Empirical Measurements of Perceived Privacy Risk

Jaspreet Bhatia and Travis Breaux
(Carnegie Mellon University)

PRIVACYCON

New Findings on Privacy Risk

New scale is introduced to control the perceived likelihood of privacy risk in user surveys. Observations of scale:

- As likelihood increases, people are less willing to share their personal data
- As societal benefits increase, people are more willing to share their personal data

When sharing personal data with the federal government, participants are more willing to share information ***about who they are*** than information ***about what they do online***

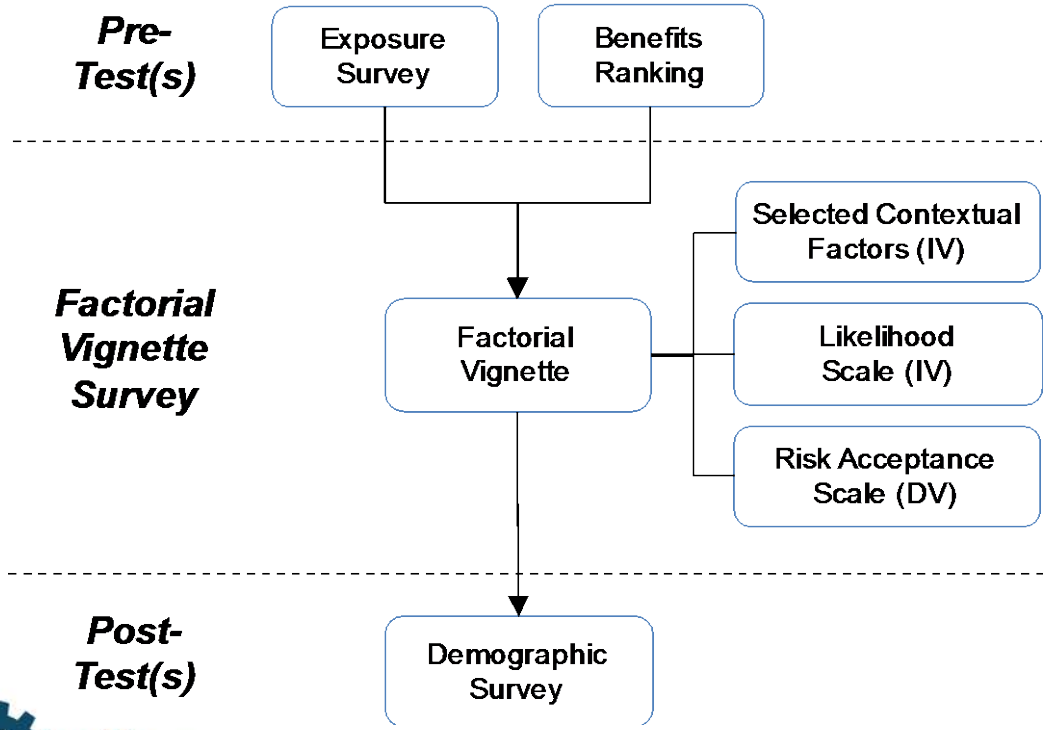
Measuring privacy risk is complex, and thus measuring “discomfort” is an easier means to approximate privacy risk

What is Risk?

- Bauer (1960) defined risk as comprised of two dimensions: uncertainty and adverse consequences
- NIST (2015) defines risk “as a function of the likelihood that an adverse outcome multiplied by the magnitude of the adverse outcome should it occur”
  - *Risk* = *Likelihood* x *Impact*
- Empirical risk research in judgement and decision sciences (1970-) distinguishes *revealed preferences*, and *expressed preferences,* or so-called *perceived risk*

R. A. Bauer “Consumer Behavior as Risk Taking,” *American Marketing Association*, R. S. Hancock (ed.), Chicago, IL: American Marketing Association, 389 398, 1960.

