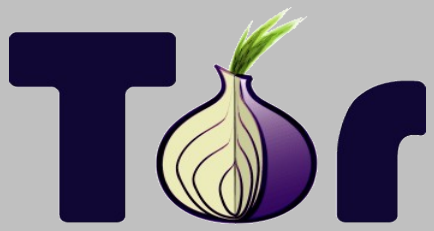
Bridging the Disconnect Between Web Privacy and User Perception

Mike Perry
W3C Identity
May 24, 2011



Enabling Privacy by Design

Users lack ability to understand tracking:

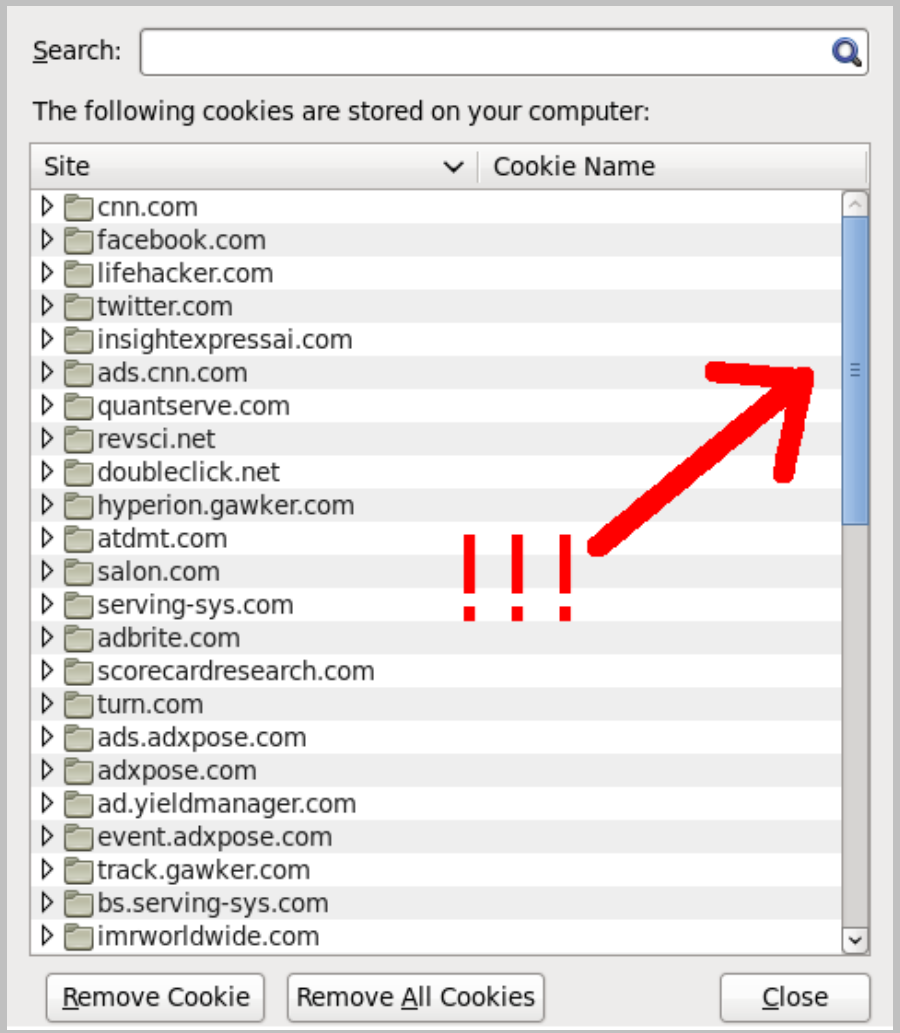
- 'Activity linkability' is invisible
  - Activity is linkable in surprising ways
  - Not just identifiers
- Omniscient entities are enabled by linkability
  - Ad networks, CDNs, Facebook
- Linkability culprits:
  - Flawed identifier transmission model
  - Poor visual cues about user identity
  - Fingerprinting

