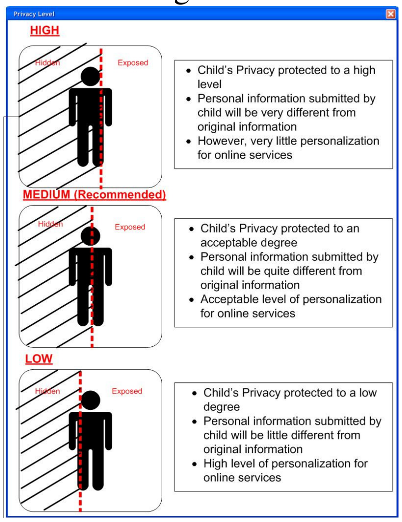
○ Low ● Medium(Recommended) ○ High

Figure **7**: Installation Step 4- Preference Setup

The final step on this screen requires the parent to select a privacy level. Privacy level here indicates the level to which the data will be perturbed before submitting it to a website. The amount of noise added to the data for a high privacy level will be greater than that for a low privacy level selection. Medium is the recommended privacy level. The higher the privacy level, the lower is the degree of personalization that a child will experience because the data submitted to the website will be very different from the true data.

We realize that privacy level may not be a very famous term amongst the non-tech-savvy parents. For this purpose we have provided an information icon which explains in simple terms the definition of the term and helps them make an informed decision. The help screen is shown in Figure **8** .

Figure **8**



Privacy Level

HIGH

- Child's Privacy protected to a high level
- Personal information submitted by child will be very different from original information
- However, very little personalization for online services

MEDIUM (Recommended)

- Child's Privacy protected to an acceptable degree
- Personal information submitted by child will be quite different from original information
- Acceptable level of personalization for online services

LOW

- Child's Privacy protected to a low degree
- Personal information submitted by child will be little different from original information
- High level of personalization for online services

Figure **8**: Privacy Level Help Popup Window

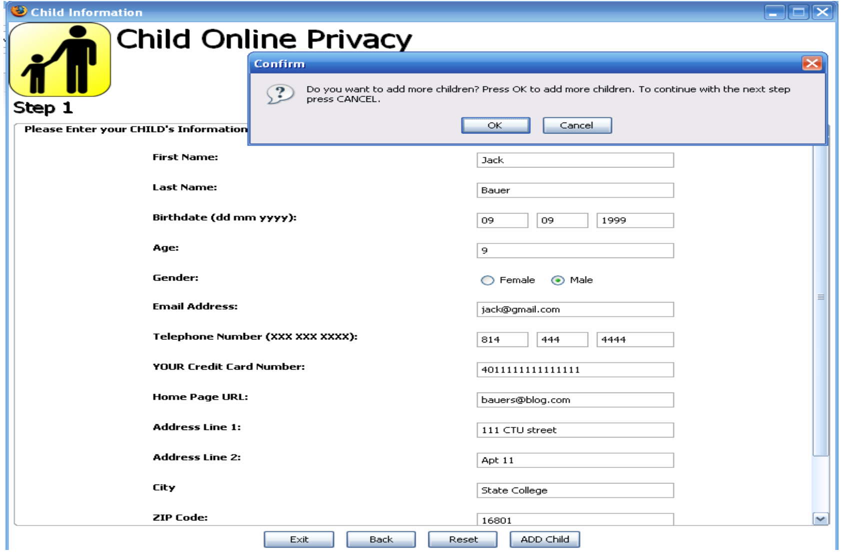
All sites by default are considered as not-trusted. Therefore, the sites indicated specifically by the parent are trusted, while all other sites are not-trusted.

Step 5: Child Information form

In this step, the parents need to fill in a form with detailed information about their children. This information is used as the reference information for blocking. Therefore, all information filled in this form is considered to be highly private and the tool prevents the child from submitting any of this true information to websites, depending on the parent's preferences.

The parent can further add information of more than one child in case the same system is accessible to more than just a single child. Thus, this tool is flexible for protecting more than one child at the same time.

Figure **9**



Do you want to add more children? Press OK to add more children. To continue with the next step press CANCEL.

Please Enter your CHILD's Information

First Name:	Jack
Last Name:	Bauer
Birthdate (dd mm yyyy):	09 09 1999
Age:	9
Gender:	○ Female ● Male
Email Address:	jack@gmail.com
Telephone Number (XXX XXX XXXX):	814 444 4444
YOUR Credit Card Number:	4011111111111111
Home Page URL:	bauers@blog.com
Address Line 1:	111 CTU street
Address Line 2:	Apt 11
City	State College
ZIP Code:	16801

Figure **9**: Installation Step 5- Child Information Form

Step 6: Notification settings

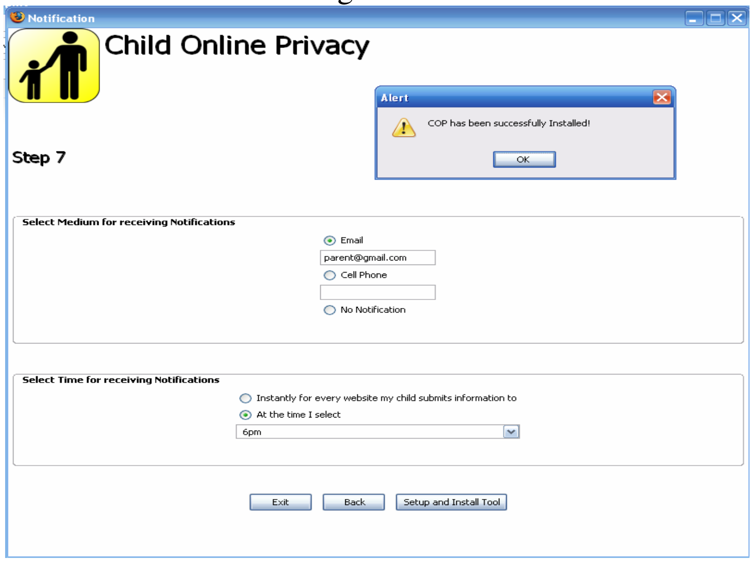
Notification is an important aspect of any system focusing on privacy. Even if the parents have set their privacy preferences and the tool enforces those, parents would feel in greater control if they are notified of what websites their children access, what information they submit to these websites and other such things. This screen allows them to set the preferences for receiving notifications.

First, the parents can select the medium of receiving the notifications. Here they have three options: either by text message or by email or no notification at all. If they select text message then they need to provide their cell phone number. For email notifications, they need to input their email address in the textbox provided. If they select to receive no notification, they are warned of its disadvantages and can then continue with the installation of the tool.

The other option that they have is to select the frequency of receiving these notifications. They can either opt to receive notifications instantly for every website that their child submits information to or all sites that the child has visited in the past period of 24 hours at a selected time each day.

This is the last step of the installation phase. Once the parents have selected the notification method, the tool will be installed.

Figure **10**



Notification

Child Online Privacy

Step 7

Alert

COP has been successfully Installed!

Select Medium for receiving Notifications

● Email

parent@gmail.com

○ Cell Phone

○ No Notification

Select Time for receiving Notifications

○ Instantly for every website my child submits information to

● At the time I select

6pm

Figure **10**: Installation Step 6- Notification Settings

Step 7: Post Installation phase

This step does not require any effort on the part of the parent. The installation phase involves the process of the parent downloading the tool, setting preferences to protect their child's privacy online and then finally installing the tool. The tool is readily made available online to be installed as an extension to the Firefox browser. It requires no technical expertise on the part of the parent.

Setting preferences is an important part of the installation phase. Cop guides the parent through the installation phase. Each step of the installation phase is explained in detail in the following section on user interface design.

On setting all required preferences, COP is installed on the client side and functions as explained in the system design overview. There are mainly four files created and stored on the client side. All information in these files is encrypted (not implemented but stated as future research goal). These 4 files are as follows;

- Parent Info File: This file stores all information relevant to the parent. This information includes the parent's name, credit card information, email address and the password for the tool.
- Child Info file: This file stores all information relevant to the child. This information includes all the details as entered by the parent during the installation phase.
- Log file: This file stores a log of all activity of the child online. It also stores the perturbed information submitted to each site.
- Preference file: This file stores the privacy preferences as set by the parent for each site. Additionally it stores the privacy level selected by the parent as well as the notification method opted for.

These files are important for the functioning of the tool as the information stored in them is used as reference for the working phase.

Additionally, the tool installs a tray icon which strictly permits password access alone. Also, there is a menu option installed within the Firefox tool menu option.

8.2.1 Preference Modification

Many a time parents may feel the need to change certain preferences. For example, after many interactions with a site, the parent may wish to include it within the list of trusted sites. Another example could be if they wish to receive the notifications at 7am as opposed to a previously indicated time. Parents indicate that they prefer technology which provides easy modification options [**26**].

In order to make such changes, parents can simply click on the tool menu option. They will be prompted to input their password after which they are presented with screen 4 where they can set preferences for trusted sites. The tool will run them through the remaining steps and the changes will be implemented immediately.

At this point, the tool sends out an email notification to the parent's email address, entered during credit card verification, informing that changes have been made. This notification is necessary as in many cases parents may make changes by mistake or unknowingly and the email can notify them of the same.

The tool menu has the following options;

- Activate/ Deactivate: This option is password protected. This ensures that the tool is disabled only by the parent and not the child.
- Edit Child Information: This option is used by parent to change the child information entered during the installation phase.
- Modify Privacy Preferences: This option enables the parent to change any preferences that were selected during the installation phase. These include the privacy level, and the preferences for each site.
- Modify Notification Setting: This option allows parents to modify the settings for the method and time for receiving notifications about their child's online activities.
- View Log: Using this option, the parents can directly view the log activity; the sites visited by the child and the perturbed information submitted to those sites. Figure **13** is a screenshot of the log window.
- About: This option opens a popup with general information about the tool such as version number, date of release, etc. Figure **12** shows the popup window.

Figure **11** is a screenshot of the Firefox menu option for COP.

Figure **11**

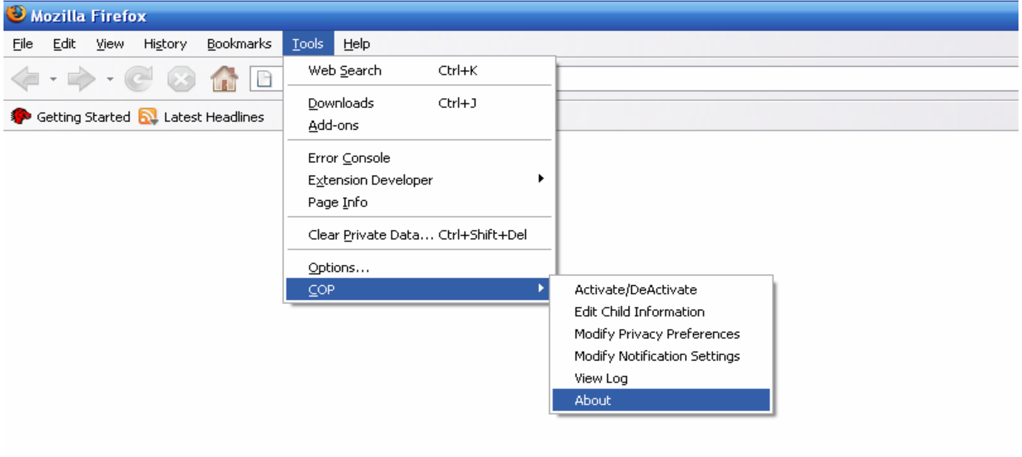
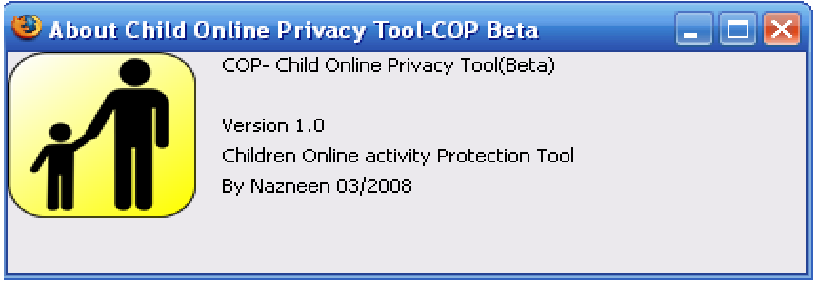


Figure **11**: Tool Menu

Figure **12**



COP- Child Online Privacy Tool(Beta)