Risk Measurement Framework



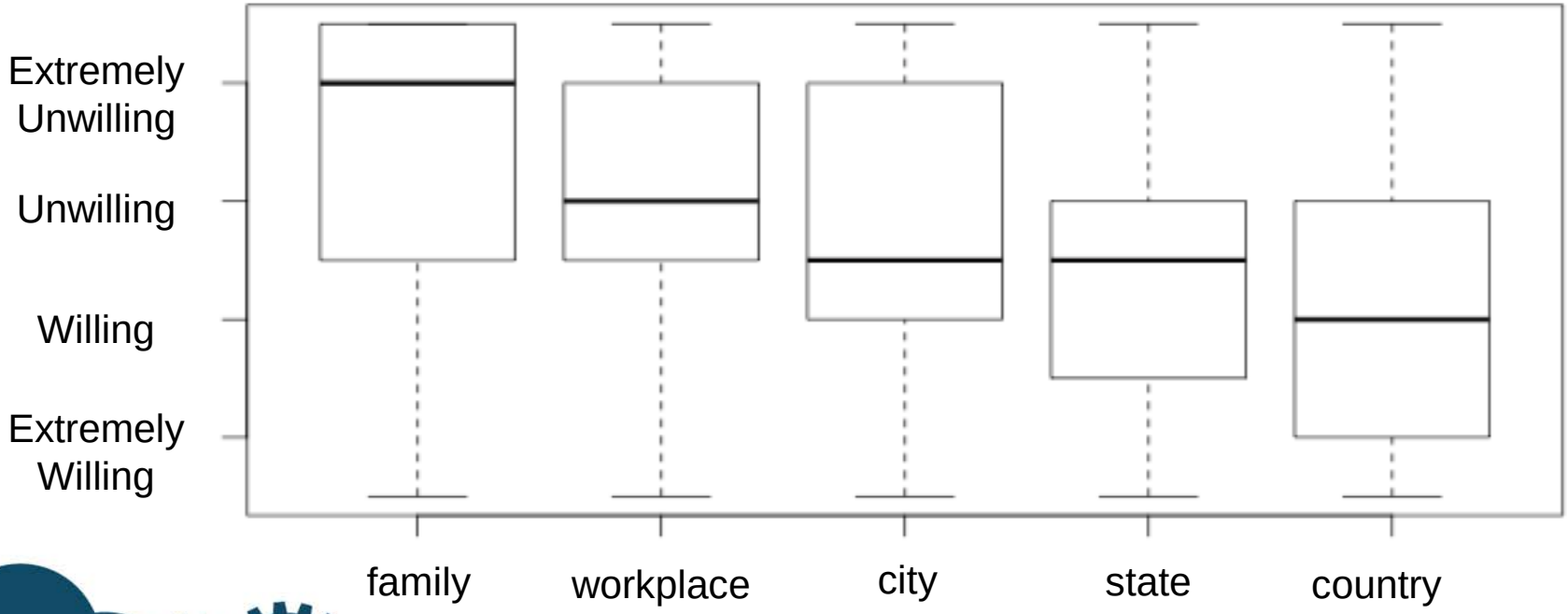
Factorial Vignette

Please rate your willingness to share your information below with the Federal government, for the purpose of **[insert data purpose]**.

