infotrust.org

INFORMATION TRUST EXCHANGE
TECHNOLOGY DEVELOPMENT TASK GROUP

Technical Description of a Privacy-by-design Customer Profile and Content Sharing Network

DISCUSSION DRAFT

Originating author:
Richard A. Lerner, Ph.D.,
CEO, Clickshare Service Corp.

Version History

Date	Export Version	Description
2016-06-24	**V0.04**	**ITE conformed-BD**
2015-12-29	**V0.031**	**Circulating DRAFT**
2015-12-17	V0.03	Expanded Description
2015-12-14	v0.02	Initial Version

This document contains two parts:

A. System Narrative Overview
B. Proof of Concept Implementation

System Narrative Overview

This document describes the technical design and operation of an Information Trust Exchange Customer Profile and Content Sharing Network (ITE) and proposes a proof-of-concept test implementation under the auspices of the **Information Trust Exchange Governing Association (ITEGA)** (in formation).

The ITE technical infrastructure (comprised of modules, interfaces and protocols) the entities that collect information about their customers to share this information in a controlled manner that ensures transparency to the customer while supporting the efforts of entities who value the information. We initially focus on media sites and advertisers, but the concept applies to any internet-based environment where customer data is valuable.

CURRENT ENVIRONMENT – MULTIPLE COOKIES/PERSONAS

In the current environment, media sites allow advertisers to use third-party browser cookies and other means to collect and aggregate large amounts of information about the media sites' customers. Multiple advertisers, exchanges and networks are free to collect and share this information entirely independently of the media sites, and out of sight of the customer whose information is being collected. These third parties "match cookies" and assemble diverse, unconfirmed "personas" of individuals. With this project, we move control of the information back to the media sites customers choose to visit and with whom the customer has some relationship. This allows the media site to work with their customers to ensure that information management is kept transparent. Media sites can not only describe their information policies, but allow their customers to see the policy's effects and usage. This allows media sites to make a stronger case to their customers of the value of information exchange to all parties.

COMMON SHARING PLATFORM

While it is possible for individual media sites to work with their advertisers or ad-networks to share data in a controlled manner, this approach does not scale well to thousands of media sites and advertisers. An "on-your-own" approach also makes it difficult for smaller media sites to leverage the full value of the customer information they maintain. The ITE solves these problems by providing a common platform for the sharing of information, along with the ability for media sites to band together, as they see fit, to combine their information, making it more valuable to advertisers. Importantly, the ITE allows for the possibility that a individual may wish to trust a media or other site with maintain a single digital identity, or "persona" for them, which can be selectively shared, as authorized by the individual, with third parties such as advertisers, exchanges, networks or other publishers. The sharing is then subject to business rules of the CPN governing use and/or retention of the user's persona.

DEFINITIONS

The entities that make up the Customer Profile Network include the media sites and other entities that collect information to be exchanged, the advertisers and other entities that wish to use this information, and new entities that link these together into one or more networks. We use the term **Data Collector** to refer to a publisher, other media site or other entity that collects information and acts as an *identity manager* for individuals. We use the term **Profile Usage Agent** to refer to an advertiser or other entity that wishes to use the information. We call the networking entities **Data Aggregators**. And we call the shared-service that authenticates users and logs events the **Authentication and Logging Service.**

The Data Aggregators form the central component of the Customer Profile Network. Each **Data Aggregator** is an independent entity that media sites and other data collectors create or join and to which they provide information. A Data Aggregator combines the information it receives from its members and makes the information available to advertisers and other entities with whom the media sites and their individual users wish to collaborate. A Data Aggregator implements standard protocols for communicating among the other elements of the system. It implements the rules dictated by its members and the overall rules of the network. These rules generally include requirements for reporting information exchanges and making this information available to its members, and through them to their customers. Data Aggregators and the Network providers also use this information to provide settlement services for payments among the elements of the network.

Each Data Aggregator is responsible for gathering the data from its member Data Collectors/identity managers and, when instructed by one of its members, for sending portions of that data to a specific Profile Usage Agent. The network comprises many Data Aggregators so

that groups of Data Collectors can participate in the network while ensuring that their customer's information is protected by the rules they set forth. A Data Collector/identity manager can choose to be its own Data Aggregator, or it can choose to join with others to create a Data Aggregator to represent their common goals, or they can choose to join an existing Data Aggregator that suits their needs. A Data Aggregator can be privately held, or run by a non-profit organization, or even run as a commercial entity. All Data Aggregators must be members of the Information Trust Exchange Governing Association, be in compliance with its exchange rules governing user privacy and financial responsibility, among other things.

