Art Into Science. 2018. Austin

Beyond Warm & Fuzzy- Ethics as a Value Prop

Erin Kenneally, M.F.S., J.D.
U.S. Dept. of Homeland Security
Science & Technology Directorate
Cyber Security Division

Homeland Security
Science and Technology

The views expressed do not reflect the official policy of the U.S. Dept. of Homeland Security

Scratching Below the Surface

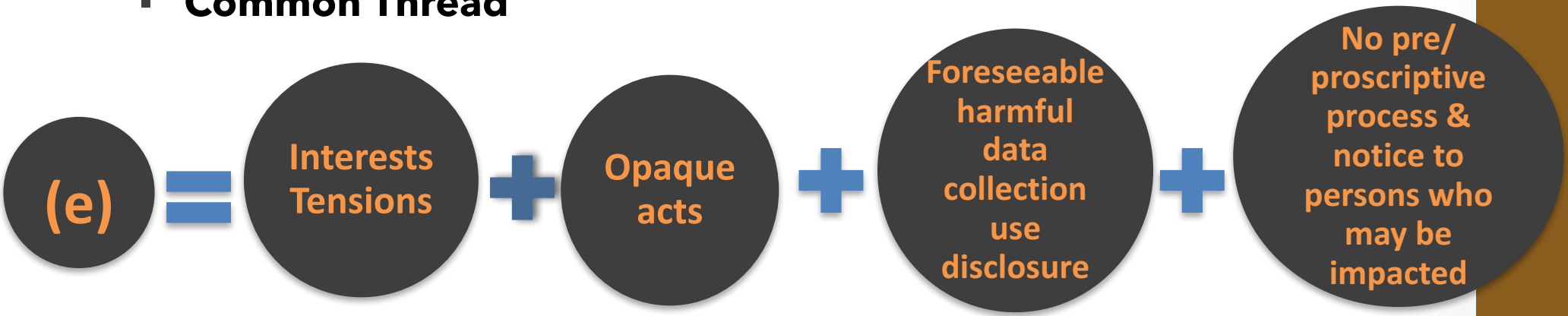
- **Ethics = Good**
- **I'm/We're a good person/company … go away**
- **Ethics can co-exist with capitalism**

- **Champion for Ethics → Attorney + Fed**

- **Proposal: Ethics as a Fundamental Ordering Force for the Evolving Technology – Law Control Plane**

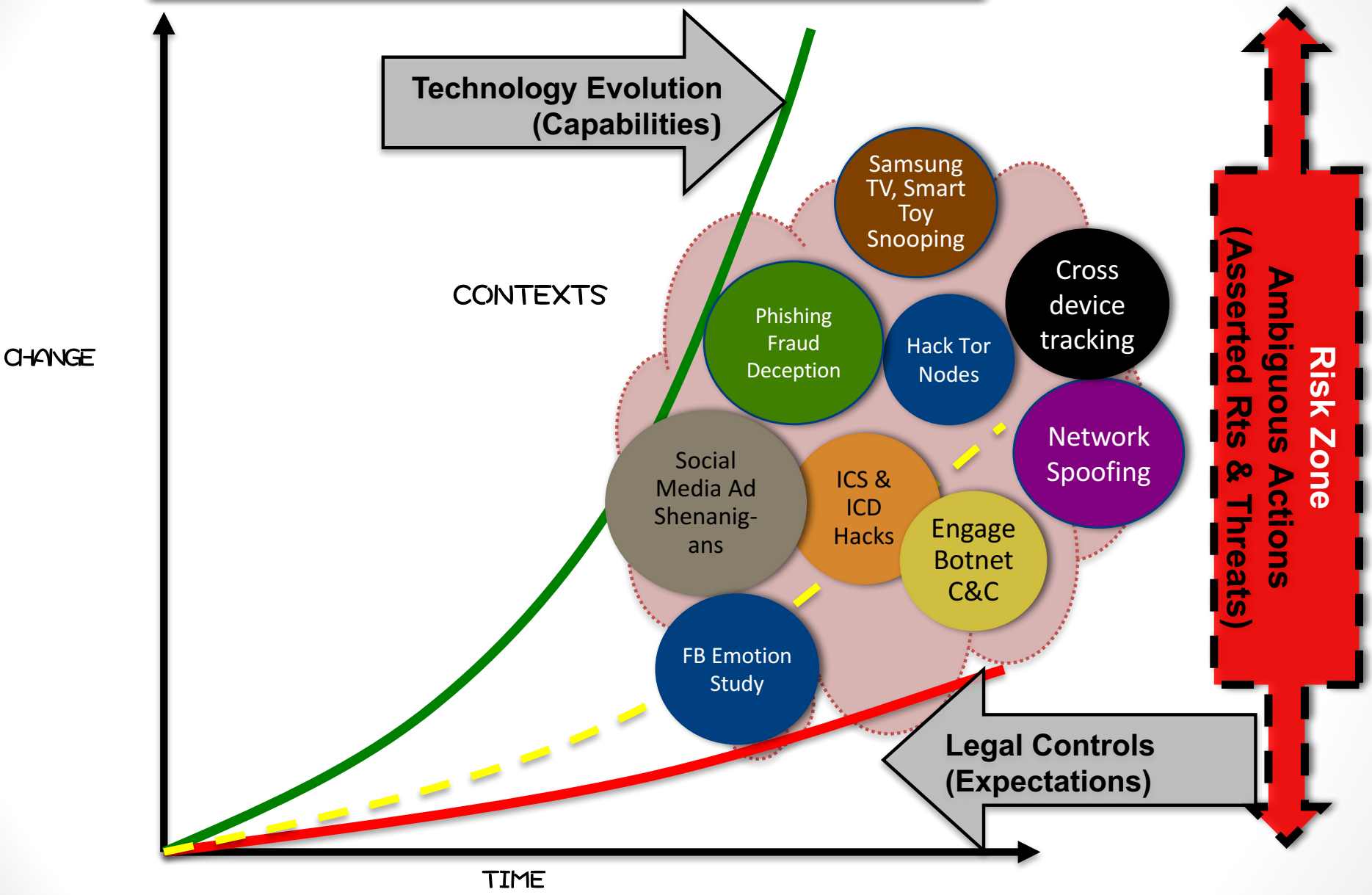
- **What characterizes anonymous observation, collection and use of sensitive data online w/o interacting with the data subject?**
  - (a) Cyber espionage and surveillance by industry & nation-states
  - (b) Online advertising and data brokering by industry
  - (c) Targeted services and content by industry
  - (d) Security R&D (honeypots, botnet recon, reverse engineering, vuln disclosure)
  - (e) AotA

