data-ppf.github.io tue mar 26, 2019

lecture 9 of 14: 'data are people too' or 'big data, old school'

chris wiggins + matt jones, columbia

themes for today

themes for today

1. tech change: data and information (last week: automating intelligence)

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy
   - (cf., Hannah Wallach in 1st week: big data is scary when it's

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy
   - (cf., Hannah Wallach in 1st week: big data is scary when it's
     - granular and

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy
   - (cf., Hannah Wallach in 1st week: big data is scary when it's
     - granular and
     - personal)

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy
   - (cf., Hannah Wallach in 1st week: big data is scary when it's
     - granular and
     - personal)
3. response: first appearance in this class of *people power* around data

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy
   - (cf., Hannah Wallach in 1st week: big data is scary when it's
     - granular and
     - personal)
3. response: first appearance in this class of *people power* around data
   - (as opposed to state + corporate power)

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy
   - (cf., Hannah Wallach in 1st week: big data is scary when it's
     - granular and
     - personal)
3. response: first appearance in this class of *people power* around data
   - (as opposed to state + corporate power)

- inconvenient truths this week: big data (still) means *labor*:

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy
   - (cf., Hannah Wallach in 1st week: big data is scary when it's
     - granular and
     - personal)
3. response: first appearance in this class of *people power* around data
   - (as opposed to state + corporate power)

- inconvenient truths this week: big data (still) means *labor*:
  - hidden, gendered (cf. Poon), "otherized" (cf. Radin) and repeatedly: devalued

themes for today

1. tech change: data and information (last week: automating intelligence)
2. immediate concern: privacy
   - (cf., Hannah Wallach in 1st week: big data is scary when it's
     - granular and
     - personal)
3. response: first appearance in this class of *people power* around data
   - (as opposed to state + corporate power)

- inconvenient truths this week: big data (still) means *labor*:
  - hidden, gendered (cf. Poon), "otherized" (cf. Radin) and repeatedly: devalued
  - quantization of complexity oft requires subjectivity

historical & social context

postwar rise of digital, personal data

- surveillance narratives (Huxley, Orwell. . .)

historical & social context

postwar rise of digital, personal data

- surveillance narratives (Huxley, Orwell. . .)
- computation accelerates the information deluge

historical & social context

postwar rise of digital, personal data

- surveillance narratives (Huxley, Orwell. . .)
- computation accelerates the information deluge
  - crib-searching -> Bush's Memex

historical & social context

postwar rise of digital, personal data

- surveillance narratives (Huxley, Orwell. . .)
- computation accelerates the information deluge
  - crib-searching -> Bush's Memex
- computation, data, and control: 'cybernetics' (in Miller, esp.)

historical & social context

postwar rise of digital, personal data

- surveillance narratives (Huxley, Orwell. . .)
- computation accelerates the information deluge
  - crib-searching -> Bush's Memex
- computation, data, and control: 'cybernetics' (in Miller, esp.)
  - recall the root of cyber is control $\kappa\nu\beta\epsilon\rho\nu'\eta\tau\eta\varsigma$ / kubernetes

contemporary context/modern day relevance

- privacy is now daily concern

contemporary context/modern day relevance

- privacy is now daily concern
  - personal

contemporary context/modern day relevance

- privacy is now daily concern
  - personal
  - corporate

contemporary context/modern day relevance

- privacy is now daily concern
  - personal
  - corporate
  - govt

contemporary context/modern day relevance

- privacy is now daily concern
  - personal
  - corporate
  - govt
- BI in all companies

contemporary context/modern day relevance

- privacy is now daily concern
  - personal
  - corporate
  - govt
- BI in all companies
  - maturation of dynamics of data+UI

contemporary context/modern day relevance

- privacy is now daily concern
  - personal
  - corporate
  - govt
- BI in all companies
  - maturation of dynamics of data+UI
  - hence “big data” as biz term

contemporary context/modern day relevance

- privacy is now daily concern
  - personal
  - corporate
  - govt
- BI in all companies
  - maturation of dynamics of data+UI
  - hence “big data” as biz term
- information deluge+IR

contemporary context/modern day relevance

- privacy is now daily concern
  - personal
  - corporate
  - govt
- BI in all companies
  - maturation of dynamics of data+UI
  - hence “big data” as biz term
- information deluge+IR
  - recall Google’s original mission+origin (stay tuned)

student observations

student observations

- government collection of data, suspicion, surveillance

student observations

- government collection of data, suspicion, surveillance
- rise of biz: not to win a war, nor to find truth, not to solve GAI

student observations

- government collection of data, suspicion, surveillance
- rise of biz: not to win a war, nor to find truth, not to solve GAI
- dynamics of public perception, esp. w.r.t. privacy

what are the new capabilities this week?

- information retrieval

leading to

what are the new capabilities this week?

- information retrieval
- surveillance

leading to

what are the new capabilities this week?

- information retrieval
- surveillance

leading to

- efficiency+control (Luhn)

what are the new capabilities this week?

- information retrieval
- surveillance

leading to

- efficiency+control (Luhn)
- discovery/disclosure+control (Igo/Miller)

what are the new capabilities this week?

