ITEGA
Information Trust Exchange Governing Association

FUNCTIONAL REQUIREMENTS FOR A "MINIMUM VIABLE PRODUCT"

Draft v1.0 bd 06-23-21

This document draws from the ITEGA document:
"Functional Specifications for User Data Sharing"
last updated June 10, 2021

ITEGA would like to support and govern core enabling technology for the sharing of users and content with one ID and one password (or other modern authentication approach) across multiple, independent ITEGA-certified websites (Member Sites), where privacy and identity are respected and managed for the user's benefit.

A. OBJECTIVES

1. Enable use privacy management through storage and updating by users of personal attributes and preferences at a most-trusted service provider (IdSP, agent or fiduciary), such as a publisher or other identity-service provider.

2. Avoid by technology, policy and audit the use, storage or aggregation of personally identifiable information, other than -- with consent -- by the user's home-base Identity Service Provider/agent/fiduciary, subject to GDPR / CPRA.

3. Support publisher/agency/advertiser real-time requests for current anonymized user profile data sourced from the user's IdSP where in compliance with GDPR/CPRA and ITEGA rules.

4. Enable authorizations, access-control, attribute sharing and network subscriptions through systematic logging of anonymous records of events where value exchange or third-party verification are required.

5. Deprecate to maximum feasibility the use of "cookies"

B. REQUIREMENTS

Build and support rapid deployment of a Proof-of-Concept prototype of one service operating within an ITEGA ecosystem. The prototype should be deployed in a matter of weeks, not many months, and can be tested by participating media organizations. Key elements of demo(s) would include at least items a through c, and ideally all, of the following.

1. SERVICES TO BE SUPPORTED OR SPECIFIED

Required:

a. FEDERATED LOGIN -- Exchange of user authorizations and recording of access events among a plurality of at least three independent-domain web services. Allow access to resources on multiple services not otherwise connected (except by the federated service) using credentials issued by a single service provider among a plurality of service providers within the federated service.

b. CONTENT ACCESS LOGGING -- Provide a shared logging function demonstrating the sharing of content among otherwise independent publishers and their users. Records access to individual resources on widely distributed hosts within a federated network, including records of the value of the resource accessed, and a unique, anonymized identity of who accessed them, such that individual records may be sorted by time, value, source or user for periodic settlement to financial, advertising or other batch networks.

c. ATTRIBUTE SHARING -- Show capability for applying open technical protocols for metadata exchange among IdSP's for *permissioned* user identity, content sharing/sale and user-opted-in advertising personalization. Stored at the user's chosen IdSP are optional, permissioned demographic values, information preferences and subscription or access-right authorities; such that the attributes may be readily viewed, accepted, rejected or possibly changed by the user.

Specified or deployed:

d. COHORT-BASED AD DELIVERY -- The user-data-sharing service enables and respects opt-in user choices regarding use of individual attributes, and will support a method to personalize advertising delivery based upon anonymous interest cohorts without identifying individuals.

e. CONTENT PERSONALIZATION -- Provide at least a concept demonstration of users supplied a personalized stream of news from hundreds of sources, where the sources may be compensated.

f. VARIABLE PRICING -- Specify enabling of a variable pricing service to real-time query a user during the process of requesting a resource across the network, allowing the user, or his or her agent, to respond to a price offer in real time based on such factors as subscription authorization, pre-established pricing preferences, variable use or credit availability.

g. VARIABLE SERVICES -- A profile-exchange service which enables access to and network sharing of user attributes for the purpose of determining types of services and their value to be provided to a user; and which is capable of varying services based upon such parameters as subscription-authorization levels and credit thresholds.

2. DATA-EXCHANGE STANDARDS REFERENCED

a. Support ITEGA-certified taxonomy for shareable user attributes

b. Support ITEGA-certified taxonomy for identifying, categorizing and tagging content

c. Support ITEGA-certified taxonomy for categorizing and routing advertising

3. COMPATIBILITY REQUIREMENT

Testing and network rules as they are established must also allow participating affiliates or publishers to continue to operate within their existing silo user and access-control services. The ITEGA protocols should be additive to these businesses.