- **Common Thread**



- **What motivates attention to these harms & differentiates acts:**
  - Law and Tech → Ordering Forces
  - When silent /unclear /gaps → risk of harms may be unattended or conflated

Enter Ethics as an Ordering Force



I. WHY ETHICS — TECH CAPABILITIES

1. **Tech = Mediating our Knowledge & Actions** … no longer providing affordances
   - Cause: Sensors, Digitization, Connectedness
   - Output: Filter, Associate, Prioritize, Classify, Measure (collection -scanners, crawlers, social media & analysis - data mining, ML, probabilistic reasoning tools)
   - Outcome: Tech decides, observes, interferes, interacts, advises FOR, ABOUT and WITH people

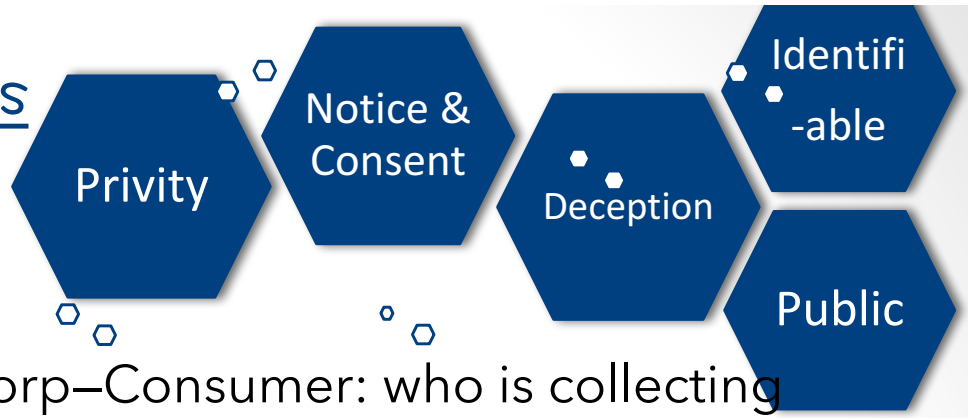
2. **Mediation → Knowledge & Action Asymmetries**
   - * **Opacity**
     * **Unilateral, Subjective Gatekeeping** (lack User control, choice, inclusion)
     * **Impact uncertain** (data use purposes emergent, physical harm is real, collateral impact, learning systems, low and slow harm)
     * **Scale & Ubiquity**
   - **Eg,** Recommender systems (FB newsfeeds, G search results) Reputation Scoring (org security, identity validity), Autonomous devices (planes, cars, weapons, agents), Classify & Predict about & for (crime, disease, employment, insurance)

