THE CENTER FOR ADVANCED STUDIES
IN SCIENCE AND TECHNOLOGY POLICY

Security and Anonymity: Rethinking the Problem Statement

Kim Taipale
Executive Director
Center for Advanced Studies

Presented at
**"In Search of J. Doe:
Can Anonymity Survive in Post-9/11 Society?"**
Woodrow Wilson Center (WWICS/AAAS/ABA)
Washington, DC May 4, 2004

Anonymity: what is it good for?

- "A White House official who spoke only on the condition of *anonymity* described Clarke's public remarks as *self-serving and politically motivated*."

  (The Washington Post, March 24, 2004)

- Context is everything (H. Nissenbaum)

Obligatory self-promotion

Technology, Security and Privacy:
The Fear of Frankenstein, the Myth of Privacy
and the Lessons of King Ludd
Yale Law School CyberCrime Conference Paper (March 2004)

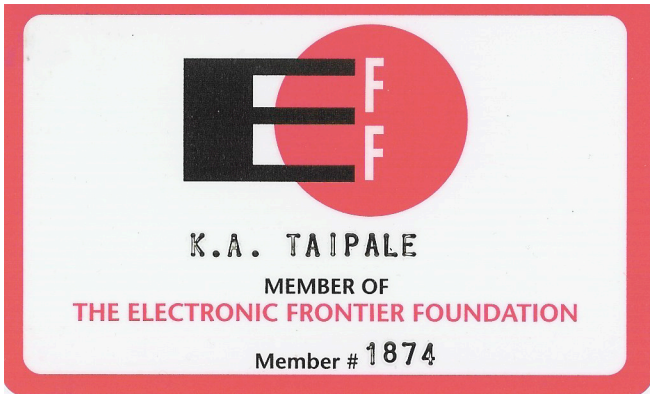
<http://www.taipale.org/papers/TSP-YLS.htm>

Data Mining and Domestic Security:
Connecting the Dots to Make Sense of Data
5 Columbia Sci. & Tech. L. Rev. 2 (December 2003)

<http://www.stlr.org/cite.cgi?volume=5&article=2>

www.advancedstudies.org

“Preemptive Defense”



(c. 1993 pre-Mosaic 1.0 release)

More self-promotion

"Identification and Domestic Security: Who's Who in Whoville"

A work in progress (Spring - Summer 2004)

Today's Discussion: Identification ≈ Anonymity

Today's presentation

- *Anonymity*
  - Identification as a privacy interest
- *Security*
  - Identification as a security interest
- *Data attribution*
  - Identification as a technical issue
- Reconciliation? (Dual obligations)

Preface: physics and metaphysics

- A new "physics" - changing base conditions - control revolution (Beniger)
- But an old metaphysical problem - the relationship between the collective interest and the individual
- *Relation of competition* between *individual rights* and *popular sovereignty* (Kant, Rousseau)
- ['balancing' individual accountability and state accountability]

The nature of the problem: A wicked problem not a balancing act

- "Wicked" Problems -- common in public policy
  - no *correct* solution, reveal additional *complexity* with each attempt at resolution, have i*nfinite outcomes* and *no stopping rule* (process ends when you run out of resources) and occur in a *social context* -- the wickedness of the problem reflects the diversity among the stakeholders in the problem
  - Resolution requires discourse (consensual compromise)

- Security and privacy also present a *measurement problem*
  - There will never be a correct amount of security or privacy only *enough* or *not enough* to satisfy certain constituencies in a particular context (~ Nissenbaum) (cf. Etzioni)
  - Cf. the metaphor of balance - "weighing" one against the other (move the fulcrum to some optimal point of balance)

Introduction: changing base conditions

- "Information Society" -- cliché but fact
  - *Digital mediation* affects the five things that gives ideas value in human activity -- their production and reproduction, storage, transmission, selection, and intelligent processing
- Creates new culture of *time* and *space* (S. Kern - modernity)
  - Data no longer transient (always available)
  - Data is proximate (available anywhere)
- Thus, the end to "practical obscurity" of data by virtue of its physical location and the end to anonymity through data transience
- The concept of privacy and privacy policy is driven by technological intrusion (collection, identification, aggregation/analysis technologies)

New information economics

- The cost of *data retention* is less than the cost of *selective deletion* (email example)
- The cost of indiscriminate *data collection* is less than the cost of *selective acquisition*
- Thus, data largely "exists" and question is under what circumstances can it be *accessed and used* (privacy) and under what circumstances can it be *attributed* (anon.) (permissibility of intrusion at point of mediation)

"Surveillance" economics

- High collective expectation of privacy
  - cannot watch *everybody* (~ Froomkin - "ocean")
- Low individual expectation of privacy
  - can watch *anybody*
- *Selective attention* ("cost" of focus to society and individual) (~ anonymity) (~ Brin) ("tools to allocate resources")
- Technical means are capital intensive not labor intensive, thus *cost per unit of information* have/will decrease
- E.g., FBI's Carnivore (analytic filter a priori) vs. NSA's Echelon (vacuum, and analyze post hoc) (logging?)

So, what is privacy?

