LEGAL AIKIDO:
A DATA-SHARING FRAMEWORK TO ADVANCE NETWORK & SECURITY RESEARCH
Talking Map

- Defining the Issue & Solution Space
- Challenges & Motivations
  - Uncertain Legal Regime
  - Incomplete Technology Solution Models
  - Data Risks
  - Under-valued Benefits of NetSec Research
- Applying Akido:
  - Self-Reg
  - Operational model: PS2 Framework
  - CAIDA
  - Kenneally (C) 2012 CAIDA | Kenneally
The Issue Space Defining the Solution

- Current posture:
  -
- (Misinformed) assumptions:
  -
  -
  -
- Muted legislative, judicial, policy drivers
  -
  -
- No widespread, standard procedures for exchange
  -
- Dynamic and normative-deficient understanding of privacy risk and research utility
  -
  -
  -
- confusion = window of opportunity
Challenges & Motivations

(1) Uncertain Legal Regime

- No legal framework explicitly prescribes, incentivizes, or forbids sharing network data for security research

- Ambiguity between tech & legal discourse re: fundamental concepts driving risk
  - PI, REP, content, URLs, IPAs, packet headers, payload...
  - Law inconsistent-functional equivalent of PII
  - E.g., is IPA 'content' and URL 'addressing' data (ECPA, 4th A. purposes)?
    - Johnson v. Microsoft (2008) - IPA does not identify persons
    - State v. Reid (2007) - REP in subscriber information attached to IPA
    - US v. Forrester (2007) - URLs may have REP-reveal communication content
    - HIPAA Privacy Rule - IPA is protected PII
    - States' data breach laws - IPA not in definition of personal information

- Social normative expectations: my IPA, URLs + search terms are digital fingerprints?
  - E.g., Tor, automated in-browser cookie and URL deletion (C) 2012 CAIDA | Kenneally
Challenges & Motivations

(2) Incomplete Technology Solution Models

- Point solutions fail to address context-dependent risks
  - Prefix-preserving anonymization subject to re-identification
    - Poster cases (Netflix, Yahoo!, Traffic injection attacks)

- Trade-off: Purely technical approaches v. research utility goals
  - E.g., Conficker (2012) CAIDA | Kenneally
Challenges & Motivations
(3) Data Sensitivity Risks

- Sources: legal compliance, ethical obligations, norms/court of public opinion
- Main categories
  - Disclosure risk
  - Misuse risk
    - False inference
    - Data confidentiality
    - Privacy

...
Challenges & Motivations
(4) Under-valued Benefits of Network Research

- Justify the Ask:
  -
  -

- Network Data sharing utility criteria
  -
  -
  -
  -
  -
  -
Solution Space: Using Aikido on Net Sec Researcher’s BFF

- **Options:**
  - Amend law
  - Aikido the law via self-reg regime

- **1st** ID the attack – on voluntary disclosure of non-content to researchers

- **2nd** Blend & harmlessly redirect attack - use the law itself to clarify the gray and embolden the exceptions
  - Consent Exception

- Provider Exception

- Provider Relevance

- Gov’t Entity

- Content
Implementing Akido

- Privacy-Sensitive Sharing (PS2) model
  - Risk – Benefit methodology
  - Enables transparency as touchstone of data sharing
  - Considers practical challenges of stakeholders
- Bottom-up
Core underpinnings:
- risks ‘contagious

1. Authorization
2.
3.
4. Purpose
5.
6. Use
7.
8.
9.
10. Audit
11.
12. Training
13.
14. Ethical
15. Disclosure minimization
PS2 Framework
Technology Component

- Disclosure Minimization/Controls
  a)
  b)
  c)
  d)
  e)
  f)
  g)
  h)
  i)
  j)

- Implementing Vehicles:
  ▸
Evaluating PS2
Addressing Data Risk & Utility Goals

- **Criteria:**
  - 1.
  - 
  - 
  - 
  - 2.
  - 
  - 

- **Conclusion:**
  - Only tech
  - Only policy
  - hybrid
  - Evaluation
  - Policy Dev tool

(C) 2012 CAIDA | Kenneally
Table 1: Privacy risks evaluated against the PS2 privacy protection components. *(Minimization refers to the techniques evaluated in Table 2.)*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Authorization</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Transparency</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Law Compliance</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access Limitation</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Use Specification</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Audit Tools</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Redress</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oversight</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Quality</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Security</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Training/Education</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impact Assessment</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Sharing</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Delete All</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Delete Part</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Anonymize</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Aggregate</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Mediate (SC2D)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age Data</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Limit Quantity</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Layer Anonymization</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 2: PS2 minimization (of collection and disclosure) techniques evaluated against utility.