- information retrieval
- surveillance

leading to

- efficiency+control (Luhn)
- discovery/disclosure+control (Igo/Miller)
- prediction+control (Poon)

power

- state power

power

- state power
  - to reform / encourage consumer protection

power

- state power
  - to reform / encourage consumer protection
  - (but also) to push back on consumer protection advocates

power

- state power
  - to reform / encourage consumer protection
  - (but also) to push back on consumer protection advocates
  - spycraft (e.g., as suggested by Watergate)

power

- state power
  - to reform / encourage consumer protection
  - (but also) to push back on consumer protection advocates
  - spycraft (e.g., as suggested by Watergate)
- corporate power

power

- state power
  - to reform / encourage consumer protection
  - (but also) to push back on consumer protection advocates
  - spycraft (e.g., as suggested by Watergate)
- corporate power
  - data brokers (e.g., FICO)

power

- state power
  - to reform / encourage consumer protection
  - (but also) to push back on consumer protection advocates
  - spycraft (e.g., as suggested by Watergate)
- corporate power
  - data brokers (e.g., FICO)
  - healthcare system (see tension between bioethics and funding)

power

- state power
  - to reform / encourage consumer protection
  - (but also) to push back on consumer protection advocates
  - spycraft (e.g., as suggested by Watergate)
- corporate power
  - data brokers (e.g., FICO)
  - healthcare system (see tension between bioethics and funding)
  - press

power

- state power
  - to reform / encourage consumer protection
  - (but also) to push back on consumer protection advocates
  - spycraft (e.g., as suggested by Watergate)
- corporate power
  - data brokers (e.g., FICO)
  - healthcare system (see tension between bioethics and funding)
  - press
- people power

power, reminder

- indirect edges among the 3

  privacy was, as Francis Sargent, the governor of Massachusetts, put it, "good politics." (and bipartisan!) (Igo)

readings: Luhn (58); Igo+Miller (71); Poon

Luhn

- “business”: who was his employer?

Luhn

- “business”: who was his employer?
- labor: what were his patents?

Luhn

- “business”: who was his employer?
- labor: what were his patents?
- impact: promoting BI to a function, centralizable, a skill

before Luhn, the deluge

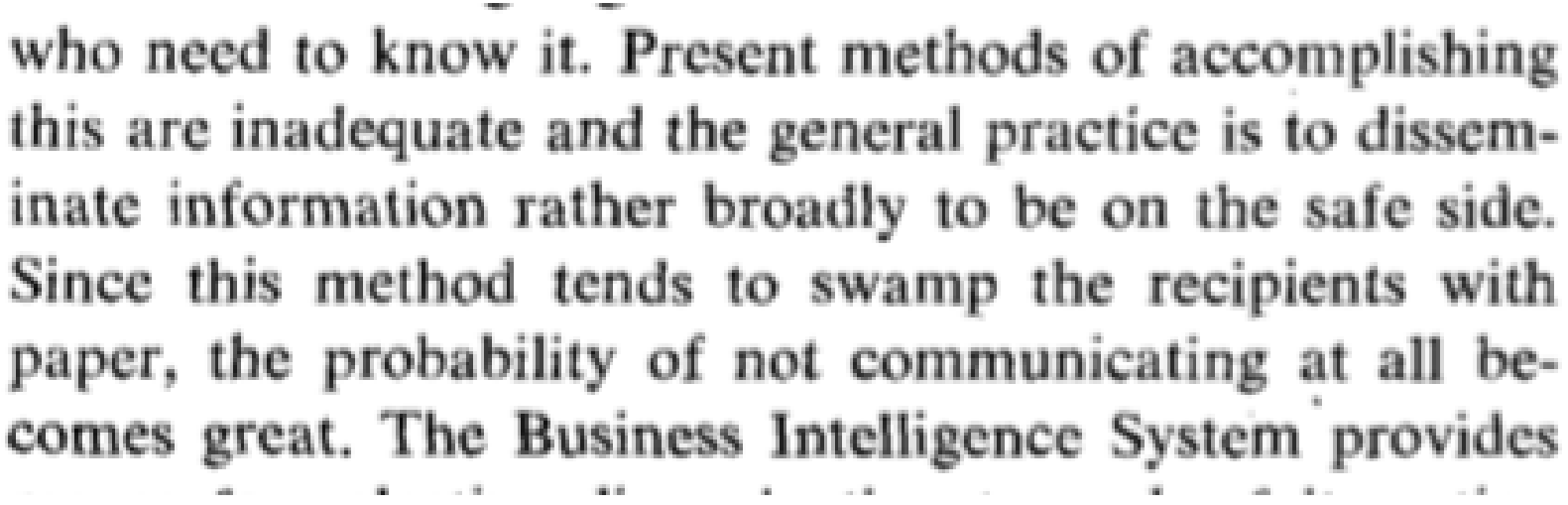

who need to know it. Present methods of accomplishing this are inadequate and the general practice is to disseminate information rather broadly to be on the safe side. Since this method tends to swamp the recipients with paper, the probability of not communicating at all becomes great. The Business Intelligence System provides