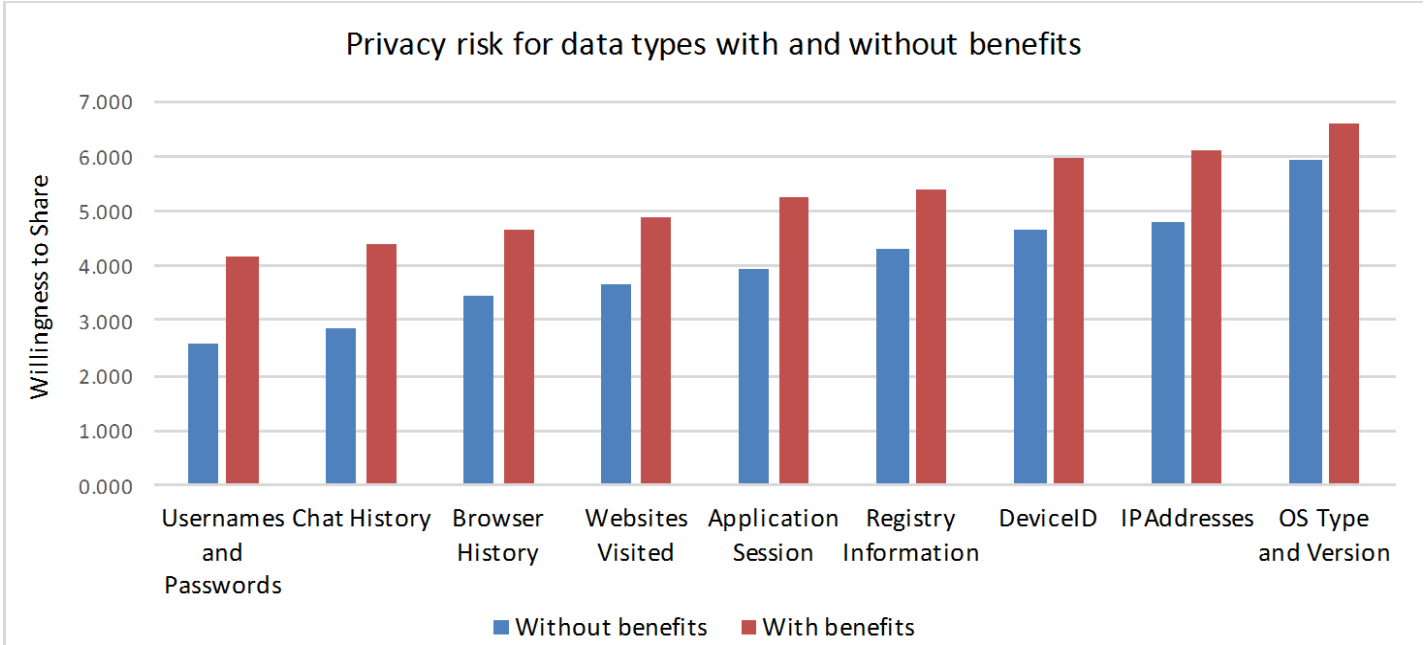
When choosing your rating for the information types below, consider the **[insert risk likelihood]** and the purpose above.

