rji DONALD W.
REYNOLDS JOURNALISM INSTITUTE

Missouri School of Journalism 🔲 University of Missouri

Infotrust
eXchange

THE INFORMATION TRUST EXCHANGE

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

REQUEST FOR PROPOSALS:
Information Trust Exchange
Proof-of-concept demonstrations

V2.0 DRAFT Bill Densmore Feb. 4, 2016

The Donald W. Reynolds Journalism Institute is inviting individuals and organizations, independently or collaboratively, to submit proposals for the of one or more proof-of-concept demonstrations or operating prototypes of services that would operate within an Information Trust Exchange framework.

Proponents are directed to the following documents for background of the framework:

- Functional specifications for user-data sharing (includes operating example)
- Information Trust Exchange Famework: Service Features and Design Specifications
- Information Trust Exchange Governing Association: Business Goals, Role and Structure

PURPOSE

For the benefit of the news industry, RJI is working independently and is seeking academic and business collaborators to create a open, non-profit, consortium for establishing and governing a trust, identity, privacy and information commerce exchange on the Internet.

RJI convened five task-group meetings from May through October, 2015, in five cities across the United States to share and edit draft proposals for the creation and operation of such an exchange. It is time to demonstrate how aspects could operate.

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. To improve the user experience, deprecate the use of third-party "cookies"and browser-based tracking code.

A Sept. 16 task-force meeting in Columbia, Mo., adopted the following:

PROOF-OF-CONCEPT MISSION / OBJECTIVES

Create a new business model for journalism through a network that helps citizens and news organizations collaborate to find, share and sell essential information that is highly personalized, relevant and essential to strengthening communities and participatory democracy.

A prototype version of the Information Trust Exchange network will be designed so that news organizations can:

- Assist public users to manage their privacy and identity "persona"; deepening user relationships
- Foster effective advertising through efficient, privacy-respecting, interest-based customization
- Enable subscription networks
- Grow audiences
- Increase revenue (monetize off-site content, higher CPMs from non-subscribers)

It will be designed so that public users:

- Efficiently find helpful and relevant information personalized to their interests/needs
- Find such relevant information faster and easier
- Have new and better control over their data and identity
- Increase their connection with geographic and topical communities
- Find the service valuable enough to pay something by subscription or per-click

In the process, news organizations will:

- Learn what it takes to aggregate content automatically and efficiently
- Collect and share user data/behavior on a "permissioned" basis
- Receive data about their users who leave their site or service
- Receive anonymous data about other's users who come to their site/service from elsewhere

The demonstration will:

- Track user data and collect analytics (not use cookies)
- Test networked subscriptions
- Personalize content
- Evaluate and analyze results

An additional goal is to demonstrate that core enabling technology exists and works for the sharing of users and content with one-ID and one password across multiple, independent websites, where privacy and identity are respected and managed for the user's benefit.

SCOPE

The Project's overall scope is to create a version of multisite registration, authentication and event-logging service which will be sufficient to prove the concept of a web/mobile/tablet user being able to access and purchase resources from multiple independent websites ("presenters" or "relying parties").

The service should use a single ID/password ("credential") issued and maintained by a home-base publisher, typically a newspaper (the "identity provider") -- but potentially any trusted service provider. Multiple such events should be shown to be able to be aggregated and settled to a single user account on a periodic basis.

The scope will include analysis of the user experience with the services, and outreach about the experiment among editors, the news industry and the public, including analysis of how presenters and publishers might share and make use of so-called "first-party" data about user activity.

TEST ELEMENTS

Test elements may include:

1. A user-management service which creates and stores on a distributed basis attributes of individual users, including 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.

2. 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; and which can "anonymize" profiles shared with third parties such that no personal identifying information is used or created.

3. A federated authentication service that allows 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.

4. An access-logging service which 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 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.

5. A variable pricing service which enables the querying of 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 or credit availability.

TEST OUTPUTS

Test or demonstration outputs might include:

1) **Confirm the capabilities of individual pieces of an operating system**.
   To illustrate with two examples:

   a. Are there single sign-on systems that provide access to content beyond the home base?
   b. Are there payment settlement processes that distribute payments for services/content to individual publishers?

