

data-ppf.github.io Jan 28 2020

lecture 2 of 14: setting the stakes

chris wiggins + matt jones, Columbia

themes for today

## themes for today

- ▶ “big data” (the view from 2012-2014)

## themes for today

- ▶ “big data” (the view from 2012-2014)
  - ▶ what are the concerns

## themes for today

- ▶ “big data” (the view from 2012-2014)
  - ▶ what are the concerns
  - ▶ ethical considerations

## themes for today

- ▶ “big data” (the view from 2012-2014)
  - ▶ what are the concerns
  - ▶ ethical considerations
  - ▶ political considerations

## themes for today

- ▶ “big data” (the view from 2012-2014)
  - ▶ what are the concerns
  - ▶ ethical considerations
  - ▶ political considerations
- ▶ “persusasion architectures”

## themes for today

- ▶ “big data” (the view from 2012-2014)
  - ▶ what are the concerns
  - ▶ ethical considerations
  - ▶ political considerations
- ▶ “persusasion architectures”
  - ▶ considerations for politics



## themes for today

- ▶ “big data” (the view from 2012-2014)
  - ▶ what are the concerns
  - ▶ ethical considerations
  - ▶ political considerations
- ▶ “persusasion architectures”
  - ▶ considerations for politics
  - ▶ past: (Bernays, etc)

## themes for today

- ▶ “big data” (the view from 2012-2014)
  - ▶ what are the concerns
  - ▶ ethical considerations
  - ▶ political considerations
- ▶ “persusasion architectures”
  - ▶ considerations for politics
  - ▶ past: (Bernays, etc)
  - ▶ present: challenge to democracy

student observations

## student observations

- ▶ “I did not know what Big Data was before doing the readings today”

## student observations

- ▶ “I did not know what Big Data was before doing the readings today”
- ▶ “I agree with Boyd and Crawford when they argue that to clean or filter data in order to extrapolate information from it and detect patterns is a subjective process”

## student observations

- ▶ “I did not know what Big Data was before doing the readings today”
- ▶ “I agree with Boyd and Crawford when they argue that to clean or filter data in order to extrapolate information from it and detect patterns is a subjective process”
- ▶ ethics (8 mentions), e.g., “informed consent”

## student observations

- ▶ “I did not know what Big Data was before doing the readings today”
- ▶ “I agree with Boyd and Crawford when they argue that to clean or filter data in order to extrapolate information from it and detect patterns is a subjective process”
- ▶ ethics (8 mentions), e.g., “informed consent”
- ▶ 3 “Cambridge Analytica” mentions, e.g., “(see: every major political election since 2008, Cambridge Analytica)”

## student observations

- ▶ “I did not know what Big Data was before doing the readings today”
- ▶ “I agree with Boyd and Crawford when they argue that to clean or filter data in order to extrapolate information from it and detect patterns is a subjective process”
- ▶ ethics (8 mentions), e.g., “informed consent”
- ▶ 3 “Cambridge Analytica” mentions, e.g., “(see: every major political election since 2008, Cambridge Analytica)”
- ▶ 2 on “mythology” in b+CA



## student observations

- ▶ “I did not know what Big Data was before doing the readings today”
- ▶ “I agree with Boyd and Crawford when they argue that to clean or filter data in order to extrapolate information from it and detect patterns is a subjective process”
- ▶ ethics (8 mentions), e.g., “informed consent”
- ▶ 3 “Cambridge Analytica” mentions, e.g., “(see: every major political election since 2008, Cambridge Analytica)”
- ▶ 2 on “mythology” in b+CA
- ▶ 9 on “subjective/subjectivity”

readings: boyd+Crawford, Wallach, Tufekci

boyd+Crawford (2012)

## boyd+Crawford (2012)

- ▶ writing from MSRBOS (later MSRNYC)

## boyd+Crawford: opening

*Technology is neither good nor bad; nor is it neutral . . . technology's interaction with the social ecology is such that technical developments frequently have environmental, social, and human consequences that go far beyond the immediate purposes of the technical devices and practices themselves. (Kranzberg 1986, p. 545)*

- Q: how does 'big data' relate to 'technology?'

