OIX OPEN IDENTITY EXCHANGE

White Paper

OIX Shared Signals Proof of Concept

The findings of an Alpha Project

"Sharing signals is as old as civilization. What is new and important is deploying signaling systems that demonstrably improve the security of existing infrastructure.

Internet Identity systems are most vulnerable when verification is spread among competing stakeholders. Sharing signals address that weakness and enables more shared security measures that increase resilience." Don Thibeau, OIX Chairman and President

By: Andi Hindle (Hindle Consulting)
Emma Lindley (Innovate Identity)
May 2016

Participants

Observers

The Shared Signals Alpha Project - Objectives, Participants and Process

In 2013, OIX published a White Paper (Nash, 2013) which identified the challenge of systemic and contagious fraud risk across the digital ecosystem caused by compromised identities; and proposed an approach to addressing this problem by sharing authoritative, privacy-protecting alerts, or 'Signals' between service providers to help identify risks and so apply earlier protection or other remediation.

A subsequent OIX UK discovery project established the potential to use such a shared signals mechanism to increase shared trust in the Identity Ecosystem (Nash, 2014). This was followed by a specific OIX UK discovery project with the GOV.UK Verify high-assurance Identity Providers ("IdPs") which identified a clear set of potential use-cases for such a system; and recommended progression to an OIX UK Alpha project to test these hypotheses in a more practical application. (Walton, 2015).

Background and context

Table of Contents

Shared Signals Background

The Shared Signals Alpha Project - Objectives, Participants and Process

Use-Cases Details

Test Details

Outcomes and Observations

Discussion and Further Questions

Future Phases

Conclusions & Next Steps

An earlier OIX Discovery Project explored whether:

signals (which share a minimum of personal data) could be shared between IdPs to better prevent fraud.

The hypothesis proved to have merit. Key signals that IdPs would find useful were identified, and a set of principles which would support IdP trust in the signals shared was developed. The recommendation was to scope an Alpha project to test the sharing of signals between some IdPs via a 'Signal Manager.'

The discovery project provided a number of agreed design and operational principles including:

- Only signals of value to IdPs should be shared between IdPs.
- Those signals need to adhere to 'quality' standards in terms of content, use and longevity
- Signals should support transaction monitoring
- Signal sharing should take advantage of open standards where they exist.

- Privacy protection is key in every instance of signal sharing
- New IdP entrants should have access to the shared signals environment as early in the on-boarding process as possible so they can manage risk

The OIX Alpha Proposal Document outlined a project to test whether:

> ***Signals (which adhere to the principle of minimal personal data sharing) can be shared between IdPs via a 'signal manager,' that IdPs have sufficient trust in, and are of sufficient quality to match and take action on.***

In discussion and agreement with the key project stakeholders (IdPs via the IDSG and OIX UK) it was agreed to take a phased approach to this alpha project in order to support a rapid timeframe for initial testing with constrained scope; to report back quickly with initial results; and to facilitate possible additional phases of testing as desirable.

This project introduced two distinct categories of project members:

- **Participants** would be active members of the initial project phase, undertaking technical work to integrate and test signal sharing in a practical manner
- **Observers** would receive regular updates on the project and be invited to provide additional input and critical feedback to the project and to the white paper; and could opt to join future project phases either in an ongoing observer capacity; or as participants.

Table 1 below provides details of the participants (and their specific roles in the project) and observers for the initial phase of work.

Participants	DigIdentity	IdP
	Post Office	IdP
	Confyrm	Signal Manager
Observers	Barclays, Experian, GBGroup, GDS, Telesign, Verizon	

Based on the process common to previous OIX Alpha projects, regular meetings were organised with all the project members to provide updates and to seek input on specific topics. These were supplemented with additional and more frequent work-stream meetings specifically with the Participants in order to help progress the technical integration and testing work.