(cont) Why Ethics — Tech Capabilities

3. **Asymmetries → Increasing Battle of Stakeholders:**
   Privacy v. Security v. Innovation v. Free Speech

4. **Battle Resolution --> Proxied by Industry**
   - Decisions & Actions impact our RIGHTS and INTERESTS
   - Many not binary rt v. wrong … depend on judgments, values, sensibilities
   - Shared judgments and assumptions uprooted (social, political)
   - Don't blame the Algos!
     - Algorithms "a series of steps undertaken in order to solve a particular problem or accomplish a defined outcome."

WHY ETHICS — LAW EXPECTATIONS



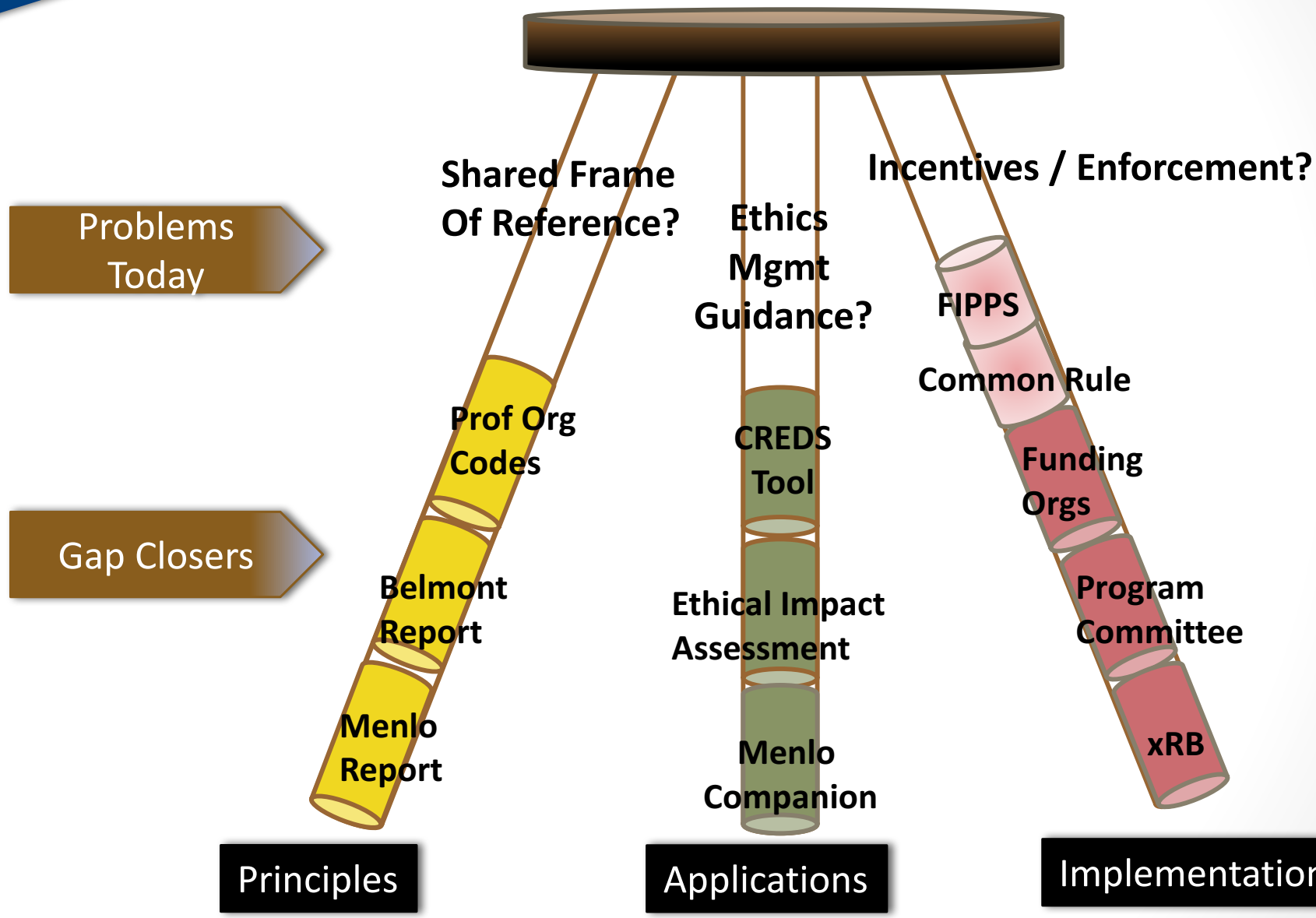
- **Privity & Causation**
  - Indirection between Corp–Consumer: who is collecting info? no direct rltnshp w/ corp, Datageddon
  - Courts conclude no "harm" or rights violation
  - Access & participation rights not triggered