2) **Create or adapt capabilities that do not exist in forms useful to the ITE**.
   We would identify these after the step #1 inventory. Examples:

   a. A content logging system
   b. Processes for gaining access to content beyond the home base

3) **Identify what participating publishers would have to do to participate in the test**.
   Examples:
   a. Content tagging protocols
   b. Content pricing protocols
   c. Marketing plan for the test among a publisher's clients

PUBLIC BENCHMARKS

Any proposed public testing of a candidate service should meeting or exceed the following:

- At least four different content web sites participate.
- At least three different service/identity providers participate.
- Evidence of a mechanism for users to "opt-in" or "opt-in" of data tracking
- Evidence of use-cases for sharing "first-party" user data among system actors
- Auto-generated reports showing discrete transactions by individual clients.
- Commercial connections to banking and credit-card networks.
- The ability to settle value and accounts among participants as needed.

CONTENT / CLIENTS

Proponents should develop relationships with one or more partners who can supply users and/or content.

TIME TABLE

Demonstrations or prototype trials should be designed to being producing meaningful, observable results not later than May 31, 2016.

DESIRED OUTCOME

At the end of the trial, RJI and its evolving collaborators expect to be able to reasonably assess the contours of the opportunity for a new user experience for news. One that allows convenient access to trustworthy information from multiple sources from a single home-base relationship and account.

Creation of the protocols, and opt-in business rules to support this experience might then be carried forward by some sort of news/information industry collaboration. This could be an important privacy and choice enhancement over what's available on the web today:

a. Services such as PayPal or Facebook Connect or iTunes must register and have credit/personal information for all users, as opposed to enabling a plurality of account registrars.

b. Each such system is "siloed" – a user of one can't access resources at another.

c. The consumer thus must chose one (or many) incompatible identity/commerce services rather than having an array of choices of competing service providers who all operate with common technical protocols and rules.

d. Aggregation and settlement of charges is not now possible across multiple content services. This means it is impossible to create a one-price packaged service of content from multiple content sources (the value that the print newspaper used to provide elegantly).

At its the Sept. 30 meeting in Portland, Ore., the ITE task group on content description, tagging, sharing and selling review a menu of 14 desired features for proof-of-concept testing, reproduced below. After ach statement, the group's consensus reply is shown in *italics* and <u>underlined.</u>

POTENTIAL REQUIREMENTS

1) User-created and updated profiles of preferences, interests and demographics
   Include ability to create sub accounts or sub profiles within a household on one payment account with individual user accounts.

2) Ability to match dynamically-specified buyer interests with colleagues, family or friends.
   Yes

3) Ability to selectively share your interests with colleagues, family or friends.
   Not important to the prototype, later in the product.

4) Discovery service focused on quality, trusted content, uniformly tagged/identified.
   Critical to include in proof-of-concept.

5) Allows content owners (publishers) to track and control access to their work.
   Yes, need to know where it is showing up. Need some control in start, these are the rules. Usage rules for content.

6) Allows each content owner to price their own content for varied uses.
   You need it but you could build it as an analytics engine, don't do anything about pricing yet. I think you need it in prototype. Content owners will make a distinction. Engineer into design but not needed in trial. Wave hands.

7) Allows sale of content on a per-click or subscription basis.
   There was not consensus here. See discussion below.

8) Allows user to be rewarded for viewing sponsored content.
   Not a function of the ITE -- not mandatory feature

9) Does not pre-empt or interfere with each publisher's own "silo" payment strategy.
   A critical given.

10) Allows content access to be variable depending upon user attributes.
   Yes

11) Allows publisher to apply their own price to digital content.
   Yes

12) Allows publisher to vary price depending up use or user or time.
   Yes

13) Allows publisher to allow access based upon variable subscription rights.
   Yes

14) Exchange does not play any role in setting pricing or commercial service offerings, just transferring data about them. (i.e., "managing the marketplace.")
   Agreed.

WHAT NEEDS TO BE IN A PROTOTYPE?

Portland participants felt that demonstrations and prototyping should seek out sources of unique, original, trustworthy, authentic content that is easily discovered or packaged on the "free" web and which can be filtered based on a user's interest. Adjectives used to describe such content included unique, original, trustworthy and authentic and user customizable. It should be adaptive to expressed and inferred user interests.

- What do we need to pull in for a protytpe? A big bucket of content and filter it based on expressed user interest.

- In the prototype put in the option to create your own topic.

- The proof of concept has to show there is a mechanism to show that it can learn and improve its effectiveness. Users are gauging content, does it adapt?