- *Secrecy* - keep data unknown
  (~ hide the footprint, or text) (McNeely 1999)
- *Anonymity* - keep data unattributed
  (~ don't reveal shoe purchase, or author)
- *Autonomy* - keep data from constraining opportunity
  (~ exclusionary rule - don't allow shoe purchase to be used as evidence unless due process procedures were followed) ("prior restraint")
- Cf. *Whalen v. Roe*, 429 US 589, 599 n. 24 (1977)

"Privacy" vs. anonymity

- Privacy (~ secrecy)
  - "right to be left alone" (keep data private)
  - withdraw from society, not be intruded on
  - property right? Warren & Brandeis, "Right to Privacy" (1890)
  - public (4thA); private (protect against others but alienable)
- Anonymity (~ protects autonomy)
  - obscure identity (keep data un-attributed)
  - allows participation without repercussions (or accountability)
  - speech right?
  - protect with strict scrutiny and due process (1stA) McIntyre et al.

Identification ≈ Anonymity

- Identification
  - Confidence that some information (identifier, identity or attribute) relates to a specific individual
  - Impacts anonymity but not privacy (w/out intrusion)
- Cf. Surveillance
  - Observation of activity
  - Impacts privacy but not anonymity (w/out attribution)
- ~ Accountability?

A brief history of privacy and anonymity

- In classical time privacy was seen as a negative
  - Greeks
    - *Demios* (public) "having to do with people"
    - *Idiotes* (private person, someone not engaged in public life)
  - Romans
    - *Publicus* (public) "that which belongs to the people as a whole"
    - *Privatus* "withdrawn from public life", and *privare* "separate or peculiar" (deprive)
  - Political philosophy of "republicanism" favors transparency -- in the public forum there is no place for anonymity (Mark Poster "The Net as a Public Sphere")

A brief history continued

- *republicanism* as political philosophy
  - Cicero, Locke, Hobbes, Machiavelli, Montesquieu, Rousseau, Kant, Arendt, Habermas, Etzioni (communitarianism), etc.
  - de Tocqueville: "individualism" results in the privatization of social life to the detriment of public life (~ Cass Sunstein, *republic.com*)
  - Kant: using lies is unacceptable, anonymity evades responsibility and disrespects the other
- Utilitarian arguments against anonymity:
  - socially inefficient (Bentham, Beccaria) (crime/discovery)
  - economically inefficient (Posner) (concealment/fraud)

"Modern" notion of privacy and anonymity

- Enlightenment, liberalism and the [French] Revolution
- Benjamin Constant (1767-1830) [I. Berlin]
  - Modern state as potential menace to individual liberty
  - Distinguished between "liberty of the ancients" based on active participation in the collective power (public life), and individual liberty and independence in the large, modern state guaranteed by political rights

Modernity continued

- Distinguished also between "*withdrawal*" (~ privacy) (protected but not privileged?) and "*obscurity*" (~ anonymity) (central to liberal freedom)
- "*Obscurity*" is the right to not be targeted individually by surveillance (~ avoid selective attention)
- Nevertheless, he recognized the obligation of the state to surveil those presenting a risk to society -- as long as there was no physical interference and the surveillance was not "felt" by the person being watched until there was the indication or beginning of a crime. (~ preemption) (~ data processing)
- Also, recognized "publicity" as positive social control, surveillance by public of state actions and new ideas (no intrusion on the private, complete transparency for the public) (~ D. Brin)

Note the changing base conditions at the time

- The Terror and the emergence of the *surveillance state* and *professional police* (spies, police agents, house numbering) (~ Beccaria)
- *Commerce* gave the private sphere of activity substance for large segments of the population (public life was the norm of the ancients, private the norm of the more modern) (~ *Idiotes)*
- The rise of the *bourgeoisie* (~ public sphere) and the *commodification of relationships* (Marx)

And now

- *Technical mechanisms of control* and balance of power between public and private (Foucault) (privatization of traditional state *data* functions)
- *Opportunity* (and duty?) for participation (and symmetric accountability) in public life (~ Brin) (~ Etzioni)
- Widening relationship between *individual rights* and *collective power* (trust vs. responsibility in a specialized national security state)
- *Changed consequences* (of a bias toward false positive or false negative)
- *Force multiplier* effect of technology: as seed value approaches the individual actor, controls trend towards impossibility and risk increases geometrically. (~ shrinking perimeter of defense)

The emerging personalized information dystopia

- Private sector: perfect personalization
  - "autonomy trap" (errors? or truth?) ("fit", G. Marx)
  - Price discrimination
  - Efficient personal service ≡ tyranny
- Public sector: perfect law enforcement
  - Existing system premised on slippage (see Froomkin) (over-criminalization and deterrence) (see Rosenzweig)
  - Automated system becomes a personalized "tax" system
  - Efficient government services ≡ tyranny
- "universal accountability"
  - control based on technical means vs. consensual commitment to the rules (Ellul) (G. Marx) (Xxxxxxx)

Security Strategies

- Access control
  - Default = "deny without permission"
  - Low cost of implementation, high cost to functionality (or freedom)
  - ~ totalitarianism
- Accountability
  - Default = "permit with accountability"
  - Lower cost to functionality (or freedom), potentially high cost to security
  - ~ liberal democratic freedom
- [~ eliminate preconditions and harden targets]

Access control strategies - confirm authorization (binary)

- May or may not require identification
  - E.g. compare airport search (violate physical dignity/privacy to disarm) vs. CAPPS (violate information privacy to establish trust)
- Requires authentication of "trust" attribute (that is, negative or positive authorization to do something) (~ "reputation")
- Raises "trusted systems" problem
  - Can never prove that you are trustworthy, only that you are not yet identified as un-trustworthy (e.g., not on watch list)
- High cost to functionality (or freedom) (& doesn't scale)