- **Notice & Consent**
  - Impracticable when interacting with 10^3 -10^5 persons behind the machines/network traffic?
  - Complexity of BD, practical limitations privacy policies
  - Inefficient (dev time, cost) eg, interfaceless devices
  - What is Public/Private?
  - Shared information issue

- **Subjectivity, Deception** may be necessary
- **"Identifiability"** Person/Human Subject? Varying linkages between data and individuals' IDs; fingerprinting innovation

- **Beneficence calculation**:
  - Short term, bells & whistles v. harm latency
  - Hard to **Quantify risks**- minimal risk? collateral harm?

II. What Ethics? State of Affairs



II. What Ethics ?

Normalizing Principles

Research - Ethics (Belmont /45CFR46/Common Rule)

Respect for Persons
- Individual Autonomy- Informed Consent

Beneficence
- Do no harm
- Minimize risk, Maximize benefits

Fairness & Justice
- Equitable selection of persons/subject
- Fair distribution of benefits & burdens

Industry- Law (FIPPs)

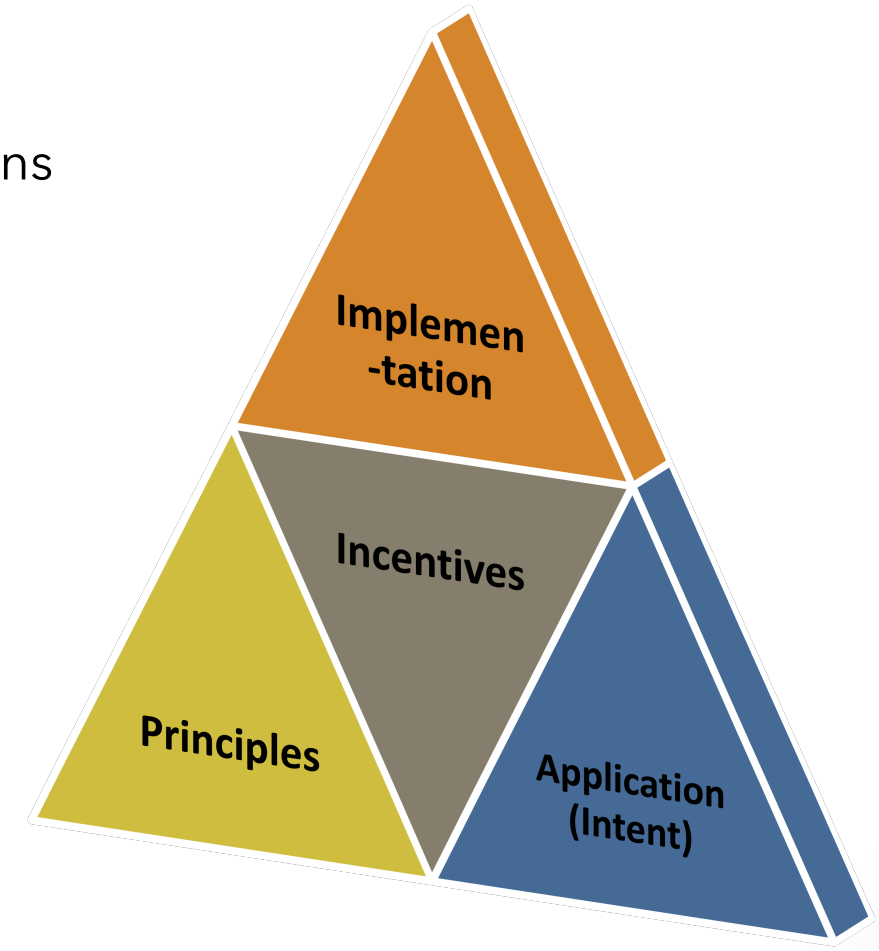
- Purpose Specification and Minimization
- Collection Limitation
- Use Limitation
- Individual Participation and Control
- Data Integrity and Quality
- Security Safeguards and Controls
- Accountability and Oversight
- Openness and Transparency
- Remedies

ISO 26000 (SR)

- Accountability
- Transparency
- Respect for Law
- Respect for International Norms
- Respect for Stakeholders Interests
- Respect for Human Rights