4. INFORMATION SHARING REQUIREMENTS

Respecting user preferences relating to the collection, use, and sharing of personal data is a major focus of ITEGA. The System envisioned for this prototype must support identity-sharing (SSO or otherwise) functionality with varying levels of information sharing among the Member Sites.

It is expected that when a user authenticates for access on a Member Site using credentials from a different Home Site the System will provide the Member Site a network identifier for the user, which might be a "hash" of the user's email address.When operating using the highest privacy settings, this identifier must not allow aggregation of data among Member Sites, nor aggregation of data relating to this user across multiple sessions.

However, when permitted by the user, the identifiers should permit aggregation of de-identified information posted to a UDEX (building the UDEX is not within this prototype scope) and permit the retrieval of data from the user's Home Site.

The System must also support exchanging other types of data between Member Sites in accordance with user preferences and in a manner that can be audited by ITEGA, Members, and users. Among information categories Respondents should consider are:

- Non-identifying permission data, such as which categories of information can be shared, privacy defaults, and similar data.
- Non-identifying access-control data, such as the user's subscriber status at their Home Site, using a set of identifiers agreed upon by the members in the network. Some examples include Complimentary access, Single Day paid access, Basic paid access, Premium paid access, and possibly some number of Network-specific access levels.
- Non-identifying interest and demographic data as defined by the network members, such as general geographic location, general age category (e.g., child or adult), and general interests.
- Internally identifying information, such as a network-user identifier.
- Identifying and protected information possibly including name, gender, race, and email address, if explicitly authorized by an individual user.

The System must allow users, through their Home Site, to explicitly authorize sharing of each information category. The System must ensure that only information explicitly authorized is shared. The System may not use any information other than as required to operate the System. Member Sites will be limited in their use of shared information by the terms of a Member Agreement (see a draft). For example, if the user permits the sharing of access-control data, the receiving Member Site can use this information to grant access to appropriate categories of content. If the user permits the sharing of demographic data, the receiving Member Site can use this data to personalize the content shown to the customer. Member Sites will not be permitted to allow third parties to use any shared information, except as allowed by an ITEGA or

sub-network Membership Agreement, and in compliance with GDPR, CPRA and other application laws.

5. TRANSACTION LOGGING SERVICES

A Transaction Logging System should allow Member Sites to log user activities that will be shared with the user's Home Site. ITEGA anticipates that later systems will access the Logging System data to support functions such as settling purchases by users among Home and Remote Member Sites, distributing membership fees among Remote Member Sites based on usage patterns, or managing billing and receipts for advertising views

6. ADDITIONAL SYSTEM REQUIREMENTS

The System ITEGA seeks must support these additional requirements.

a. No participating Member should be required to upload its user account data to the network servers in order to participate. We envision that each participating Member will be responsible for authenticating their users and managing their users' account data. When necessary, a participating Member will use a System-provided or authorized API to indicate that a user has been authenticated, and will respond to data requests according to its understanding of their own user's privacy settings. In particular, authentication credentials (e.g., usernames and passwords) must never be shared with the System's servers. We anticipate that participating Members will use a variety of authentication technologies including some based on blockchains, such as Self Sovereign Identity.

b. The System must provide a robust reporting protocol that will allow ITEGA to gather the data necessary to perform experiments that compare different data-sharing policies and different remote-content access policies. In addition, each participating Member must be able to generate or receive periodic reports that show their users' activity levels and compare these to network-wide activity levels.

c. ITEGA must be able to configure groups of participating Members into separate, independent networks. All data managed by the System for one network must be kept separate from data from any other networks. The proposed System can limit a participating Member to joining a single network. However, Respondents can submit Systems that allow a participating Member to join multiple networks, or allow networks to link together, provided that this does not allow data from one network to "leak" to other networks, except as permitted by the participating Members of each network.

7. PROJECT DELIVERABLES

The proof-of-concept project should deliver the following components:

- An API specification that allows Network Members to integrate with the network and a reference implementation that can be used to set up test Network Members.

