Dismantling the Twelve Privacy Purposes

Sabah Al-Fedaghi
Computer Engineering Department , Kuwait University
P.O. Box 5969 Safat 13060 Kuwait
sabah@eng.kuniv.edu.kw

Abstract. *Purpose* appears in all privacy guidelines, codes, policies, and legislations. It plays a central role in many privacy-related systems such as P3P, Hippocratic databases, EPAL, and XACML. We show that the P3P 12 standard purposes mix uses of personal information with acts on personal information and mix uses of personal information privacy with other states of affairs that have several interpretations. Some purposes are not even strongly privacy-related purposes. In this paper, P3P is singled out as the object of study; however, the implication applies similarly to other projects. We propose to use chains of information handling that let the user exercise more control on the use of his/her PI and allow the personal information gatherer to excise more control on the processing and accessing of information in its procession.

1. Introduction

The privacy landscape is rich with privacy-enhancing technology in response to concern about privacy erosion caused by the increased appetite for personal information in all aspects of life. The Platform for Privacy Preferences (P3P) provides means for policy privacy specification and exchange [13]. The Enterprise Privacy Authorization Language (EPAL) concentrates on privacy authorization in enterprise-internal privacy policies [7]. The eXtensible Access Control Markup Language (XACML) supports directly-enforceable policies both for privacy and for access control in general [10]. Hippocratic databases have been introduced as systems that integrate privacy protection within relational database systems [1].

We claim that in spite of these impressive systems, insufficient attention is directed to fundamental terms of informational privacy. In this paper, we single out P3P since it is the oldest of these projects that is supposed to reach a mature foundation of specification. We direct our efforts on the most important notion in P3P and other systems: purpose.

Please use the following format when citing this chapter:

Al-Fedaghi, S., 2007, in IFIP International Federation for Information Processing, Volume 238, Trust Management, eds. Etalle, S., Marsh, S., (Boston: Springer), pp. 207–222.

2. Problem

Purpose commonly is defined in general terms as how the collected data can be used, or the intended use of the data element, or a description of the reason(s) for data collection and data access [8]. According to Thibadeau,

> Because P3P is an outstanding work, it deserves serious critique. It is essential to know what it does, and what it does not do. For a period of time, P3P will be a work in progress. There is opportunity to hone the edge on this knife so beautifully made [15].

One edge to be honed is more specificity in declaring the purpose. Purpose is defined in the 2006 W3C Working P3P Draft as "The reason(s) for data collection and use." Reasons are given in response to why questions. Why do you collect my personal information? Because I want to use it in "telemarketing," "delivery," etc. This is analogous to "Why do you want to take my money?" "Because I want to use it in trading, investing, etc."

(1) I need to know how

However, there remains the equally important how question:
How do you use my money for this purpose?
To answer this question, you don't give me reasons but actions. For example,
- I will use it to buy and sell stocks, or
- I will buy with it old houses to renovate and sell for profit.

I would be foolish if I were satisfied with only the answer to the why question.
- Why do you want my money?
- To invest it for you.
- OK, here it is.

This is approximately the logic of personal information exchange in P3P. We will propose a mechanism to specify the answer to the how and why questions concurrently.

(2) Separating the why from the how

We investigate the semantics of the P3P 12 purposes and show that their specifications sometimes reflect the answer to the why question rather than reasons that answer the why question. Take, for example, the P3P purpose "I collect personal information 'to determine the habits, interests, or other characteristics of individuals and combine it with identified data to make a decision that directly affects that individual'" [15]. The determination of habits, interests, or other characteristics of individuals, and combining them with identified data, is an answer to the how question, while making a decision is subject to the answer to the why question. As we will see later, this separation is important because there are a limited number of ways of how to use personal information; hence, the answer to the why question can be specified in a precise manner.

(3) Several interpretations of the same purpose

The interpretation of the 12 P3P purposes is overly verbose. According to Thibadeau,

> We could have hundreds of very specific purposes. For people who know about the science of human intentionality, it makes sense to be able to list many specific purposes...and the writers of the 1.0 working draft specification...understand that a purpose or intent is actually a simple thing to state and evaluate [15].

Answering the *how* question uncovers multiple interpretations of the answer to the question "Why are you collecting and using my personal information?"

(4) Is this a privacy-related purpose?

The 12 P3P purposes sometimes sway away from privacy-related situations. A P3P purpose, "Information may be used to...without tying identified data," doesn't deal with personal information defined as personally-identifying information. If these purposes are necessary, then they should not be mixed in the same basket with personal information use purposes. This point will be discussed in section nine

3. Background

This section summarizes published works that give the definition of personal information (PI) and its flow model [3] [4]. The purpose is to make the paper a self-contained work since these publications are very recent.