CHOICE OF DATA AGGREGATORS

We anticipate that most Data Collectors/identity managers will want to join a Data Aggregator with other Data Collectors in order to make the information they collect more valuable to Profile Usage Agents. (However, it is also possible that some individuals will want to affiliate with a Data Collector/identity manager who adopts a business strategy of not sharing data, because that is the preference of the individual). A Data Aggregator with many sources can create a more complete profile of each customer by combining the information it receives from it sources. Each Data Collector can choose how tightly they want to control the information they collect; balancing the value of keeping their data closely held against the value of combining their data with that of other Data Collectors.

In typical usage, a Data Collector/identity manager first determines that it wants to release certain information to a Profile Usage Agent (such an advertisers or ad exchange which are ITE members). It then instructs its Data Aggregator to make this information available and the Data Aggregator returns a token that the Profile Usage Agent can use to retrieve the data. The Data Collector passes this token to the Profile Usage Agent. The Profile Usage Agent sends the token to a Data Aggregator and receives the profile data, encrypted using its public key. If the profile data to be delivered is small, the Data Aggregator can return the encrypted data rather than a token so that the Profile Usage Agent can decrypt the data immediately without having to request the data from a Data Aggregator.

The ITE permits Data Aggregators to communicate with one another to simplify the process of passing data from one Data Aggregator to any of thousands of potential Profile Usage Agents. While the profile data may flow through multiple Data Aggregators to a Profile Usage Agents, the profile data is encrypted so that an intermediary Data Aggregator is not able to read the data passing through it.

COOKIES, DO NOT TRACK AND ANONYMIZING DATA; THE AUDIENCE PROFILE BOOK

It is commonly understood that the use of so-called “third-party cookies” is associated with methods of opaque compiling of profiles of specific users, often without their knowledge or consent. It is an objective of the ITE to reduce or eliminate such practices. In particular, the ITE infrastructure is intended to support respecting of “Do Not Track” signals sent by users through their browsing software or other means.

Some users may elect not to have specific attributes shared in a way that will allow them to be individually demographically, interest or location targeted by advertising or other content. Some Use Profile Agents may wish not to expose their users to such targetting, for business or other reasons. Some Data Collectors may adopt a policy of inhibiting such targetting. In such instances, ITEGA supports the development of **Audience Profile Books (APBs)** in which users are other targeted because they are within an attribute cohort of a size large enough to

obfuscate their individual attributes. APBs are discussed HERE.
https://docs.google.com/document/d/1APHGo8zrOFPKeFuoK_oURvk4GqJFTxVCfm-ZMEcds-A/edit

KEY SYSTEM GOAL – DATA CONTROL CLOSE TO INDIVIDUAL

A Profile Usage Agent can use the profile data it receives for any purpose, subject to the rules of the ITEGA. Some of these rules are enforced administratively and others are enforced by the implementation of the system. In general, a Profile Usage Agent is limited in how long it can keep the user profile information and in what ways it can be combined with other information. The goal of the ITE is to provide a Profile Usage Agent with the information it requires to perform its tasks, while keeping the information under the control of the Data Aggregator providing the information, which has a direct relationship with the individual whose personal data is to be shared

Data Aggregators send three types of data to Profile Usage Agents.

a. A unique individual identifier
b. Short-term profile data
c. Long-term profile data

IDENTITY PERSISTENCE DISCUSSED

A Profile Usage Agent may need to be able to associate *its own* data with an individual, over periods of time beyond single sessions. For example, an advertising platform may want to record which advertisements it has delivered to an individual during the last week or month, to ensure the desired distribution of advertisement delivery *("frequency capping")*. This requires that the Data Aggregator deliver to the Profile Usage Agent, the same unique identifier for the authenticated individual each time it delivers profile data to the Profile Usage Agent. In order to prevent Profile Usage Agents from pooling information, the unique identifier for an individual will be different from different Data Aggregators (in the event an individual has relationships with more than one Data Aggregator/identity manager) and each Profile Usage Agent may receive a different unique identifier for the same individual. Furthermore, the unique identifier may change periodically, allowing the Profile Usage Agent to collect information on an individual for a period of time, say a week or a month, but not indefinitely.

A Profile Usage Agent can use and record the unique individual identifier as an index to its own information about the individual, **Private Information**, such as what advertisements have been delivered in the past. This Private Information might include how the Profile Usage Agent previously classified the customer, with the usage restrictions described below for short-term profile data.

RULES ABOUT SHARING DATA

The short-term profile data can be used to assist in the delivery of content to the customer for the current session, but, as an ITEGA business rule, may not be stored or aggregated for future use.

The long-term profile data can be used to generate aggregated private data recorded using the unique individual identifier. The actual data may or may not be allowed to be stored long term.

By rule, in no circumstances shall a Profile Usage Agent share the data it receives from a Data Aggregator with other third parties. Most importantly, a Profile Usage Agent may not share with others the unique individual identifier it receives or an identifier assigned by the Profile

Usage Agent that corresponds to a unique individual identifier. A Profile Usage Agent may share private data, including data aggregated from the long-term profile data.