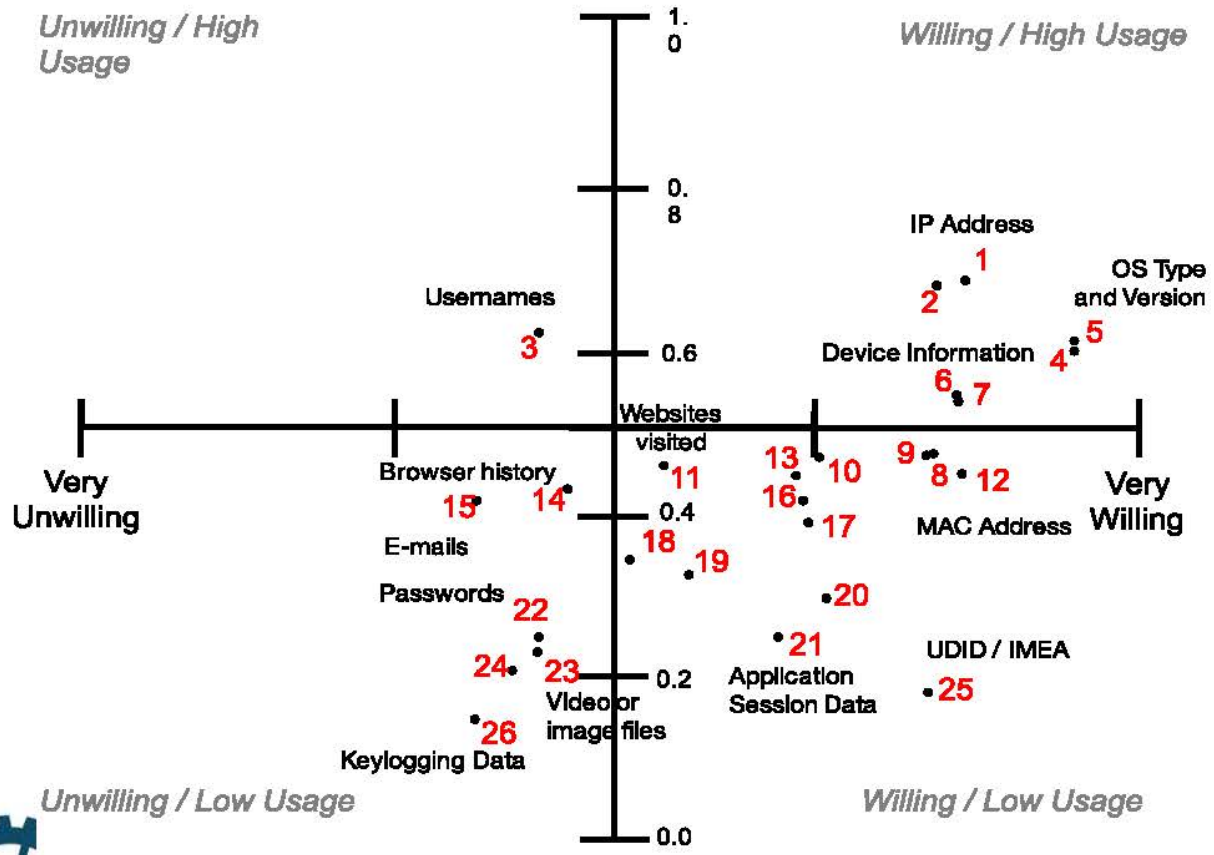
	Extremely Willing	Very Willing	Willing	Somewhat Willing	Somewhat Unwilling	...
Age Range	○	○	○	○	○	
Home Address	○	○	○	○	○	