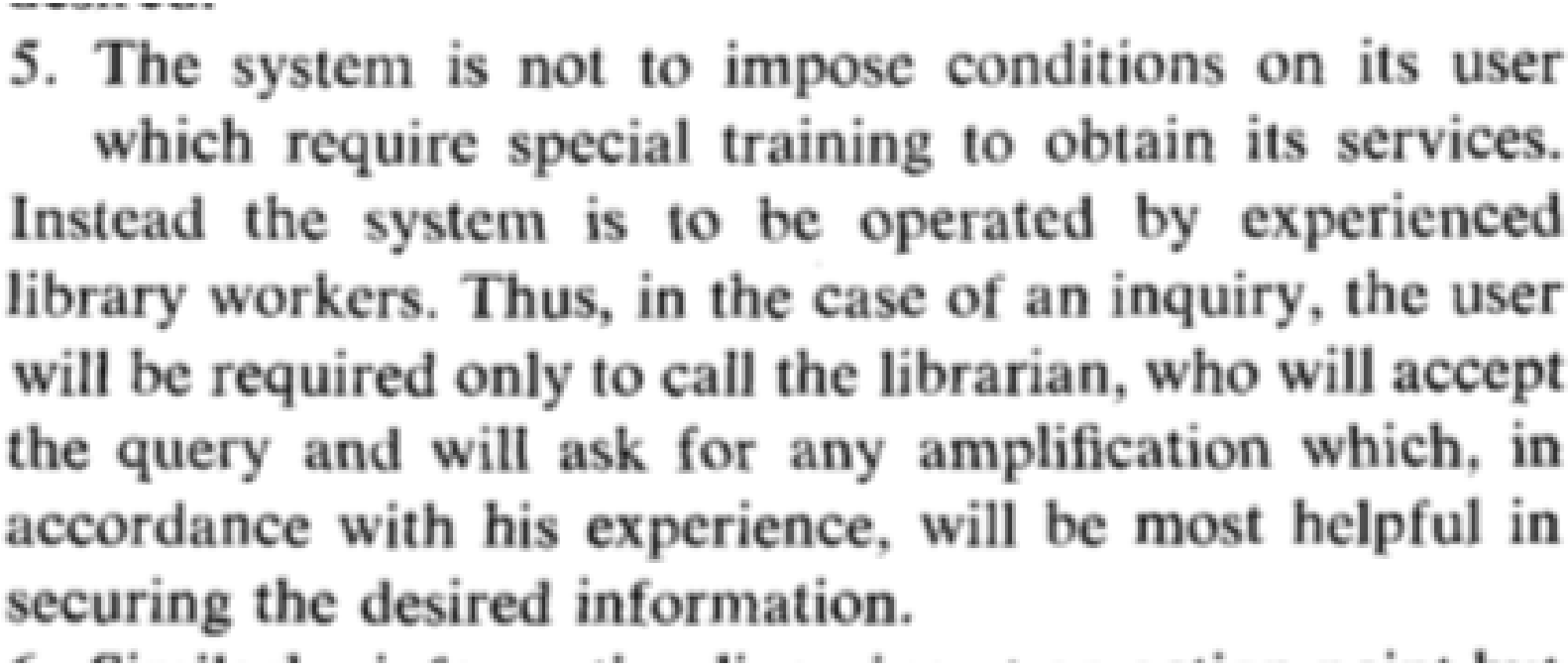
Figure 1: Human problems

infosec as afterthought

There are many details which might have to be provided to adjust the general form of the system to specific applications. One such requirement might be classification, by an editor, of documents with regard to security, proprietary interests and proper utilization of information.

Figure 2: move fast, break things

partial democratization



5. The system is not to impose conditions on its user which require special training to obtain its services. Instead the system is to be operated by experienced library workers. Thus, in the case of an inquiry, the user will be required only to call the librarian, who will accept the query and will ask for any amplification which, in accordance with his experience, will be most helpful in securing the desired information.

Figure 3: “librarian” but “not to impose. . . special training”