III. How Ethics

- **Quiz #2: What's the difference between "Norm-building," "Influence Operations," "Propaganda," and "Advertising"?**
  - (a) Underlying Principles
  - (c) Implementation
  - (c) Underlying Applications
  - (d) Incentives
  - (d) All of the above



How Ethics: Minding the Gap Options

- **(1) Bottom-Up**
  - "Ethically-Defensible" Research & Commerce
    - **Tool Building**: Decision support capabilities, Notice & Consent, Disclosure Control
      - **Education & awareness**
      - **Self Governance**; community consensus & oversight; market differentiation
      - **Enlist expertise**
- **(2) Top-Down**
  - Stick/Carrot :
    - Dreaded **"R"; xRBs**
    - **Tie to funding**, publication; reward ethical behavior
- **(3) Sideways**
  - Getting *New York-Times'd*
  - **Reputation** lever

/Example/ Trickle-Up Ethics: Cyber-risk Ethics Decision Support (CREDS) Tool

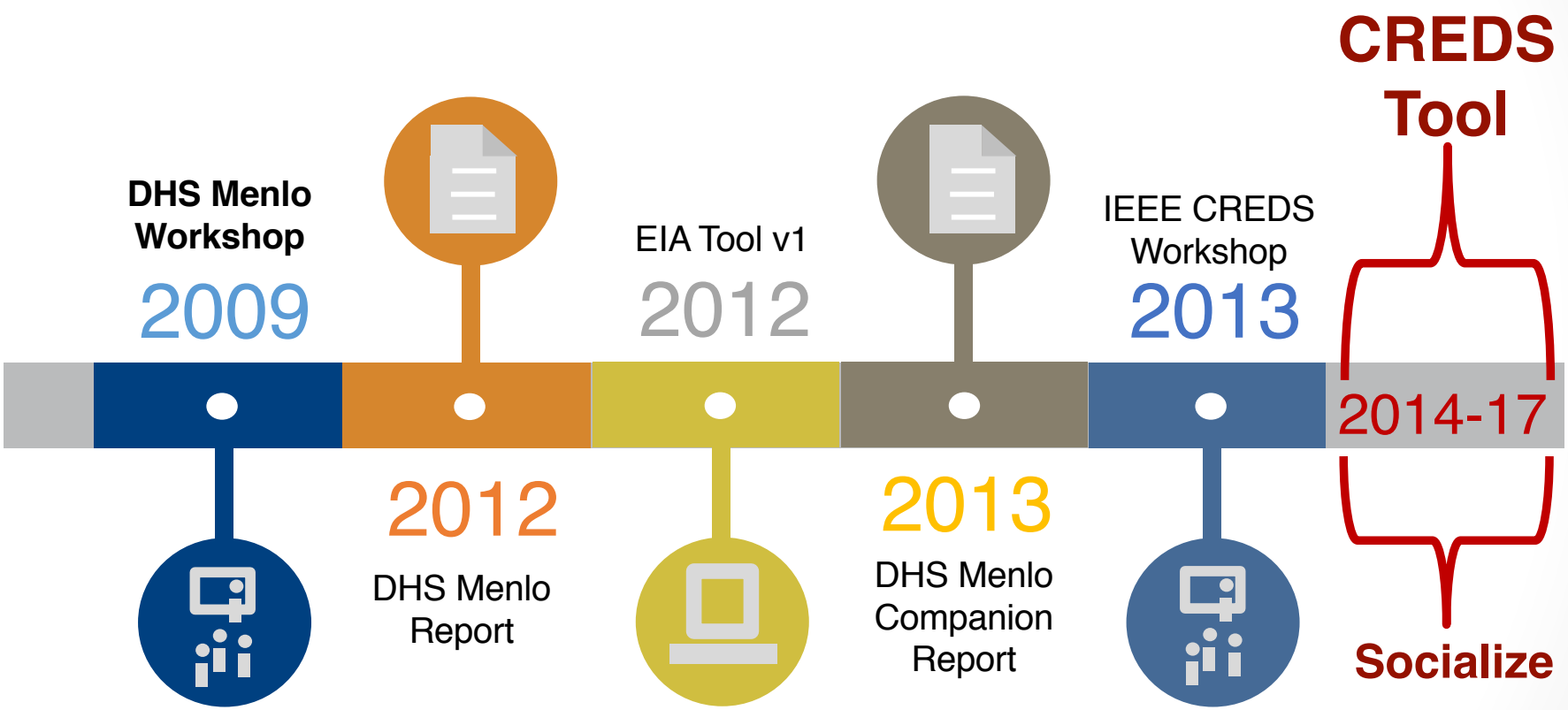
- **Objective:** Operationalize a decision support conceptual framework + methodology into a tool that codifies ethical and legal principles

- **Goals**:
  - o identify and communicate ethical uncertainty and risk;
  - o estimate potential ethical impacts of technology;
  - o measure and improve human judgment and reasoning.