- An API specification that allows Network Members to log activities to be shared among Remote Member Sites and the Home Member Sites.

- Technology that implements the server-side of the APIs to provide the necessary functionality. This can be either technology that ITEGA operates or technology that the entities operate as an agent of ITEGA.

- Evidence of a mechanism for users to “opt-in” or “opt-out” of data tracking consistent with GDPR and/or CCPA and/or web-browser consent dialogues.

- Auto-generated reports showing discrete transactions by individual end-users who’s PII can only be known to their “home-base” IdSP.

- Specifications for how value and accounts among participating IdSPs will be logged, processed and settled.

8. OTHER REFERENCE MATERIALS

The following research and working documents provide background on the ITEGA ecosystem frameworks. These materials are advisory and not definitive, but are provided in an effort to clarify or invite questions.

- Summary of proposed requirements and objectives for ITEGA federated SSO
- Single Sign On Network Description
- Functional specifications for user-data sharing (includes operating example)
- Proof of Concept Prototype Element: ADVERTISING SERVICE
- Proof of Concept Rationale Narrative and user experience A-and-A
- Product Requirements Document (Draft)

APPENDIX A

SUPPLEMENTAL SERVICE/DESIGN ATTRIBUTES PROPOSED FOR REFERENCE

Excerpted from:

- Information Trust Exchange Framework: Service Features and Design Specifications

A-1. User identity and profile attributes

The ITEA facilities the transfer of the following identifiers for each request made by a user for resources across the network:

Network-level attributes (accompany all requests)

1. UserID – A globally unique attribute which includes the user's home-base host ID. This is the minimum attribute necessary to log access records for payment or credit and is analogous to a credit-card number.

2. One or more customer-group codes to identify user assignment to specific groups for publisher- or service-provider proprietary purposes.

3. A service-class to define eligibility of the user for specific levels of pricing, content or services

4. The content server ID of the publisher supplying content and optionally requesting a royalty payment ("PubMbrID")

Preference-level attributes (accompany and constraint all requests)

5. Other flags regarding preferences for: (a) privacy (b) parental control (c) advertising viewing preference (d) do-not-track

Identity attributes (optionally shared with request)

6. Identity attributes available for sharing (or not) depending upon privacy preference (above), include user-supplied nickname, email, fullname, date of birth, genderl, postal code, country, language and timezone

Business attributes (optionally supplied with end-user permission)

7. A vending publisher may request other business attributes of the person and the person's home base may or may not supply the attributes based upon the user's expressed privacy preferences. The attributes may be stored and supplied in formats developed by Schema.org (http://schema.org/Person)

EduPerson attributes (optionally supplied with end-user permission)

8. A vending publisher may request other Internet2 "eduPerson" attributes of the person and the person's home base may or may not supply the attributes based upon the user's

expressed privacy preferences. The attributes may be stored and supplied in formats developed by Internet2:

http://www.internet2.edu/media/medialibrary/2013/09/04/internet2-mace-dir-eduperson-201203.html

Interest identities and topics

9. A vending publisher/marketer may request from the user's home-base service provider attributes related to any topical "interests" and "identities" stored in the form of key words or phrases depending upon the user's privacy preference.

A-2. Digital content tags for pricing or royalty management

The Proof-of-Concept should support a schema for vending publishers to XML-tag royalty- or price-identified content which will be recognized and respected by user service providers, and logged as necessary for financial settlement. **Thus content can reside anywhere on the network and the rights owner will be paid for use.** Among basic content attributes are:

1. The creation date/time in YYYYMMDDHHMMSS format.
2. An expiration date supplied by the original content producer in the same format.
3. The PubMbrID of the creator or publisher entitled to royalty or payment.
4. A optional Digital Object Identifier (http://doi.org)

C. Tracking/settlement of value exchange

Finally, the Proof of Concept should support a schema enabling the negotiation and computation of value exchange. The table invoked will depend upon whether the resource is chargeable content, or sponsored content (including advertising).

5. A variable table of royalty payments (or a key to a master royalty-payment schedule) used to compute the charge to the user's service provider upon the digital vending of the resource depending upon use, service class and other custom factors.