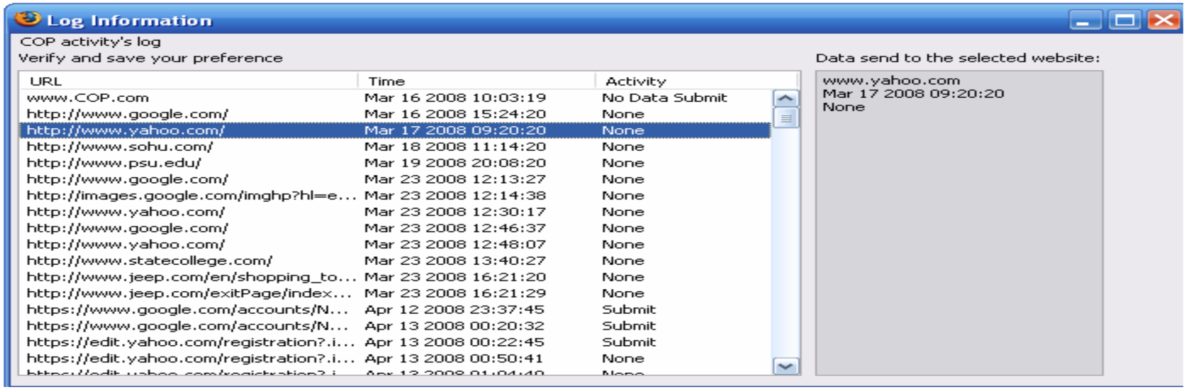
Version 1.0

Children Online activity Protection Tool

By Nazneen 03/2008

Figure **12**: About Menu Option

Figure **13**



COP activity's log

Verify and save your preference

URL	Time	Activity
www.COP.com	Mar 16 2008 10:03:19	No Data Submit
http://www.google.com/	Mar 16 2008 15:24:20	None
http://www.yahoo.com/	Mar 17 2008 09:20:20	None
http://www.sohu.com/	Mar 18 2008 11:14:20	None
http://www.psu.edu/	Mar 19 2008 20:08:20	None
http://www.google.com/	Mar 23 2008 12:13:27	None
http://images.google.com/imghp?hl=e...	Mar 23 2008 12:14:38	None
http://www.yahoo.com/	Mar 23 2008 12:30:17	None
http://www.google.com/	Mar 23 2008 12:46:37	None
http://www.yahoo.com/	Mar 23 2008 12:48:07	None
http://www.statecollege.com/	Mar 23 2008 13:40:27	None
http://www.jeep.com/en/shopping_to...	Mar 23 2008 16:21:20	None
http://www.jeep.com/exitPage/index...	Mar 23 2008 16:21:29	None
https://www.google.com/accounts/N...	Apr 12 2008 23:37:45	Submit
https://www.google.com/accounts/N...	Apr 13 2008 00:20:32	Submit
https://edit.yahoo.com/registration?.i...	Apr 13 2008 00:22:45	Submit
https://edit.yahoo.com/registration?.i...	Apr 13 2008 00:50:41	None
https://edit.yahoo.com/registration?.i...	Apr 13 2008 01:04:40	None

Data send to the selected website:

www.yahoo.com
Mar 17 2008 09:20:20
None

Figure **13**: Log Window

8.3 Technical Challenges

During the development of COP, there were a number of technical challenges that arose and had to be handled in order to successfully implement the tool. A few of these technical challenges have been highlighted below.

1. Parse Data Field: To perturb outgoing data, a prerequisite is the capability of the tool to understand the meaning of each data field. For example, how does COP know a string in the http request (using either GET or POST method) represents 'First name'? It is easy to know it if the name attribute of the input field in the http request

message is represented using a standard name such as “First name”, but this is not always true. Based on the study of HTML source code of registration forms from many well-known websites, including Google, Yahoo, MySpace, MSN, Amazon, etc., it is seen that most of them assign very meaningful, easy-to-understand names to the input fields, but MSN uses strange names such as "pff00000000010001" for last name and “pff00000000010007” for gender. To partially address this problem, we may also look at the value attribute to identify the meaning of the input field. For example, if COP knows that the first name of the child is Bob (the parent provides it to COP) and finds “Bob” in an input field, it will know that this is a ‘First name’ field. This may however not work when dealing with values such as age. An HTTP request may have many other numerical values making it almost impossible to find the age value accurately and then perturb it. Another solution is to parse the HTML source code of the registration form. For example, to identify the age field, we may first locate it in the HTML source by searching for the keyword “age” because it is the text usually displayed to the client in the browser. We can then check for something similar to <Input type= “text”…>, representing the input field for age. Values selected by the child for fields related to the control types of “radio”, “checkboxes”, “menus” could also be very difficult to identify, and hence perturb, by looking at the http request alone. For instance, some websites, instead of obtaining the user’s age through the ‘text’ or the ‘textarea’ element, provide a ‘menu’ control, listing all the possible ages. Again, parsing the HTML source would help in identifying the real meaning of the data. One of the greatest technical challenges in COP was to implement a robust HTML parser.

However, the parser has been successfully implemented. It is based on a keyword search. The keywords represent the child’s information such as first name, age, etc. that are filled in by the parent during the installation phase. The parser works in two steps;

Recognition Phase: In this phase, all input fields in the HTML page are recognized. For the experimentation and testing phase, all input fields’ color was changed to blue. It was seen that the parser successfully recognized all input fields in all tested registration forms available online.

Perturbation phase: The inputs as entered by the child in the form are perturbed using the rules of perturbation and the parent's set preferences.

2. Perturb Special Data: Disclosure of email address is seen as a major compromise to the child's privacy as indicated by a focus group study [**26**]. For this reason, disclosing the email address of the child, without perturbation, is not an option. Perturbing the email address submitted by the child may not be simple in the sense that it may lead to many issues and conflicts in certain cases, for example, (i) a web site may send a link via email and the child is required to click the link to complete his registration; (ii) when the child requests for a forgotten password, it is sent through email; (iii) when the child subscribes to e-newsletters, e-magazines, which are delivered to the provided email address. In all these three cases, email address perturbation is a major challenge. As argued, disclosing the true email address is not an option because most parents are not comfortable with that. A completely perturbed random email would not work either. At this stage, we consider the best option is to replace the child's email address with the parent's email address. Parents can forward these messages to their children.

It has been shown that a large portion of the US population can be re-identified using a combination of 5-digit ZIP code, gender, and date of birth [**66**]. Hence disclosure of these three items at the same time should be considered a threat to privacy. COP can warn parents who opt to allow all these three items to be disclosed to websites when setting preferences for their children.

Another challenge arises with the practice of using security questions to help users recover their forgotten passwords. The user has to provide an answer to the question he selects. Often, the questions ask for some private information of the user, which potentially leads to a privacy breach. One solution is that COP could record users' answer and replaces it with a random string. During the phase of password recovery, COP will compare the answer provided by the child with the recorded answers. If a match is found, it will replace the answer in the web request with the random string.

3. Parental Verification and Automated Parental Consent: COPPA requires that, before a website can collect information from a child below 13 years of age, they should obtain verifiable parental consent. Obtaining this consent every time that information is requested from a child may get cumbersome both, for the child as well as the parent. Therefore we need a way of delegating this responsibility to the tool for approving website policies. Note that since COP will be enforcing the privacy preferences set by the parent, irrespective of the operator's policy, COP can safely approve of any website policy. This brings the issue of the tool obtaining the parent's consent and providing it to the operator in an automated way. For this purpose, COPPA suggests few methods including an electronic consent form, a toll free number or credit card verification. To enable automated consent, credit-card-based verification seems to be the best choice. In this case, COP also has to function as a network service by listening to a well-known port number, so that the web server may automatically fetch the credit-card number for age verification. This is obviously dangerous due to the lack of authentication of the remote web server. An attacker may exploit this to steal credit card numbers. To address this problem, we propose to apply the technique of identity-based signature (IBS) [**67**][**68**]. Specifically, any trusted third party could work as a key generator to generate a public/private key pair for the parents, where the public key is simply their email address or phone number. These keys are stored in COP. As an automated authorization service, upon request, COP will digitally sign a consent form with the private key. The signed form together with the public key parameters is sent to the merchant site, which then verifies the authenticity of the form. This thus provides automated verifiable parental consent to the site.

4. Handling Security Attacks: If a malicious website is aware of the fact that the client is using COP to perturb the outgoing registration data, it may attempt to bypass the detection of COP. As we discussed above, accurately parsing data field is a challenging task and with such potential attacks, it becomes even more difficult. For

example, the website may obfuscate the representation of the http GET/POST request to thwart the analysis of COP. In the registration form, the text “age” may be replaced with an “age” image. Another challenge could be caused by JavaScript (or Jscript). In normal cases, a button control is embedded in a form; once clicked, the web browser composes a http request by putting together all user input data and sending it in a request (using either GET or POST method). In this case, our plug-in implementation of COP is able to intercept this standard request. However, with JavaScript, a webpage may link a button to an embedded JavaScript code, for example, <input key="send" onclick="SendData()" value="sendsecretly" type="button">, where SendData() is a JavaScript function. This function can easily read all the data the user has provided so far, encode it in a specified secret way understandable to the web server alone, and finally send the data to the server. Thus, the browser (and COP) will not see the original data fields as the values look random. This in turn could make perturbation of the data an impossible task. To address this attack, we consider automatically adding a hook JavaScript method before the JavaScript data submission method. For example, we may implement COP inside the web browser layout engine (e.g., Gecko for Firefox). When it detects the above HTML source code, it can modify the code to <input… onclick= “ICheckFirstHook(); SendData()”…> instead. Here ICheckFirstHook() is a JavaScript function, added by COP to the HTML source to check/perturb the user input data before it is passed to the original JavaScript function of by the web server.

5. Different Data Types: One challenge in that our goal is to distort several different data types. For this purpose, we can categorize the data into three groups; Numerical, String and Enumeration. The same data perturbation technique cannot be used as each data type has different properties that we are trying to preserve during the perturbation procedure.

6. Preserve Meaning of Data: Another challenge is that instead of randomly perturbing the data, we need to ‘forge’ data as we face the special online case of registration, where in most cases verification is performed on the data the user submits. Example of such cases would be credit card number (validated in accordance

with card type), zip code (checked for existence), etc. We also need to preserve statistical characteristics of the original data after perturbation. This implies that from the website's point of view, the distribution of its users' ages (for example) can still be evaluated from perturbed data without any knowledge of their true ages. Our model would allow websites to build statistical models required for their business purposes while protecting the child's privacy.

7. Linking Data Items: Finally, it has been shown[9] that a large portion of the US population can be re-identified using a combination of 5-digit ZIP code, gender, and date of birth. Hence disclosure of these three items at the same time should be considered a threat to privacy of the child.

This problem has been classified as a future research problem and time will be spent to identify all major security attacks and overcome those.

[9] Uniqueness of simple demographics in the U.S. population. L. Sweeney.

Chapter 9

COP User Evaluation Study

We investigated results around user's attitude towards COP as a solution for enforcing COPPA. We present results from the trial of a preliminary version of COP. COP clearly achieved higher acceptance among users as a way of protecting children's online privacy. The user responses to survey questions highlighted certain flaws which are currently being addressed.

9.1 Data Collection

We presented COP to a group of 20 users, all of whom are familiar with technology but not necessarily experts. The users were drawn from different academic backgrounds including law and policy, information technology and science, security and risk analysis, with all users having basic knowledge about privacy and privacy related laws. By presenting the users with a working version of the tool, we aimed to gain insight into the behavior and feelings that the participants had in regards to the usability of COP, user's comfort with trusting and using the tool, user's comments on its technical implementation and whether users regarded COP to be a good solution for the issue of child online privacy. One challenge was to gain this insight without intruding or providing them with additional guidance that can influence behavior. Therefore, we equipped the users with certain background knowledge and a brief introduction to the tools usage.

The users were initially provided with an introduction to COPPA and its clauses, and a summary of existing solutions for protecting online privacy of users and children in particular. This information was considered as background information, necessary for the user to provide relevant and well-informed feedback about COP. The COP system design overview was explained to them to enable them understand the overall working of the tool and understand its role in actually implementing COPPA.

The study lasted for 75 minutes at the end of which, the participants were asked to fill out a survey about their experience with using the tool. We were also able to collect participants' thoughts about COPPA and existing solutions throughout the study, and record their opinions about specific aspects. The initial part of the study was in the form of a discussion where the users were asked questions and were required to provide their comments and thoughts.

9.2 Data Analysis

After gathering the data and transcribing the recorded comments, we analyzed it by dividing the 75 minute session into the following categories;

- Users' thoughts on the feasibility of achieving COPPA's clauses
- Users' attitude towards existing solutions
- Their attitude towards COP's preliminary technical implementation
- COP's usability
- COP as an effective solution to protect children's online privacy

We analyzed these findings against our understanding of the participants' academic backgrounds. We first looked at their general comments on the design overview as a whole in order to establish initial flaws in the system if any and then focused on its technical implementation at the client side. We further divided the users' feedback on the technical implementation into two phases; (i) Installation phase, and, (ii) Working phase. This helped us clearly establish the areas that require further work and finally establish an outline of our future goals for improvising COP.