A number of general signal definition/content principles derived during the Discovery Project phase were incorporated into the Alpha project, including:

- Sharing a minimum amount of personal data to optimise privacy and security.
- Sharing sufficient data in order for a signal recipient to confirm a match.
- IdPs are able practicably to publish signals to and consume signals from a Signal Manager
- Published events must contain sufficient information for a Signal recipient to take appropriate action
- The value of a signal in the IdP-to-IdP context will need to include enough data for the IdP to be confident to make a decision to take action.
- Signal content should include minimum 'hashed' data, and content will be different depending on the particular signal (where personal information is shared, the service should seek to minimise sharing and use privacy-protecting mechanisms wherever possible).

Consistent with OIX Alpha project guidelines, it was also agreed that no real user data would be used during the testing phase.

In initial technical meetings, and informed by the outcome of the earlier Discovery Project, the Participants agreed specifically to test two signals, in support of two distinct use-cases. Each of these use-cases identifies a specific type of cross-IDP fraud risk, sharing the following characteristics:

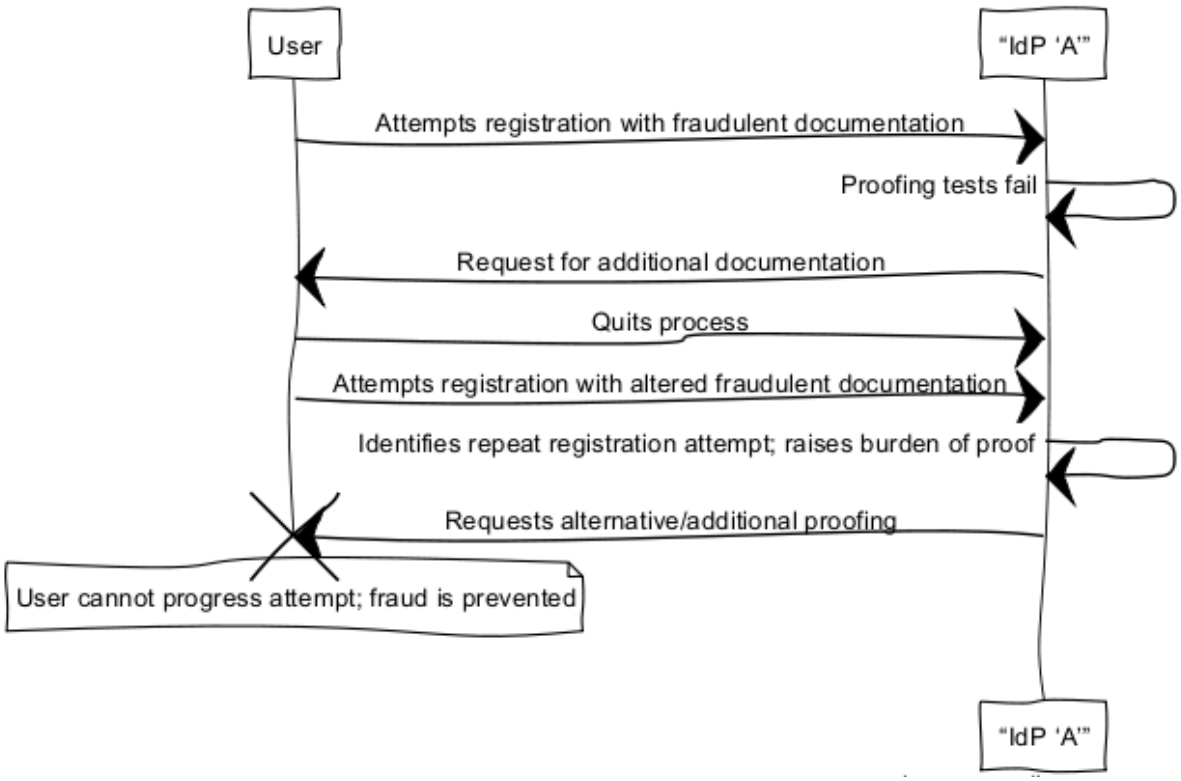
- Seen actively in the wild
- Relevant signals can be practicably triggered by an IDP based on operational events or operator intervention
- Potential for significant impact to systematically affect the integrity of the high-assurance process; and to the individual user affected
- With no clear mitigation possible **without** the sharing of information between IdPs.