3.1 Personal Information

Personal information theory assumes two fundamental types of entities: *Individuals* and *Non-individuals* [6]. *Individuals* represents the set of natural persons and *Non-individuals* represents the set of non-persons. Personal information (PI) is any linguistic expression that has referent(s) in *Individuals*. There are two types of PI:

(1) *Atomic* personal information is an expression that has a single human referent (e.g., *John is 25 years old*, *Bob is a poor guy*). "Referent," here, implies an identifiable (natural) person.

(2) *Compound* personal information is an expression that has more than one human referent (e.g., *John loves Mary*).

The relationship between individuals and their own atomic personal information is called *proprietorship*. If p is a piece of atomic personal information of $v \in$ *Individuals*, then p is proprietary personal information of v, and v is its *proprietor*. An Example of non-personal information is *Spare part 123456 is in store XYZ*. Any compound personal statement is privacy-reducible to a set of atomic personal statements. Personal information privacy involves acts on personal information in the context of creating, collecting, processing, disclosing, and communicating this type of information.

3.2 Personal Information Flow Model (PIFM)

The personal information flow model divides the functionality handling PI into four stages that include informational privacy entities and processes, as shown in Figure 1.

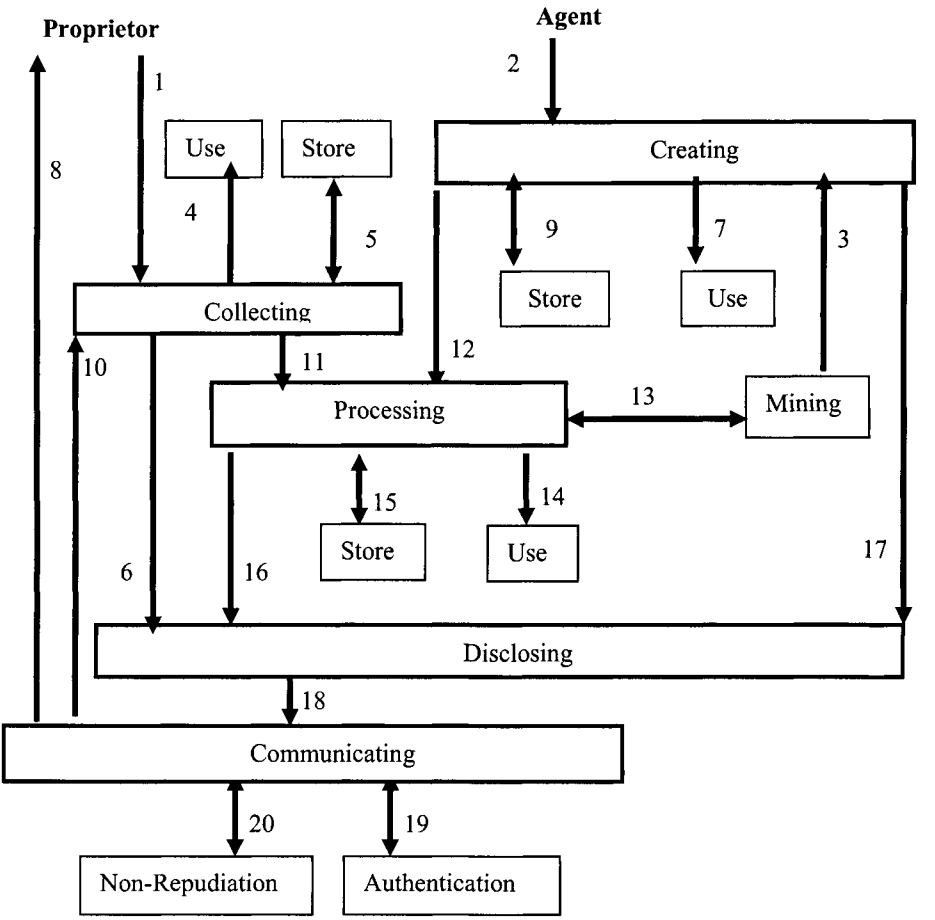

Figure 1. The PI Flow Model (PIFM).