Accountability strategies - confirm adherence to rules (variable)

- Generally requires some form of identification, authentication or traceability (~ vetting for pre-acc.)
- Anonymity vs. pseudonymity (cf. token)
  - Anonymity - data/activity cannot be attributed
  - Pseudonymity -- data/activity cannot be attributed *in the ordinary course* (~ control through due process)
- Surveillance and accountability
  - Overt surveillance -- preempt/chill (Panopticon, 1984, beat cop)
  - Covert surveillance -- defend/accountability (Constant)
  - Data analysis -- non-selective processing (DMDS)

Negative impact of accountability

- "Chilling effects" on culture of freedom
  - Inhibits exercise of protected rights
  - "potential knowledge is present power" because "people act differently if they know their conduct *could* be observed" (TAPAC, 2004) (~ Constant)
  - Contextual to social structure and efficacy (~ trust)
  - "When is accountability a bad thing?" (Xxxxxxx)

Chilling effects as noise abatement

- Bateson rule and free speech
  - A systems stability is related to noise and available bandwidth (when does noise interfere with signal?) Implement noise controls.
  - But, "all that is not information, not redundancy, not form and not restraint [i.e, not orthodoxy] -- is noise, the only possible source of *new* patterns."
  - Fundamental problem is distinguishing "good" noise ("new signal" or learning) from "bad" noise (= or ≠ speech)
  - Can we subject noise control to free institutional constraints? ("Squelch" control) (how and who) ("lesser evils" NYT)

Security problem in counterterrorism

- Political requirement for *preemptive* action, not *reactive* law enforcement (political and social consequences of security breach potentially destabilizing) (~ probability neglect?)
- Obvious insufficiencies of after-the-fact accountability to control suicide attackers
- "In-liers" (cf. outliers or deviants) (common attributes vs. shared attributes)
- [Privacy problem in counterterrorism: next event leads to martial law (T. Franks)]

Intelligence vs. law enforcement

- Intelligence is a *preemptive* strategy based on probabilities and disruption (act on suspicion) (system bias to eliminate false negatives)
- Law enforcement is a *reactive* strategy based on evidence and conviction (act on proof beyond reasonable doubt) (system bias to eliminate false positives)
- Cf. community policing (responsibility vs. trust)

NB: in counterterrorism the network itself is the problem

- Social network theory and social engineering
- Communication network itself has allowed a "critical mass" of malicious actors to act in concert (or at least mutually reinforce) (al-Q as organization vs. movement)
- And, technology acts as force multiplier
- Thus, identifying and disrupting these sub-networks is the key (disrupt the paths of infection) (cf., reactive LE)
- *Immunization strategy* vs. cure (intel vs. LE) (prior restraint)

Thus, the ideological divide

- Is anonymity the line between:
  - *freedom and totalitarianism* (requiring absolute secrecy of data for its own sake) (Rotenberg, Steinhardt), or
  - *freedom and anarchy* (based on accountability under constitutionally recognized due processes in which autonomy is protected through selective revelation of identity subject to defined constraints and controls) (Constitutional law)?

Privacy lobby has a fetish for secrecy

- Premised on an unchallenged assumption of a constitutional right to *absolute anonymity* rather than examining whether any particular intrusion is a permissible burden under strict scrutiny (all noise is good, all suppression is bad) (Rotenberg, Steinhardt, etc.) (impossibly high standard for technology and not achieved or desired in the real world)
- Nevertheless, the privacy lobby (like the NRA with gun control) must defend an absolute position for institutional reasons (slippery slope, fear and fundraising, raison d'etre)
- Fail to distinguish between communication, transaction and record anonymity, and won't recognize alienability of privacy (Gmail?)
- Is government the greater evil? Even if so, do we leash it or blind it?

Constitutional law

- "Anonymity" (or forced identity) cases
  (aren't these really pseudonymity cases?)
  - McIntyre, Talley, ACLF v. Buckley, Village of Stratton, etc.
  - Strict scrutiny - is identification necessary to achieve a compelling state interest (no less intrusive alternative)
  - All these cases recognized compelling interests but suggested alternative accountability strategies under old physics (~ "uncontrolled leakage" of identification data is a condition of physical encounters, G. Marx)
  - Cf. Hiibel v. Nevada (give name during *Terry* stop?) and Gilmore v. Ashcroft (ID to travel on commercial airline?)

Constitutional Law II

- "Clear and present danger" cases
  - Schenck, Debs, Abrams, Brandenburg
  - Abrams dissent: "silly leaflet by an unknown man … poor and *puny* anonymities" (Holmes)
  - Brandenburg opinion: "the threats were often loud, but always *puny*" (emphasis added) (Douglas)
  - "imminent lawless action" (wrong physics?) (see CyberSemiosis)

Rethinking current base conditions

- Changing nature of compelling state interest and the balance of power
  - No longer "I am weak, the state is strong"?
  - Asymmetric threats no longer "puny"?
  - Force multiplier effect of technology (seed input)
- Changing nature (and availability) of alternative strategies
  - Technology enables "true" anonymity? (cryptography)
  - "No court order can break strong encryption"
  - Allow vs. insist on anonymity?
- Changing nature of the consequences of false positives and false negatives (thus, rethink systems bias strategies)