Improving the Origin Model

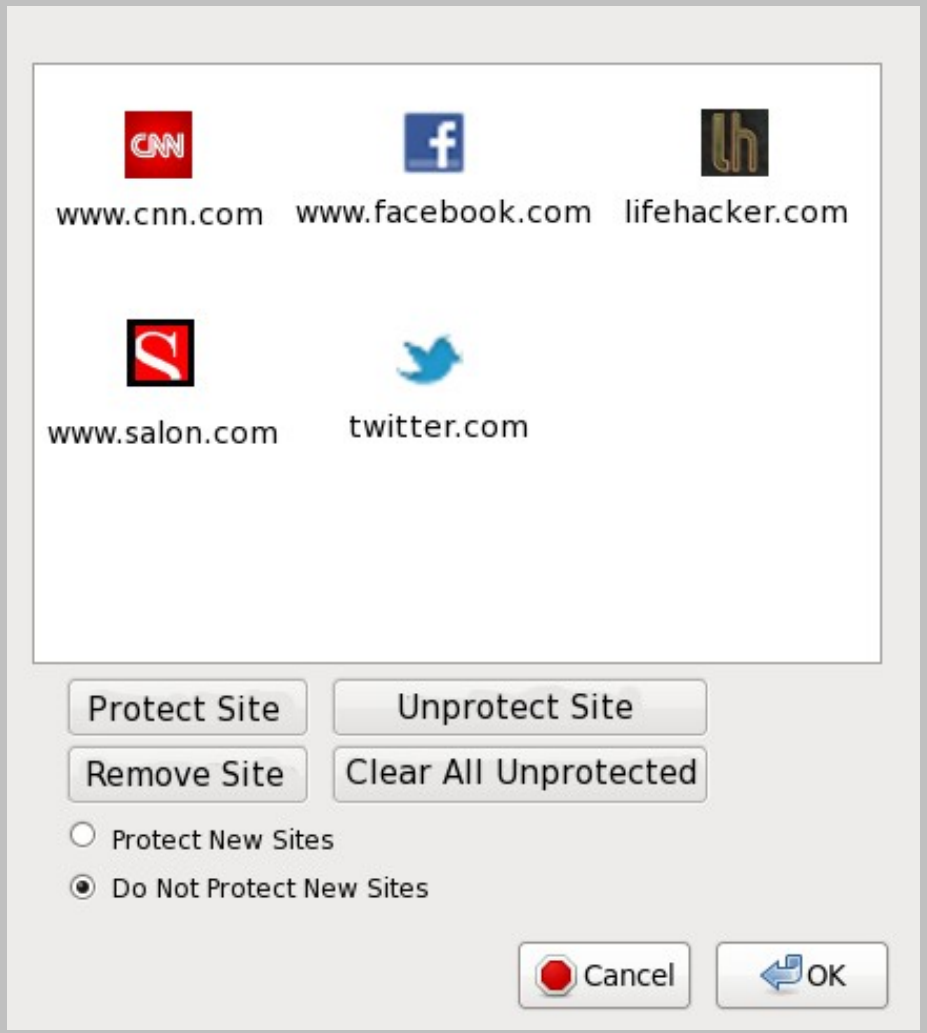
- “Double-key” all new and existing identifiers
  - Use top level domain + 3rd party domain
  - This includes the cache, too (See: SafeCache)
  - Fits users' mental model of website relationships
- Can use to satisfy RFC 2965 sec 3.3.6 (finally!)
- OpenID and OAuth logins still possible
  - Allows concept of “Logging into CNN via Facebook”
- May need whitelist request for (some) cookies
  - Xauth is primary example of a protocol w/ issues
    - Alternatives available (Eg: User-launched Popups)

Top-Level Origin Privacy UI

New model allows simpler UI for the same site data:



Vs



Representing Identity

Users should be given cues about who the web thinks they are

- Browser-level pseudonym, avatar icon, or Firefox Persona
- Password protected browser state storage
  - Users should be able to log in and out of identities
  - Mozilla Weave
- Private Browsing Mode is a special case with no storage and low linkability to saved identities
  - Should have user-selectable persona/cue, too

Fingerprinting

- Ad networks are already moving beyond cookies
  - Banks too
- Unique browser attributes identify users
  - Installed plugins, fonts, user agent, resolution, time..
- Can be measured by Panopticlick Entropy Metric
- Solutions can be origin-based or identity-based
  - Origin solutions/restrictions may be more effective
- New web features need evaluation
  - This can get tricky..
  - May need to rely on simulations or intuition

Summary

1. Improve identifier origin model
   - Greatly simplifies privacy UI
   - All site data managed together (including history)
2. Provide (better) identity cues to users
3. Address Fingerprinting Linkability
   - Fingerprinting Linkability between origins
   - Fingerprinting Linkability between identities