New PI is created at Points 1, 2, and 3 by proprietors or non-proprietors (e.g., medical diagnostics by physicians) or is deduced by someone (e.g., data mining that generates new information from existing information). The created information is utilized either at Point 4 (e.g., use: decision-making), Point 5 (stored), or Point 6, where it is disclosed immediately. Processing the personal information stage involves acting (e.g., anonymization, data mining, summarizing, translating) on PI that includes using and storing processed PI (Points 14 and 15). The disclosure stage involves releasing PI to insiders or outsiders (Points 18, 19, and 20). The "disposal"

or disappearance of PI can happen anywhere in the model, such as in the transformation to an anonymous form in the processing stage. "Store" in Figure 1 denotes both storing and retrieving operations.

Using the PI flow model, we can build a system that involves a proprietor on one side and others (other persons, agencies, companies, etc.) who perform different types of activities in the PI transformations among the four stages of PIFM. We will refer to any of these as PI agents. PI agents may include anyone who participates in activities over PI.

How many ways to act on PI? Types are distinguished as acts on PI as follows:

- Gathering personal information from: (1) proprietor him/her, (2) an agent who possesses the personal information.
- Storing/retrieval of personal information: (5) raw (as collected) data, (15) processed data, (9) created data.
- Processing personal information: (11) non-mining processing of collected PI, (12) non-mining processing of created PI, (13) mining that produces implied PI, (3) mining that creates new PI (e.g., *John is risk*).
- Creating personal information: (3) automatically (mining), (1) manually by proprietor, (2) manually by non-proprietor.
- Disclosing personal information: (6) gathered (collected) data, (16) processed data, (17) created data, (8) disclosing to proprietor, (10) disclosing to non-proprietor.
- Use: (4) raw (as collected) data, (14) processed data, (7) created data.
- Communicating personal information: (18) sending through the communication channel, (19) and (20) characteristics of communication channel.

These acts form ordered sequences or *chains*, as will be discussed later.

4. Purposes and P3P

In P3P, we find 12 declared standard purposes: current, admin, develop, tailoring, pseudo-analysis, pseudo-decision, individual-analysis, individual-decision, contact, historical, telemarketing, and other-purpose. The purpose element in P3P contains one or more of these pre-defined values and can be qualified with values such as opt-in, opt-out, and always. These purposes suffer from the following shortcomings:

- Not specific, since it is possible to produce an infinite number of these purposes.
- Mixing uses of personal information with acts on personal information.
- Mixing uses of personal information privacy with other states of affairs that have several interpretations.

In order to dismantle these purposes, we need to construct a framework for the semantics of acts and uses.

5. Framework

Acts perform an action on something, while *Uses* refers to putting something to a particular purpose. Consider the case of acts and uses with respect to grapes:

(1) Acts on grape: Plant it; Eat it; Collect it, Store it, Dry it ...

(2) Uses of grape: Medical treatment of a person, Decorating cakes (eyes in a face), Celebrating [I/others], Teaching students addition and subtraction, Fueling cars (bioethanol fuel).

To distinguish between acts and uses, we adopt the structure of agent/action/patient shown in Figure 2. It includes an agent who acts on a patient. "Patient" is a term used in ethics to refer to the object that receives the action. For *acts*, this agent/action/patient becomes actor/acts-on/patient, as shown in 3. For *uses*, the model involves a third entity: the usee, as shown in Figure 4. The usee is the one used by the user to act on a patient. For example, a physician uses information to treat a patient.

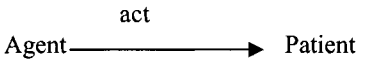

Figure 2. Basic agent/patient.

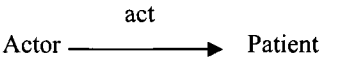

Figure 3. Binary relationship in acts.

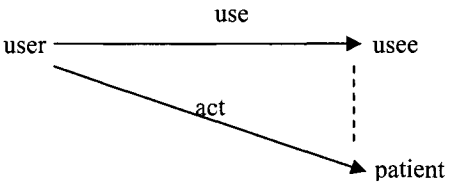

Figure 4. Ternary relationship of uses.

Here, we have a feature that distinguishes acts on personal information from its uses. In acts, the patient is personal information, while in uses, the patient is not personal information.

6. Dismantling "CURRENT"

According to P3P, the purpose "current" refers to:

> Completion and Support of Activity For Which Data Was Provided: Information may be used by the service provider to complete the activity for which it was provided, whether a one-time activity such as returning the results from a Web search, forwarding an email message, or placing an order; or a recurring activity such as providing a subscription service; or allowing access to an online address book or electronic wallet [16].

We show that this purpose:

- Mixes uses and acts
- Displays uses that have several interpretations (several possible chains)
- Displays acts that have several interpretations (several possible chains)

Mixing Uses and Acts

The definition of P3P purposes mixes acts and uses, as shown in Table 1.

Table 1. Acts and uses in purpose: current.