Luhn the inventor + tinkerer

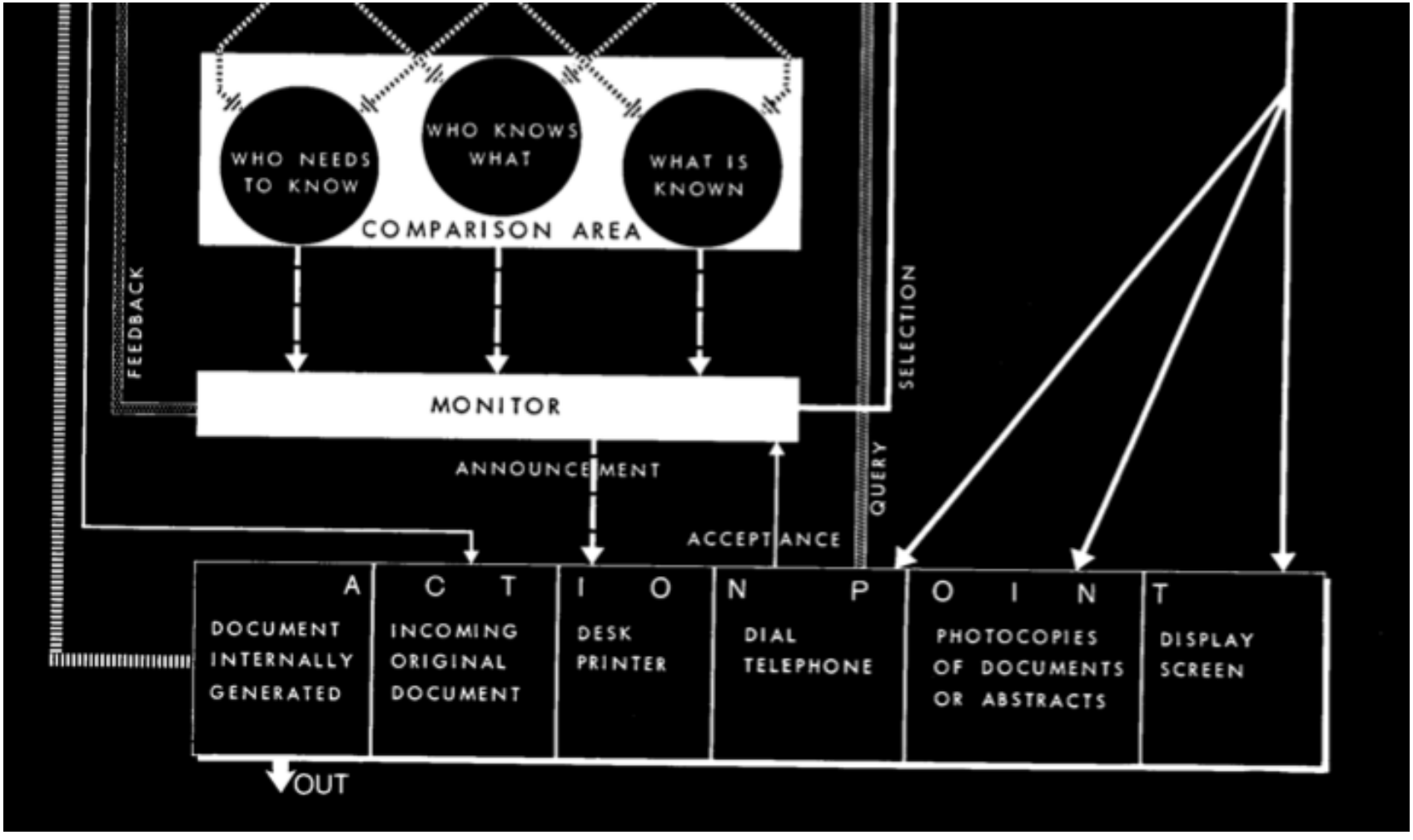


Figure 4: people are part of the information machine

Igo

Igo: Woke was in

- Feb 23, 1974

  But until the day comes when science finds a way of installing a conscience in every computer, we must develop human, personal safeguards that prevent computers from becoming huge, mechanical, impersonal robots that deprive us of our essential liberties.

Igo: Woke was in

- Feb 23, 1974
- Nixon "Radio Address About the American Right of Privacy.

  But until the day comes when science finds a way of installing a conscience in every computer, we must develop human, personal safeguards that prevent computers from becoming huge, mechanical, impersonal robots that deprive us of our essential liberties.

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975
  - Social Security Number Task Force of 1970

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975
  - Social Security Number Task Force of 1970
  - Fair Credit Reporting Act of 1970

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975
  - Social Security Number Task Force of 1970
  - Fair Credit Reporting Act of 1970
  - Privacy Act of 1974

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975
  - Social Security Number Task Force of 1970
  - Fair Credit Reporting Act of 1970
  - Privacy Act of 1974
  - The Family Educational Rights and Privacy Act of 1974

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975
  - Social Security Number Task Force of 1970
  - Fair Credit Reporting Act of 1970
  - Privacy Act of 1974
  - The Family Educational Rights and Privacy Act of 1974
  - Video Privacy Protection Act of 1987

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975
  - Social Security Number Task Force of 1970
  - Fair Credit Reporting Act of 1970
  - Privacy Act of 1974
  - The Family Educational Rights and Privacy Act of 1974
  - Video Privacy Protection Act of 1987
  - Electronic Communications Privacy Act

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975
  - Social Security Number Task Force of 1970
  - Fair Credit Reporting Act of 1970
  - Privacy Act of 1974
  - The Family Educational Rights and Privacy Act of 1974
  - Video Privacy Protection Act of 1987
  - Electronic Communications Privacy Act
  - Computer Matching and Privacy Protection Act of 1988

Igo: state as reformer

- "sectoral approach" (U.S. privacy law)
- acts and commissions:
  - FOIA 1966/1967
  - "Church" (Select Committee to Study Governmental Operations with Respect to Intelligence Activities of the United States Senate) 1975
  - Watergate hearings
  - Rockefeller Commission 1975
  - Pike Committee 1975
  - Social Security Number Task Force of 1970
  - Fair Credit Reporting Act of 1970
  - Privacy Act of 1974
  - The Family Educational Rights and Privacy Act of 1974
  - Video Privacy Protection Act of 1987
  - Electronic Communications Privacy Act
  - Computer Matching and Privacy Protection Act of 1988
  - Drivers Privacy Protection Act of 1994

Igo: principle of *transparency* (in addition to privacy)

FERPA

- then: parent access to student records to correct errors

Igo: principle of *transparency* (in addition to privacy)

FERPA

- then: parent access to student records to correct errors
  - "Unlocking School Files: The Buckley Amendment," Washington Post, November 17, 1974"

Igo: principle of *transparency* (in addition to privacy)

FERPA

- then: parent access to student records to correct errors
  - "Unlocking School Files: The Buckley Amendment," Washington Post, November 17, 1974"
- now: non disclosure of student records to non-parents

Miller 1971

- opens with Lamport poem

Miller 1971

- opens with Lamport poem
- please take note of TOC

Miller 1971

- opens with Lamport poem
- please take note of TOC
  - dumpster fire in ch 1,2,3,4

Miller 1971

- opens with Lamport poem
- please take note of TOC
  - dumpster fire in ch 1,2,3,4
  - law in ch 5,6

Miller 1971

- opens with Lamport poem
- please take note of TOC
  - dumpster fire in ch 1,2,3,4
  - law in ch 5,6
  - solutions in ch 7

Miller Ch 2 on privacy: is it this?

- what even is privacy?

  Privacy is a special kind of independence, which can be understood as an attempt to secure autonomy in at least a few personal and spiritual concerns, if necessary in defiance of all the pressures of modern society. . [I]t seeks to erect an unbreachable wall of dignity and reserve against the entire world. The free man is the private man, the man who still keeps some of his thoughts and judgments entirely to himself, who feels no over-riding compulsion to share everything of value with others, not even those he loves and trusts.

Miller Ch 2 on privacy: is it this?

- what even is privacy?
  - independence?