## boyd+Crawford: themes

1. Big Data changes the definition of knowledge

## boyd+Crawford: themes

1. Big Data changes the definition of knowledge
2. Claims to objectivity and accuracy are misleading

## boyd+Crawford: themes

1. Big Data changes the definition of knowledge
2. Claims to objectivity and accuracy are misleading
3. Bigger data are not always better data



## boyd+Crawford: themes

1. Big Data changes the definition of knowledge
2. Claims to objectivity and accuracy are misleading
3. Bigger data are not always better data
4. Taken out of context, Big Data loses its meaning

## boyd+Crawford: themes

1. Big Data changes the definition of knowledge
2. Claims to objectivity and accuracy are misleading
3. Bigger data are not always better data
4. Taken out of context, Big Data loses its meaning
5. Just because it is accessible does not make it ethical

## boyd+Crawford: themes

1. Big Data changes the definition of knowledge
2. Claims to objectivity and accuracy are misleading
3. Bigger data are not always better data
4. Taken out of context, Big Data loses its meaning
5. Just because it is accessible does not make it ethical
6. Limited access to Big Data creates new digital divides

b+C: NB

- ▶ attention to PII

b+C: why we chose

Wallach (Dec 2014)

## Wallach: background

- ▶ NeurIPS 2014 1st **FATML** meeting

## Wallach: background

- ▶ NeurIPS 2014 1st **FATML** meeting
- ▶ Wallach is NLP/CS PhD and prof -> MSRNYC



## Wallach: themes

- ▶ Defining big data

## Wallach: themes

- ▶ Defining big data
- ▶ Concerns: “ethics. . . outside the expertise”

## Wallach: themes

- ▶ Defining big data
- ▶ Concerns: “ethics. . . outside the expertise”
- ▶ “Computational Social Science” / cross-disciplinary benefits

## Wallach: themes

- ▶ Defining big data
- ▶ Concerns: “ethics. . . outside the expertise”
- ▶ “Computational Social Science” / cross-disciplinary benefits
- ▶ biases: “even scientists”

## Wallach: why we chose

- ▶ impact of FATML

# Wallach: why we chose

- ▶ impact of FATML
  - ▶ fairness (tech field)

# Wallach: why we chose

- ▶ impact of FATML
  - ▶ fairness (tech field)
  - ▶ fatACM

# Wallach: why we chose

- ▶ impact of FATML
  - ▶ fairness (tech field)
  - ▶ fatACM
- ▶ Q: where do F,A,T sit in rights,harms,justice?



# Wallach: why we chose

- ▶ impact of FATML
  - ▶ fairness (tech field)
  - ▶ fatACM
- ▶ Q: where do F,A,T sit in rights, harms, justice?
- ▶ taxonomy

# Wallach: why we chose

- ▶ impact of FATML
  - ▶ fairness (tech field)
  - ▶ fatACM
- ▶ Q: where do F,A,T sit in rights, harms, justice?
- ▶ taxonomy
  - ▶ Wallach: prediction, explanation and exploration

# Wallach: why we chose

- ▶ impact of FATML
  - ▶ fairness (tech field)
  - ▶ fatACM
- ▶ Q: where do F,A,T sit in rights, harms, justice?
- ▶ taxonomy
  - ▶ Wallach: prediction, explanation and exploration
  - ▶ Gartner: describe, predict, prescribe

Tufekci

## Tufekci: themes

- ▶ *past*

# Tufekci: themes

- ▶ *past*
  - ▶ Bernays

# Tufekci: themes

- ▶ *past*
  - ▶ Bernays
  - ▶ 1968 (and earlier)

# Tufekci: themes

- ▶ *past*
  - ▶ Bernays
  - ▶ 1968 (and earlier)
- ▶ *future*



# Tufekci: themes

- ▶ *past*
  - ▶ Bernays
  - ▶ 1968 (and earlier)
- ▶ *future*
  - ▶ 2016 etc.