Example given by P3P	Type
Returning the results from a Web search	use
Forwarding an email message	act
Placing an order	use
Providing a subscription service	use
Allowing access to an online address book or electronic wallet	use

Example: Consider the phrase "Completion and Support of Activity For Which Data Was Provided." Analogously, we can introduce the following scenario:

- I am taking your money to complete and support activities for which you give me your money.
- I give you money to buy laptop from you.
- I am taking your money to complete and support delivering the laptop to you (use).

In this case, *acts* on money can include paying money to my employees, paying money for others (DHL, manufacturer), charging money, converting money ...

Table 2 shows the five examples given in P3P purpose and the equivalent money-scenario actions. In (2), "Forwarding an email message" and "Transferring money" are *acts* where PI and money are patients, respectively. Forwarding an email message is a communicating act because the message is the patient, i.e., the object of forwarding. In contrast, in (1), "returning the results from a Web search" and "delivering laptop," the PI and money are used to perform non-PI actions.

This discussion shows that P3P purpose "current" mixes uses and acts.

Table 2

	P3P Examples	Money examples
1	Returning the results from a Web search	Delivering laptop
2	Forwarding an email message	Transferring money
3	Placing an order	Placing an order for laptop
4	Providing a subscription service	Providing a maintenance service for laptop
5	Allowing access to an online address book or electronic wallet	Allowing access to workshop

***Uses* have several interpretations**

In P3P's purpose "current": *uses* have several interpretations. Figure 5 shows one possible interpretation. PI is collected and then used without processing it or disclosing it. Yet, another interpretation is possible in another stage.

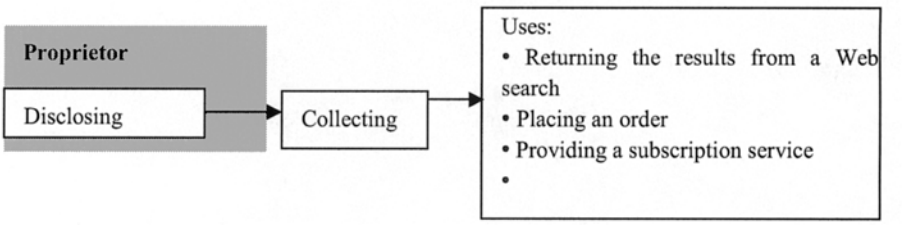

Figure 5. Uses on col

***Acts* have several interpretations**

The P3P's example "Forwarding an email message" ((2) in table 2)) depends on whether the email contains PI or otherwise. "Forwarding a non-PI email message" is a mix of use and a chain of acts, as shown in Figure 6.

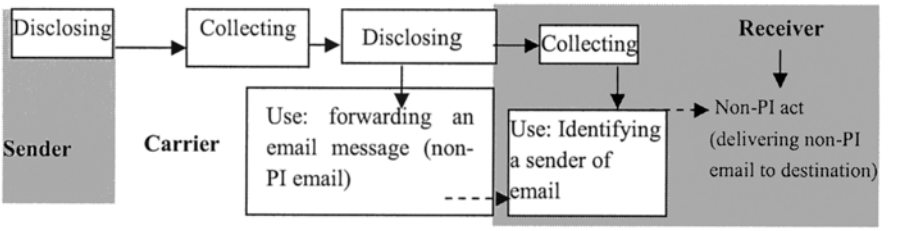

Figure 6. A mix of use and a chain of acts.

If the email contains PI, then the mail itself is part of the PI flow, as shown in Figure 7. P3P "forwarding an email message" hides important differences related to the type of email. When I say *give me your PI in order to forward an email message*, then this may mean:

(1) Forwarding that involves personal information.

(2) Forwarding that involves non-personal information

P3P's purposes mix these two different actions.

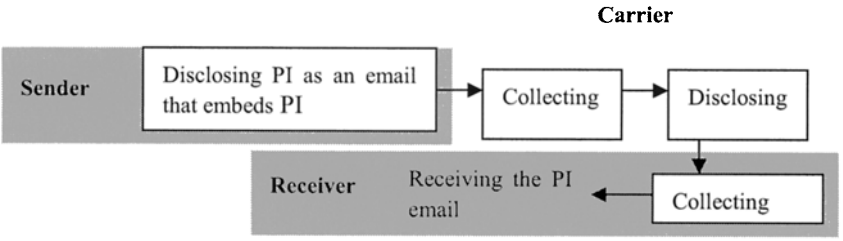

Figure 7. A chain of acts.