Defensible perimeter, a shifting paradigm

- Dan Geer - Computer and system security
  - As the risk increases the defensible perimeter contracts
  - Thus, in cyber security:
    - risk WAS GOING TO BE system-mediated => AOL (closed) vs. Internet (open)
    - risk WAS trust-mediated => the firewall
    - risk IS application-mediated => the code scanner
    - risk WILL BE data-mediated => tracking & synchronization
- ~ Counter-terrorism security
  - line at the border (system)
  - line at the airport (application)
  - line in the data -- (individual)

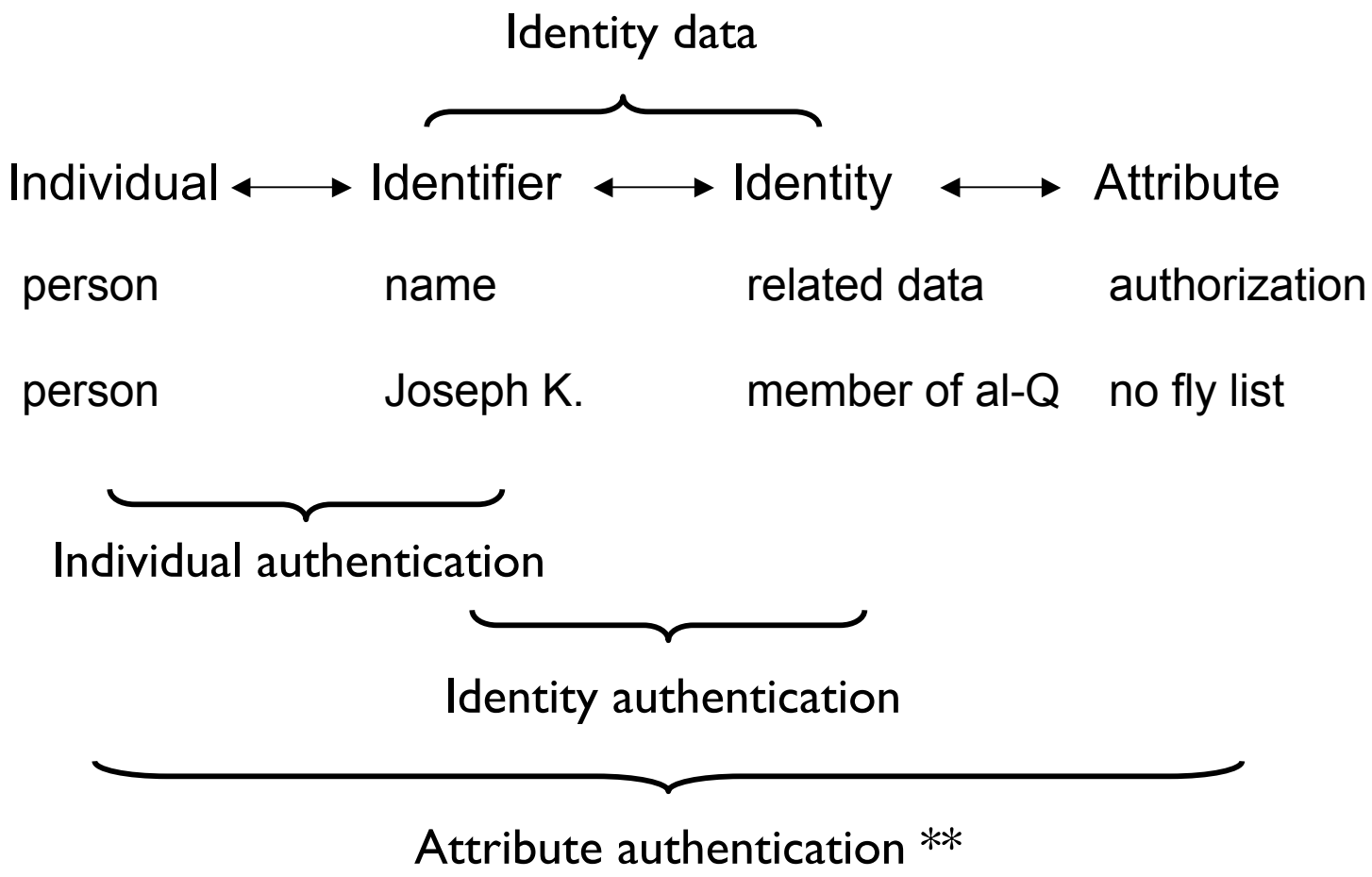
Data attribution
(identification as a technical issue)

- Three types of attribution (NAS)
  - Individual authentication (~identification)
    - *Confidence* that an identifier refers to a specific individual
  - Identity authentication (~indexing)
    - *Confidence* that an identifier refers to an identity
  - Attribute authentication (~authorization) **
    - *Confidence* that an attribute applies to a specific individual
    - ~ "reputation" attributes