9.2.1 COPPA and Existing Solutions

In this section we analyze the discussion on COPPA, what users thought was essential to achieve, what they thought was technically challenging to implement and their attitude towards existing solutions.

Each user was given a handout with a summary of COPPA. Further, the goals of COPPA, clauses relevant to our study, the sliding scale approach and the significance of the term “verifiable parental consent” were explicitly explained to the users. Users unanimously agreed that verifiable parental consent was difficult to realize technically. They were then introduced to three recommendations of the sliding scale approach, (i) Credit Cards, (ii) Print and sign consent form, and, (iii) Toll free number and questioned about the effectiveness of each option.

Users were of the opinion that using offline methods such as print and sign consent form and a toll free number could significantly slow down the process of registration and access for the child, have the overhead of hiring extra personnel and add to unnecessary financial costs for the company. These two methods were discarded after some discussion and using credit card verification to obtain verifiable parental consent was seen as an effective option. One user pointed out that credit card again could only confirm adult consent and not essentially prove the relationship of the child to the credit card holder. However, we are at the mercy of available technology and adult consent seems to be the best option. Certain users were concerned about children getting access to credit cards of parents in an unethical manner. This certainly is not an issue for our tool since credit card number is required only during the initial phase of the installation of COP and is stored to provide future automatic consent. We do not see any reason why a child would be interested in installing COP. Finally these issues are beyond the scope of our project and tool.

We thought that introducing the users to existing solutions for online privacy would be a good way to make users aware of the other attempts to protecting privacy. This would essentially help them give informed opinions about COP and compare it to other solution and tools which exist. While doing so it was necessary not to influence their opinions about the tools in any manner.

Disabling cookies as a solution was immediately thought as not being effective to achieve most requirements of COPPA and as a rightly user pointed out that "*the child could still give information*". Anonymizers were the second solution but again as a user indicated, "*it still does not prevent a child from being like this is my name, this is my address, this is my phone number. It just prevents direct traffic tracing... it doesn't really address the issue of [explicit] information disclosure*". The third solution, having privacy policies for websites, was regarded as not sufficient because "*having a policy doesn't mean the company enforces it or follows it*" as a user said. Trust seals were introduced to users as a fourth alternative. 1 out of the 20 respondents claimed to check for trust seals when making a purchase. This implies that trust seals alone cannot be entrusted with the task of protecting a child's privacy online.

Following these general solutions, a few child-specific solutions were presented including blocking software, content filters, and OS parental controls such as in Microsoft's Windows Vista. Users were of the opinion that "*It works sometimes but then also a lot of sites like when new sites popup filters are a little bit lagging behind them in catching them right away.*" Thus using such software to protect a child's privacy is not recommended.

POCKET (Parental Online Consent for Kids' Electronic Transactions) is a solution which has a similar goal as that of COP and we thought it could be used to set the groundwork for introducing COP. The system design of POCKET was explained in detail to the users and their opinion of the tool for being used as a comprehensive solution to child online privacy was sought. Users saw it as a solution with "*unbalanced weight put on companies to change their existing internet infrastructure*" and saw it having "*nasty system integration problem*". They were of the opinion that the fact that "*sites are going to have to change to fit the predetermined format for this one database and not the other way around*" was infeasible. Also, it would imply "*financial overhead to switching over to this system*" and "*every company would have to call consultants*".

The users saw POCKET as having lot of responsibility vested on the server side. One of the participants was of the opinion that it "*only addresses child privacy on sites that are targeted towards children…because like amazon.com, their target audience is not 5-13 years olds, their target audience is more like 18-35 year olds*". Therefore, sites such as amazon.com have no or very little incentive to be POCKET-compliant. Users were also concerned that building trust in the TTP would take significant amount of time and were not convinced with the fact that all sites need to be POCKET compliant for the child to disclose information to them.

9.2.2 COP- System Design and Technical Implementation

A step by step approach was adopted in order to present COP to the users. The system design overview was explained initially to make users understand the high level function of the tool. The second step included encouraging users to install the tool using only the handout of instructions and then try using the tool to register with various sites.

Not many users had comments about the setup of the system design. They only expressed few concerns including how the tool would handle websites using https, all of which has already been addressed by our tool and that was explained to the users.

9.2.3 Installation Phase and Usability

Following the discussion on system design overview, users were required to install COP. A handout of instructions to install the tool was given to the participants. No additional guidance was provided as one major goal for us was to get feedback on the usability of the tool and the ease of installing the tool.

As noticed by our team, most of the users were able to easily install the tool without requiring additional guidance. They found it easy to follow instructions and install the tool independently. One of users expressed concern that "*storing the parent's information in plain text files is not good*" while another user thoguth that "*encryption of* *the data will be critical*". However, this is certainly a concern we are aware of and future versions of the tool will encrypt this sensitive data before storing it on the client side. One of the users commented that the tool was "*easy to install and doesn't get in the way of the browser*". This reconfirmed that users did not have a problem while installing the tool.

9.2.4 Working Phase

COP mainly works on the concept of data perturbation. We intended to get user feedback on the method which they thought would be most suitable for anonymizing the child's data before submitting it to websites. For this purpose users were introduced to the available techniques of data anonymization including perturbation, blocking, aggregation, sampling and swapping. There was unanimous agreement on the impossibility of using aggregation, sampling and swapping as they would require a larger population of data as compared to the data record of only a single child. The technique of blocking was discussed but discarded after understanding that sites may not allow registration using blocked data.

Finally, the concept of data perturbation was clarified and participants were allowed to fill in their comments about this method as an answer to an online survey question. They were asked if they "think data perturbation is the best option for protecting the personal information". The response statistics are recorded in Table **3**.

Most users who selected 'maybe' thought that it was not the 'best' solution but it was certainly better than the other options we have. One user said that he was "*not completely sure about this*" while one commented that "*I didn't quite understand what perturbation was doing*". This can be related more to the users not having sufficient background knowledge on the subject matter of data perturbation.

Some users had concerns such as the fact that some sites may be able to bypass the data perturbation process or the tool might change certain data that is not really private information. However, we have already accounted for these factors while developing the tool. No site can bypass the data perturbation process as it is enforced at the client side and data is perturbed before it is submitted. Therefore, sites have no scope of bypassing such a process. Secondly, the tool changes only the information that is stored on the client side and highlighted by parents as private information. Therefore the tool would not in any way change data that has not been indicated as private information by the parent.

Two of the users who opted for 'no' thought that spreading awareness about child privacy is a better option than a technical solution. However, we argue that the tool can be used in conjunction with spreading awareness to achieve a guarantee for protecting child privacy. The other users thought that blocking is a better technique. However, as we discussed earlier, blocking would not allow the child to register with most sites.

Some users added additional comments saying that data perturbation is "*the best solution we have now*", and "*it does seem like a good solution for many sites such as Google, Yahoo and other popular sites*". One of the comments was more specific to COP and it said that "*it is a good way to involve parents in children's online activities*".

Analyzing these comments it is clear that user's may not be very sure about using data perturbation for protecting privacy but are convinced that compared to other solutions it is the best available option. We have tested COP against sites other than popular sites such as Google and Yahoo and it can be used effectively in all cases. All other concerns that users had have been addressed by our tool and no specific comment seemed to highlight the inaccuracy in the technique of data perturbation as an anonymization technique. Awareness is certainly an important aspect as studies have shown that most parents are not aware of the existence of COPPA [Crossler et.al. 2007] but it can be seen more as a necessary condition rather than a sufficient condition. COP can be used in addition to educating parents about the importance of protecting their child's online privacy.

9.2.5 COP- An effective solution?

The final part of the survey focused on COP as a solution for protecting children's online privacy. Users were required to fill out an online survey answering two questions. The responses and analysis of the responses to both these questions are discussed here.

The first question focused on whether users thought COP is in accordance with COPPA requirements. The response statistics of this question are recorded in Table **3**. As it can be seen 83.33% of the users strongly believed that COP is in accordance with COPPA and as a user mentioned “it follows COPPA implemented requirements”. Users were of the opinion that COP is a well thought out solution and tries to follow COPPA to the best it can, given limitations of existing technology.

Table **3**

Table **3**: User Response Statistics of COP survey

Question	**User Reponses (%)**		
	Yes	**No**	**Maybe/ Other**
Do you think data perturbation is the best option for protecting personal information?	8.33%	33.33%	58.33%
Do you think COP is in accordance with COPPA requirements?	83.33%	16.66%	0%
Do you think COP is a comprehensive solution for attaining the goal of child online privacy?	41.66%	33.33%	25%

Only 16.66% of the users disagreed but their justifications for believing so were not convincing. For example, one user mentioned that “ *No, COPPA is intended to protect children via the server side, this is a client side solution, and the lack of widespread usage means websites still need to be COPPA compliant*”. However, in our defense COPPA does not mandate the solution to be a server side solution and a client side solution would more so empower parents with the responsibility of protecting their children rather than having to depend on websites.

Next we focused on users general views about the tool and its effectiveness for the purpose it was developed. Participants were required to state if they think COP is a comprehensive solution for attaining the goal of child online privacy. Their response statistics are recorded in Table **3**.

41.66% of the users indicated that COP is a comprehensive solution to enforce COPPA and protect children's online privacy. One user responded saying it was not too intrusive or cumbersome. Another user said it is a "step in *the right direction*" while yet another called it "*a very good framework for a solution to child privacy*". 25% of the users opted for 'maybe'. The reason indicated by them was that though it seemed like a good solution, it was not very comprehensive as yet and might require parental awareness for it to be used effectively. These factors are very true and we are working towards adding additional features to make it more comprehensive.

All of the users who did not agree with the question (33.33%) were only concerned about the very stringent software requirements and the fact that it works with a particular browser. At this stage COP is implemented as a Firefox extension and cannot be used with other browsers. However, we plan to make COP a cross browser solution by implementing it at the OS level. This would remove the dependency of COP on any of the software making it a flexible tool. Our work in this direction would address all the reasons provided by users who said that COP is not a comprehensive solution for protecting child online privacy.

After studying all these user comments, it is clear that COP is a unique and good solution towards achieving child privacy. The user interface was well received by users and there was not much effort required to familiarize the user with the tool. The only major flaw of COP remains that it is dependent on the browser with which it can be used. This flaw can be accorded to the fact that it simply the first test version of COP, and future work would be focused towards making it a cross browser, platform independent solution.

Chapter 10

Analysis of COP

10.1 Mapping COPPA Requirements to COP

COPPA spells out very specific requirements for protecting children's online privacy. Each clause of COPPA suggests certain measures that need to be taken to abide by the law and ensure that children's online privacy is protected. A clause-by-clause mapping of COPPA to COP [refer to Table **5** for detailed description] reveals that COP has been designed while keeping in mind the requirements of the law and is a comprehensive solution to enforce COPPA.

The first clause of COPPA defines the categories of personal information that need to be covered and protected to guarantee that a solution does not violate a child's privacy. COP not only perturbs all mentioned categories of data but provides additional flexibility to block further information which may be private to particular parents. COPPA describes requesting personal information from a child through online medium as a prohibited way of collecting of data from the child. COP perturbs personal data before submitting it to websites and is thus compliant with this requirement. Also, COP is a tool which works for all kinds of websites and covers all categories of operators stated by COPPA.

Verifiable parental consent is an important requirement of COPPA. It is mandatory for the operator to make any reasonable effort to ensure that before personal information is collected from a child, a parent receives notice of the operator's personal information collection, use, and disclosure practices and provides authorization. Very specific care has been taken to accommodate these requirements into COP. COP uses credit card verification procedure along with a one-time PIN number sent to parent's email address, to make sure that it is the parent who is setting preferences for the child's online activity. Parents provide consent for disclosing certain categories of information during the installation phase. During interaction (registration process), the website is expected to send its policy for approval to the client side. COP intercepts this communication as it already has the consent of the parent for disclosing specific categories of information. Further, COP sends notifications to parent via email/text message for websites the child visits and submits information to and maintains a log which can be used as a means for being aware of what information was submitted by child to which particular website.

The design of COP is done responsibly and the requirement of COPPA against children providing consent is met. Communication of website's policy from the operator's end to the client-side is intercepted by the tool and approval is sent out automatically, making it impossible for a child to provide consent. Credit card verification process during setup confirms that the parent has provided the consent and not the child. In addition to all clauses mentioned so far, COPPA has requirements spelt out for parent's access rights to their child's information. COP stores values of data fields submitted to all websites by the child using a log service. Parents can also, at any time, modify preferences to refuse further collection of data by any website.

10.2 Mapping ITP Framework Design Principles to COP Design

Design Principles were derived from the study of the ITP framework and these design principles were categorized as degree of convenience, degree of concern and degree of protection. COP design very specifically accounts for all requirements of the ITP framework to develop a comprehensive privacy enhancing solution [refer to Table **6** for detailed description].