- **Target**: Researchers, Product Developers, Overseers (ERB, PC, Funders)

- **Methodology**:
  - Derive rights and obligations/responsibilities ethics & laws tenets, organizing principles, best practices
  - Transform EIA logic and methodology into an online decision support tool (CREDS)
  - Test and improve with real world, case-based scenarios and consultation with a range of stakeholders

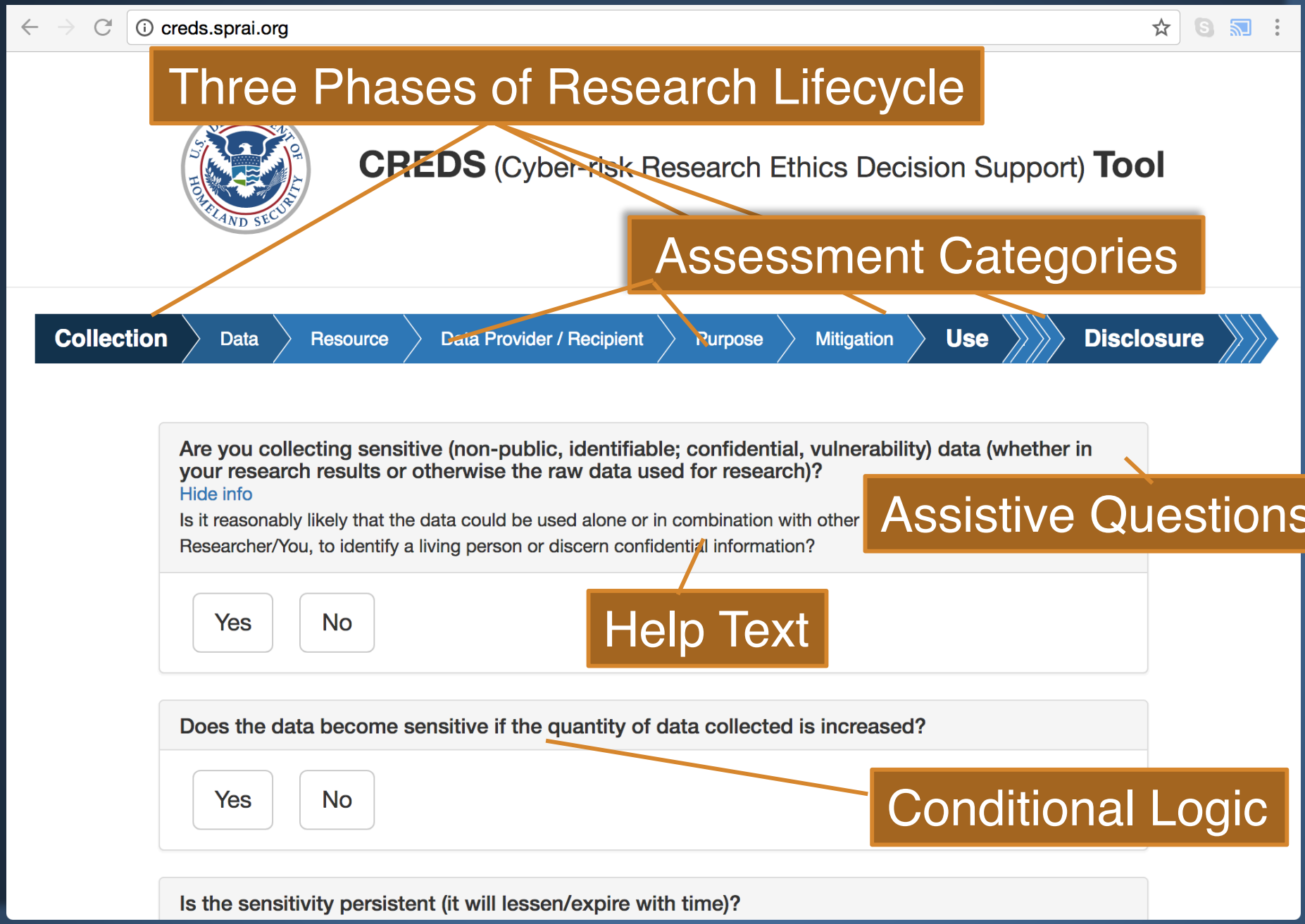
From Whence It Came…



CREDS Tool – Ethics Logic

ETHICAL IMPACT ASSESSMENT			
Research Lifecycle	**Ethical Principles**	**Risk Factors**	**Assistive Questions**
(1) Research Collection **(2) Research Use & Management** **(3) Research Disclosure** **(1)**	**Respect for Persons -** (Identification of	**Nature of the Data**	
		Sensitivity: non-public, identifiable; confidential	
	Beneficence (Minimizing risk to individuals; Maximizing benefit to society; Mitigating realized harms))		
		Nature of the Resource/System	
