The protection of digital identities in the context of big data analytics

Harry Halpin
OECD Technology Foresight Forum
23 October 2012
There are real-life consequences

Of data analytics and regulations
Technical Solutions do not work by themselves
- Mathematically impossible to truly anonymize large data
- Signals for user-intent (Do-Not-Track) require co-operation of the analytics
- How to even retrieve your personal data from the cloud?

Legal solutions have to be based in technology
- Cost is very high to remove content from large data.
- Users often can't understand their consent to terms-of-service legal contracts, and these contracts often cannot be enforced technically.

We can't turn back the clock..
- Large data has revolutionary consequences for individual users and wider society beyond just advertisements
- Carbon emissions, health, scientific discovery
Policy must become technological

Technology then becomes policy
Standards are the Key!

1. The current identity standards fragmentation (OpenID Connect, Mozilla Personae, Facebook Connect)

2. The need for secure digital signatures and better cryptography for high-value data: See W3C Web Cryptography API.

3. The need to secure the Web in general: W3C Web Application Security Working Group and IETF Web Security/JOSE groups.

1. Need to co-ordinate national-level attempts (NSTIC, India) with global open standard-bodies
The value of data is inherently collective, not individual.

The Web becomes the Web of Data
Enabling the protection of identity while allowing value to be created

Think Answers to Questions, Not Collecting Large Data