While studying the degree of concern, the main design principles listed included making the solutions customizable with giving control of personal information to the user. If the solution is configurable to various situations, it makes the solution more acceptable. The COP installation phase is completely customizable and provides the user (parent) with flexibility to block information according to personal choice. Further, it provides for a choice of perturbing "other" information and the flexibility of blocking any information which the parents think is private. It categorizes websites into “trusted” and “not-trusted” thus accounting for the context of disclosure and the party to which information is being disclosed. Client side solution gives user complete control of their personal information. These considerations make COP a solution that tries to mitigate the degree of concern that different people may have in regards to their information.

Degree of protection demands that the solution should achieve information balance, maintain user’s identity and ensure a degree of reliability and safety with security being an important feature. In COP, the user’s (child’s) identity in the form of age is maintained and at the same time private information is protected through perturbation. It ensures information balance by expecting websites to provide privacy policies and is safe since it does not solely depend on the website for protection. All parent information is kept secure using encryption on client side. COP tries to thus achieve the highest degree of protection possible.

Users demand a degree of convenience while using technology and the last category of design principles caters to this very requirement. It recommends considering the environment of collection and actor relations while avoiding overfit and underfit technologies. It strongly suggests that the technology should provide notice on disclosure, choices for disclosure and access after disclosure. Feedback and control are essential part of the design of any technology. To address all these principles, COP is developed to exactly achieve what is required and does not include extra features to prevent it from being over/under fit. Notification is an important part of the tool and the parent is kept notified of the child’s online activities. Feedback is provided in the form of maintained logs and parents have complete control of their child’s information.

Following the mapping of the ITP framework design principles to the design of COP, two things become clear essentially; one that the ITP framework can provide a way of critiquing a technology or solution; and second that COP is a socially responsible and technically viable solution to protecting children’s online privacy.

10.3 Mapping Value Sensitive Design Principles to COP Design

The value sensitive design approach presents three paradigms namely, social complexity, psychological control and legal requirements. It can be shown easily that COP has, in a justified manner, been developed using the design approach [refer to Table **7** for detailed description]. Studying the social aspect it is apparent that it is recommended to protect children's online privacy and at the same time preserve their ability to access content. COP is not visible to the children being protected and thus children can easily access all sites that they wish to. However, simultaneously, data perturbation enables children's flow experience of online browsing with minimal content blocking.

From the psychological perspective of this design approach, user control should be maximally empowered. COP is developed as a complete client side solution and a TTP is not involved in any transactions. This makes the user the one who is in complete control of the technology and in turn, his information.

Lastly, from the legal perspective, COP should comply with COPPA on collection, parental consent, and access. As seen in section **10.1** , parents are constantly kept updated of their child's online activities and it provides for automatically obtaining verifiable parental consent. COP can hence be seen as a solution that is successfully designed using the value sensitive design approach.

Chapter 11

Conclusion and Future Direction

The objectives set out for this study have been stated in Chapter 1 and it can be seen that each of these objectives have been successfully achieved. The main objectives were to analyze existing literature on privacy and research a solution for protecting children's privacy online. The study was carried to achieve each of the objectives and this report is a summary of the study.

Following an in-depth study of the literature on information privacy, the need for a framework to analyze all the existing works in the field of privacy was felt. The I-T-P framework is thus introduced and it provides a holistic view of the existing works in the field. Information, technology and people are the three main players or factors that are affected when trying to protect privacy on the internet. The interactions between information, technology and people have been evaluated to arrive at design principles which essentially need to be accounted for while developing privacy aware solutions. These design principles have been categorized as; first, the design principles which can alleviate the degree of concern people have towards their information; second, design principles that contribute to enhancing the degree of protection of information with regards to technology; lastly, the design principles that can enhance the degree of convenience of usage of a technology for people.

The three categories of design principles summarize essentially the factors that would enhance the trust and experience of people using the technology. The I-T-P framework is thus a critical lens for analyzing works on information privacy. It can be used as guidance for future research in the field and a standard to analyze the effectiveness of any solution developed for the purpose of protecting information privacy.

Further, based on the review of existing works, child online privacy was recognized as a problem requiring attention and is researched as the topic for the case study. The motivation for the study is from the fact that increasing number of children between the ages 6-13 years are accessing the internet and there is an urgent need to protect them in this unsafe online environment. COPPA is the law which governs the actions related to online privacy of children. The study of each clause of COPPA further established the requirements in detail.

Child Online Privacy (COP) tool has been successfully implemented and tested as a solution to the acute problem of protecting children's privacy online. All design principles that were recognized, during the in-depth literature review, as essential for the development of a privacy enhancing solution have been accommodated while developing COP. COP utilizes the value sensitive design approach that adopts a tripartite methodology by systematically integrating and iterating on conceptual, technical and empirical investigations of online privacy.

Results from the user evaluation study and the tests performed to understand the viability of COP are listed. The positive response from the users confirms that COP is a promising and easy-to-use tool. The evaluation studies have also highlighted a few shortcomings of COP. COP has been implemented as a Firefox extension and it perturbs outgoing data. This is under the assumption that most people opt to use Firefox as their browser. If this assumption is not true then it is possible for the child to bypass the tool. For example, using another browser such as Internet Explorer to access the internet, the COP tool will not be effective as it is developed as an extension to Firefox browsers only. Therefore, to overcome this limitation, COP needs to be a cross browser solution and be implemented at the OS level to ensure that all methods of bypassing the tool are blocked. For the current version of COP, parents can use Firefox as their default browser and uninstall and other browsers but future research may look at the possibility of implementing COP at the OS level, thus making it a more versatile solution. If COP is successfully implemented at the OS level, it can perturb all outgoing data irrespective of the browser being used.

More detailed evaluation studies can be performed with a larger and more diverse user group to recognize technical flaws if any as well as any usability issues. COP can be presented to parents who are the potential users of the tool. Learning their perspective would help in establishing COP as a popular solution. Efforts can thus be directed to making it a tool that is usable by parents who are not well versed with technology.

Another aspect which can prove the viability of any solution is the quantification of its efficiency level. COP is a tool for protecting privacy and no single method of quantifying privacy exists. A study to quantify the level of privacy achieved by COP could establish a measure for judging the tool and its efficiency. Based on these measures, concrete studies to increase the level of privacy achieved by COP can be performed.

COP is a promising solution to the real problem of protecting the online privacy of children below the age of 13years and can be established as a full fledged solution to alleviate the concerns of parents. The ultimate goal could be to educate people about the dangers that children face in the online world, importance of privacy, the existing laws governing child online privacy and further encourage them to use solutions such as COP to protect their children from falling prey to any mishaps due to lack of privacy online.

Child online privacy is an issue that calls for immediate consideration. Through this study we hope to encourage more active research towards finding solutions to this grave problem and encourage researchers to make COP and other similar solutions available to assist parents and assuage their concerns.

Bibliography

1) Retrieved May 2008, from http://www.netlingo.com/right.cfm?term=Internet%20era

2) Coffman, K. G; Odlyzko, A. M. (1998-10-02) "The size and growth rate of the Internet" AT&T Labs Retrieved on 2007-05-21

3) Philip E. Agre, M. R. (1998). Technology and Privacy: The New Landscape. Cambridge MA, USA, MIT Press

4) Coffman, K. G; Odlyzko, A. M. (1998-10-02) "The size and growth rate of the Internet" AT&T Labs Retrieved on 2007-05-21Chapter 2

5) Burleigh, S. (May 2001). Retrieved May 2008, from http://www.livinginternet.com/i/ip_growth.htm

6) Westin, A.F. "Privacy and Freedom", Atheneum, New York, 1967

7) Kobsa, Alfred, (2007) "Privacy-Enhanced Web Personalization," In P. Brusilovsky, A. Kobsa, W. Nejdl, eds.: The Adaptive Web: Methods and Strategies of Web Personalization. Berlin, Heidelberg, New York: Springer Verlag, pp. 628-670.

8) Harris, Louis and Associates and Westin, Alan, (1991), Harris-Equifax Consumer Privacy Survey 1991. Atlanta, GA: Equifax Inc.

9) Lederer, Scott (2003) "Designing Disclosure: Interactive Personal Privacy at the Dawn of Ubiqitous Computing," Working paper, Computer Science Division, University of California at Berkeley, December 2003.

10) Milberg, Sandra J., Burke Sandra J., Smith H. Jeff and Ernest A. Kallman (1995) "Values, personal information privacy, and regulatory approaches," Communications of the ACM (38:12), pp.65-74.

11) Armstrong, M. J. C. a. P. K. (1999). "Information Privacy Concerns, Procedural Fairness, and Impersonal Trust: An Empirical Investigation." INFORMS - Organization Science 10(1): 104-115.

12) H. Jeff Smith, S. J. M. a. S. J. B. (1996). "Information Privacy: Measuring Individuals' Concerns about Organizational Practices." MIS Quarterly 20(2): 167-196.

13) Cranor L., Reagle J. and Ackerman M. (1999), Beyond Concern: Understanding Net Users' Attitudes About Online Privacy, AT&T Labs Technical Report.

14) Phelps, J., Nowak, G., and Ferrell, E. (2000) "Privacy Concerns and Consumer Willingness to Provide Personal Information," Journal of Public Policy & Marketing (19), pp. 27-41.

15) Ackerman, M. S. (2004). "Privacy in pervasive environments: next generation labeling protocols." Personal and Ubiquitous Computing 8(6): 430-439.

16) Burkert, H., (1997), "Privacy-enhancing technologies: typology, critique, vision", Technology and Privacy – The New Landscape, The MIT Press, USA, 1997, Chapter 4, p. 125-142.

17) Hockheiser, H., (2002), "The platform for privacy preference as a social protocol: an examination within the US policy context", ACM Trans Internet Technology 2(4):276–306.

18) Goffman, E. (1961), "The Presentation of Self in Everyday Life", New York: Anchor-DoubleDay.

19) Ng-Kruelle, G., P. A. Swatman, et al. (2002). "The Price of Convenience : Privacy and Mobile Commerce." Journal of Electronic Commerce 3(3): 273-285.

20) Williams, D. H. (2006) "LBS Development - Determining Privacy Requirements" Direction Magazine.

21) Junglas, I. A. and R. T. Watson (2003) "U-commerce: A conceptual extension of e-commerce and m-commerce." International Conference on Information Systems 104(9): 744-755.

22) Jiang X., Hong, JI and Landay, JA, (2002), "Approximate Information Flows: Socially-based Modeling of Privacy in Ubiquitous Computing", Proceedings of Ubicomp, 2002.

23) Langheinrich, Mark, (2001) "Privacy by design—principles of privacy-aware ubiquitous systems", Ubiquitous Computing, Ubicomp 2001, pp. 273–291.

24) FTC (2006). Xanga.com to pay $1 Million for Violating Children's Online privacy Protection Rule.

25) FTC (2008). Imbee.com Settles FTC Charges Social Networking Site for Kids Violated the Children's Online Privacy Protection Act; Settlement Includes $130,000 Civil Penalty.

26) Robert Crossler, F. B., Janine Hiller, Payal Aggarwal, Karthik Channakeshava, Kaigui Bian, Jung-Min Park, and Michael Hsiao (2007) "Parents and the Internet: Privacy Aareness, Practices, And Control.", 2007.

27) Thierer, A. (2007). Social Networking and Age Verification: Many Hard Questions; No Easy Solutions. The Progress and Freedom Foundation.

28) EPIC (2000). Pretty Poor Privacy: An Assessment of P3P and Internet Privacy.

29) Cranor, L. (2003). P3P: making privacy policies more useful. IEEE Security and Privacy magazine: 50-55.

30) Robert Crossler, F. B., Janine Hiller, Payal Aggarwal, Karthik Channakeshava, Kaigui Bian, Jung-Min Park, and Michael Hsiao (2007) "The Development of a Tool to Protect Children's Privacy Online.", 2007.

31) Camp, L.J., and Connelly, K. (2007) "Privacy in Ubiquitous Computing," in: Digital Privacy: Theory, Technologies and Practices, Alessandro Acquisti, Sabrina De Capitani di Vimercati, Stefanos Gritzalis and C. Lambrinoudakis (eds.), Taylor & Frances, New York, NY, 2007

32) Friedman, B. (2004) "Value Sensitive Design. Encyclopedia of human-computer interaction," Berkshire Publishing Group, Great Barrington, MA, 2004, pp. 769-774.

33) Friedman, B., Kahn, P.H., Jr., and Borning, A.(2006) "Value Sensitive Design and information systems," in: Human- Computer Interaction and Management Information Systems: Foundations, P. Zhang and D. Galletta (eds.), M E Sharpe, Armonk, NY, 2006.

34) NRC (2007) Engaging Privacy and Information Technology in a Digital Age, National Academies Press, Washington, DC, 2007.

