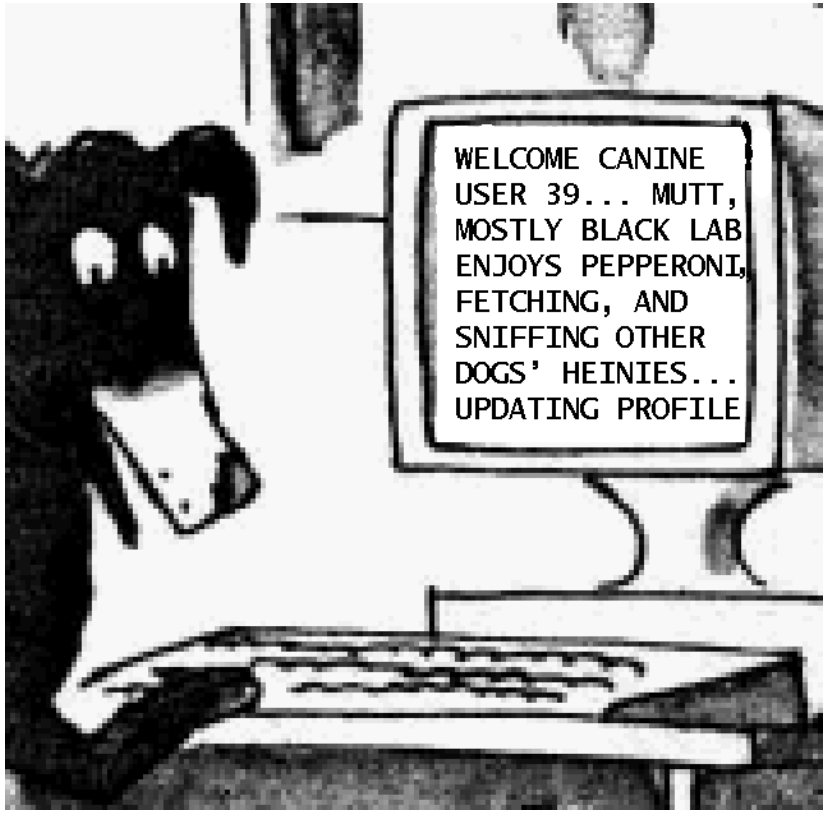
Internet Privacy: Then and Now

"On the Internet, nobody knows you're a dog."



Right cartoon obtained from `http://www.unc.edu/depts/jomc/academics/dri/idog.html`.
Left cartoon by Peter Steiner has been reproduced from page 61 of July 5, 1993 issue of The New Yorker.

WPI

Internet Privacy: Leakage, Linkage and Lifetime of Data

Craig E. Wills
Computer Science Department
Worcester Polytechnic Institute
cew@cs.wpi.edu

Privacy is a Big Topic

Potential concern for private information loss to many entities:

- Commercial/Business
- Governments/Defense & Security Agencies
- Neighbors
- ...

Focus of this excerpted presentation is on individual control of private information to commercial entities on the Internet.

More importantly on the Leakage, Linkage and Lifetime of Information.

Leakage, Linkage and Lifetime

Unwanted dissemination of private information to a third party requires the presence of three conditions:

1. leakage of information,
2. linkage of information from different sources, and
3. a non-trivial lifetime for the information.

Elimination of any of these three conditions prevents potential harmful privacy loss from occurring.

So what can be done?

Elimination of Data Lifetime

What if lifetime of data can be controlled?

- "Ephemeral messages are incredibly freeing and make people communicate more authentically and freely with their friends."

  Sep'13 Communications of the ACM

- Snapchat as an example where pictures last 3-10 seconds.

- Timed revocation (or expiration) of private data using keys from trusted server.

Problems: local caching/snapshots, data replication. Is data really ephemeral? Requires trust in a shared entity.

Elimination of Linkage

Many vectors:

- Heavy use of tracking cookies by most sites.
- Browser fingerprinting.
- IP Addresses
- Globally unique identifers: email addresses, usernames, social network identifiers (enabling linkage across devices)

Allows linkage of information across sites and applications.

Must control all vectors, not just one.

Elimination of Leakage

First-party sites are often in the best position to prevent leakage of information.

Users can block content, but existing tools are one-size-fits-all (or used as such) where users do not understand what is being blocked.

Ad blockers do not work for ads from site itself.

Location data is of interest to sites/applications for personalization, but may be particularly sensitive from privacy perspective.

What Can a User Do?

1. Tension between users wanting to protect information and aggregators wanting to collect it—more information leads to better ad targeting.
2. Ads pay for content, this is a way for users to see more relevant ads.
3. First-party sites are often in the best position to prevent leakage of information.
4. Users can make it harder for aggregators by refusing their cookies and blocking their content—can be done via browser settings and extensions, but may impact page display/function.
5. Need to look at semantic solutions where users can better understand and control what happens with their information.
6. Users should minimize information given to sites and applications—cannot leak what they do not know!