Use-Cases Details

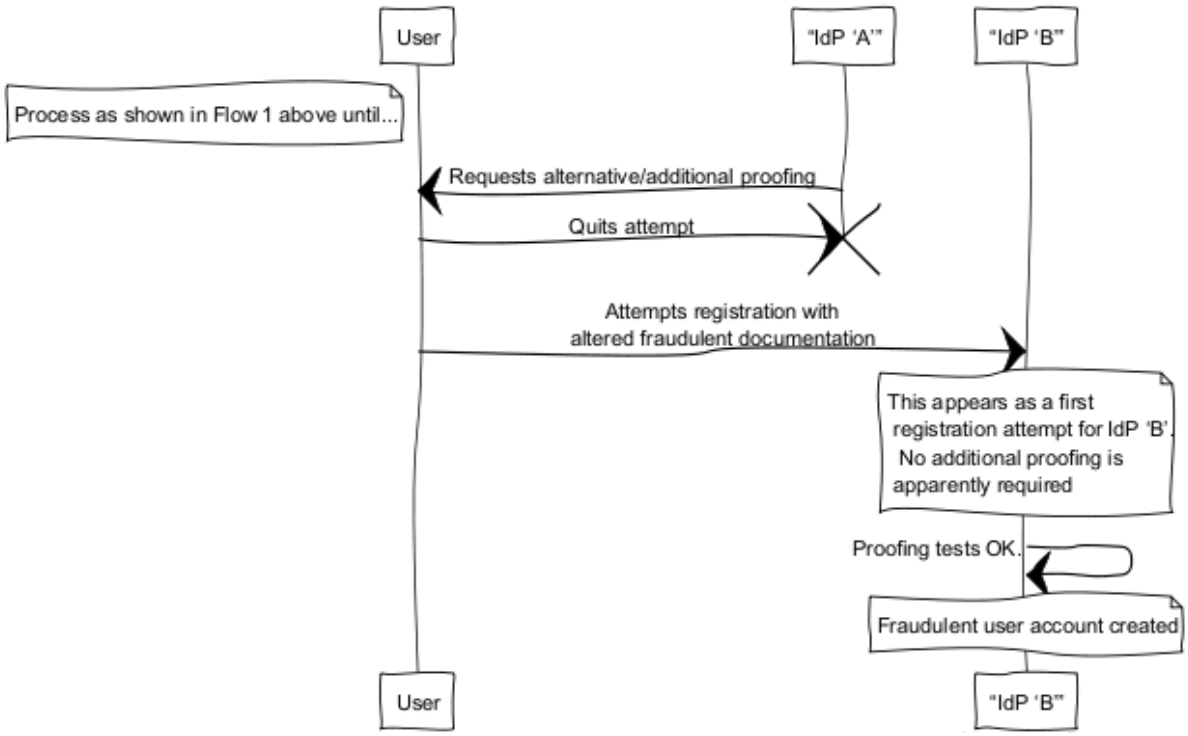
Registration Velocity

The Gov.UK Verify operating guidelines (OPS, 2014) require the identity proofing threshold to be increased if a user attempts to register, fails and retries multiple times. This problem was described by the IdPs during the Discovery project as a 'Velocity Check'.