- One of the requirements is a way for a user to characterize themselves for the system.   A simple elegant way for somebody to quickly inform the system. Existing systems do that through what are the word clouds of the things you tweet.

- As that profile is being built, you go to that personalized page and there are all these pages checked dynamically.

THOUGHTS ABOUT PRICING

At Portland, some discussion centered on methods to enable direct user payments for content (subscriptions, per-click, royalty pools, etc.) There was no consensus on what should be demonstrated in a pilot. On participant argued direct payments "would not scale" and an alternative would be a shared royalty pool based on usage.

Another participant said this would not work because publishers would not agree on the relative value of their works. A third participant suggested focusing in prototyping not on a specific business model but on the user value proposition. A fourth participant asserted: "You need a pay environment otherwise it is not a test of anything. This is a test of a system of high enough value that people will pay."

> The Portland meeting highlighted the need for more work to develop consensus

The Portland meeting highlighted the need for more work to develop consensus on how the ITE-sanctioned networks would facilitate pricing of content. Prototype efforts will need to demonstrate flexibility to enable a variety of approaches to pricing – subscription, per click, free, priced by the presenter or by the producer.

There was significant consensus on other attributes. We reviewed 14 potential requirements for a proof-of-concept demonstration / pilot. After each statement, the group's consensus reply is in *italics and underlined.*

QUESTIONS AND ANSWERS

To faciliate an understanding of the user experience sought, the proponent is advised to read the following question-and-answer and also review this post on the RJI website:

http://rjionline.org/privacypersonalizationpayment/part02

This Q-and-A document describes the user experience to be operated as a proof-of-concept for networked purchasing of topical and local news content in a shared-user network.

After the narrative Q-and-A there follows an Appendix A which provides a chart and more technical description of what is going on.

What will be demonstrated?

Working under a contract with RJI, technical partners and pilot publishers will deploy a service which allows a user from any one of several of its existing customer websites (each with their own subscription databases), to access "protected content" on web, tablet or mobile devices of other customers' websites.

Why does this matter?

Information is increasingly being consumed in bits and pieces from multiple independent sources by users. This happens because aggregators and their friends refer them to it that way. It also happens because users tend to have a habit of "browsing" the web to discover new experiences and information. The more traditional usage pattern – going to a single home base (like a newspaper front page) and relying upon that sole starting point and only its proprietary, packaged content – is only one part of the common use case.

As a result, publishers who want to maintain a deep relationship with users must learn how to help to find information the need anywhere on the web – information without walls. RJI wants to test the hypothesis that a network of websites can be affiliated only by a thin protocol of common authentication and tracking of their shared users – and present a compelling, trustworthy user experience.

In this phase, it is unlikely that the number of participating publishers will reach a critical mass sufficient to warrant users paying a premium amount for the network services. The objective is to merely test the technology of authentication, logging and settlement, and the user interface to it.

How will public users find out about this test?

All of the users involved will be registered users at their "home-base" publisher. When they "log-in" at their home base they will be directed to a version of the LifeStream news personalization service – a customer home page just for them.

What are some of the offers?

- ■ Limited day passes to digital editions
- ■ Day passes to websites or counted access to the website articles
- ■ Special resources not available to the general public
- ■ PENDING – A test opportunity to earn loyalty points

Is there money involved?

Not for the consumer at this stage. The system enables the tracking of access to resources at participating websites for individual charges, or as part of subscription bundles. We will demonstrate the reporting of access and charge records among and between content providers. However in this test, *no consumer end users will be charged anything other than what they might already be paying at their host "home base" publisher website.*

What will users be told about payment?

Users will be told by their "home based" publisher that they are receiving free access to special, premium content not generally available on the open web. A part of the RJI research will be to

survey these users for their attitudes about the use *and value* of the resources they are enjoying access to.

One way to think of what is enabled is an "easy pass" to content at multiple websites. In this proof-of-concept experiment all network resources will be free to consumers as an added benefit offered by their home-base service.

What happens when the customer arrives on a remote site?

The system recognizes that this customer is a network customer and puts up an information page promoting the free access to other-wise restricted resources. Depending how users react in the test, the system may direct users to internal resource links on participating websites without any intermediate explanatory page. This ability to access resources on remote websites, with network login credentials issued by only by any one of participating home-base sites, (rather than a central registration service) is unique to the Clickshare environment.

What happens if a customer arrives first at a remote site?