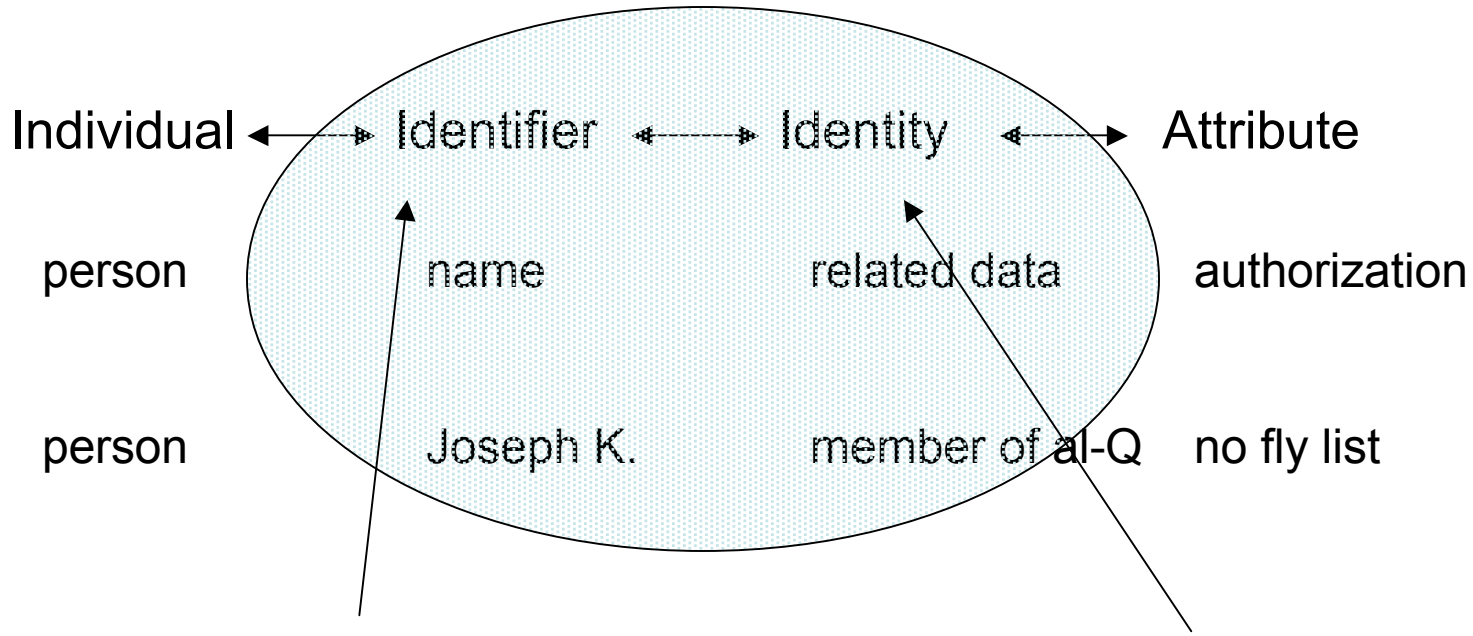
Using seven types of identity data

- legal identifier (name, SS#, drivers license #, etc.)
- traceable pseudonyms (trace/track, persistent/temp)
- untraceable pseudonyms (~ anonymity) (track?)
- address (place or node)
- patterns (data mining) (traffic analysis)
- social categorization (~ address)
- authorizing tokens (verified ID as hall pass)

Data attribution



Data Attribution - additional problems



Multiple Identifiers:
Name, SS#, Driv Lic #, arbitrary, etc.
As well as aliases

Multiple Identities:
Professional, family, social, etc.

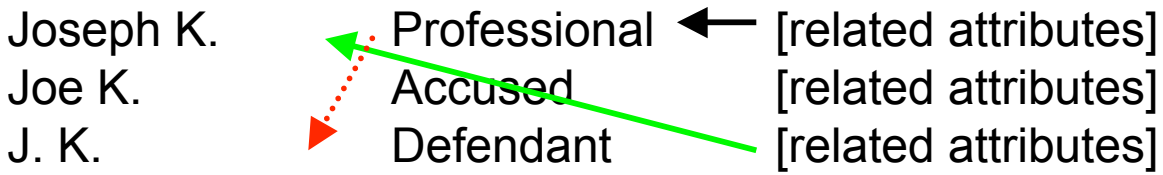
Entity resolution

Individual ⟷ Identifier ⟷ Identity ⟷ Attribute

Same individual	Joseph K.	Bank Clerk	[related attributes]
	Joe K.	Grubach's tenant	[related attributes]
	J. K.	Leni's lover	[related attributes]
Same place	123 Main Street	Postal address	[related attributes]
	Main and Broad	Intersection	[related attributes]
	Courthouse	Functional	[related attributes]

