**Title:**

**Vision Paper: Enabling Privacy for the Paranoids.**

**What is the research problem?**

Based on history/experience, resting complete control/trust on a third party (corporations) to protect end-users privacy may not work. Knowingly or accidentally cases have been registered where such violations have happened. This paper talks about re-thinking/re-modelling so as to give the end-user the control on his/her data.

**Overview/main points of the proposed approach/architecture**

* P3P (Platform for Privacy Preferences) are standards that allow an organization to declare its privacy policy, which the end-user can go through and if he/she agrees than they can conduct business with the organization.
* Even with the P3P framework, there are loopholes and the end-users data might result with an unwanted third party, which the end-user might not be comfortable/protected about.
* Also, there might be cases were the organization itself might go bankrupt and its record might be subjected to scrutiny, in which case such data might be exposed.
* P3P as well Hippocratic Databases puts complete control/trust on the organization for safeguarding the end-user data.
* The papers proposes a better way to put back the control of the data to its rightful owner. The level of control can be defined/set by the user themselves.
* To retain control at the desired level following the papers suggests the below initial measures:
* Permission: Every time the set data is used, permission should be seeked from the end-user.
* No Copies: The user data is provided in a non-repeatable format.
* Supervision: End-user has the right to supervise how his/her data is used.
* No Integration: Data provided should not allow integration with any Third party, as a join condition.
* Retaining Control, Example Email address: Using temporary/generated email address, which are mapped to the actual email address is recommended. Here instead of exposing the address to the entire organization, the user has to rely/trust only the agent, who generates/maps the temporary address to the original user address.
* Retaining Control, Credit Card Number: The same idea discussed above can be applied to credit card info as well. However in this case the agent can be within the bank or outside. The temporary card number here is referred to as *pseudonum.* For further control the agent can impose on the *pseudonum* as well, such as, Timeout, Limited use, Restricted use, Invalidation, Isolation etc.
* Generalizing this idea, the paper talk about dividing the personal information into four categories and furthermore discussion the detail of each
* Local identifiers: User identifies numbers like SSN or telephone number are in most cases used as keys in the local database. Instead, a private identifier can be generated for the user, allowing the end-user to not expose the true sensitive data.
* Challenges: Duplicate identifier generated by the end users are a strong possibility for which the organization has to be prepared with a conflict resolution technique.
* Foreign-key identifiers: In this case the individual identifiers are not just used as local identifiers, but can also be used as foreign-key to allow legitimate integration.
* The controlling organization needs to have framework in place to identify/locate just foreign-key relations and operate accordingly.
* Value Predicates: Giving a simple example of a user trying to avail a senior citizen discount, for this the end-user is required to disclose his/her age “x”. However in this case the end user provides his/her age to the agent who in-turn gives a value predicate p(x), which in turn just gives true/false, based on the senior citizen criteria. Just in this case the end-user can safe-guard his/her personal information, yet achieve the required.
* Here either the agent can cheat by modifying the predicate or the end-user can also cheat by giving false values.
* Notary protocol/Trusted Third party protocol can be used here (weaker though)
* Multi-source value predicates: For data/predicates which information from multiple sources, which may or may not be trusted. Cryptographic ways or having a third-party which in turn is trusted by all others is a way to go in this case.
* P4P: Paranoid Platform for Privacy Preference, a framework/set of information type where an individual control his/her information.

Private information can be classified along three dimensions

1. Ownership
2. Function
3. Desired level of control

* It a simple attribute-entity model.

1. Ownership: The paper includes an example to interaction between two entities, an individual and an organization. In this case. A part of the data included in this transaction such as the purchase item detail, bill should be owned be owned by the organization, however the detail linking this purchase to the individual should be owned by the end-user.
2. Function: The level of control on the information depends on what type of role it plays in the transaction. The types are defined as below:
3. Identifier: Which is used to identify an entity like phone number
4. Service handler: A attribute which provides path to the service like email address
5. Input to the predicate: based on this value, the other entities can evaluate.
6. Copy: Attributes can be copied, to distinguish between the primary and secondary copies.
7. Desired level of control: how the user want to control the data
   1. Complete Privacy: the information should not be revealed at all.
   2. Limited Time/Use: The information can be used for limited time/use.
   3. No Predicate: cannot be used as an input to other evaluations/predicates
   4. No Integration: Integration prohibited with another subset/entity
   5. Accountable: Information shared, but the other entity is responsible if misused.
   6. Sharable: Information can be shared/further processed based on agreed terms.

* Challenges for P4P framework/model:

1. Interfaces for the Agents and entities, such generic interface needs to be implemented with a lot of details
2. Human interface: to enable people to describe their privacy goals and select policies for their agents.

**Conclusion:**

This paper emphasizes on investing and research to devise technology which give the end-user control over its data rather than relying on third-party. It recommends a framework for such process that would enable/give control to its end users.