		Platform	
		Network	
		Nature of the Data Provider, Data Recipient, Data Subject	
		Stakeholders rights and interests	
		Nature of the Data Collection Purpose	
	Justice (Fairness & Equity in selection of subjects and distribution of research benefits)	**(2)** **Harm Mitigation**	**(3)**
	Respect for Law and Public Interest (Compliance with Law; Transparency & accountability of actions)	Collection controls (operational (access type), data (filtering, anon), legal/policy agreements))	
		Data Protection	
		Stakeholder consent	
		Legal Exception	

CREDS Tool — Operationalizing Ethics



CREDS Tool — Ethics Risk Heat Map

CREDS (Cyber-risk Research Ethics Decision Support) Tool

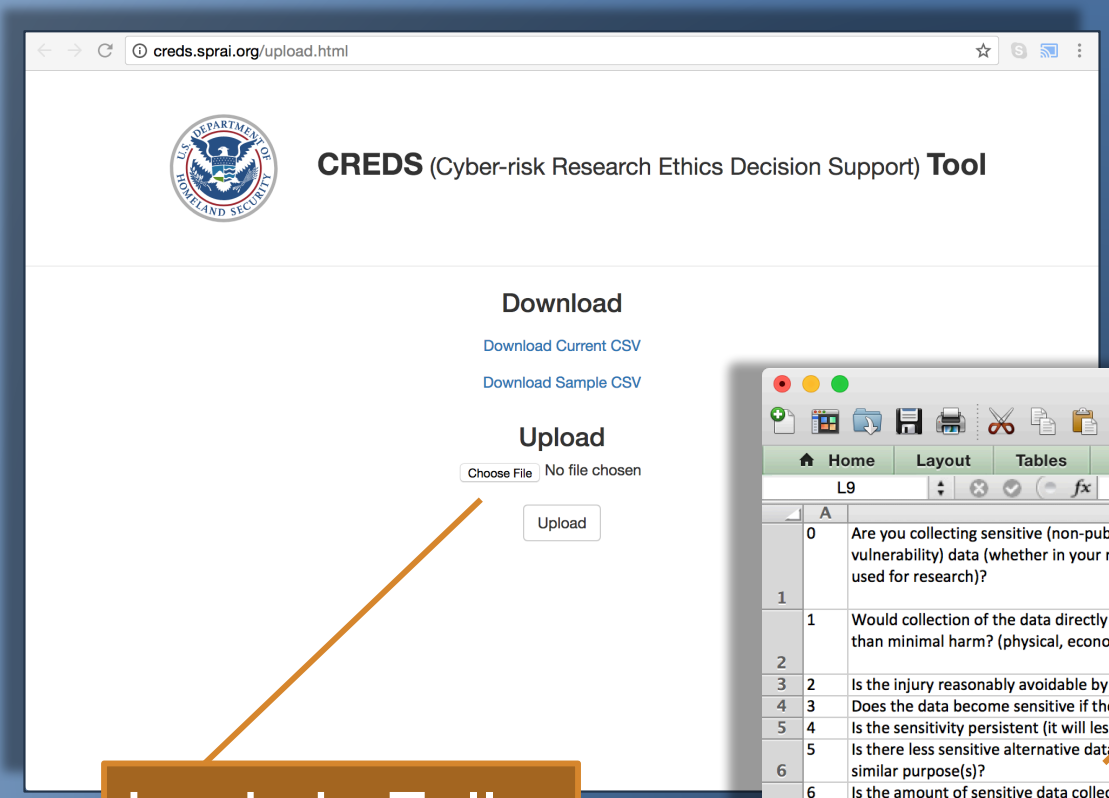
Heatmap

Results Summary

	Data	Resource	Data Provider/Recipient	Purpose	Mitigation
Collection	4 / 9	1 / 2	1 / 2	1 / 3	3 / 5
Use	1 / 2	3 / 5	2 / 3	4 / 8	5 / 9
Disclosure	7 / 13	1 / 1	2 / 4	0 / 0	0 / 0