7. Dismantling "ADMIN"

P3P "Admin" purpose refers to:

> Web Site and System Administration: Information may be used for the technical support of the Web site and its computer system. This would include processing computer account information, information used in the course of securing and maintaining the site, and verification of Web site activity.

This would include
(1) Processing computer account information,
(2) Information used in the course of securing and maintaining the site,
(3) Verification of Web site activity by the site or its agents.

This method of description juxtaposes acts and uses. It can be written (or graphed) systematically thus: PI is *gathered*, *processed*, and *used* [acts on PI] for [uses of PI]: (1) The technical support of the Web site and its computer system
(2) Securing and maintaining the site
Notice how such a statement reflects the subgraph in the PIFM: gathering → processing →using →different types of usage. The term "processing" here may be interpreted to involve mining. In this case, the wording will be:

PI is *gathered*, *processed*, *mined*, and *used* for...

Item (3) raises doubt about the meaning of "its agents." If these agents are different entities than the collecting entity then the PI in the PIFM crosses borders to another region of PIFM through *disclosing*.
This purpose, in addition to its juxtaposing description, is also vague.
Example: According to P3Pbook.com [9],

> We ... collect ... the information contained in standard web server logs ... The information in these logs will be used only by us and the server administrators for website and system administration and for improving this site. It will not be disclosed unless required by law. We may retain these log files indefinitely.

But "will be used only by us and the server administrators for website and system administration and for improving this site" can mean anything except disclosing the information to others. The chain (1)(4)(5)(11)(13)(3)(9) means that we will collect your information, process it, mine it, and generate new information about you to be stored indefinitely. We can see that the current P3P method of specification of purpose expresses little to the user. In [5], *chains* is used to replace "business purpose."

8. Dismantling "DEVELOP"

P3P "develop" purpose refers to:

> Research and Development: Information may be used to enhance, evaluate, or otherwise review the site, service, product, or market. This does not include personal information used to tailor or modify the content to the specific individual nor information used to evaluate, target, profile, or contact the individual.

Using PI "to enhance, evaluate, or otherwise review the site, service, product, or market" can have two types of interpretation: good and bad. These two interpretations are shown in Figures 8 and 9. The exceptions in the statement of the P3P purpose try to avoid the bad interpretation. However, the exceptions by themselves do not exclude the possibility of disclosure to a third party.

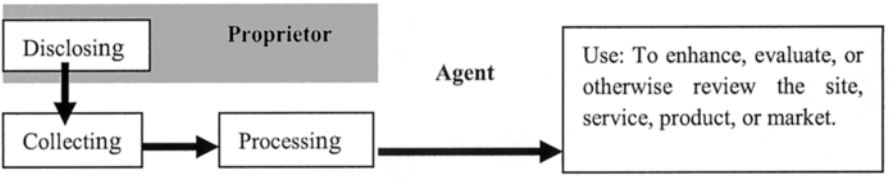

Figure 8. Good interpretation of purpose "develop."

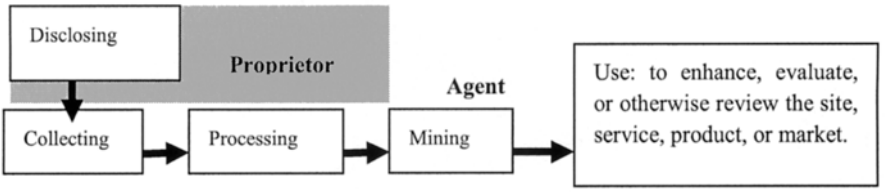

Figure 9. Bad interpretation of purpose "develop."

We can see how the fragmented P3P method of specifying purposes is where exceptions are specified in a non-systematic way. PIFM forces the specification of all of the trail of flow from source of creating PI to the destination where PI is used.

9. Where is the Personal Information?

The P3P purpose "Pseudonymous Analysis" refers to:

> Information may be used to create or build a record of a particular individual or computer that is tied to a pseudonymous identifier without tying identified data (such as name, address, phone number, or email address) to the record...

If the requested personal information will be merely anonymized, then why not asking for de-identified information in the first place. This purpose does not involve personal information. The situation is as collecting information about the shopping of a completely veiled woman. You do not need any personal information to accomplish that. Thus, the collected information is not covered by the PI flow model. This purpose is mixing privacy with ethics and etiquettes.

10. Telemarketing

The P3P purpose "Telemarketing" refers to

> Contacting Visitors for Marketing of Services or Products Via Telephone: Information may be used to contact the individual via a voice telephone call for promotion of a product or service.