Privacy is a special kind of independence, which can be understood as an attempt to secure autonomy in at least a few personal and spiritual concerns, if necessary in defiance of all the pressures of modern society. . [I]t seeks to erect an unbreachable wall of dignity and reserve against the entire world. The free man is the private man, the man who still keeps some of his thoughts and judgments entirely to himself, who feels no over-riding compulsion to share everything of value with others, not even those he loves and trusts.

Miller on Ch 2: is it that?

seemingly unrelated or inconsistent. Of late, however, lawyers and social scientists have been reaching the conclusion that the basic attribute of an effective right of privacy is the individual's ability to control the circulation of information relating to him—a power that often is essential to maintaining social relationships and personal freedom.[2] Correlatively, when an individual is

Figure 5: privacy is control of disclosure

- Presages Helen Nissenbaum 2010

Miller Ch 2 on privacy

- what does “private” describe?

Miller Ch 2 on privacy

- what does "private" describe?
  - a fact F?

Miller Ch 2 on privacy

- what does "private" describe?
  - a fact F?
- a fact F and a set S of people who know?

Miller Ch 2 on privacy

- what does "private" describe?
  - a fact F?
- a fact F and a set S of people who know?
- a fact F, a person D who discloses, and a set S of people disclosed to?

Miller Ch 2 on privacy

- what does "private" describe?
  - a fact F?
- a fact F and a set S of people who know?
- a fact F, a person D who discloses, and a set S of people disclosed to?
  - cf., Helen Nissenbaum 2010

Miller Ch 2 on privacy

- what does "private" describe?
  - a fact F?
- a fact F and a set S of people who know?
- a fact F, a person D who discloses, and a set S of people disclosed to?
  - cf., Helen Nissenbaum 2010
- a F, D, S, and the intent ι of D and S?

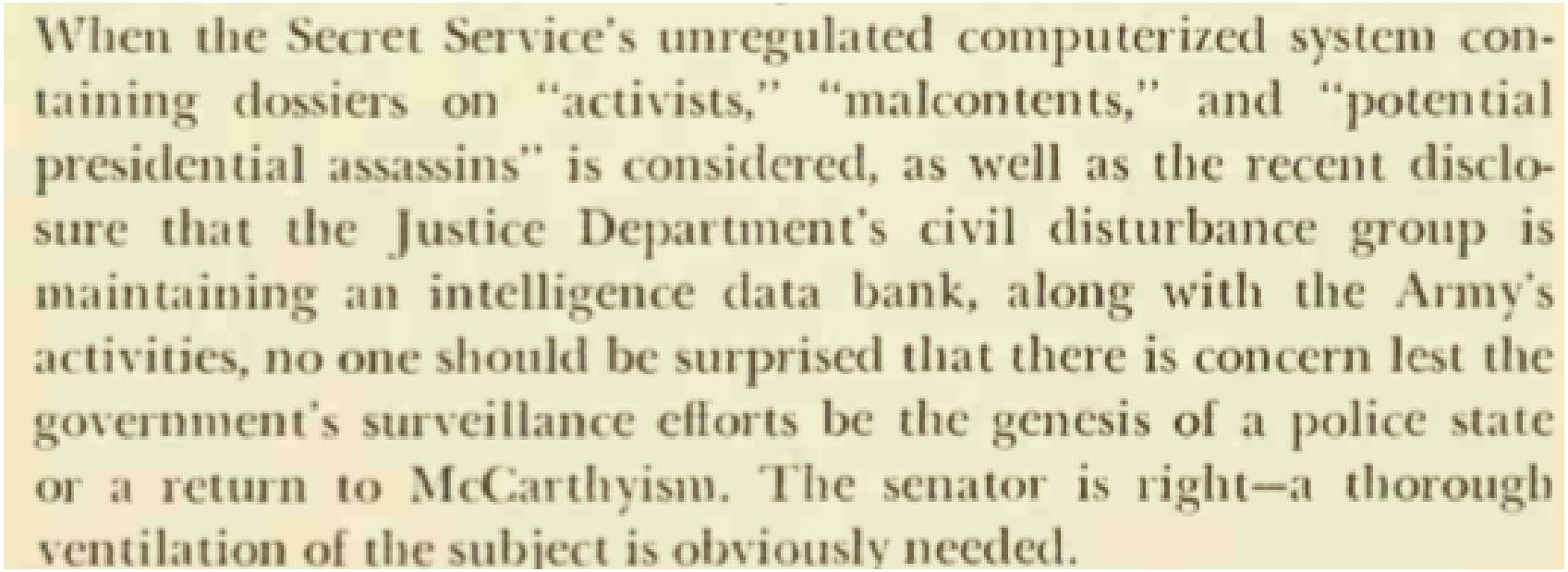
Miller Ch 2 on privacy

- what does "private" describe?
  - a fact F?
- a fact F and a set S of people who know?
- a fact F, a person D who discloses, and a set S of people disclosed to?
  - cf., Helen Nissenbaum 2010
- a F, D, S, and the intent ι of D and S?
- ...

Miller (and Igo)

- ‘dossier’ implies criminal

miller on McCarthyism



When the Secret Service's unregulated computerized system containing dossiers on "activists," "malcontents," and "potential presidential assassins" is considered, as well as the recent disclosure that the Justice Department's civil disturbance group is maintaining an intelligence data bank, along with the Army's activities, no one should be surprised that there is concern lest the government's surveillance efforts be the genesis of a police state or a return to McCarthyism. The senator is right—a thorough ventilation of the subject is obviously needed.