## Tufekci: past

- ▶ Bernays

# Tufekci: past

- ▶ Bernays
  - ▶ **Engineering of Consent** 1947/1955 “engineering approach. . . scientific principles. . . getting people to support ideas and programs”

# Tufekci: past

- ▶ Bernays
  - ▶ **Engineering of Consent** 1947/1955 “engineering approach. . . scientific principles. . . getting people to support ideas and programs”
- ▶ For more info, enjoy

# Tufekci: past

- ▶ Bernays
  - ▶ **Engineering of Consent** 1947/1955 “engineering approach. . . scientific principles. . . getting people to support ideas and programs”
- ▶ For more info, enjoy
  - ▶ **Century of the Self** (2002, 4 hrs)

# Tufekci: past

- ▶ Bernays
  - ▶ **Engineering of Consent** 1947/1955 “engineering approach. . . scientific principles. . . getting people to support ideas and programs”
- ▶ For more info, enjoy
  - ▶ **Century of the Self** (2002, 4 hrs)
  - ▶ **These Truths** and “**Politics and the New Machine**” from **Jill Lepore** (esp. on **Simulmatics**)

## Tufekci: Bernays as backdrop

*Intelligent men must realize that propaganda is the modern instrument by which they can fight for productive ends and help to bring order out of chaos*

- Bernays
  - ▶ role of experts in science vs morals

## Tufekci: Bernays as backdrop

*Intelligent men must realize that propaganda is the modern instrument by which they can fight for productive ends and help to bring order out of chaos*

- Bernays
  - ▶ role of experts in science vs morals
  - ▶ dark money, 1920s style



## Tufekci: more on Bernays

*There was one basic lesson I learned in the CPI—that efforts comparable to those applied by the CPI to affect the attitudes of the enemy, of neutrals, and people of this country could be applied with equal facility to peacetime pursuits. In other words, what could be done for a nation at war could be done for organizations and people in a nation at peace.*

- Bernays, 1891-1995, quoted in **Cutlip**, 1994

Tufekci: why we chose

power and principles

how did new capabilities 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

appendix

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism



## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- ▶ 2020-03-10 : 8 of 14 birth and death of AI

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- ▶ 2020-03-10 : 8 of 14 birth and death of AI
- ▶ 2020-03-24 : 9 of 14 big data, old school (1958-1980)

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- ▶ 2020-03-10 : 8 of 14 birth and death of AI
- ▶ 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- ▶ 2020-03-31 : 10 of 14 data science, 1962-2017

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- ▶ 2020-03-10 : 8 of 14 birth and death of AI
- ▶ 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- ▶ 2020-03-31 : 10 of 14 data science, 1962-2017
- ▶ 2020-04-07 : 11 of 14 AI2.0

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- ▶ 2020-03-10 : 8 of 14 birth and death of AI
- ▶ 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- ▶ 2020-03-31 : 10 of 14 data science, 1962-2017
- ▶ 2020-04-07 : 11 of 14 AI2.0
- ▶ 2020-04-14 : 12 of 14 ethics

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- ▶ 2020-03-10 : 8 of 14 birth and death of AI
- ▶ 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- ▶ 2020-03-31 : 10 of 14 data science, 1962-2017
- ▶ 2020-04-07 : 11 of 14 AI2.0
- ▶ 2020-04-14 : 12 of 14 ethics
- ▶ 2020-04-21 : 13 of 14 present problems: attention economy+VC=dumpsterfire

## appendix

- ▶ 2020-01-21 : 1 of 14 intro to course
- ▶ 2020-01-28 : 2 of 14 setting the stakes
- ▶ 2020-02-04 : 3 of 14 risk and social physics
- ▶ 2020-02-11 : 4 of 14 statecraft and quantitative racism
- ▶ 2020-02-18 : 5 of 14 intelligence, causality, and policy
- ▶ 2020-02-25 : 6 of 14 data gets real: mathematical baptism
- ▶ 2020-03-03 : 7 of 14 WWII, dawn of digital computation
- ▶ 2020-03-10 : 8 of 14 birth and death of AI
- ▶ 2020-03-24 : 9 of 14 big data, old school (1958-1980)
- ▶ 2020-03-31 : 10 of 14 data science, 1962-2017
- ▶ 2020-04-07 : 11 of 14 AI2.0
- ▶ 2020-04-14 : 12 of 14 ethics
- ▶ 2020-04-21 : 13 of 14 present problems: attention economy+VC=dumpsterfire
- ▶ 2020-04-28 : 14 of 14 future solutions