But which telemarketing? Telemarketing use of gathered (raw) data, processed data, mined data ...? The P3P purpose "Telemarketing" specifies the end point of several possible chains in the PIFM. An important issue in this context is the completeness of specification of acts and uses. The telemarketing purpose is an example of a chain without acts but with use. The following example gives acts without use.
Example: Consider the following sentence from 2002 Doubleclick's policy: "DoubleClick does use information about your browser and Web surfing to determine which ads to show your browser." According to Hogben,

> P3P would cover the preceding sentence with the Element <customization/> and possibly <individual-decision/> and <tailoring/> however it is not clear from any of these, and it cannot be expressed, that it is for the purposes of advertising third-party products [12].

In PIFM, "processing" involves many standard processing of information such as tailoring, anonymization, modification, translation, summarization, and generalization. The "patient" in each case is the personal information. The problem in Doubleclick's statement is that the chain of acts on PI is incomplete where the chain does not end in a *use*. The processing of PI "to determine which ads to show your browser" in Doubleclick's policy informs the proprietor of an act (mapping PI to ads) on his/her PI without completing the chain to such use as "advertising third-party products" or without completing the chain through crossing the boundary of disclosure to another advertising agent. We can say that the issue here is a matter of specifying (complete) chains. The specification of chains of acts on PI forces the agent to fully acknowledge all of its acts on and uses of PI.

11. Alternative Approach

The PIFM provides a foundation for developing an alternative approach. Each purpose can be translated to a set of chains of acts/uses. Chains of acts on PI are chains of *information handling* that start with one of the following acts:
- A proprietor discloses his/her PI. This act also can be described as an agent collecting PI from a proprietor.
- An agent collects PI from another agent. This act may be preceded by the act of a disclosing agent to indicate where the PI is coming from.
- A non-proprietor creates PI.

These three acts are the only sources that supply any agent with PI. Suppose that a company has a piece of personal information. This piece of information is collected either from its proprietor, from another agent, or created internally by the agent. Starting with any of these sources, that piece of PI flows into the PI information handling system (manual or automatic) subjected to different acts such as processing, utilization, mining, and so forth. This track of acts can be traced through *chains*.

In our system, we envision a Proprietor Agent (called PRAG) as an agent that examines the policy of those who request collecting or gathering the personal information. PRAG represents the proprietor in order to reach a decision about whether he/she should disclose his/her personal information. We also represent any of the others as a PI agent (called PIAG). PIAG represents anyone who participates in activities over PI except the proprietor. Figure 10 illustrates this mode of operation on PI. Of course, a proprietor can have the role of an agent (his/her own agent) and any of the others also can be its own PI agent.

Figure 10. PRAG and PIAG relation.

Both PRAG and PIAG know the PIFM. Thus, they can communicate with each other according to this knowledge. General Procedure for the dialogue between PRAG and PIAG is as follows:

1. PIAG requests PI from PRAG and sends a subgraph representing the set of acts and uses that would be performed on the requested PI.

2. PRAG examines the subgraph, and a dialogue is initiated with PIAG regarding different acts in the subgraph.

3. The results of such a dialogue may lead to: agreeing to and accepting the transaction, refusing to complete the transaction, or hanging some parts of the subgraph according to the proprietor preferences.

Example: Consider the following dialogue between PIAG and PRAG
- PIAG requests certain PI from PRAG.
- PRAG asks for the set of acts/uses on the requested PI.
- PIAG sends a sub-graph.

We assume here, without loss of generality, that there is one sub-graph associated with the requested PI. Also, we assume that the sub-graph contains the specific acts and uses. For example: Use is "delivery," Processing is "statistical analysis," and Disclosure is "to XYZ crediting company."
- PRAG then asks whether it is possible to have another sub-graph that indicates that there is objection to disclosing the PI to any third party.
- PIAG answers YES if you (PRAG – its user) pay in e-cash "because we disclose your PI to check your credit."
- Etc.

From this dialogue, we see that PRAG knows PIFM. So, in general, PRAG can ask any question about PIFM. Notice that we present here the general methodology of using chains to specify the operations on PI performed on the enterprise side. The issue of user friendliness reflects a different problem that concerns the type of communication between PRAG and the proprietor. Our approach does not only allow the user to excise more control on the use of his/her PI, which he/she may elect not to do, but also allows the PI gatherer to excise more control on the processing and accessing of PI in its procession. The PIFM, for instance, can be used to define access control policy to the PI databases [2].