At the point where they hit the remote site's login page, they will see a Clickshare network logo and be invited by that site to login with their home-base login. For details of how this will work, see the Appendix.

Why should publishers want to participate in this RJI research?

For the benefit of the news industry, RJI would like to develop a "proof of concept" for a networked system of news/information sales, by item or subscription, that is convenient for clients and allows news organizations to maintain and grow their account and trust relationship with clients. For more than 15 years, commentators have called for creation of a one-account, one-ID, one-bill access service to information resources. AOL once offered this inside a proprietary "silo." Sites like Facebook and Google+ are now doing so within web-facing silos.

The idea of the RJI-ITE demonstration is to show a hybrid "open silos" service where a plurarily of publishers and information service providers can maintain preferred relationships with their subscribers but allow those subscribers to access premium resources of other services.

The point of the ITE technology specification is to describe and test the efficient authentication of users across multiple sites and the aggregation of reports about their access to resources.

So why is it necessary to "track" access and what do you mean by that?

Compare this approach to the wired and wireless phone networks. Calls are made – and tracked – across multiple carriers. These carriers use the call-detail reports in background to settle their financial exchange, to cover the cost of their network services. So that's one purpose – to allow an efficient clearing house for content access and purchasing records.

But there's another service possible – optional, opt-in sharing of demographic and preference information among users and sites. For example, a user's age or geo-location or content preferences could be passed – with the user's permission – across the network in a uniform way. This could allow for:

- More customized editorial services
- Advertising targeted to the intentions and location of the user

- The ability to make offers – and even pay consumers for their attention

Why is RJI doing this?

It is in keeping with RJI's mission to test with a group of cooperating publishers a potential way to sustain valuable journalism-related service. In addition this trial will allow RJI to:

(a) Have a platform for early stage academic research on the willingness of individuals to purchase news content when it is convenient and economical to do so; and,

(b) Facilitate the creation of the non-profit Information Trust Exchange by showing that suitable core enabling technology exists and works.

What do you hope to accomplish?

At the end of the trial, we expect to have demonstrated the contours of the opportunity for a new user experience for news. One that allows convenient, personalized access to trustworthy information from multiple sources from a single home-base relationship and account. A service that allows the user's "home-base" publisher to begin a one-to-one relationship which can optionally protect user privacy while offering access to valuable content and opt-in commercial offers.

<div align="right"><u>**APPENDIX A**</u></div>

A Proposed ITE Network Protocol

This document describes in semi-technical terms the flow of user requests and data among the service points within a network proposed for the Information Trust Exchange pilot. The key points explored in this technical description:

1. The place where a customer/user has their "home base" – where their demographic and preference information is stored or managed – is independent of where they go to access information. There can be multiple "home bases" within the network (rather than a single one such as only Facebook or only Google).

2. In order to reach resources at independent content services within the network, the customer/user has to be "authenticated" (or "logged in.") This log-in process, from the user's point of view, occurs at whatever site within the network they visit to access information. But in reality the login prompts are linked only to their "home base" service, so only their "home base" knows who they are with any detail. This is an important privacy and choice enhancement over what's available on the web today:

    a. Services such as PayPal or Facebook Connect or iTunes must register and have credit/personal information for all users.
    b. Each such system is "siloed" – a user of one can't access resources at another.
    c. The consumer thus must chose one (or many) incompatible identity/commerce services rather than having an array of choices of competing service providers who all operate with common technical protocols and rules.
    d. Aggregation and settlement of charges is not now possible across multiple content services. This means it is impossible to create a one-price packaged service of content from multiple content sources (the value that the print newspaper used to provide elegantly).

3. It is a goal of the ITE pilot to demonstrate that these limitations of the current Internet can be overcome.

4. The description uses as an example two content services – *Automotive News* and Crain's Detroit Business – in common ownership but on different service platforms. In practice, they interaction could be among information services of completely indepent providers – such as the Concord, N.H., Monitor and Crain's, or the Rutland Herald and Nashua Telegraph.

The next page begins with a diagram and description, followed by the process flow summary.

Key observation

It is important to note that the content servers have no knowledge of the authentication occurring behind the scenes as part of the ITE network. Each content server defers authentication to its associated registrar. The registrar either authenticates the customer locally, or uses the ITE network to handle remote authentication. This model avoids any requirement to add software to the myriad content servers in use by each publication.