THE INFORMATION TRUST EXCHANGE

http://www.infotrust.org

**Trust, identity, personalization,
content and user sharing for the news industry**

Functional specification for user-data sharing

(draft v2.1 -- 12-29-15-BD)

This WORKING document proposes functional specifications for exchange of permissioned user data to support customized service of digital content – advertisements, stories or other services. Organizations undertaking one or more proof-of-concept demonstrations or prototyping should refer to the "Request for Proposals: Information Trust Exchange Services" as well as the broader Service Design Specifications. The Consumer Use Case is also instructive. The definition of "Personally Identifiable Information" in this specification is intentionally broad to inspire discussion about what the advertising-industry needs and does not need. An attempt has been made to make terms used consistent with those in the document "Testing a Customer Profile Network" (v0.031 12-29-15) from Clickshare Service Corp.

■ *Bill Densmore*

A. OBJECTIVES

1. Support storage and updating by users of personal attributes and preferences at a most-trusted service provider, such as a publisher or other identity-service provider.

2. Support publisher/agency/advertiser real-time requests for current unique user profile data.

3. Centrally log events for value exchange and third-party verification

4. Avoid use, storage or aggregation of personally identifiable information.

5. Deprecate the use of "cookies"

B. DEFINITIONS

The following terms are used as defined below. Terms in *italics* are taken from and used in the document, "Testing a Customer Profile Network". (citation above).

1. "Advertising Service" – A advertiser, marketer, agency, advertising network or advertising exchange, or any other entity involved in presenting Commercial Messages to Users. If doing so

across the ITE, they must be Members. An Advertising Service may operate as a Profile User Agent.

2. "Attributes" – A set of characteristics stored in a database which are unique to a particular user's account, such as zipcode, service class, subscription start/stop, demographic or preference information.

3. "Commercial Messages" – Advertisements, sponsored content and other information objects created by Advertisers for the purpose of influencing Users, and often presented to Users in collaboration with Publishers.

4. "Cookie" – A persistent HTML cookie storied on an end-user's device.

5. *"Customer Profile Network" – A common platform for the sharing of individual User profile information for the purpose of improving the User experience and managing Privacy.*

6. *"Data Aggregator" – a central point of a Customer Profile Network. An independent entity, such as the Information Trust Exchange, which media sites and other Data Collectors and Profile Usage Agents join.*

7. *"Data Collector" – A media site or other entity that collects information and acts as an identity manager for individual Users.*

8. "Event" – A single, unique completion of an HTTP request for a resource, such as a page, photo, video, advertisement or other content item.

9. "Exchange" – One or more services governed by the Information Trust Governing Association which provides such services as authentication, authorization, transfer of tokens and keys and logging to members of the ITGA. The Exchange functions as a "Data Aggregator."

10. "Key" – An alphaneumeric set of values, typically encrypted, which one Member can provided to another Member for purposes of linking to a unique set of stored Attributes.

11. "ITE Mod" – An ITE-compliant modification module to an Apache or other web server to allow it to perform ITE required actions before serving responses to an HTTP "GET" request.

12. "Log" – A database which contains records of individual accesses to resources and which may include such things as date, time, URL, UserID, Resource Value Class, Markup Ratio or other standard attributes of an Event.

13. "Member" – A business entity which has been admitted into ITGA membership and which may be a publisher, advertiser, agency, author, producer, identity-service provider, technical service provider or other exchange participant or operator.

14. "Personally Identifiable Information" (PII)– Any representation of information that permits the identity of an individual to whom the information applies to be reasonably inferred by either direct or indirect means. Further, PII is defined as information: (i) that directly identifies an individual (e.g., name, address, social security number or other identifying number or code, telephone number, email address, etc.) or (ii) by which an agency intends to identify specific individuals in conjunction with other data elements, i.e., indirect identification. (These data elements may include a combination of gender, race, birth date, geographic indicator, and other descriptors). Additionally, information permitting the physical or online contacting of a specific individual is the same as personally identifiable information. This information can be maintained in either paper, electronic or other media. (source: U.S. Dept. of Labor). Also, information which can be used to distinguish or trace an individual's identity, such as their name, social security

number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name. Also any **information** that can be used to distinguish one person from another and can be used for de-anonymizing anonymous data can be considered **PII**. (Also see, Wikipedia entry).

15. "Presenter" (or "Home Base") – An entity which manages user accounts for the purpose of providing services, which might include subscriptions, resource-access control or custom services tailored to a unique user, such as personalized information, including advertisements. *Also known as a "Data Collector."*

16. *"Profile User Agent" – An advertiser or other entity that wishes to use profile information acquired from a Data Collector, which in turn manages it in behalf of one or more Users.*

17. "Provider" (or Publisher) – A provider of information services, which also provides positions within their web services for the placement of advertisements by third parties for viewing by a user. *Also functions as a Profile Usage Agent.*

18. "Session" – A continuous period of time during which a unique Token is assigned to a User by the Exchange in order to facilitate customization of services by Members. It is anticipated that ITGA rules will forbid cross-Session linking of Tokens and Attributes for the purpose storing a permanent unique identification of an individual, so as to constitute PII.

19. "Token" – An alphaneumeric set of values, which may or may not be encrypted, which is provided by one member to another by a common transfer Protocol. A token may contain a Key.

20. "Transfer Protocols" – One or more set of technical and business rules for the communicating of Tokens, Keys, Log reports, Attributes and other required requests, responses and services among and between Members of the ITGA.

21. "User" – An individual consumer end user

22. Information Trust Governing Association (ITGA) – A non-profit, public-benefit collaboration of Members (see below) organized to facility trust, identity, privacy and information commerce.