35) Patil, S., and Kobsa, A.(2008) " Privacy Considerations in Awareness Systems: Designing with Privacy in Mind," in: Awareness Systems: Advances in Theory, Methodology and Design, P. Markopoulos, B.d. Ruyter and W. Mackay (eds.), Springer Verlag, Berlin, Heidelberg, New York, 2008.

36) Nissenbaum, H. (2004) "Privacy as Contextual Integrity," *Washington Law Review* (79:1) 2004.

37) Sheehan, K.B. "Toward a typology of Internet users and online privacy concerns," *Information Society* (18:1), Jan- Feb 2002, pp 21-32.

38) Altman, I.(1975) The Environment and Social Behavior: Privacy, Personal Space, Territory, and Crowding, Brooks/Cole Publishing, Monterey, CA, 1975.

39) Palen, L., and Dourish, P.(2003) "Unpacking "privacy" for a networked world," Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press, Ft. Lauderdale, Fl., 2003, pp. 129-136.

40) Burkert, H.(1997) "Privacy-enhancing technologies: typology, critique, vision," in: *Technology and Privacy: the New Landscape,* P. Agre and M.Rotenberg (eds.), MIT Press, Cambridge, MA, 1997.

41) Altman, I.(1977) "Privacy Regulation: Culturally Universal or Culturally Specific?," *Journal of Social Issues* (33:3) 1977, pp 66-84.

42) Johnson, C.A.(1974) "Privacy as Personal Control," in: *Man-Environment Interactions: Evaluations and Applications: Part 2,* D.H. Carson (ed.), Environmental Design Research Association, Washington, D.C., 1974, pp. 83-100.

43) Laufer, R.S., Proshansky, H.M., and Wolfe, M.(1973) "Some Analytic Dimensions of Privacy," Paper presented at the meeting of the Third International Architectural Psychology Conference, Lund, Sweden, 1973.

44) Margulis, T.S. (1974) "Privacy as Behavioral Phenomenon: Coming of Age (1)," in: *Man-Environment Interactions: Evaluations and Applications: Part 2,* D.H. Carson (ed.), Environmental Design Research Association, Washington, D.C., 1974, pp. 101-123.

45) Margulis, S.T. (1977) "Conceptions of Privacy - Current Status and Next Steps," *Journal of Social Issues* (33:3) 1977, pp 5-21.

46) Margulis, S.T. (2003) "Privacy as a social issue and behavioral concept," *Journal of Social Issues* (59:2) 2003, pp 243-261.

47) Proshansky, H.M., Ittelson, W.H., and Rivin, L.G. (1970) *Environmental Psychology: Man and His Physical Setting*, Holt, Rinehart, and Winston, New York, 1970.

48) Stone, E.F., Gueutal, G.H., Gardner, D.G., and McClure, S. (1983) "A Field Experiment Comparing Information-Privacy Values, Beliefs, and Attitudes Across Several Types of Organizations," *Journal of Applied Psychology* (68:3) 1983, pp 459-468.

49) Wolfe, M., and Laufer, R.S.(1974) "The Concept of Privacy in Childhood and Adolescence," in: *Privacy as a Behavioral Phenomenon, Symposium Presented at the Meeting of the Environmental Design Research Association,* S.T. Margulis (ed.), Milwaukee, 1974.

50) Dinev, T., and Hart, P. (2004) "Internet Privacy Concerns and Their Antecedents - Measurement Validity and a Regression Model," *Behavior and Information Technology* (23:6) 2004, pp 413-423.

51) Goodwin, C. (1991) "Privacy: Recognition of a Consumer Right," *Journal of Public Policy and Marketing* (10:1) 1991, pp 149-166.

52) Nowak, J.G., and Phelps, J. (1997) "Direct Marketing and the Use of Individual-Level Consumer Information: Determining How and When "Privacy" Matters,," *Journal of Direct Marketing* (11:4), Fall 1997, pp 94-108.

53) Phelps, J., Nowak, G., and Ferrell, E. (2000) "Privacy Concerns and Consumer Willingness to Provide Personal Information," *Journal of Public Policy and Marketing* (19:1) 2000, pp 27-41.

54) Sheehan, K.B., and Hoy, G.M. (2000) "Dimensions of Privacy Concern among Online Consumers," *Journal of Public Policy and Marketing* (19:1) 2000, pp 62-73.

55) Culnan, M.J., and Armstrong, P.K. (1999) "Information Privacy Concerns, Procedural Fairness and Impersonal Trust: An Empirical Investigation," *Organization Science* (10:1), Jan-Feb 1999, pp 104-115.

56) Yamaguchi, S. (2001) "Culture and Control Orientations," in: *The Handbook of Culture and Psychology,* D. Matsumoto (ed.), Oxford University Press, New York, 2001, pp. 223-243.

57) Xu, H. (2007) "The Effects of Self-Construal and Perceived Control on Privacy Concerns," in: *Proceedings of 28th Annual International Conference on Information Systems (ICIS 2007)*, Montréal, Canada, 2007.

58) Culnan, M.J., and Bies, J.R.(2003) "Consumer Privacy: Balancing Economic and Justice Considerations," *Journal of Social Issues* (59:2) 2003, pp 323-342.

59) Edelman, B. (2006) "Adverse Selection in Online “Trust” Certifications," Harvard University, 2006.

60) Kargupta, H. D., S. Wang, Q. Krishnamoorthy S (2003). "On the privacy preserving properties of random data perturbation techniques." Data Mining, 2003. ICDM 2003. Third IEEE International Conference: 99-106.

61) Kim, J.J. and W.E. Winkler (2001): Multiplicative Noise for Masking Continuous Data, unpublished manuscript.

62) Krishnamurty M., D. B. a. P. J. K. (1995). "Accessibility, Security, and Accuracy in Statistical Databases: The Case for the Multiplicative Fixed Data Perturbation Approach " JSTOR- Management Science 41(9): 1549-1564..

63) Chong K. Liew, Uinam J. Choi, and Liew, C.J. (1985) A data distortion by probability distribution. ACM Transactions on Database Systems, 10, 3 (1985), 395-411.

64) Charu C. Aggarwal, and Yu, P.S. (2008) Privacy Preserving Data Mining: Models and Algorithms. Kluwer Academic Publishers, 2008.

65) Bayardo R., Agrawal R. (2005) "Data Privacy through Optimal k-Anonymization," icde, pp. 217-228, 21st International Conference on Data Engineering (ICDE'05)

66) Sweeney L. (2000) Uniqueness of simple demographics in the U.S. population, LIDAPWP4, Carnegie Mellon University, Laboratory for International Data Privacy, Pittsburgh, PA: 2000. Forthcoming book entitled, The Identifiability of Data.

67) Cha, J. C. (2003). An Identity-Based Signature from Gap Diffie-Hellman Groups, Springer Berlin / Heidelberg.

68) Shamir, A. (1985). Identity-based cryptosystems and signature schemes, Santa Barbara, California, United States, Springer-Verlag New York, Inc.

69) Trochim W. (1997), W.M.K. Knowledge Base: An Online Research Textbook: Introduction to Evaluation. Cornell University, 1997

70) Trochim, W. (2000) The Research Methods Knowledge Base. Cincinatti, OH USA: atomicdogpublishing.com, 2000

71) Somekh, B., and Thaler, M. (1997) Contradictions of management theory, organizational cultures and self. Educational Action Research, 5, 1 (1997), 141-160.

72) McNiff, J. (1988) Action research: Principles and Practices. London: Routledge, 1988.

73) Morgan, D.L. (1997) Focus groups as qualitative research. London: Sage, 1997

74) Kreuger, R.A. (1988) Focus groups: a practical guide for applied research. London: Sage, 1988.

75) CDT. (2008) Child Safety and Free Speech Issues in the 110 Congress. Center for Democracy and Technology.

76) Zwick, D., and Dholakia, N. "Models of Privacy in the Digital Age: Implications for Marketing and E-Commerce,"Research Institute for Telecommunications and Information Marketing (RITIM), University of Rhode Island.

77) WWK. (2007). Teen Research Unlimited: Cox Communications Teen Internet safety Survey Wave II. http://www.webwisekids.org/index.asp?page=statistics.

78) Pinto, R., Ishitani, L., Almeida, V., Júnior, M.W., Fonseca, A.F., and Castro, D.F (2004) "Masks: Managing Anonymity while Sharing knowledge to Servers," in: Proc. of IFIP International Federation for Information Processing, Springer Boston, 2004, pp. 501-515.

79) Milne, G.R., and Culnan, M.J.(2004) "Strategies for reducing online privacy risks: Why consumers read(or don't read) online privacy notices," Journal of Interactive Marketing (18:3) 2004, pp 15-29.

80) Bauer, M.(2003) "New Covert Channels in HTTP: Adding Unwitting Web Browsers to Anonymity Sets," in: Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2003), Washington, DC, October 2003.

81) BBC. (2004). Net blamed for rise in child porn. http://news.bbc.co.uk/1/hi/technology/3387377.stm.

82) Wang, H. W. a. M. K. O. L. a. C. (1998). "Consumer privacy concerns about Internet marketing." Commun. ACM 41(3): 63-70.

83) Udi Manber, A. P., and John Robison (2000). "Experience with personalization on Yahoo." Commun. ACM 43(8): 35-39.

84) Travis D. Breaux, M. W. V., Annie I. Antón (2006). "Towards Regulatory Compliance: Extracting Rights and Obligations to Align Requirements with Regulations." 14th IEEE International Requirements Engineering Conference (RE'06) 0: 49-58.

85) Travis D. Breaux, A. I. A. (2008). "Analyzing Regulatory Rules for Privacy and Security Requirements." IEEE TRANSACTIONS ON SOFTWARE ENGINEERING 34(1).

86) Pang-Ning Tan, V. K. (2001). "Discovery of Web Robot Sessions based on their Navigational Patterns." Data Mining Knowledge Discovery 6(1): 9-35.

87) Olson, J. S., J. Grudin, et al. (2005). A study of preferences for sharing and privacy. CHI '05 extended abstracts on Human factors in computing systems Portland, OR, USA, ACM Press New York, NY, USA: 1985 – 1988.

88) Cranor, L. R., S. B., and Kormann, D. (2003). An Analysis of P3P Deployment on Commercial, Government, and Children’s Web Sites as of May 2003.

89) Cranor, L. R. (1998). Putting it together: Internet privacy: a public concern, ACM. 2: 13-18

90) Mulligan, D., A. Schwartz, et al. (2000). "P3P and Privacy: An Update for the Privacy Community." Retrieved April 2000, 2000, from http://www.cdt.org/privacy/pet/p3pprivacy.shtml

91) Bradley, K., R. R., Barry Smyth (2000). Case-Based User Profiling for Content Personalisation. AH '00: Proceedings of the International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, London, UK, Springer-Verlag.

92) Goldberg, I., D. Wagner, et al. (1997a). Privacy-enhancing technologies for the Internet. Proceedings of IEEE COMPCON, San Jose, CA, USA.

93) Claessens, J., B. Preneel, et al. (1999). Solutions for anonymous communication on the Internet. Security Technology, 1999. Proceedings. IEEE 33rd Annual 1999 International Carnahan Conference on.

94) Chung, G. G., S.M. (2005). Cool Hunting the Kids’ Digital Playground: Datamining and the Privacy Debates in Children’s Online Entertainment Sites. System Sciences, 2005. HICSS '05. Proceedings of the 38th Annual Hawaii International Conference. Hawaii, IEEE: 194a-194a.

95) Boyan, J. (1977). The Anonymizer: Protecting User Privacy on the Web CMC.

96) Rezgui, A., M. O., Athman Bouguettaya, Brahim Medjahed (2002). Preserving Privacy in Web Services. WIDM '02: Proceedings of the 4th international workshop on Web information and data management, McLean, Virginia, USA, ACM

97) Kumaraguru, P. and L. F. Cranor (2005). Privacy Indexes: A Survey of Westin’s Studies. Institute for Software Research International, Carnegie Mellon University.

98) Lederer, S. (2003). Designing Disclosure: Interactive Personal Privacy at the Dawn of Ubiquitous Computing. Computer Science. Berkeley, University of California. Master's: 196.

99) Clarke, R. (2001). Introducing PITs and PETs: Technologies Affecting Privacy. Privacy Law & Policy Reporter 7: 181-183, 188.

100) Tavani, H. T. and J. H. Moor (2001). " Privacy protection, control of information, and privacy-enhancing technologies." SIGCAS Comput. Soc. 31(1): 6-11.

101) Sawyer, S. and H. Huang (2007). "Conceptualizing Information, Technology, and People: Comparing Information Science and Information Systems Literatures." Journal of the American Society for Information Science and Technology 58(10): 1436 - 1447.

102) Moor, J. H. (1997). "Towards a theory of privacy in the information age." ACM SIGCAS Computers and Society 27(3): 27-32.

103) Brunk, B. D. (2002). "Understanding the Privacy Space." First Monday 7(10).