Detailed Q&A Breakdown

Lifecycle	Risk Factor	Question	Response
Collection	Data	Are you collecting sensitive vulnerability) data (whether i data used for research)?	No
		Does the data become sensitive if the quantity of data collected is increased?	Yes
		Is the sensitivity persistent (it will lessen/expire with time)?	No



User Defines a Configurable Decision Tree

Logic is Fully Customizable By User

	A	B	C	D	E	F	G	H	I	J	K
1	0	Are you collecting sensitive (non-public, identifiable; confidential, vulnerability) data (whether in your research results or otherwise the raw data used for research)?	Is it reasonably likely that the data could be used alone or in combination with other reasonably available data, by Researcher/You, to identify a living person or discern confidential information?	1	1	1	1	Yes	1	No	3
2	1	Would collection of the data directly or indirectly expose an individual to more than minimal harm? (physical, economic, legal, reputation, or psychological)		1	1	1	1	Yes	2	No	3
3	2	Is the injury reasonably avoidable by the affected person?		1	1	1	1	Yes	3	No	3
4	3	Does the data become sensitive if the quantity of data collected is increased?		1	1	1	1	Yes	4	No	4
5	4	Is the sensitivity persistent (it will lessen/expire with time)?		1	1	1	1	Yes	5	No	5
6	5	Is there less sensitive alternative data available that would serve substantially similar purpose(s)?		1	1	1	1	Yes	6	No	6
7	6	Is the amount of sensitive data collected proportional to what is necessary for research?		1	1	1	1	Yes	7	No	7
8	7	Is the purpose of the collection a legitimate research/educational use or commercial?		1	1	1	1	Yes	8	No	8
9	8	Does any federal/state privacy or data protection law/regulation restrict/prohibit collection of all/part of the data by the Researcher?		1	1	1	1	Yes	9	No	10
10	9	If so, Is there an exception to the law/reg that allows the Researcher to collect?		1	1	1	1	Yes	10	No	10
11	10	Does any law/reg allow/prescribe collection of all/part of the data by the Researcher?		1	1	1	1	Yes	11	No	11
12	11	Is there an intellectual property right/interest in the data that would be infringed or waived if the data is disclosed?	Does it contain patentable content that is intended to be filed or published but has not? Does it expose information that would otherwise be protected by trade secret? Does it expose information that would negatively impact the Researcher's interests viz. competitive position, reputation, etc.? Are the data subject to a restrictive license? Do the data relate to an invention report the Researcher has or intends to file with its organization?	1	1	1	1	Yes	12	No	12
13	12	Was data collection dependent on using sensitive information about a research platform or system (e.g., security or proprietary -related aspects of a network or system)?		1	1	2	1	Yes	13	No	13
14	13	Did you have permission to collect the data from the platform it was collected?		1	1	2	1	Yes	14	No	14
15	14	Did Researcher collect the data?		1	1	3	1	Yes	21	No	15
16	15	Does another entity(s) have ownership rights/interests in the data?		1	1	3	1	Yes	16	No	16
17	16	Does the source collector/owner allow research use of the data		1	1	3	1	Yes	17	No	17
18	17	Does the source collector/owner of the data allow further disclosure of the data (via a release statement or otherwise)?		1	1	3	1	Yes	18	No	18

creds (2).csv
Normal View
Ready
Sum=0

Ethics and Art Into Science- Parting Thoughts

- **We have Technical and Legal Models … It's ~about Priorities:**
  - Developers: no lack of attempts to create shared & connected experiences (nike app), auto decisions (situation awareness)
    - Is notice/consent a shared experience to be similarly engineered ?
    - /eg/ POS, within set-up wizards, privacy dashboard, UMA
  - Consumer-Users : measure-everything world (speed, retweets, mentions, likes)?
    - Do we need distractions in record time?

- **"Post-Truth"**
  - Oxford Dictionary 2016 WOTY:

'relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief'.

- **(1) Principles are drivers**
- **(2) People will ignore facts when they lack Trust**

Help Us Help You:

- **CREDS Alpha: http://creds.sprai.org/**
  (NOTE: real Alpha will be usable in February)
- **CREDS T&E page: http://credststst.sprai.org/**
- **GitHub Code & Issue Tracker (tag me for access):**
  https://github.com/teamnsrg/creds/tree/master/
- **All sorts of links and info on the project:**
  https://www.impactcybertrust.org/ethos

- **Erin.Kenneally@HQ.DHS.gov**
- **erin@elchemy.org**