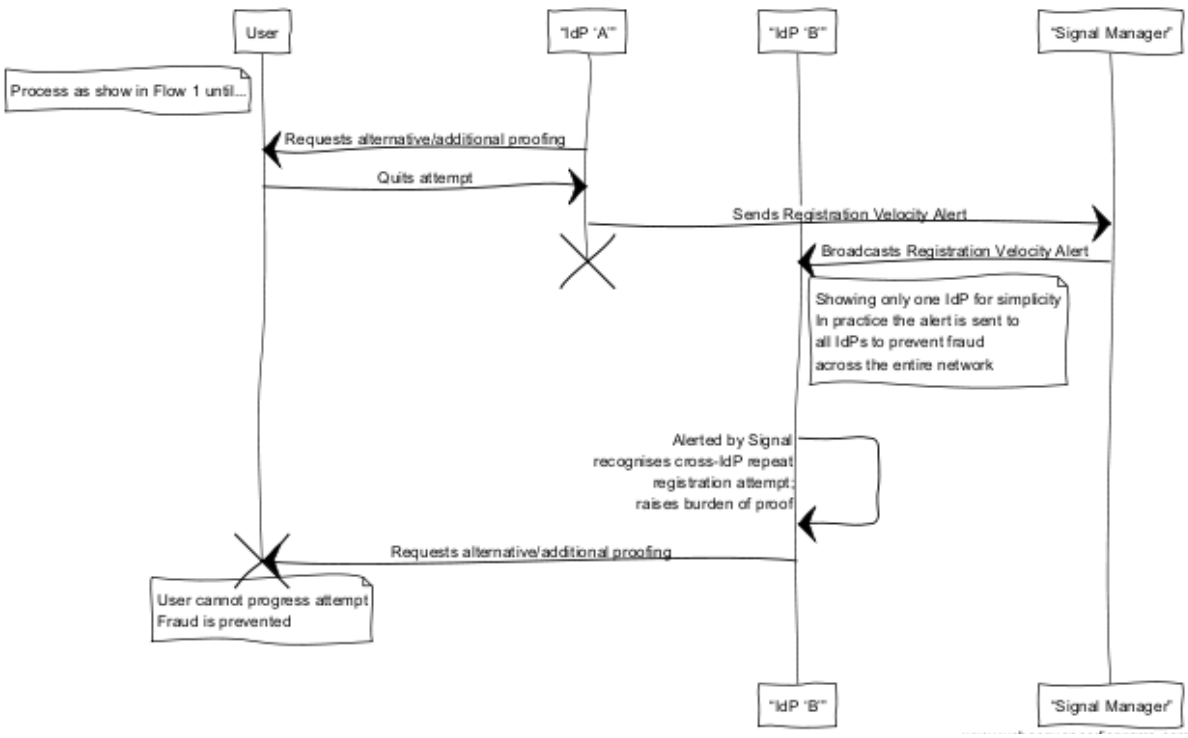
IdPs implement solutions to identify this behaviour and to mitigate against it when multiple registration attempts occur on their own service, as illustrated in Figure 1.



However, a smart fraudster might attempt this process across multiple IdPs. As there is currently no way for information about failed registration attempts to be shared between IdPs, such an attempt would circumvent any IdP-specific protection (Figure 2).



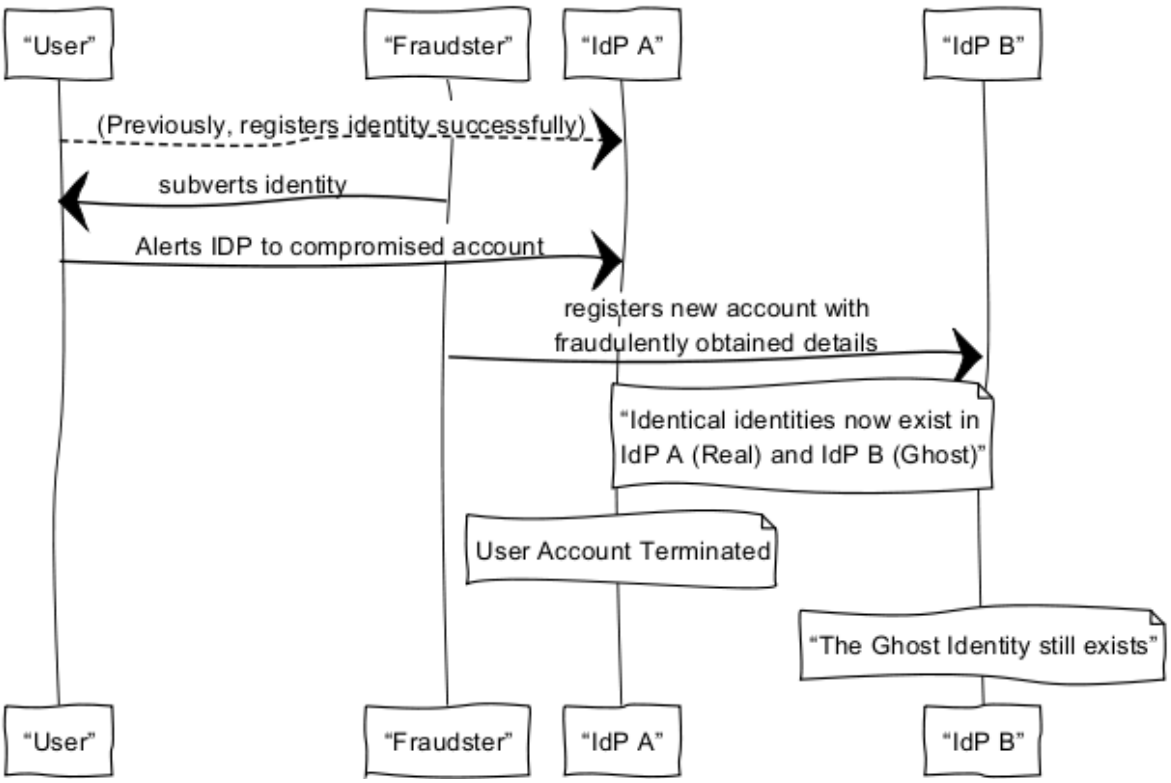
The Registration Velocity use-case for Shared Signals, then, proposes to mitigate this risk by providing a mechanism to share a signal between IdPs when a registration failure occurs, so that IdPs can be aware of a user coming to register at their site trying to game the system after failure at another IdP (Figure 3)