6. A variable table of credits paid to user's service provider upon the end user's viewing of a digital resource, depending on level of use or interaction.

7. A retail "Markup Ratio" in use by the User Service Provider which is provided to the content-serving publisher in real-time so that if the end-user is to be shown the object's price before purchase, the price show will be "retail" not "wholesale." *(See Appendix B)*

APPENDIX B

PRICING – WHOLESALE-RETAIL

The description below is from the 2015 white paper, "From Personal to Payment," of the Donald W. Reynolds Journalism Institute. It describes a solution for web content pricing that was patented by Clickshare Service Corp. The patent reached term and expired in Oct., 2020, so the approach is no longer proprietary. It is more fully described in the public patent document.

A frequent question posted by interviewees involves pricing. If news organizations are going to share users, and share content, who is going to be in control of pricing? *(See Exhibit O)* The answer: No one person or entity. Some examples:

- Airlines benefit from a common air-traffic control system and they share airports. They fly similar aircraft made by the same companies. But they compete on pricing, many routes, and most aspects of service.

- Thousands of companies float their stock on major exchanges. The price of their stock is subject to near absolute competition for investors' dollars. Yet they also use common bidding, trading and settlement systems.

- Online advertising exchanges work in milliseconds with demand-side and sell-side platforms to match willing advertisers with willing publishers and aggregators to deliver "impressions" to interested consumers. Prices range dramatically, as do the content and form of the advertisements.

But what if you added to the mix the idea of wholesale-retail pricing, just like in the real world? If General Electric Co. makes a toaster oven and sells it to Wal-Mart Stores Inc., Wal-Mart then decides how to price the toaster. Think of the Internet market for information as like a Wal-Mart store. The retailer – your preferred publisher or service provider – is responsible for billing you and paying for what you buy from his or her store. Then, they go pay the originating publisher – the wholesaler – for the items you purchased -- to make up your personalized information bundle. And imagine, as with the advertising exchanges, that this happens instantly. The originating publisher, if it knows something about you, might vary the offer (price and terms). Your home-based publisher, the retailer, might chose to give you some of the items as part of a package, and ask you to pay for other pieces a la carte. Unlike Wal-Mart, the inventory of a digital information retail store doesn't need to be shipped or stored in bricks-and-mortar fashion. It can be sought, priced, sold and consumed in milliseconds.

All that's needed to make such a system work is a standardized method – a set of protocols – for exchanging information about users and logging -- to a common place -- the cost of what is purchased. A useful feature might be the ability to aggregate many small purchases that are charged periodically – making efficient use of financial-transaction networks like the bank Automated Clearing House (ACH) networks and avoiding relatively steeper credit-card interchange fees.

Imagine this scenario: *The New York Times* might send you an email and say for an extra $1 a month, you get 10-15 clicks per month from a set of French language publications. It's just $1 a month and you'll have that Francophile bonus. What would happen when you click to an article at *Le Figaro*? They would have some price they had set on that article – maybe it is five cents (converted from Euros). When you click on that article as a *New York Times* user, the exchange should record a payment to *Le Figaro* of five cents and record a charge to *The New York Times* of five cents. But whether you as a consumer ever pay anything other than that extra $1 -- ought to be up to *The New York Times*.

If you have a system where the parties on a business-to-business basis agree to pay the cost of people

~~Apple is not going to play in a~~ new ITE ecosystem if that ecosystem requires the company to shut down the

surfing within the system, then all it becomes is a strategic business exercise how much *The New York Times* should charge you per month. *The Times* might do this for awhile and find they are losing money by just charging you $1 a month, so they might come back to you and raise the package to $2 a month. Or maybe it has a cap on it of 30 clicks per month -- then you have to pay more.

One can't presume to guess how all those things will work out. What we need to create is a system that enables all of that and then allows the free market to operate as it does so well –- which is to have pricing and packages find their equilibrium. What is described is a free market for digital information – a economic libertarian's delight! But don't we need to start by enabling those kinds of capabilities?