Figure 6: Dossier. . . McCarthyism. . . senatorial investigation "is needed"

elaboration by Igo

ties during the McCarthy years, the word "dossier" conjured up a suspect citizen who had through his own activities earned the attention of authorities. Merely having a dossier implied having something to hide; indeed, possessing "a record" was one feature distinguishing the bad citizen, or the noncitizen, from the good. For this reason, keeping tabs on law-abiding Americans, no matter that it was a routine practice of credit card companies and the IRS alike, could generate considerable unease. The headline of one negative report on the National Data Center, "There's a Dossier on You," tapped into just this sense of disquiet.[23]

Figure 7: Recording implied you did something wrong

counterarguments abound, e.g., in Tech, e.g., Schmidt

If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place.

('Nothing to hide argument')

It is possible that [Google's] information could be made available to the authorities."

much greater transparency and no anonymity. . . true anonymity is too dangerous.

Miller (and Igo)

- labels are compressions and subjective

Miller (and Igo)

- labels are compressions and subjective
  - e.g., “arrest”

Miller (and Igo)

- labels are compressions and subjective
  - e.g., “arrest”
- iterations

Miller (and Igo)

- labels are compressions and subjective
  - e.g., “arrest”
- iterations
  - model-reality iteration,

Miller (and Igo)

- labels are compressions and subjective
  - e.g., “arrest”
- iterations
  - model-reality iteration,
  - product-market iteration,

Miller (and Igo)

- labels are compressions and subjective
  - e.g., “arrest”
- iterations
  - model-reality iteration,
  - product-market iteration,
  - “understand” vs “control conduct”

quote (miller)

To some degree, the widescale use of computers to determine consumer appetites and voter attitudes adds new dimensions to the study of human dynamics. Unfortunately, the name of the game is not necessarily to give the citizenry what it wants; often these surveys are intended to divine a method of making palatable what industry or government already has decided to offer the public. One New Jersey firm is developing a data bank on doctors in order to enable drug companies to promote their products in a way that is suited to the habits and personality of individual doctors.[32] As this illustrates, the line between the use of cybernetics to understand an individual and its use to control or affect his conduct and beliefs is shadowy at best and one that is likely to be transgressed with some frequency.

Figure 8: "understand an individual" vs "control..conduct and beliefs"

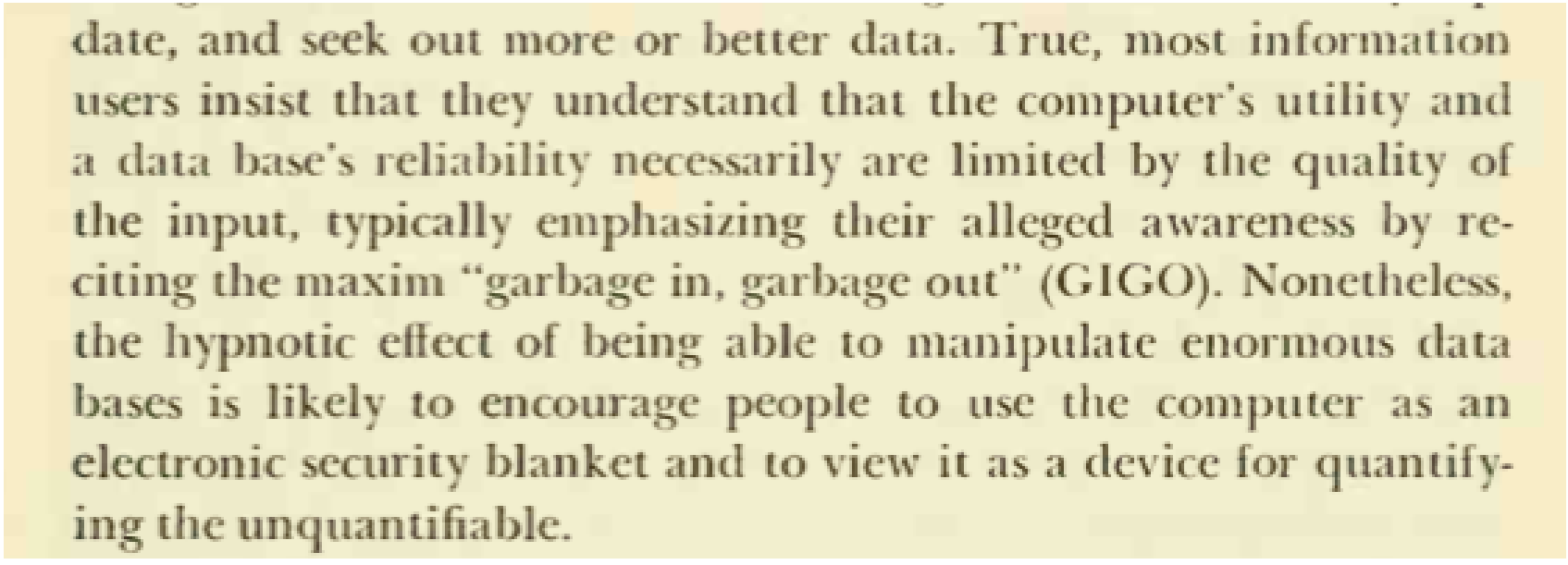
iteration and contestation

imagined the writer, might beget a whole new form of individualism. Basing his discussion on the targeted mailing lists then beginning to make an appearance, Asimov reasoned, "What [the consumer] receives will be so likely to be of interest to him and to be slanted to his particular needs that, even if he does not buy, he will feel that someone has gone to the trouble of knowing what he might want."[77] Indeed, it was perhaps only in a thoroughly computerized world that a person *could* be fully individualized. Until that day, "his wants and needs are unknown to anyone but himself and his immediate associates," making him a "faceless nothing."[78] The known citizen, in his futuristic scenario, was a well-cared-for citizen.