Ghost Identity

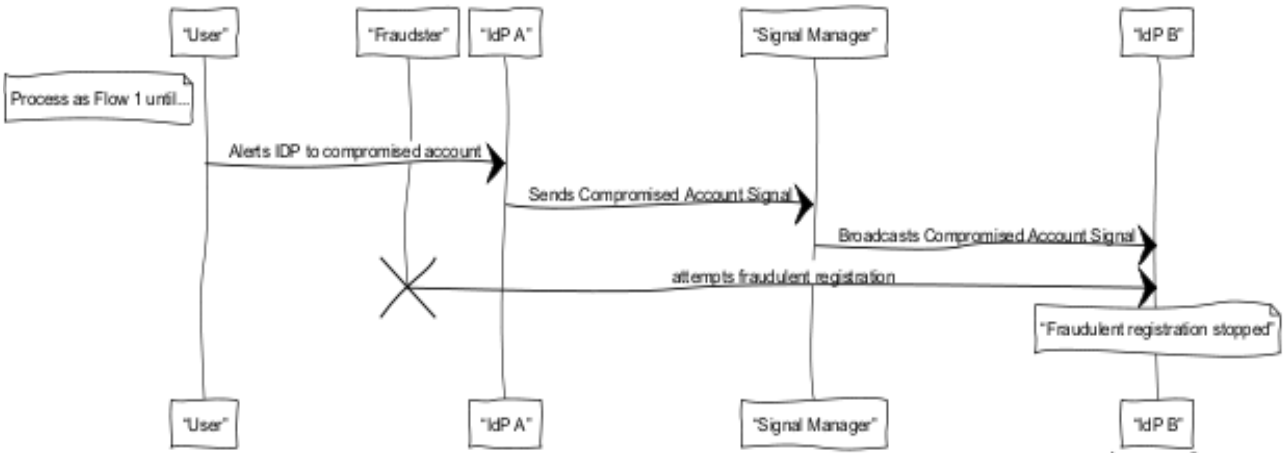
If an identity is subverted at one IdP, then the information about that Identity could be used at a different IdP (by a 'black hat' user) to register a new, fake user with that IdP.

At best, assuming the compromised is identified, there would now be a fraudulent account at the second IdP. At worst, there would exist 2 verified identities, one real and one 'fake'. Figure 4 illustrates one potential flow; others exist, with similar outcomes.



IdP's will typically have either internal activity monitoring systems which can flag potentially compromised accounts; and/or user self-reporting mechanisms that can alert the IdP to a compromised account. However, today, there is no mechanism for one IdP to share information quickly and easily about a potentially compromised account with other IdPs. Without alerts, resolving the Ghost Identity scenario is difficult, and potentially very costly.