104) Boritz, J. E. and W. G. No (2006). "Internet Privacy: Framework, Review and Opportunities for Future Research." Available at SSRN: http://ssrn.com/abstract=908647.

105) Chen, C.-S., J. Filipe, et al. (2006). The 'Right to be let alone' and Private Information. Enterprise Information Systems VII, Springer Netherlands: 157-166.

106) Micarelli, A., F. Gasparetti, et al. (2007). Personalized Search on the World Wide Web. The Adaptive Web, Springer Berlin / Heidelberg. 4321/2007: 195-230.

107) Jendricke, U. and D. Tom Markotten (2000). Usability meets Security-The Identity-Manager as your Personal Security Assistant for the Internet. Computer Security Applications, 2000. ACSAC '00. 16th Annual Conference. New Orleans, LA, USA, IEEE.

108) Roger Dingledine, N. M., Paul Syverson (2004). Tor: the second-generation onion router. USENIX Security Symposium archive Proceedings of the 13th conference on USENIX Security Symposium USENIX Association Berkeley, CA, USA.

109) Cranor, L.R., S. E., Jason Hong, Ponnurangam Kumaraguru, Cynthia Kuo, Sasha Romanosky, Janice Tsai, and Kami Vaniea (2005). Foxtor. Pittsburgh, CMU, C. U. Privacy and S. Laboratory: 19.

110) Levine, B.N., C. S. and (2002). "Hordes: a multicast based protocol for anonymity." Journal of Computer Security 10(3): 213-240.

111) Khosrowpour, M. (1994). Information Technology and Organizations: trends, Issues, Challenges. Hershey, PA, IGI Publishing.

112) Kuflik, B. S. a. Y. E. a. A. M. a. T. (2005). "PRAW- The Model for PRivAte Web: Research Articles." J. Am. Soc. Inf. Sci. Technol. 56(2): 159-172.

113) Marc Rennhard, B. P. Introducing MorphMix: PeertoPeer based Anonymous Internet Usage with Collusion Detection. WPES '02: Proceedings of the 2002 ACM workshop on Privacy in the Electronic Society, Washington, DC, ACM.

114) Maschiach, Y. E. a. B. S. a. A. (2002). A New Privacy Model for Web Surfing. NGITS '02: Proceedings of the 5th International Workshop on Next Generation Information Technologies and Systems, Springer-Verlag.

115) Sharad Goel, M. R., Milo Polte, Emin Gun Sirer (2002). "Herbivore: A Scalable and Efficient Protocol for Anonymous Communication."

116) Srinivasan, R. S. B. B. A. (2002). "P5: A Protocol for Scalable Anonymous Communication." IEEE Security and Privacy

117) Yuval Elovici, B. S., and Yael Spanglet3 (2005). Hidden-Web Privacy Preservation Surfing (Hi-Wepps) Model. Privacy and Technologies of Identity, Springer US: 335-348.

118) Zorzo, R. E. D. G. a. S. e. r. D. (2006). Privacy protection without impairing personalization by using the extended system MASKS and the extended contextualized P3P privacy policies. WebMedia '06: Proceedings of the 12th Brazilian symposium on Multimedia and the web, ACM.

Appendix A

Children's Online Privacy Protection Act (COPPA)

Table **4**

Table **4**: Summary of COPPA Law

Issue	**Position in United States**
*1. **Law Applicable***	The Children's Online Privacy Protection Rule (16 C.F.R. $312 et. seq.)(the "Rule"), effective April 21, 2000, implementing the Children's Online privacy Protection Act of 1998 (15 U.S.C. 6501 et. seq.). On March 8, 2006, after performing the review required by the Act to occur within five years of the Rule's effective date, the Federal Trade Commission ("FTC") voted unanimously to retain the Rule without modification.
*2. **Scope of the Law***	
a) Personal Information	The Rule primarily applies to the online collection of Personal Information from a child under the age of 13 ("Child" or "Children"). Personal Information" is defined expansively and includes: • First and last name; • A home or other physical address including a street name and name of a city or town; • An email address or other online contact information, including but not limited to an instant messaging user identifier or a screen name that reveals an individual's email address;

	• A telephone number; • A Social Security Number; • A persistent identifier, such as a customer number held in a cookie or a processor serial number, where such identifier is associated with individually identifiable information; • A combination of a last name or photograph of the individual with other information such that the combination permits physical or online contacting; and • Information concerning the Child, or the parent or legal guardian of that Child ("Parent"), that the Operator collects online from the Child and combines with an identifier described above.
b) Personal Information Collection	The Rule applies to the "collection" of Personal Information from a Child. "Collection: is broadly defined and applies to the online gathering of any Personal Information form a Child, including; • Requesting that a Child submit Personal Information online; • Enabling a Child to make Personal Information publicly available through a chat room, message board, or other means, except where the Operator deletes all individually identifiable information from postings by a Child before that are made public, and also deletes such information from the Operators records; or • The passive tracking or use of any identifying code linked to an individual such as a cookie.
c) Collection by Operator	The Rule applies to any operator of a website or online service that is directed to Children or to any other operator with actual knowledge that it is collecting or maintaining Personal Information of a Child.

	An "Operator" is any person (or entity) who operates a website or an online service and who collects or maintains Personal Information from or about the users of or visitors to such website or online service, or on whose behalf such information is collected or maintained and the website or online service is operated for commercial purposes. [19] A website "directed to Children" means a commercial website, or portion thereof, that is targeted to Children. A website does not necessarily meet this definition, however, solely because it refers or links to a commercial website or online service directed to Children by using information location tools. [20]
d) Jurisdiction/ Territoriality	The Rule applies to the collection of Personal Information about a Child by an Operator in situations that involve commerce: (a) Among the States of the US or with one or more foreign nations; or (b) In any territory of the US or in the District of Columbia, or between any such territory and 1) Another such territory; or 2) Any State or foreign national; or (c) Between the District of Columbia and any State, territory, or foreign nation.
e) Sensitive Personal Information	Not applicable.
f) Employee Personal Information	Not applicable.

3. ***Consent Requirements***	
a) General	None.
b) Sensitive Personal Information	Not applicable.
c) Parental Consent	Prior to collection, use, and/or disclosure of Personal Information about a Child, an operator must obtain from a Parent of the Child verifiable Parental consent. In addition, the Operator must provide a Parent with the option of consenting to the disclosure of the Child's Personal Information to the Operator but not to a third party. The Rule explains that obtaining "verifiable Parental consent" means that the Operator must make any reasonable effort (taking into consideration available technology) to ensure that before Personal Information is collected from a Child, a Parent: • Receives notice of the Operator's Personal Information collection, use, and disclosure practices; and • Authorizes any collection, use, and/or disclosure of Personal Information. The Operator must also take steps to ensure that the person providing the consent is actually the Child's Parent. Acceptable forms of consent include a consent form signed and returned by the Parent by mail or facsimile, the use of a credit card by the Parent in conjunction with the transaction, a call to a toll-free number provided by the Operator and staffed by trained personnel, or the electronic forms of consent discussed below.

d) Minor Consent	A Child cannot consent to the collection of his or her Personal Information. This consent must instead be obtained from a Parent of the Child. The Rule does not address consent requirements for minors who are 13 or over.
e) Employee Consent	Not applicable.
f) Online/ Electronic Consent	Electronic consent will suffice if the Parent consents using a digital certificate based on public key technology or an e-mail using a PIN or password obtained using a digital certificate or another appropriate verification method. In accordance with the Rule's sliding scale approach to Parental consent, which was extended indefinitely on April 21, 2006 by the FTC, if the Operator is not releasing the Personal Information to a third party, consent may be obtained by using e-mail coupled with additional steps to provide assurances that the person providing the consent is the Parent. Acceptable additional steps to obtain these assurances include sending a delayed confirmatory e-mail to the Parent following receipt of consent, or obtaining a postal address or telephone number from the Parent and confirming the parent's consent by letter or telephone call.
g) Exceptions to prior Consent/ requirements	In certain situations, the Operator is not required to obtain Parental consent *before* collecting and/or disclosing Personal Information about a Child. The situations include: • Where the Operator collects the name or online contact information of a Parent or Child to be used exclusively for obtaining Parental consent or providing Parental notice (the Operator must delete this information after a reasonable time if there is no response);

	• Where the Operator collects online contact information from a Child for the sole purpose of responding directly on a one-time basis to a specific request from the Child, and the Operator immediately deletes that online contact information immediately after responding to the Child; • Where the Operator collects online contact information from a Child to be used to respond directly more than once to a specific request from the Child; • Where the Operator collects a Child's name and online contact information to the extent reasonably necessary to protect the safety of a Child participant on the website. In all of the situations described above, except for here the Operator deletes the information after responding on a one-time basis to the child, the Operator must provide and seek Parental consent after the Personal Information has been collected. Moreover, the Rule generally requires that the Operator delete all contact information after the relevant transaction has been concluded. An Operator can also provide notice and seek consent after the fact to the extent reasonably necessary to protect the security or integrity of its website, to take precautions against liability, to respond to judicial process, to provide information to la enforcement agencies, or for an investigation on a matter related to public safety.
4. ***Information / Notice Requirements***	The Operator must provide two types of notice. The Operator must post a conspicuous link to a notice of its information practices on its website's homepage as well as any area where Personal Information is collected from Children. The notice must provide the

following information:

- The contact information (name, address, telephone number, and email address) for all Operators collecting information about Children on the website;
- The types of Personal Information collected from Children and the manner of collection (passive v. active);
- How such Personal Information is or may be used by the Operator;
- Whether the Personal Information is disclosed to third parties, (and the types of businesses engaged in by such third parties, the purposes for which the Personal Information is used, and whether such parties are subject to agreements to protect the information);
- That the Parent has the option to consent to the collection and use of Personal Information without consenting to its disclosure to third parties;
- That the Operator is prohibited from conditioning a Child's participation in an activity on the Child's disclosing more Personal Information than is reasonably necessary to participate in the relevant activity; and
- That the Parent can review and have deleted his/her Child's Personal Information and also refuse to permit collection or use of the Child's Personal Information (the notice must also specify the corresponding procedures for doing so).

In most instances, the Operator also must provide notice directly to a Parent of the Child from whom it seeks to collect Personal Information *before* it collects such information. This notice must contain the

	information listed above. In certain limited situations, the notice may be provided after the information is collected.
5. ***Processing Rules***	In general, the Rule prohibits unfair or deceptive acts or practices in connection with the online collection, use, and/or disclosure of the Personal Information of a Child.
6. ***Safe Harbor***	An Operator will be deemed to comply with the Rule if it complies with self-regulatory guidelines that are issued by representatives of the marketing or online industries, or by other persons, and are approved by the FTC. Industry groups must file a request with the FTC for approval of self-regulatory guidelines that meet the standards set out in the Rule; such requests are subject to notice and comment requirements prior to approval.
7. ***Rights of Individuals***	
a) Parent Access rights	The Parent of any Child who has provided Personal Information to an Operator has the right to request access to such information. Upon receiving such a request, the Operator is required to provide the Parent with the following information: • A description of the specific types or categories of Personal Information collected from the Child by the Operator, such as name, address, telephone number, e-mail address, hobbies, and extracurricular activities; • The opportunity at any time to refuse to permit the Operator's further use or future online collection of Personal Information from that Child, and to direct the Operator to delete

	Personal Information collected from the Child; and - A means of reviewing any Personal Information collected from the Child.
b) Child's Rights	An Operator is prohibited from conditioning a Child's participation in a game, the offering of a prize, or another activity on the child's disclosing more Personal Information than is reasonably necessary to participate in such activity.
c) Additional Rights	No specific requirements apply.
8. ***Registration / Notification Requirements***	No specific requirements apply.
9. ***Data Protection Officers***	Not applicable.
10. ***International Data Transfers***	Not applicable.
11. ***Security Requirements***	An Operator must establish and maintain reasonable procedures to protect the confidentiality, security, and integrity of Personal Information collected from Children.
12. ***Special Rules for the Outsourcing of Data Processing to Third Parties***	Persons or entities who delegate or outsource the responsibility for collecting and maintaining Personal Information from a Child are subject to the Rule.
13. ***Enforcement and Sanctions***	Violations of the Rule are considered to be unfair or deceptive acts prohibited by the Federal Trade Commission Act and, consequently, are

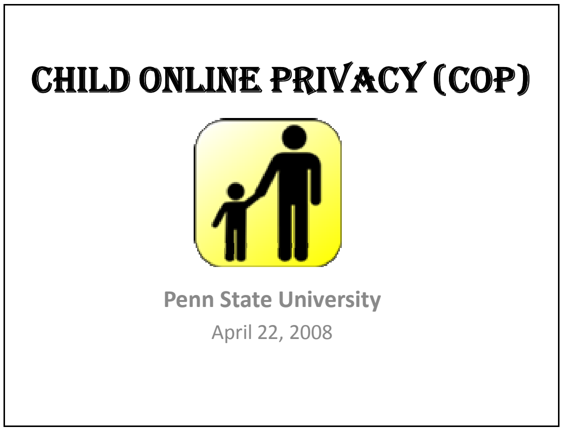
	subject to FTC enforcement actions and/or financial penalties ($11,000 per violation). COPPA also gives States and certain other federal agencies authority to enforce compliance.