To compare the chains approach with the proposed W3C Platform for Privacy Preferences [16], we utilize the following scenario used in the W3C Working Draft:

> Claudia has decided to check out a store called CatalogExample, located at http://www.catalog.example.com/. Let us assume that CatalogExample has placed P3P policies on all of their pages, and that Claudia is using a Web browser with P3P built in. Claudia types the address for CatalogExample into her Web browser. Her browser is able to automatically fetch the P3P policy for that page.... Then she proceeds to the checkout page. The checkout page of CatalogExample requires some additional information: Claudia's name, address, credit card number, and telephone number. Another P3P policy is available that describes the data that is collected here and states that her data will be used only for completing the current transaction, her order [16].

Assuming that the credit card has been issued by different company, the phrase "her data will be used only for completing the current transaction, her order" means on the face the chain $(1_x)(4_x)(5_x)(6_x)(10)_z(11)_z(16)_z(6)_z(10)_x$ where the subscript x refers to CatalogExample and z refers to the crediting company (RECIPIENT element). The parenthesis in the chain denotes a don't care sequence. The chain represents a well-understood series of acts. In English, this chain expresses the following: *Your personal information will be stored and used for delivery by us and disclosed to your credit company, which solely will process it to check your credit and return OK/not OK for us. Accordingly, your merchandise will be delivered to you by us and your PI will be kept as a record of the transaction for (say) a year.*

The chain method is an explicit specification of this instance of acting on personal information (supplemented with retention period, etc.), while "her data will be used only for completing the current transaction, her order" is ambiguous specification. "Completing the current transaction" can mean for CatalogExample

many things that cover different chains in CatalogExample's region of actions and the credit card company's region of acts and beyond these two companies. According to the W3C, "P3P declarations are positive, meaning that sites state what they do, rather than what they do not do," [16] simply because it is impractical to list "what they do not do." In contrast, the PI flow model represents a "closed system" that excludes what is not specified. Thus, the specified chains are the permitted acts on PI, while the rest of the chains are not permitted.

A policy specification and its "privacy statements" can be made in chain language instead of an imprecise list of items. The PI flow model is simple to understand with a limited number of acts on personal information that can be used in designing a privacy preference language.

According to the W3C Working Draft's scenario,

> Claudia's browser examines this P3P policy. Imagine that Claudia has told her browser that she wants to be warned whenever a site asks for her telephone number. In this case, the browser will pop up a message saying that this Web site is asking for her telephone number and explaining the contents of the P3P statement. Claudia then can decide if this is acceptable to her. If it is acceptable, she can continue with her order; otherwise, she can cancel the transaction.

But how can Claudia decide? The telephone number can be used in many chains that can be interpreted as "the current transaction." The usual behavior is obeying the maximum entropy law (uncertainty means 50% opportunity for misuse; hence, cancel the transaction). However, if she is given explicit information that her phone will be used only in the chain (1)(4)(5) (store: until delivery time and use: guarantee delivery), she probably would be more willing to complete the transaction. The basic thesis here is that the clearer picture people have regarding the fate of their personal information, the more they are willing to expend their privacy. The chain method provides the user with a full general description of what is being performed on his/her PI. According to Hogben, "P3P cannot guarantee that the promise matches the practice and presents a solution that can be compared to the solution adopted by restaurants, who wish to make clients trust their hygiene practices. They put the kitchen in full view of their customers. In the same way, given a sufficiently standardized system, perhaps based on P3P..." [12]. The PIFM certainly improves the transparency of PI handling and puts "the kitchen in full view of their customers"; nevertheless, it is not specific for particular circumstances.

According to the W3C Working Draft's scenario,

> Alternatively, Claudia could have told her browser that she wanted to be warned only if a site is asking for her telephone number and was going to give it to third parties and/or use it for uses other than completing the current transaction. In that case, she would have received no prompts from her browser at all, and she could proceed with completing her order.

Again, "giving it to third parties" and "use it for uses other than completing the current transaction" are descriptive specifications that can mean many things. "Third party" may mean the credit company that already has Claudia's number or a

marketing company. The method of specification of a different third party is ambiguous. Even if these third parties are specified, what type of acts will be performed on personal information? The phrase "use it for uses other than completing the current transaction" does not specify whether the uses involved are informational acts or non-informational acts.

12. Privacy and Secrecy

In P3P, you "enumerate the types of data or data elements collected and explain how the data will be used" [16]. According to the W3C Draft,

> Identified data is information in a record or profile that can reasonably be tied to an individual... The P3P specification uses the term "identified" to describe a subset of this data that reasonably can be used by a data collector *without assistance from other parties to identify an individual.*

This means that the equation $a^2=b^2+c^2$ is "personal information" of Pythagoras because it "reasonably can be tied" to him. In another passage, it is stated that:

> IP addresses are not considered identified even though it is possible for someone (e.g., law enforcement agents with proper subpoena powers) to identify the individual based on the stored data... However, if a Web site collects IP addresses but actively deletes all but the last four digits of this information in order to determine short-term use, but insure that a particular individual or computer cannot be identified consistently, then the data collector can and should call this information non-identifiable.