The second use case, then, is to share a signal between IdPs when a potential compromised account is identified, so that IdPs can take appropriate action either to guard against a fraudulent registration; or to validate whether an existing matching account has already been registered and apply additional verification to the account to establish if that account is indeed fraudulent. Figure 5 illustrates a user self-reporting flow; with a different trigger, the same flow would apply for IdP-identified account compromise. As with Flow 1, different flows with similar outcomes exist:



Test Details

The Confyrm Signal Manager provides a RESTful API for event distribution and subscription. Initial integration work for testing purposes required DigIdentity to implement calls to this API to broadcast and to receive signals. Once shown to be successful with a simple 'hello world' test signal, this integration work was duplicated in the Post Office system. The end-to-end solution was then available for proper testing work.

For the purposes of this Alpha test, potential triggers for each signal were identified and documented; and initiated manually for signal distribution.

The use-cases under investigation in this phase of work required different signal treatment, both in terms of the content of the signal, and in terms of the distribution mechanism.

Registration Velocity: Signal Details

Unlike other signals discussed in the Discovery Project, for Registration Velocity we do not know who the user is (since they have not registered yet); and so we cannot use the Minimum Data Set to identify the user.

The Alpha project, therefore, tested the use of the email address as the identifying 'subject' for the signal.

The signal contents are therefore:

```
Subject: hash of email address & postal code
Alert: Suspect Registration
Timestamp: {ISO-8601 compliant}
```

Signal Details: Compromised Identity Alert

For the purposes of the Alpha project, the initial trigger for the test was assumed to be user self-asserted subversion (in technical meetings during the project, IdPs reported real-world events of this type). As discussed in the Discovery Project, the likely identifier for the account will be the email address & postal code. The signal contents are therefore:

```
Subject: hash of email address & postal code
Alert: Account Takeover
Timestamp: {ISO-8601 compliant}
```

Outcomes and Observations

The tests showed that signals can be effectively shared between IDPs utilizing a Signal Manager, based on an identified and manually initiated trigger, by either IdP, via the Signal Manager, and received by either IdP.

The tests illustrated that a received signal, in both use-cases, could be used as described to help prevent the fraud risks identified. The specific process for doing this would differ per IdP and (potentially) per Signal. Some cases would require a manual intervention step; others might be handled with further automation at the signal recipient.

For the purposes of this Alpha test, the Subject was hashed. The Participants discussed and agreed that in production (with live data) a pre-shared key deployment would be recommended so that this data could be fully encrypted end-to-end during distribution.

The Signal Manager as initially provided for testing assumed existing knowledge at both the signal producer and the recipient of the account details. (This mechanism supports better protection for individual users, since recipients will only see Signals for those individual accounts in which they have a specific, legitimate and pre-registered interest). Policy controls at the Signal Manager supported creation of a closed group of IDPs to ensure that pairwise and more complex relationships can be established on a per IdP and signal type to limit signal distribution to appropriate IdP recipients.

In the Registration Velocity use-case, however, the signal recipient does not have existing details for the account. This was identified during testing as a blocking issue; the signal manager system was therefore adapted to support an alternative method of signal distribution for this use-case, so that testing could be completed. This mechanism is appropriate for 'closed' groups, such as the GOV.UK Verify IdPs, who are operating within strictly defined and audited guidelines - particularly from a privacy perspective.

Discussion and Further Questions

This phased project was designed specifically to rapidly test the Alpha hypothesis, and to identify potential areas for further investigation. These were collected from the Participants during the testing phases and the technical meetings; and from the Observers during the project meetings. They are listed here in the order in which they were collected, grouped into general, and use-case specific categories, but with no particular prioritisation.

General

- Is the subject identified described in the use-cases above sufficient; or is additional data required?
- What is the best method for key sharing to support data encryption
- Can/should general guidelines be agreed by all IdPs for a given signal trigger? And/or should any such guidelines be signal or use-case dependent?
    - IdPs might carry out their own analysis with a risk-engine to self-generate alerts
    - Should this alert carry a different weighting? Or even be a different signal type?
- Consent for information sharing
    - Does the incoming GDPR directive have any import, in particular given that some use-cases require passing an email address and postcode (albeit hashed) between IDPs. If 'yes', what, and how could this be mitigated?