Attribute resolution

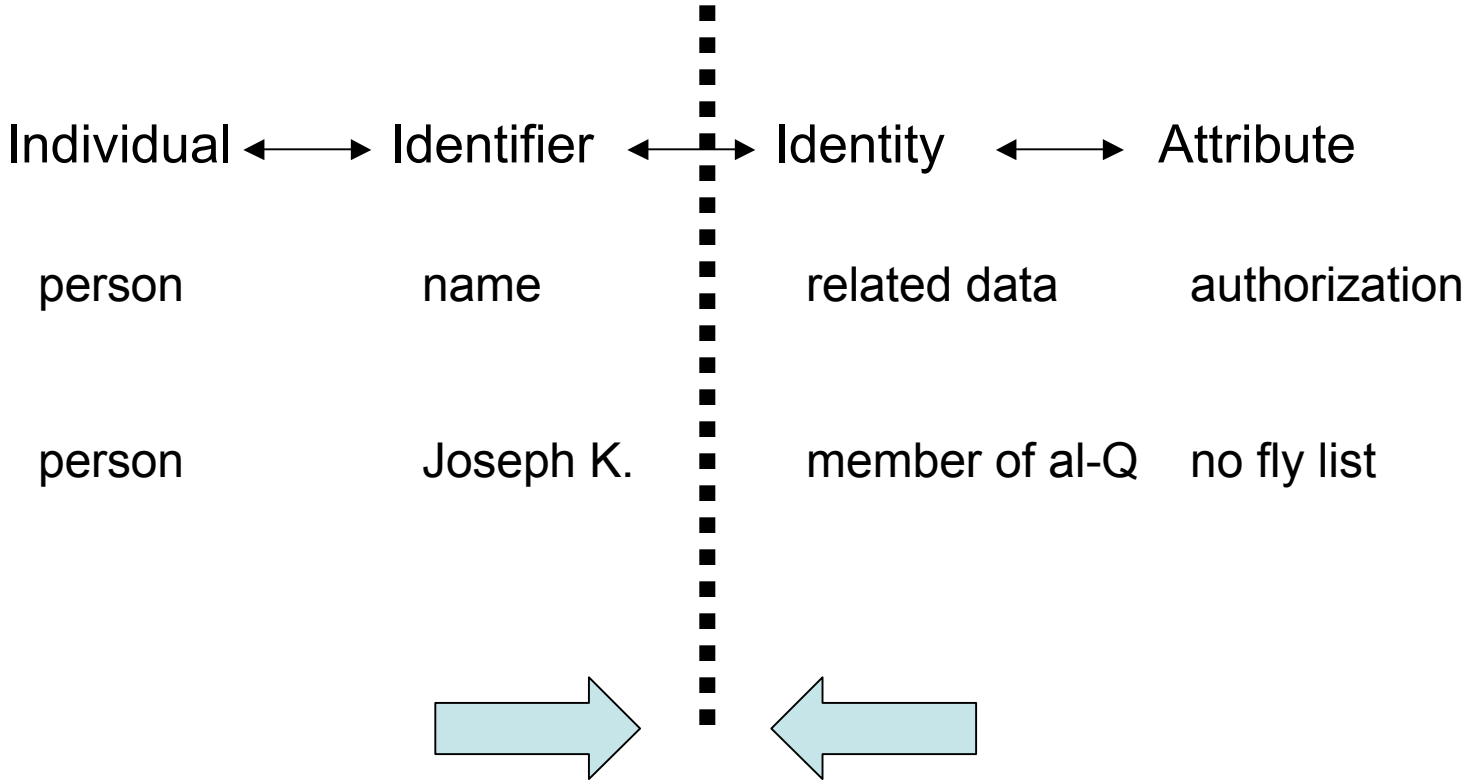
Individual ⟷ Identifier ⟷ Identity ⟷ Attribute



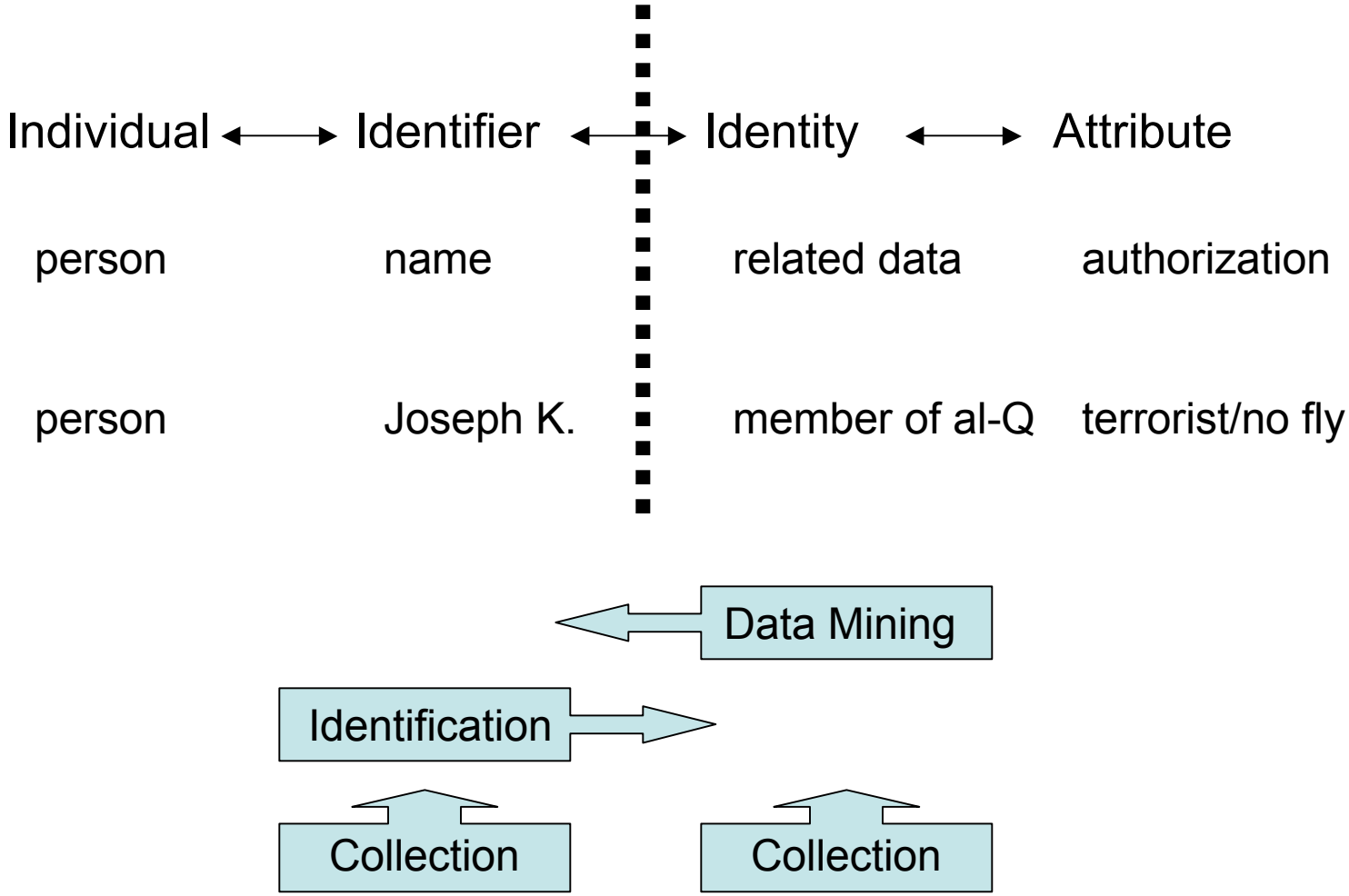
Attributes can ‘belong’ to the individual, to the identifier or to the identity and can be time, place, manner dependant