This approach generates confusion between the definition of personal information and subsets of, restrictions on, and exceptional situation of this information. The definition involves, in addition to previous criticisms, ambiguity. What about the case of *John has bought Mary's laptop*? Is it *John's identifiable information* or *Mary's information*? Consider the information *John F. Kennedy is a very busy airport.* Is it identifiable information of John F. Kennedy?

Our definition of personal information provides a better formalism to specify this type of information. With its foundation, it is possible to add certain restrictions to make the information suitable in certain applications. According to the P3P Draft,

> The Working Group decided against an identified or identifiable label for particular types of data. However, user agent implementers have the option of assigning these or other labels themselves and building user interfaces that allow users to make decisions about web sites on the basis of how they collect and use certain types of data.

So, any data that you have can be "personal information" if you choose to call it so. Such an approach mixes personal information with non-personal (but may be

personally owned) information. If I have a proof that P=NP, then this is not personal information. Personal information *refers* to its proprietor.

13. Conclusion

The personal information flow model or similar theoretical framework ought to be given more attention in order to build a foundation for personal information handling policies and systems. Many issues remain to be addressed, including concerns related to syntax specification, mapping to a user's purposes, effects on access control privacy negotiation, and privacy policy enforcement.

REFERENCES

[1] Agrawal, R. Kiernan, J. Srikant, R. and Xu, Y. (2002). Hippocratic databases. In The 28th International Conference on Very Large Databases (VLDB), Hong Kong, China, August.

[2] Al-Fedaghi, S. (2007). Beyond Purpose-Based Privacy Access Control. The 18th Australasian Database Conference, Ballarat, Australia, January 29th - 2nd February.

[3] Al-Fedaghi, S. (2006a). Anatomy of Personal Information Processing: Application to the EU Privacy Directive, Inter. Conf. on Business, Law and Technology (IBLT 2006), Copenhagen, December..

[4] Al-Fedaghi, S. (2006b). Aspects of Personal Information Theory, 7th, The Seventh Annual IEEE Information Assurance Workshop (IEEE-IAW), West Point, NY: US Military Academy, June 20-23.

[5] Al-Fedaghi, S. (2006c). Personal Information Model for P3P, W3C Workshop on Languages for Privacy Policy Negotiation and Semantics-Driven Enforcement, 17 and 18 October 2006, Ispra/Italy.

[6] Al-Fedaghi, S. (2005). How to Calculate the Information Privacy, The Third Annual Conference on Privacy, Security and Trust, St. Andrews, New Brunswick, Canada.

[7] Ashley P., Hada S., Karjoth G., Powers C., and Schunter, M. Enterprise Privacy Authorization Language, W3C Submission 10 November 2003. http://www.w3.org/Submission/EPAL/.

[8] Byun, J. Bertino, E. and Li, N. (2005). Purpose Based Access Control of Complex Data for Privacy Protection, SACMAT'05, June 1–3, 2005, Stockholm, Sweden.

[9] Cranor, L.F. Web Privacy with P3P, 2002, O'Reilly & Associateshttp://p3pbook.com/examples.html.

[10] Cover, R. (Editor), Extensible Access Control Markup Language (XACML), October 10, 2006. http://xml.coverpages.org/xacml.html#v20CD.

[11] EU Directive (1995). DIRECTIVE 95/46/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, 24 October. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:EN:HTML.

[12] Hogben, G. A technical analysis of problems with P3P v1.0 and possible solutions, "Future of P3P" workshop, Virginia, USA, 12-13 November, 2002. http://www.w3.org/2002/p3p-ws/pp/jrc.html.

[13] OECD (1980). Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, http://www.oecd.org/document/18/0,2340,en_2649_34255_1815186_1_1_1_1,00.html.

[14] P3P (2002). The Platform for Privacy Preferences 1.0 (P3P1.0) Specification, The World Wide Web Consortium, April 16, 2002, http://www.w3.org/p3p/.

[15] Thibadeau, R., A Critique of P3P: Privacy on the Web, Aug 23, 2000 (Postscript, April 20, 2004). http://dollar.ecom.cmu.edu/p3pcritique/#postscript.

[16] W3C Working Draft 10, The Platform for Privacy Preferences 1.1 (P3P1.1) Specification, February 2006. http://www.w3.org/TR/P3P11/.