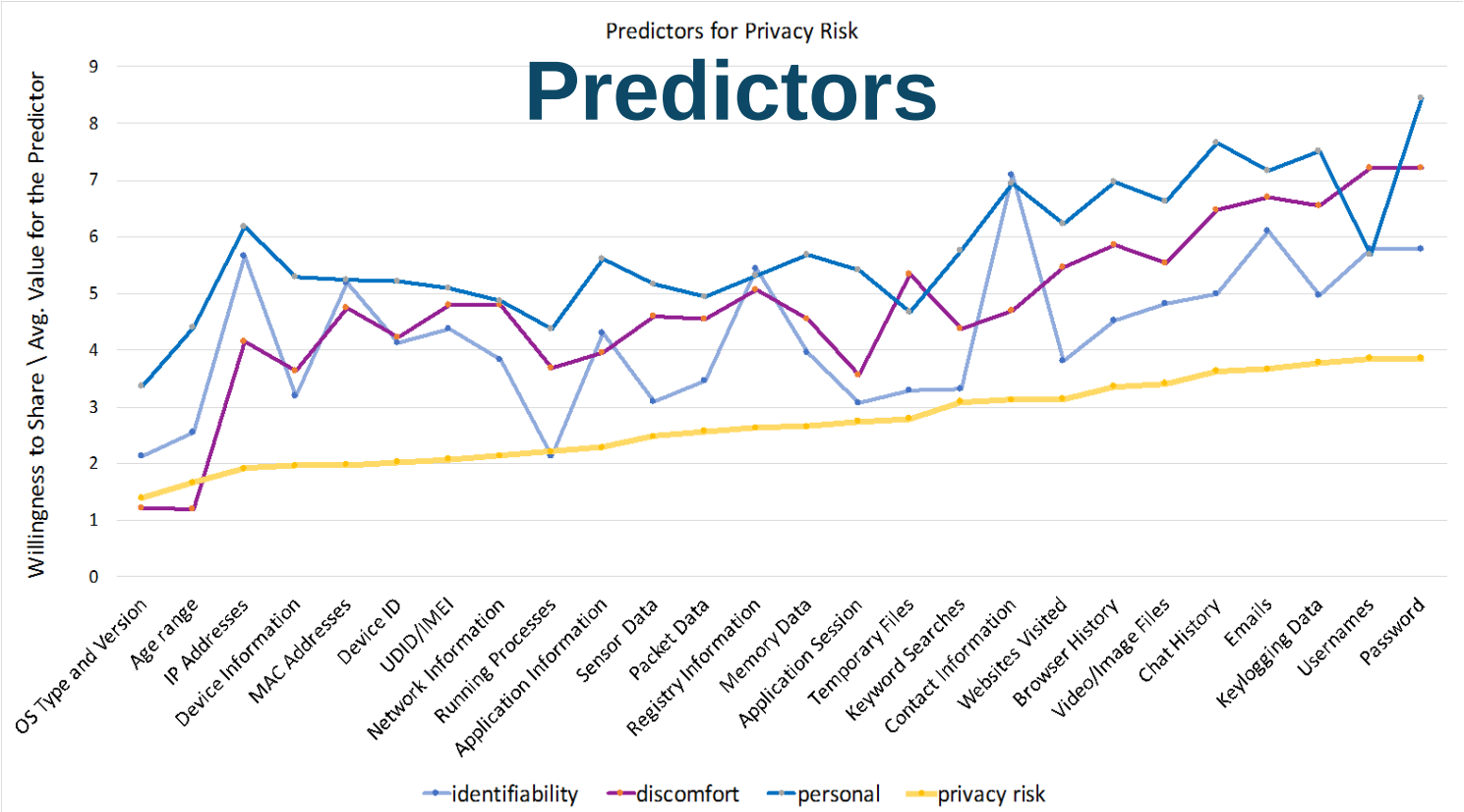
Risk Likelihood



Benefits


Privacy risk for data types with and without benefits
Willingness to Share
7.000
6.000
5.000
4.000
3.000
2.000
1.000
0.000
Usernames and Passwords
Chat History
Browser History
Websites Visited
Application Session
Registry Information
DeviceID
IPAddresses
OS Type and Version
Without benefits
With benefits


Unwilling / High Usage
Willing / High Usage
1.0
0.8
0.6
0.4
0.2
0.0
IP Address
1
2
OS Type and Version
5
4
Usernames
3
Device Information
6
7
Websites visited
9
13
10
8
12
11
16
Very Unwilling
Browser history
15
14
17
Very Willing
MAC Address
E-mails
18
19
Passwords
20
22
21
UDID / IMEA
24
23
Application Session Data
25
Video or image files
26
Keylogging Data
Unwilling / Low Usage
Willing / Low Usage


Predictors for Privacy Risk
Predictors
Willingness to Share \ Avg. Value for the Predictor
0
1
2
3
4
5
6
7
8
9
OS Type and Version
Age range
IP Addresses
Device Information
MAC Addresses
Device ID
UDID/IMEI
Network Information
Running Processes
Application Information
Sensor Data
Packet Data
Registry Information
Memory Data
Application Session
Temporary Files
Keyword Searches
Contact Information
Websites Visited
Browser History
Video/Image Files
Chat History
Emails
Keylogging Data
Usernames
Password
identifiability
discomfort
personal
privacy risk

Impact of Privacy Risk Metric

Facilitates “Privacy by Design” by helping software designers identify high-risk data types associated with their business practices

Helps regulators and policy authors identify high-risk items so that related data practices can be described clearly in policies

Assists users in making decisions related to using online services

Supports the standardization of the Privacy Impact Assessments

This research was funded by:
NSF Award CNS-1330596, NSA Award #141333
and ONR Award #N00244-16-1-0006