Appendix B

COP Evaluation Material

B.1 Tutorial Slides

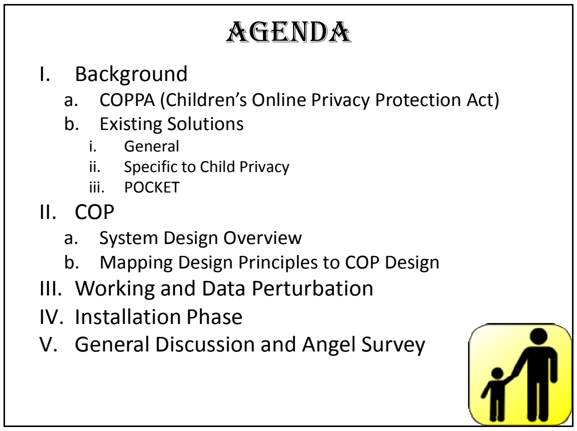
Figure **14**



CHILD ONLINE PRIVACY (COP)

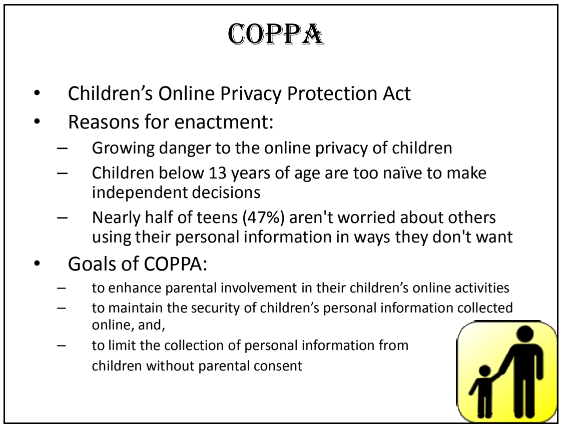
Penn State University

April 22, 2008



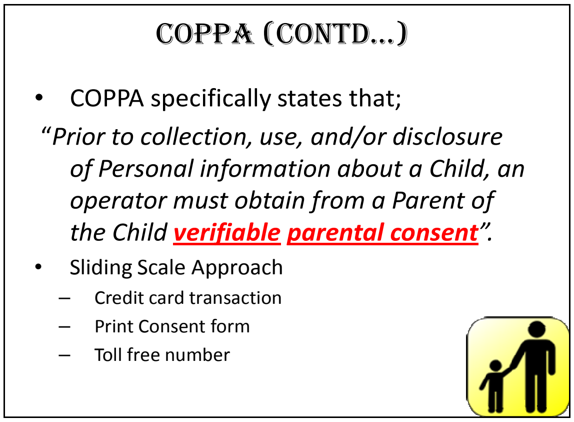
AGENDA

I. Background
   a. COPPA (Children's Online Privacy Protection Act)
   b. Existing Solutions
      i. General
      ii. Specific to Child Privacy
      iii. POCKET
II. COP
   a. System Design Overview
   b. Mapping Design Principles to COP Design
III. Working and Data Perturbation
IV. Installation Phase
V. General Discussion and Angel Survey



COPPA

- Children's Online Privacy Protection Act
- Reasons for enactment:
  - Growing danger to the online privacy of children
  - Children below 13 years of age are too naïve to make independent decisions
  - Nearly half of teens (47%) aren't worried about others using their personal information in ways they don't want
- Goals of COPPA:
  - to enhance parental involvement in their children's online activities
  - to maintain the security of children's personal information collected online, and,
  - to limit the collection of personal information from children without parental consent



COPPA (CONTD…)

- COPPA specifically states that;

"Prior to collection, use, and/or disclosure of Personal information about a Child, an operator must obtain from a Parent of the Child ***verifiable parental consent****".*

- Sliding Scale Approach
  - Credit card transaction
  - Print Consent form
  - Toll free number

EXISTING SOLUTIONS- GENERAL

1. Cookie blocking/ removing software (browser option)
2. Anonymizer
3. Privacy Policies
4. Trust Seals
   - TRUSTe (http://www.truste.org/),
   - BBOnline reliability (http://www.bbbonline.org/),
   - Hacker Safe (https://www.hackersafe.com/)

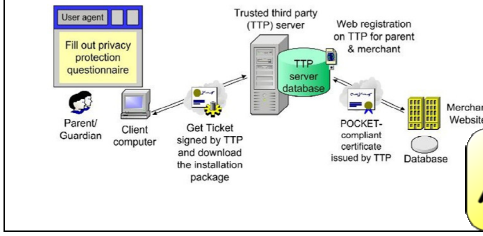
CHILDREN TARGETED SOLUTIONS

1. Blocking software and content filters
   - Filters out "bad" content
   - Blocks access to certain sites
   - E.g. Net Nanny, Parental control toolbar extension
2. OS parental controls (e.g. Windows Vista)

CHILDREN TARGETED SOLUTIONS (CONTD…)

3. POCKET (Parental Online Consent for Kid's Electronic Transactions)

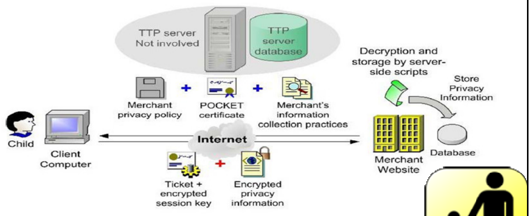
➢ Registration Phase


User agent
Fill out privacy protection questionnaire
Parent/ Guardian
Client computer
Get Ticket signed by TTP and download the installation package
Trusted third party (TTP) server
TTP server database
Web registration on TTP for parent & merchant
POCKET-compliant certificate issued by TTP
Merchant Website
Database

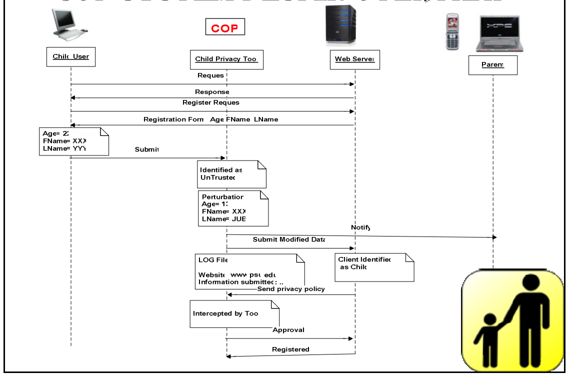
CHILDREN TARGETED SOLUTIONS (CONTD…)

3. POCKET

➢ Transaction Phase


TTP server Not involved
TTP server database
Merchant privacy policy
+
POCKET certificate
+
Merchant's information collection practices
Decryption and storage by server-side scripts
Store Privacy Information
Child
Client Computer
Internet
Ticket + encrypted session key
+
Encrypted privacy information
Merchant Website
Database

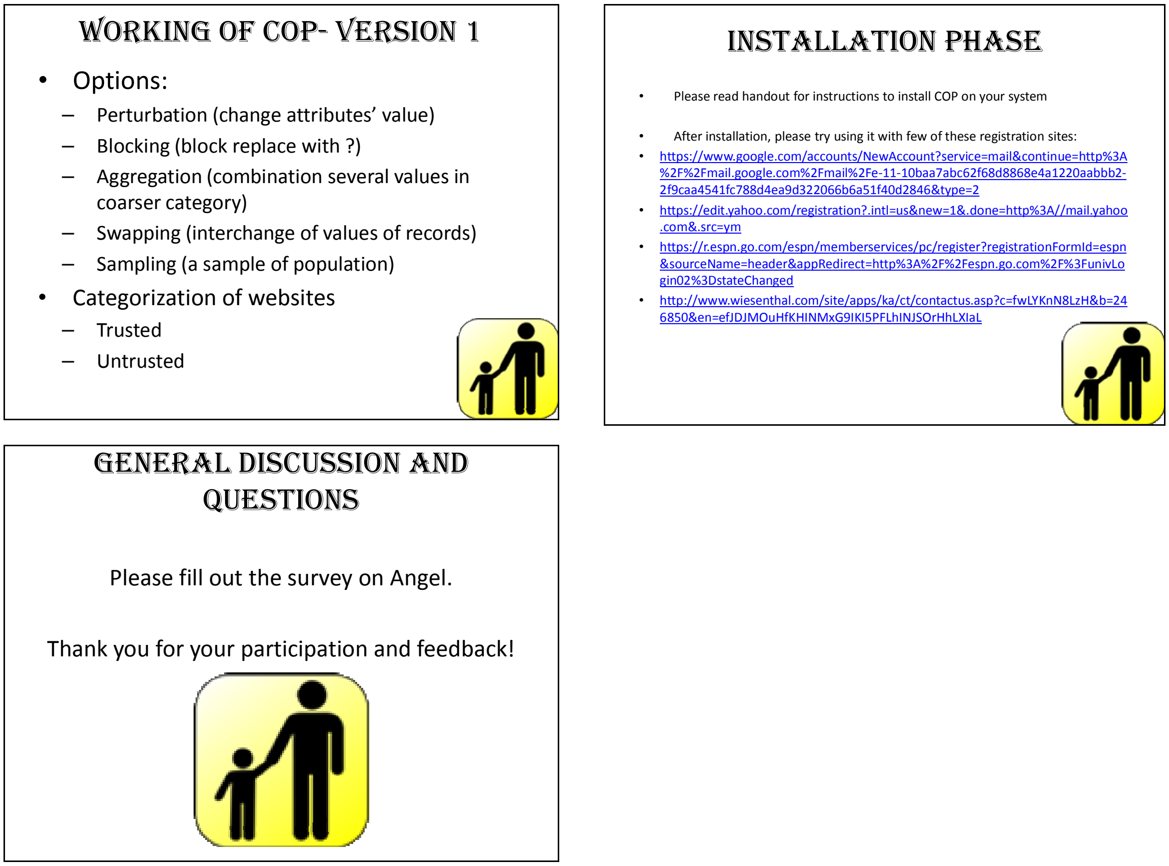
COP SYSTEM DESIGN OVERVIEW



COP DESIGN PRINCIPLES

Table . Mapping Design Principles to the COP Design

Paradigms	Design Principles	COP Design
Social complexity	Protecting children's online privacy *versus* preserving their ability to access content	• Transparent to the children under protection; • Data perturbation to enable children's flow experience of online browsing with minimal content blocking.
Psychological control	Maximally empower user control	• TTP is not involved; • Client side solution to empower parental control.
Legal requirements	Comply with COPPA on collection, parental consent, and access	• Parents are constantly kept updated of their child's online activities; • Automatically obtaining verifiable parental consent



WORKING OF COP- VERSION 1

- Options:
  - Perturbation (change attributes' value)
  - Blocking (block replace with ?)
  - Aggregation (combination several values in coarser category)
  - Swapping (interchange of values of records)
  - Sampling (a sample of population)
- Categorization of websites
  - Trusted
  - Untrusted

INSTALLATION PHASE

- Please read handout for instructions to install COP on your system
- After installation, please try using it with few of these registration sites:
- https://www.google.com/accounts/NewAccount?service=mail&continue=http%3A%2F%2Fmail.google.com%2Fmail%2Fe-11-10baa7abc62f68d8868e4a1220aabbb2-2f9caa4541fc788d4ea9d322066b6a51f40d2846&type=2
- https://edit.yahoo.com/registration?.intl=us&new=1&.done=http%3A//mail.yahoo.com&.src=ym
- https://r.espn.go.com/espn/memberservices/pc/register?registrationFormId=espn&sourceName=header&appRedirect=http%3A%2F%2Fespn.go.com%2F%3FunivLogin02%3DstateChanged
- http://www.wiesenthal.com/site/apps/ka/ct/contactus.asp?c=fwLYKnN8LzH&b=246850&en=efJDJMOuHfKHINMxG9IKI5PFLhINJSOrHhLXIaL

GENERAL DISCUSSION AND QUESTIONS

Please fill out the survey on Angel.

Thank you for your participation and feedback!

Figure 14: Tutorial Slides

B.2 Handout- Installation Steps

Installation Phase: Follow each step carefully to successfully install the tool.

PART 1:

1. Unzip the .txt files (Userinfo, ParentInfo, Preference, PIN, log) into your T:/ drive. (Make sure you keep it in the root folder itself and not in any sub-folder)
2. Double click on the .xpi file to install the extension. Click on 'Install Now' and then Restart Firefox. (You should see a pop-up with the Welcome screen of the installation phase)

[Other Notes:

- Exiting (using 'Exit' button) before completing the installation steps might cause problems with the working of the tool.