[on-duty/off-duty, apparent authority, tokens, etc.]

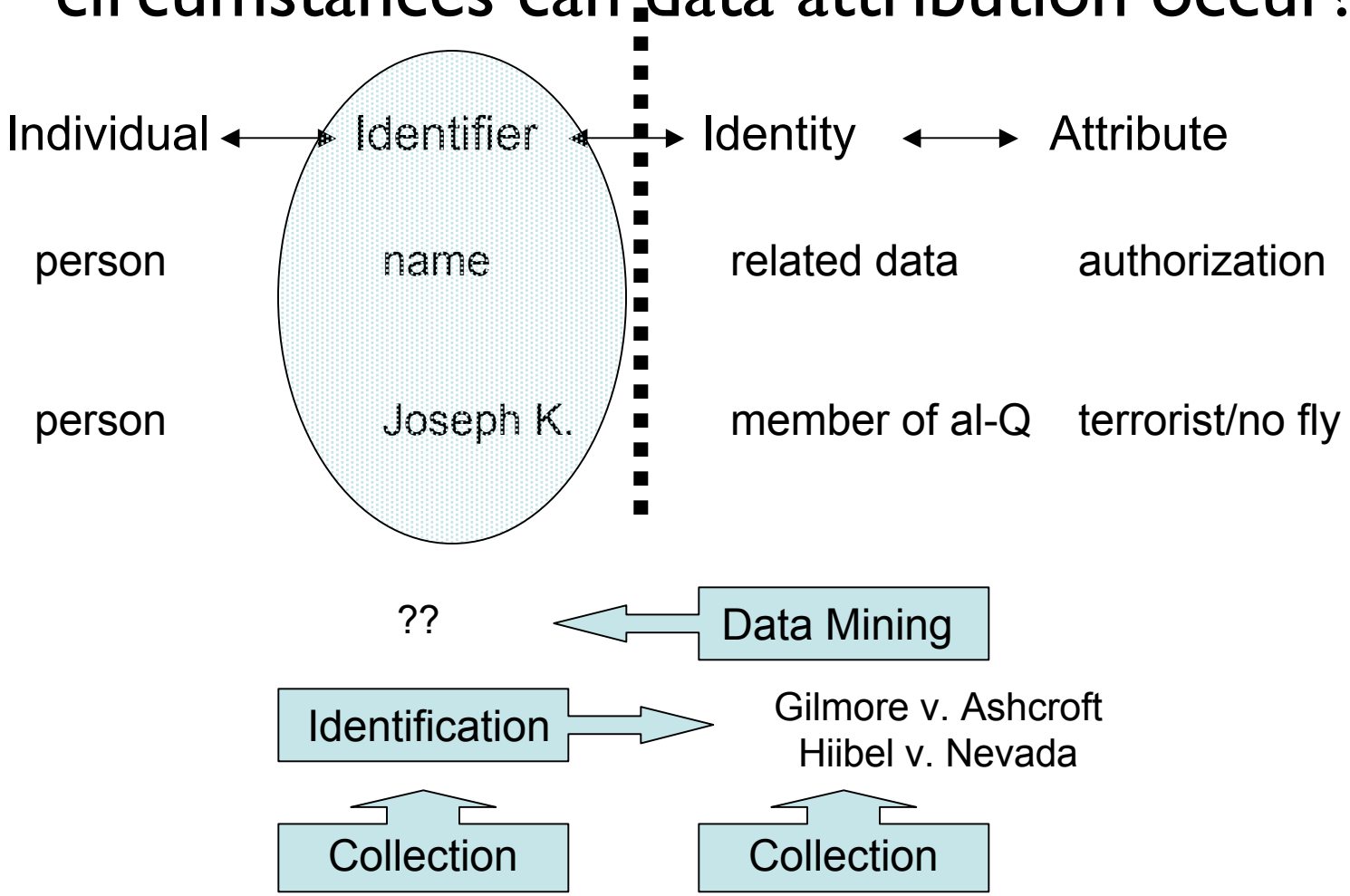
Data Attribution and the Privacy Divide



Intrusive Technologies



Policy/legal Intervention: Under what circumstances can data attribution occur?