Registration Velocity Use-Case

- What happens if the user attempts to use a different email address at IdP 'B'? Is there a way a shared signals system could help mitigate such a fraud attempt?

Ghost Identity Use-Case

- What other flows and/or triggers potentially exist for this use-case. Should these (if any) be tested further?
- The potential exists with this use-case for the system itself to be exploited by an IdP to drive net new account creation. Is this something which can be mitigated either:
    a. Contractually (is it already?); and/or
    b. Using data flow analysis at the Signal Manager to look for unusual signal patterns (so to monitor and prevent abuse of the system)?; and/or
    c. Through implementation of collective agreement and/or standards imposition in updated good practice guidelines.

Future Phases

In addition to the questions raised directly in relation to the testing carried out during this phase of the alpha project, a number of items arose during the technical and project meetings which could be considered for a future phase of testing; or would benefit from further discussion in the group.

- The GPG requires (as an auditable item) that email addresses presented by users for registration should be checked against an 'email watch-list'. No such watch-list currently exists. A shared signals system could help with this by:

    a. Distributing alerts from large email providers (such as Google); and/or

    b. Providing a mechanism to help the IdPs compile and maintain their own 'group-specific' watch-list

- Should GDS (via the GPG?) 'mandate' the use of Shared Signals; and/or provide other guidelines on usage of such a system? A shared signals system has particular value if all IdPs agree to co-operate in specific and well-defined ways.

- An IdP reported a real-world use-case (encountered on two separate occasions) where they found a fraudulent passport during verification. The document ID was correct, and the document was not on the Interpol watch-list, but the document was clearly fraudulent based on other tests. Although the document should eventually appear on the Interpol list, this introduces a time delay during which the document could potentially be used to create a fraudulent account at another IdP; or potentially in other ways (for example, to create a bank account or obtain a credit card). This would appear to be similar to the 'registration velocity' case, though it might require additional detail to be shared; and it might be desirable to share this information with other institutions who are **not** IdPs.

- One of the use-cases discussed in the Discovery project was around communication paths being subverted (email, SMS, others). Is there still interest in a practical test of this?

- Is there a case for testing the interoperability of multiple signal managers?

- Are there other use-cases or questions from the original Discovery project that should be investigated during a subsequent phase?

Conclusions & Next Steps

This initial phase of the alpha project has established that it is indeed possible to share signals between IdPs; and that these signals are useful in providing information which the IdPs otherwise would not have to help detect and prevent fraud.

The work has identified two areas for further investigation:

1. More detailed questions around the use-cases that formed part of this phase of testing.
2. New use-cases which could be tested

It seems clear that an additional testing phase is warranted - either to test additional use-cases, or to explore the more detailed questions arising from the original use-cases. Some of the 'observers' from this current phase may also wish to become active participants.

The focused approach we have taken in this project to date - a smaller number of active participants with tightly-defined objectives - supports rapid testing and conclusion, and seems a sensible methodology to pursue. Involvement of a wider group of observers allows conclusions to be shared and discussed, and resulting feedback to be incorporated into the white paper, and should facilitate rapid progress over time.

The first step of any subsequent phase should therefore be to agree a specific objective for testing, and a small number of participants interested in those specific objectives to carry out the tests.

Works Cited

Nash, Andrew. "The Shared Signals Model." OIX, October 2013. (http://oixuk.org/wp-content/uploads/2014/04/The-Shared-Signals-Model-1.pdf)

Nash, Andrew. "Protecting the Identity Ecosystem" OIX, November 2014. (http://oixuk.org/wp-content/uploads/2014/11/Protecting-the-Identity-Ecosystem.pdf)

OPS. "IPV Operations Manual v2.3.1". Cabinet Office and Government Digital Service, December 2014

Walton, Sarah. "Reducing Fraud and Improving Online Safety Through IDP Signal Sharing." OIX, July 2015. Web 29 March 2016. (http://oixuk.org/wp-content/uploads/2014/05/OIX_SHAREDSIGNALS-3_IDP_WP_WITH-LOGO.pdf)