- At any stage of the installation phase, the ‘Back’ button can be used to go to previous step.]

PART 2:

STEP 1: Welcome Screen (Information not included in COP v1.0)

- Provides an introduction to the tool and also an overview of the tool’s functions.
- Click on ‘Continue’ to continue with installation.

STEP 2: License Agreement (Information not included in COP v1.0)

- Provides user with a license agreement for the copyright of the tool.
- Click on ‘Agree’ to agree with the agreement and to continue with installation.
- Clicking on ‘Do not agree’ will cause the installation process to abort.

STEP 3: Privacy Policy (Information not included in COP v1.0)

- This is the privacy policy of the tool.
- Click on ‘Accept’ to accept the policy and to continue with installation.
- Clicking on ‘Do not accept’ will cause the installation process to abort.

STEP 4: Parental Verification

- This step is useful for obtaining “Verifiable Parental Consent”.
- Enter your Name, Credit Card Number (CCN), and Email address. Click ‘Submit’ button to verify CCN against name (required to make sure you are the parent installing the tool. Your credit card will NOT be charged)
- (Once verified) a unique one-time PIN number is sent to your email address provided.

(Note: For COP v1.0, the PIN number is simply displayed as a popup once Submit is clicked. It is also stored in the PIN.txt file on your T:/ drive)

STEP 5: Password Screen

- Enter the PIN number.
- Enter a new password.
- Click on ‘SET PASSWORD’ to continue.
- This will be used to prevent children from accessing/uninstalling the tool.

STEP 6: Preference Setup

- This step allows parents to set preferences for certain sites. All sites by default are untrusted.

- To change settings for any site that is already in the table (not allowed for the 2 entries of ‘Trusted’ and ‘Untrusted’. Those are just as examples);
    - Double click on the site name
    - Check/uncheck boxes at the bottom of the screen

(Note: Checking (Y) implies that the information will be perturbed before being sent out to the web server)

-
    - Click on ‘Change’. You can see the values in the table (Y/N) against the site name being changed
    - Click on ‘Save’ to save the changes and go to next screen

- To add a new site to the table;
    - Input the site url in the textbox provided below the table
    - Check/ Uncheck boxes to select the information to be blocked

(Note: Checking (Y) implies that the information will be perturbed before being sent out to the web server)

-
    - Click on ‘add new site’.
    - Click on ‘Save’ to save the changes and go to next screen
- You also need to select a “Privacy Level”. **(Not implemented in COP v1.0)**

STEP 7: Child Information

- Input all the information of the CHILD that you wish to be blocked from revealing to websites.
- Once the form has been completed;
    - Click on ‘ADD child’. You will see a pop-up
    - Click on ‘OK’ to add another child or click on ‘Cancel’ to continue to next step

STEP 8: Notification

- Notification is an important part of any privacy solution.
- Select the medium via which you wish to be informed of your child’s activities.
- Second, select whether you wish to be informed instantly when your child reveals information, or whether you want it daily at a selected time.

(Note: In COP v1.0, the log.txt file is the only way of viewing the child’s activity)

B.3 Interactive Session Questionnaire

The following questions were part of the interactive session and were directed towards the users. Their opinions have been recorded in the evaluation study section of this report.

1. What aspect of the COPPA law would you say is the most difficult or infeasible to implement using a technical approach?
2. Can any of the following solutions be used independently to enforce COPPA?
   a. Cookies
   b. Anonymizer
   c. Privacy Policies
   d. Trust Seals
   e. Blocking software, content filters and OS controls
3. Do you think POCKET is a complete and good solution for protecting children's online privacy or do you see any flaws?
4. What do you think of a TTP as opposed to a client side solution?
5. Do you think that the mapping between COP and Value sensitive design method is justified or does it seem inaccurate in any way?
6. Which of the following options do you think is the best for protecting data?
   a. Blocking
   b. Perturbation
   c. Sampling
   d. Aggregation

e. Swapping

7. Consider the case where a child registers for an e-newsletter. If the email address is perturbed during registration, the child will never receive the newsletter. Do you have any suggestions for that? How about providing the parent's email address instead?

8. What do you think about the COP logo? Is it misleading in any way?

B.4 Web-Based Post-Tutorial Questionnaire

The following questions were presented to the users after they tested the COP tool.

1) Do you think COP is in accordance with COPPA requirements? If not then please specify how.

2) Do you think COP is a comprehensive solution for attaining the goal of child online privacy? If not, do you have suggestions for features that should be included or features that we should get rid of?

3) Do you think data perturbation is the best option for protecting the personal information?

4) What would you suggest is a good way of measuring the privacy achieved by the tool?

5) Other comments

Appendix C

Mapping COP to Given Requirements

C.1 Mapping COPPA Requirements to COP

Table **5**: Mapping COPPA Requirements to COP

Table **5**

Issue	**Position in United States**	**Method of implementation**
Personal Information	The Rule primarily applies to the online collection of Personal Information from a child under the age of 13 ("Child" or "Children"). Personal Information" is defined expansively and includes: • First and last name; • A home or other physical address including a street name and name of a city or town; • An email address or other online contact information, including but not limited to an instant messaging user identifier or a screen name that reveals an individual's email address;	1. Achieved through **perturbation** of the following categories of Personal Information; • First Name • Last Name • Address/Zip Code • Email Adress • Age • Telephone Number • Social Security Number • Picture • Gender • Credit card number

	• A telephone number; • A Social Security Number; • A persistent identifier, such as a customer number held in a cookie or a processor serial number, where such identifier is associated with individually identifiable information; • A combination of a last name or photograph of the individual with other information such that the combination permits physical or online contacting; and Information concerning the Child, or the parent or legal guardian of that Child ("Parent"), that the Operator collects online from the Child and combines with an identifier described above.	• Homepage URL • Hobbies/Interests/ Habits • Other (custom data) 2. Level of perturbation (low, medium, high) can be selected
Personal Information Collection	The Rule applies to the "collection" of Personal Information from a Child. "Collection: is broadly defined and applies to the online gathering of any Personal Information form a Child, including;	• Personal Information submitted by a child using a form (during registration) is perturbed before submission • Operator has no

	• Requesting that a Child submit Personal Information online; • Enabling a Child to make Personal Information publicly available through a chat room, message board, or other means, except where the Operator deletes all individually identifiable information from postings by a Child before that are made public, and also deletes such information from the Operators records; or The passive tracking or use of any identifying code linked to an individual such as a cookie.	access to real information of Child besides what is explicitly permitted by parent
Collection by Operator	The Rule applies to any operator of a website or online service that is directed to Children or to any other operator with actual knowledge that it is collecting or maintaining Personal Information of a Child. An "Operator" is any person (or entity) who operates a website or an online service and who collects or maintains Personal Information from or about the users of or visitors to such website or	COP works for all websites whether it is targeted towards children or not. It prevents submission of real personal information by child to any website (during registration).

	online service, or on whose behalf such information is collected or maintained and the website or online service is operated for commercial purposes. A website "directed to Children" means a commercial website, or portion thereof, that is targeted to Children. A website does not necessarily meet this definition, however, solely because it refers or links to a commercial website or online service directed to Children by using information location tools.	
Jurisdiction/ Territoriality	The Rule applies to the collection of Personal Information about a Child by an Operator in situations that involve commerce: (d) Among the States of the US or with one or more foreign nations; or (e) In any territory of the US or in the District of Columbia, or between any such territory and 3) Another such territory; or 4) Any State or foreign national; or	Applicable to forms from any website

	Between the District of Columbia and any State, territory, or foreign nation.	
Parental Consent	Prior to collection, use, and/or disclosure of Personal Information about a Child, an operator must obtain from a Parent of the Child verifiable Parental consent. In addition, the Operator must provide a Parent with the option of consenting to the disclosure of the Child's Personal Information to the Operator but not to a third party. The Rule explains that obtaining "verifiable Parental consent" means that the Operator must make any reasonable effort: • Receives notice of the Operator's Personal Information collection, use, and disclosure practices; and • Authorizes any collection, use, and/or disclosure of Personal Information. The Operator must also take steps to ensure that the person providing the	• COP uses credit card verification procedure along with a one-time PIN number sent to parent's email address, to ensure that the Parent is the one setting preferences for the child's online activity. • Parents provide consent for disclosing the categories of information during the Trusted and Un-Trusted steps of installation phase. • During interaction (registration process), the website may send its policy for approval to the client side. COP intercepts this communication as it already has the consent of the parent for disclosing certain information (depending on whether trusted/un-trusted) and the remaining information is

	consent is actually the Child's Parent. Acceptable forms of consent include a consent form signed and returned by the Parent by mail or facsimile, the use of a credit card by the Parent in conjunction with the transaction, a call to a toll-free number provided by the Operator and staffed by trained personnel, or the electronic forms of consent discussed below.	perturbed. • COP sends notifications to parent via email/text message for websites the child visits and submits information to. • COP also maintains a log which can be used as a means for being aware of what information was submitted by child to which particular website.
Minor Consent	A Child cannot consent to the collection of his or her Personal Information. This consent must instead be obtained from a Parent of the Child. The Rule does not address consent requirements for minors who are 13 or over.	• Communication of Policy from website to client is intercepted by tool and approval is sent out making it impossible for child to provide consent. • Credit Card Verification process during setup confirms that the Parent has provided the consent.
Online/ Electronic Consent	Electronic consent will suffice if the Parent consents using a digital certificate based on public key technology or an e-mail using a PIN or	Credit card verification in addition to a one-time Pin number sent to parent's email address.

	password obtained using a digital certificate or another appropriate verification method.	
Processing Rules	In general, the Rule prohibits unfair or deceptive acts or practices in connection with the online collection, use, and/or disclosure of the Personal Information of a Child.	Not possible for Operator to collect information using any unfair/deceptive means as information will be perturbed for any sort of data submitted.
Parent Access rights	The Parent of any Child who has provided Personal Information to an Operator has the right to request access to such information. Upon receiving such a request, the Operator is required to provide the Parent with the following information: • A description of the specific types or categories of Personal Information collected from the Child by the Operator, such as name, address, telephone number, e-mail address, hobbies, and extracurricular activities; • The opportunity at any time to refuse to permit the Operator's further use or future online collection of Personal Information from that Child, and to direct the Operator to	• COP stores values of data fields submitted to particular websites using Log • Parents can, at any time, modify preferences to refuse further collection of data.

	delete Personal Information collected from the Child; and • A means of reviewing any Personal Information collected from the Child.	
Special Rules for the Outsourcing of Data Processing to Third Parties	Persons or entities who delegate or outsource the responsibility for collecting and maintaining Personal Information from a Child are subject to the Rule.	Applicable to all websites or sources.

C.2 Mapping ITP Framework Design Principles to COP Design

Table **6**: Mapping ITP Framework Design Principles to COP Design

Table **6**

Paradigms	**ITP Framework Design Principles**	**COP Design**
Degree of Concern	• Customizable to accommodate for different kinds of people. • Users should be given control over their data. • Configurable to for data of different sensitivity levels. • Context of disclosure and the party to which information is being disclosed are important	• The COP installation phase is completely customizable and provides the user (parent) with flexibility to block information/ websites according to personal choice. • Further, it provides for a choice of "Other" information for parents to enter any information which they think is private. • It categorizes websites into "trusted" and "untrusted" thus accounting for the

		context of disclosure and the party to which information is being disclosed. • Client side solution gives user complete control of their information.
Degree of Protection	• Understand the ideal category of PET. • Achieve information balance, maintain user's identity and ensure a degree of reliability and safety. • Security is an important feature	• The user's (child's) identity in the form of age is maintained and at the same time private information is protected through perturbation. • It ensures information balance by expecting websites to provide privacy policies and is safe since it does not solely depend on the website for protection. • All parent information is kept secure sign encryption on client side.
Degree of Convenience	• Consider environment of collection and actor relations. • Avoid overfit and underfit technologies. • Provide notice on disclosure, choices for disclosure, access after disclosure. • Feedback and control are necessary.	• It is developed to exactly achieve what is required and not include extra features to prevent it from being over/under fit. • Notification is an important part of the tool and the parent is kept notified of the child's online activities. • Feedback is provided in the form of maintained logs. • Parent's have control of their child's information.

C.3 Mapping Value Sensitive Design Principles to COP Design

Table **7**: Mapping Value Sensitive Design Principles to COP Design

Table **7**

Paradigms	**Design Principles**	**COP Design**
Social complexity	Protecting children's online privacy *versus* preserving their ability to access content	• Transparent to the children under protection; • Data perturbation to enable children's flow experience of online browsing with minimal content blocking.
Psychological control	Maximally empower user control	• TTP is not involved; • Client side solution to empower parental control.
Legal requirements	Comply with COPPA on collection, parental consent, and access	• Parents are constantly kept updated of their child's online activities; • Automatically obtaining verifiable parental consent