Hiibel, et al., may not matter

- Bentham (1788)
  - "a new nomenclature … a proper name borne only by himself"
- Froomkin (2004, yet so 20th C.)
  - *"The Uneasy Case for National ID Cards"*
  - Social security numbers, drivers license, etc.
- J. Jonas (2004 and beyond)
  - Entity resolution "a solved problem"
  - Identity is discernable from data analysis
- Emerging ID technologies (DNA sniffers, ID at a distance, gait recognition, facial recognition, etc.)

Documentation of identity and the anonymity continuum

- G. Marx (B. Traven, *The Death Ship)*

  "You ought to have some papers to show who you are."
  "I do not need any papers. I know who I am."
  "Maybe so. But others are also interested in who you are."

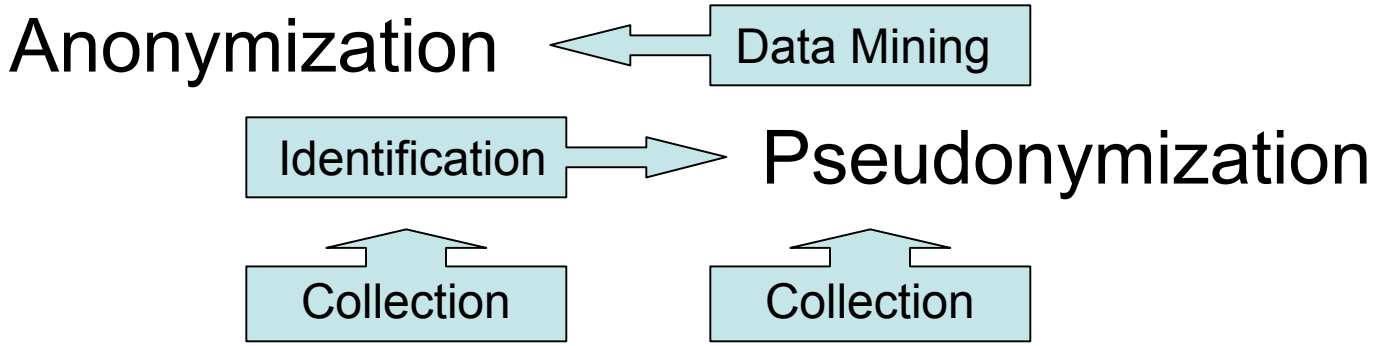
- Brazil (Gilliam, 1985)

  "Do you want to see my papers?"
  "No need, sir"
  "But I could be anyone."
  "No you couldn't, sir, this is information retrieval."

- Anonymous donors (predictable patterns, G. Marx)

Technology intervention to preserve 'anonymity'

Individual	Identifier	Identity	Attribute
person	name	related data	authorization
person	Joseph K.	member of al-Q	terrorist/no fly



Anonymity and data analysis

- Use anonymizing technologies to allow for non-attributable data processing
- Share and match anonymized data
- Selective revelation on increasing predicate
- Retain data control with original party
- Cf. data hashing with key encryption

Pseudonymity and identification

- Use technology to match the data demand to the transaction requirement (smart card with encryption and segmentation)
- That is, *only reveal attributes relevant to the particular authentication required to complete the specific transaction* (Lessig "certification")
  - Traffic stop -- authorized to drive
  - Commercial transaction -- authorized credit
  - Neither transaction requires transfer of "identity"
  - *Verification with 'anonymity'*
- Use SALT (shared key) to control/limit search (DB, time period, etc.)
- Persistent (alias/nym) vs. temporary; trackable vs. traceable (cookies)

Provide protective mechanisms for favored applications

- Anonymous remailers (stripping, chaining, encrypting)
- Web surfing anonymizers (proxies and firewalls)
- Based on "escrowing" identity at point of mediation
- Protect specific *activity* with additional statutory protection (e.g., whistleblower, medical, etc.)
  (cf. *data*: video/cable records, tax, HIPAA, etc.)
- Encryption (but KSL, Magic Lantern) (biometric)

Another caveat: traffic analysis

- Communication patterns themselves reveal significant evidence of organization
- And "chatter" in known networks may mean activity
- Social network analysis (power laws)
  ID leadership, organization
- Don't need access to "content" (wiretap v. pen/tap)
- Don't need "identification" (dataveillance)

Technical features to enable due process interventions (see DMDS)

- Database architecture (Markle second report, 2003)
  - Centralized (warehouse) v. distributed architecture
- Rule-based Processing ([IAIT 2003])
  - DRM; intelligent agents and "smart data" (labeling, etc.)
- Selective Revelation (Anonymization/Pseudonymization)
  - Due process intervention (build in institutional resistance and checks and balances) (Rosenzweig)
- Authentication and Audit ("watch the watchers") (~ Brin)
  - Control abuse/misuse (custody of logs as policy issue)

Development drivers

- Privacy as competitive advantage (?)
  - consumer demand vs. marketing demand
  - "one man's pirate is another man's broadband customer"
- Requirement for anonymous data sharing
  - Public sector: sources and methods, and liability
  - Private sector: trade secrets, competitive advantage and liability
  - Shared infrastructure, federated identity
- Online voting requires pseudonymity

Conclusion

"Must a government, of necessity, be too strong for the liberties of its own people, or too weak to maintain its' existence?" A. Lincoln

Security and Privacy are dual obligations