C. OPERATING EXAMPLE

It one intention of the ITGA (in formation) to facilitate the delivery of Commercial Messages to Users in order to facilitate and sustain one business model for delivering news and other information in the public interest. The delivery methods should, as a priority, be respectful of User privacy, choice and time. The following example is intended to meet that intention, and priority.

1. A User using HTTP protocol on an open network transmits a URL to a web or mobile service provider (Provider or Presenter) which invokes a GET request to receive a resource in HTML (or variants).

2. If the GET request is to a page which is part of the ITE (content or ads are offered by providers who want to personalize services by obtaining User Attributes), a modification in the web server withholds momentarily serving the page.

3. The server's ITE Mod looks at the submitted URL to see if it has a query-string appended which

includes an ITE Member/User ID.

a. If it does, the server tags all the URLs on the page about to be served with the same ITE Member/User ID and then serves the page.

b. If it does not, the server sends an ITE-unique dialog page back to the User's browser containing a pixel request to authenticate.infotrust.org and perhaps some visible text. The authentication pixel request includes the URL of the overall resource requested by the User.

4. Authenticate.infotrust.org first creates and tags a temporary session identifier to the requested URL. It then checks to see if there is an infotrust "cookie" on the User's device. If there is, it reads the cookie to see if it contains the unique identifier of the User's Presenter service provider (home base), as well as the User's ITE-globally-unique User identifier (these may be combined in one alphanumeric sequence such PublisherUserID1234).

5. If the ID is contained in the cookie, authenticate.infotrust.org next invokes a unique ITE Protocol to contact the User's Presenter publisher to confirm the User's unique ID and obtain one or more Attributes of the User which the user has previously authorized to be shared with the Exchange *(a "Data Aggregator")* . The ID and Attributes are returned to authenticate.infotrust.org

6. If there is no cookie, authenticate.infotrust.org sends a log-in request to the User's browser, with a list of Presenter/Publishers, asks the User to select theirs, then redirects their browser to that Presenter/Publisher for login. The request from authenticate.infotrust.org incudes a temporary session identifier which links the User's request uniquely to the underlying URL to be served by the Provider/Publisher. When the Presenter/Publisher logs the user in, it also uses ITE Protocol to send authenticate.infotrust.org the User's unique ID, a set of Attributes, and also returns the temporary session identifier.

7. The Exchange stores the User's Attributes in a dynamic session database at authenticate.infotrust.org. It then matches the original session identifier with the Provider/Publisher's original URL request. Using ITE Protocol, the Exchange then uses the temporary session identifier to determine which Provider/Publisher to send the User's ID and temporary session identifier back to the Provider/Publishers ITE Mod webserver.

8. Upon receiving the User's ID, and session identifier, the Provider /Publisher's ITE Mod webserver now serves the original URL request, first dynamically appending to all the URLS on it a ITE-complaint query string which includes the User's ID and session identifier.

9. Using ITE Protocol, the Provider/Publisher sends an Log report of the HTTP service event to the Exchange, containing standard HTTP log fields, but including the User's ID, the session identifier, and the Page Value Class, and Markup Ratio, the Rights Owner if the page is being exchanged for value -- and any other elements determined by the ITE to be useful.

10. As the User's browser paints the page resource, it makes calls to each Advertising Service selected by the Publisher to provide Commercial Messages as part of the HTTP resource gotten by the User's browser.

11. Each Advertising Service receives a GET request with the appended ITE-complaint query string which includes the unique UserID for the person logged into the device and (presumably) reading the page as well as the temporary session identifier.

12. The Advertising Service, using ITE Mod on its webserver, next takes the ITE User ID and session identifier and, using ITE Protocol, asks the Exchange for the temporarily stored attributes of the User.

13. The Exchange examines the temporary session identifier, determines that it is valid and current, and returns any Attributes of the User that the User has prospectively authorized for release.

14. The Advertising Service using the temporary session identifier, now matches the User Attributes received from the Exchange with the Advertising position on the original URL, determines which ad to serve from inventory based on the Attributes, and returns that HTML code and images to the User's browser.

15. The Advertising Service, using ITE Protocol, then sends a Log report of the service of the ad to the Exchange, containing standard HTTP log fields, but including the User's ID, the session identifier, any Page Value Class for the advertisement, if the ITE is optionally expected to aggregate and clear a payment from The Advertising Service to either the user's Presenter/Service Provider or to the Provider/Publisher. MemberIDs for each are included in the Log report.

D. PROOF OF CONCEPT DEVELOPMENT

To execute a proof-of concept demonstration of this ITE user-sharing service the following principal efforts are required:

1. Code a modification to Apache and other common web servers which (1) intercepts GET requests and performs actions on them before other services are invoked, (2) performs URL-query-string appends as the last step before a page is served and (3) Communications with a designated ITE Exchange service via ITE Protocol.

    a. One version can work for Providers/Publishers only.

    b. One version cal work for Advertising Services only.

    c. One version of the Provider/Publisher mod code can also have additional functions required of a Presenter (manages user accounts and serves User Attributes) as well as Provider/Publisher.

    d. The simplest versions do not have to support Page Value or Markup Ratio services where personalization, not payment is all that is required.

2. Develop and activate an ITE Exchange service which (a) uses ITE Protocol to communicate with the ITE Modules of Provider, Presenter and Advertising Service web services (b) Temporarily stores User IDs and Attributes, safeguarding and deleting them according to ITE business rules (c) Responds to HTTP requests from ITE Members and (d) Accepts and stores Log reports of ITE network events, for offline processing and aggregating for payments, for analysis by a User's home-base Presenter/Provider or for anonymous, aggregated analytics.