Figure 9: quote, igo

Asimov loves persuasion architectures

Miller on fallibility of data



date, and seek out more or better data. True, most information users insist that they understand that the computer's utility and a data base's reliability necessarily are limited by the quality of the input, typically emphasizing their alleged awareness by reciting the maxim "garbage in, garbage out" (GIGO). Nonetheless, the hypnotic effect of being able to manipulate enormous data bases is likely to encourage people to use the computer as an electronic security blanket and to view it as a device for quantifying the unquantifiable.

Figure 10: Miller on GIGO

Miller on data as authority

Nor should we ignore the real possibility that prejudice to the individual will not end with the decision to seek an indictment. Commenting on this particular computer application, a lawyer observed: "[T]he . . . computer can tell you where the stars are going to be a million years from now. Do you think a jury is not going to believe that it can tell you where a bookie is in the Bronx?"[24]

Figure 11: astronomy ergo predictive policing

Poon on FICO (why saved for last. . .)

- end with FICO because it's a living fossil

Poon on FICO

- credit became statistical (learned coefficients)

They compare the USA [and] Russia. . . [where] credit cards are distributed on the basis of subjective evaluations, on social networks and on 'trust' (ie, demonstrable friendships, kinship ties and employment), while in the USA, a mature market, institutions such as credit bureaus make calculative credit scoring practices feasible.

Poon on FICO

- credit became statistical (learned coefficients)
- was led by biz not academia

> *They compare the USA [and] Russia. . . [where] credit cards are distributed on the basis of subjective evaluations, on social networks and on 'trust' (ie, demonstrable friendships, kinship ties and employment), while in the USA, a mature market, institutions such as credit bureaus make calculative credit scoring practices feasible.*

Poon on FICO

- credit became statistical (learned coefficients)
- was led by biz not academia
  - cf. death tables, Gaunt, etc

They compare the USA [and] Russia... [where] credit cards are distributed on the basis of subjective evaluations, on social networks and on 'trust' (ie, demonstrable friendships, kinship ties and employment), while in the USA, a mature market, institutions such as credit bureaus make calculative credit scoring practices feasible.

Poon on FICO

- credit became statistical (learned coefficients)
- was led by biz not academia
  - cf. death tables, Gaunt, etc
- was not immediately adopted relative to craft

They compare the USA [and] Russia. . . [where] credit cards are distributed on the basis of subjective evaluations, on social networks and on 'trust' (ie, demonstrable friendships, kinship ties and employment), while in the USA, a mature market, institutions such as credit bureaus make calculative credit scoring practices feasible.

Poon on FICO

- credit became statistical (learned coefficients)
- was led by biz not academia
  - cf. death tables, Gaunt, etc
- was not immediately adopted relative to craft
  - cf. death tables, Gaunt, etc

They compare the USA [and] Russia. . . [where] credit cards are distributed on the basis of subjective evaluations, on social networks and on 'trust' (ie, demonstrable friendships, kinship ties and employment), while in the USA, a mature market, institutions such as credit bureaus make calculative credit scoring practices feasible.

Poon on Labor

This was a two-stage process performed by housewives working at piece rates of a few cents per sample application out of their homes. The meticulous work of 'the home- coders' was the backbone of the scorecard since it was their job to interpret the writing on the ledger cards and reliably convert it into the standardized numerical codes demanded by the analytic process. Codes were transferred to paper, reviewed for accuracy by a woman assigned as a 'checker', and subsequently transferred to punch cards so that the data could be read by machine. As one of the women who headed coding described it, a punch card machine is 'like a typewriter, you put your IBM cards in — they're about five by seven — and you have to sort them. If we punched a certain digit that would mean [occupation]: housewife' (Senior coder A).

Poon on Labor, cont.

Because of its repetitive and mechanical nature, coding was considered a mundane task in the company. Yet upon scrutiny it is clear that the work that was done involved its own form of skilled decision-making that was far from obvious. A former coder made clear to me that '[t]here was some interpretation on all of this. You couldn't just copy it. That was the hard part, coding it. [. . .] They didn't just say he's been three times thirty days late in nice English' (Senior coder B). Another drew attention to the fact that, "We had to read these logs of payments and every company didn't do the same thing, and we'd get so confused' (Senior coder C).

The Wells system involved teams of women working with boxes of punch cards and 42 pound Fridan [sic] calculators.

Poon: role of state

> *[if you] wonder why such flawed techniques continue[d] to prevail in the lending industry, the author points to the role of the state in sanctioning these methods,*

- cf., e.g., p-values

Poon: iteration

each time statistical practices have been introduced to a problem in a substantive domain, expertise, networks of associations, technical objects and even new interpretations of probability must be formed to accommodate this extension (Desrosiéres, 2000; Gigerenzer et al., 1989; Hacking, 1975; Porter, 1988). If statistical theories change as they travel, and if the places they go must be rebuilt and rearranged to fit to accommodate them in practice, then it is from the details of this mutual refitting that novel calculative effects must emerge.

- where have we seen this before?