BILLING OF PROFILE USAGE

Profile Usage Agents can be billed based on the quantity and types of profile data delivered. The Data Aggregators are responsible for recording use of profile data. The Profile Usage Agent may receive a profile package that it can decrypt itself, or it may receive a token it needs to send to the Data Aggregator to receive the profile data. In the former case, the Data Aggregator records each profile package it generates in response to a request by a Data Collector. The Profile Usage Agent can reconcile these with a record it maintains internally of profile packages it receives. If the Profile Usage Agent receives a token, the Data Aggregator records the use of the token when sending profile data to the Profile Usage Agent. It also seconds a record of the profile-data usage, as an Event, to its appropriate **Authentication and Logging Service** (ALS).

Each **ALS** either operates its own **Settlement Service** or uses a shared Settlement Service. A settlement service, based upon usage data recorded over time by a Data Aggregator and reported to its ALS, computes from ALS-provided event-log records the fees due from each Profile Usage Agent and the fees due to each Data Collector and settles them periodically across existing financial networks

ROLE OF A DATA COLLECTOR (publisher/indentity provider)

A Data Collector maintains a relationship with its customers (individuals) that includes the gathering of customer information. This implies that the Data Collector has some means for identifying individual customers. For web services, this is most often accomplished by having individuals log in or otherwise authenticate with the service. Once identified, the Data Collector can associate the information it gathers with the proper individual. A Data Collector can gather information about an individual both by direct communication and by inferring information from context or the actions a customer performs. Direct communication may take the form of a profile the customer can complete. Inferred information may include what pages the customer requested, purchases the customer makes, etc.

In order for a Data Aggregator to combine information from multiple sources, the Data Collectors must agree on some minimum attributes to be collected and sent to the Data Aggregator so that data on an individual from each source can be matched. Each Data Aggregator can set its own standards beyond minimum standards established by the ITEGA. However, it is likely to include personally identifying information such as an email address, name, address, phone, etc. The Data Aggregators are responsible for protecting this information as required by the originating user, the ITEGA and any law or regulation

SINGLE-SIGN ON FOR SUBSCRIPTIONS

In some cases the Data Aggregator may provide Single Sign On services for its members or other means (e.g., cookies) for identifying an individual across all of its member sites for services such as subscriptions, access rights or event-billing or payment. This cross-site authentication will be managed by an **Authentication and Logging Service**.

Proof of Concept Implementation

1) Pre-request Data/Control flow

Data Collectors generally send static profile update information to their Data Aggregator whenever new profile information or updates become available. Under ITEGA rules, the Data Collector directives to update information about their users through additions, deletions or changes must be followed by the Data Collector(s) in order that the Data Collector's user profiles can be trusted as authoritative.

2) Real-time Data/Control flow

a) An individual initiates a request to a content site (e.g., a browser request for an article).

b) The content site requests, from its Data Aggregator (using a direct server call), a profile package for a particular individual and Profile Usage Agent. This request can also include additional "dynamic" information, such as the content or content category of the individual's request. The Data Aggregator records the request and returns a value the Profile Usage Agent uses to obtain the profile data. Business rules will determine whether the data returned to the Profile Usage Agent is all known profile data about an individual, or only a segment of data depending on the type of Profile Usage Agent, or whether they are deploying Audience Profile Books). This value is either delivered in the content returned to the individual's device to be sent along with a subsequent request to the Profile Usage Agent for content (e.g., to request an advertisement in a particular location). Or, the content site can send the value directly to a Profile Usage Agent in a request for content to be delivered to the individual.

c) The Profile Usage Agent uses the value to obtain the profile data for the individual. The value can be either a data payload encrypted using the Profile User Agent's encryption key. Or, the value can be an identifier the Profile Usage Agent sends to the Data Aggregator, in a server call, to retrieve the profile data.

d) The Profile Usage Agent determines the content to deliver, updates its internal data as allowed, and delivers the content.

3) Proof of Concept Implementation

The following data flows will be enabled in a Proof of Concept implementation:

a) At least two reference Data Collectors will acquire profile data from users, and, with user permission, normalize send it to a Data Aggregator service to create an authoritative, session-limited file for each each active user.
b) The Data Aggregator service where authorized, will send user data back to a user-owning Data Collector in order to maintain the Data Collectors files as authoritative.
c) A reference Profile Usage Agent will be created to make profile requests to a Data Aggregator, receive profiles in response, and to send event records to the Data Aggregator for periodic settlement.
d) Create one or more end-user services, applications or interfaces, which will permit users to create, view, modify, delete, share or protection a profile or attributes and monitor their use. The may include demographic flags, interest identities and commercial intentions.
See: https://docs.google.com/spreadsheets/d/1i-7tEBGwqa7IUyFoworLEl4xIg1QeK_ryfVELS7NCbE/edit#gid=487804185
Such services function as agents of aData Collector, or may act as a Profile Usage Agent.

-- END OF DOCUMENT --