Poon: subjective design choices

the very distinction between 'good' and 'bad' was flexible. . . . [e.g.,] one missed payment, two missed payments in 12 months, not paying at all for three months. . .) and policies on what was con- sidered an account in default varied between firms.

just how many files, from how narrow and recent a time period in a firm's history were necessary to build a representative model.

contemporary article(s) on the theme

your thoughts:

- big data in business (cf. Luhn)

contemporary article(s) on the theme

your thoughts:

- big data in business (cf. Luhn)
- privacy

contemporary article(s) on the theme

your thoughts:

- big data in business (cf. Luhn)
- privacy
- credit scoring dumpster fires

contemporary article(s) on the theme

your thoughts:

- big data in business (cf. Luhn)
- privacy
- credit scoring dumpster fires
- algorithms of oppression

power and principles

how did this capability rearrange power? who can now do what, from what, to whom?

role of rights, harms, justice?

foreshadowing data for Thursday

reminder of themes/big main takeaways

themes

up next

- what even is

up next

- what even is
  - privacy

up next

- what even is
  - privacy
  - ethics

up next

- what even is
  - privacy
  - ethics
    - how we define

up next

- what even is
  - privacy
  - ethics
    - how we define
    - how we design

up next

- what even is
  - privacy
  - ethics
    - how we define
    - how we design
- what happens when data, intelligence, and money meet?

Appendix

Radin (optional reading)

- "invisible labor" of big data, hidden even the word 'data'!

  researchers who donated the data were primarily concerned with attribution — making certain that their work and funders were credited properly — not about compensating patients, whose data had been made anonymous and was therefore regarded as protected.

Bonus quotes

quote 2 (miller)

As recently as a decade ago we could smugly treat Huxley's *Brave New World* and Orwell's *1984* as exaggerated science fiction having no relevance to us or to life in this country. But widespread public disclosures during the past few years about the new breed of information practices have stripped away this comforting but self-delusive mantle. It is now common knowledge that whenever a citizen files a tax return, applies for life insurance or a credit card, seeks government benefits, or interviews for a job, a new dossier is opened on him and his informational profile is sketched. Indeed, there are precious few things left in life that will not leave distinctive electronic tracks in the memory of a computer—tracks that can tell a great deal about our activities, habits, and associations.

Figure 12: quote 2, miller

quote 8 (igo)

The very term "dossier," not to mention a "giant dossier bank," as the *New York Times* characterized one of the proposals for centralizing federal record keeping, had already acquired a bad reputation.[22] Linked to those

with a criminal record or who had engaged in subversive political activities during the McCarthy years, the word "dossier" conjured up a suspect citizen who had through his own activities earned the attention of authorities. Merely having a dossier implied having something to hide; indeed, possessing "a record" was one feature distinguishing the bad citizen, or the noncitizen, from the good. For this reason, keeping tabs on law-abiding Americans, no matter that it was a routine practice of credit card companies and the IRS alike, could generate considerable unease. The

quote 9 (igo)

The very term "dossier," not to mention a "giant dossier bank," as the *New York Times* characterized one of the proposals for centralizing federal record keeping, had already acquired a bad reputation.[22] Linked to those

with a criminal record or who had engaged in subversive political activities during the McCarthy years, the word "dossier" conjured up a suspect citizen who had through his own activities earned the attention of authorities. Merely having a dossier implied having something to hide; indeed, possessing "a record" was one feature distinguishing the bad citizen, or the noncitizen, from the good. For this reason, keeping tabs on law-abiding Americans, no matter that it was a routine practice of credit card companies and the IRS alike, could generate considerable unease. The

quote 10 (igo)

78. Isaac Asimov, "The Individualism to Come," *New York Times* advertising supplement, sec. 11, January 7, 1973, pp. 12, 13.

Figure 15: quote 10, igo

quote 11 (igo)

37. "Retail Credit's Day in Court," *Privacy Journal* 1: 2 (December 1974): 1; "Retail Credit Co. on Trial," *Privacy Journal* 1: 11 (September 1975): 4. Intriguingly, Retail Credit attempted to use FCRA against the Federal Trade Commission, arguing that the federal agency should not be privy to its customer files. On the Retail Credit Company, see the exposé "Anything Adverse?" *New Yorker*, April 21, 1975, as well as *Roemer v. Retail Credit Co.* of 1975. The company would soon, as a result of all this bad publicity, rename itself Equifax.

Figure 16: quote 11, igo

quote 12 (igo)

Still, expectations that a computerized world would usher in tremendous benefits were alive and well. The faith that computer systems would help government run better, deliver public services more effectively, and

engineer new solutions to social problems animated discussions of expanding electronic networks. The municipal government of New Haven,

Figure 17: quote 12, igo

quote 13 (igo)

8. This last was the Government Accounting Office's blueprint for a center called FEDNET. Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford, CA: Stanford Law Books, 2010), 39. See also Robert Ellis Smith, *Ben Franklin's Website: Privacy and Curiosity from Plymouth Rock to the Internet* (Providence, RI: Privacy Journal, 2000), 309–312. "A detectable trend toward increased public criticism of the plans of organizations and their personal records" was also obvious in Britain, according to Malcolm Warner and David Stone, *The Data Bank Society: Organizations, Computers and Social Freedom* (London: George Allen & Unwin, 1970), 20.

Figure 18: quote 13, igo

quote 14 (igo)

dividual civil and bodily right to privacy. In the very years that French theorist Michel Foucault's writings on classification and distributed discipline were starting to be translated and read in the United States, a homegrown analysis of social surveillance was crystallizing. This analysis called attention to the ways citizens were captured or hemmed in by the many social institutions that "knew" them, whether through a rap sheet, a school dossier, or a credit report. Even if the term itself would have been unfamiliar to many, Foucault's description of "panopticism"—a system of control that worked because individuals never knew when they were being watched—shared much with discussions about an emerging surveillance society in the United States. Foucault had borrowed the concept

Figure 19: quote 14, igo