ARC23Jan

ETHICS OF BIG DATA

Susan Castro
Department of Philosophy

WICHITA STATE UNIVERSITY

OUTLINE FOR TODAY'S TALK

- Definition of terms: Big Data, Ethics
- 5 ethical theory types
  - Application of ethical considerations
- Basic strategies w examples
- Reading lists
- Q&A (go ahead and ask along the way!)

Slides will be available at wichita.edu

WHAT DOES "BIG DATA" INCLUDE?

1. BIG data (not a well-defined term)
   - Everything on the internet; all the data that your phone and other devices collect; data sets measured in terabytes or more (quettabytes!)
     - The world generates so much data that new unit measurements were created to keep up (NPR Science November 19, 2022)

Parameters of concern:

- How it's collected, from what/whom,
- How it's stored, protected, for how long

How to store data for 1,000 years

Wichita State University

WHAT DOES "BIG DATA" INCLUDE?

2. What we do with it at scale
   - Machine Learning (ML) tools/applications trained on big data sets using algorithms that find correlations and patterns
     - Deep learning tools (ML w 3+ layers or hidden layers)
   - Artificial 'Intelligence' (we're still *far* from the singularity)

Parameters of concern:

- How data is shared, how responsibilities transfer
- Generative AI
- Predictive AI
- Transparency
- Verification, algorithmic drift

WSU Wichita State University

WHAT DOES "BIG DATA" INCLUDE?

3. What it does to us at scale
   - (Familiar concerns) Human Plasticity and Development
     - Skills, knowledge lost/gained, e.g. navigation, attention span
     - Sociality transformed, Mental health impacts
   - Environmental Ethics
     - Power use/carbon footprint
     - Water use

   The Staggering Ecological Impacts of Computation and the Cloud

   Data centers, backbone of the digital economy, face water scarcity and climate risk

Q: WHAT IS "ETHICS"?
A: PHILOSOPHY OF THE GOOD

5 Ethical Theory Types

1. Agent-centered
   a. Kantian ethics/deontology
   b. Virtue Ethics
2. Outcome-centered
   a. Consequentialism
   b. Utilitarianism
3. Relationship-centered
   a. Care ethics
4. Public/justice-centered
   a. Human rights
   b. Social contract
   c. Social justice
5. Community/culture-centered
   a. Confucian ethics
   b. African ethics

These are not mutually exclusive categories!

APPLICATION OF ETHICAL CONSIDERATIONS

1. **Agent-centered ethical theories**
   a. Kantian ethics/deontology
      Conduct: Good will, duty/moral obligation, respect, autonomy
   b. Aristotelian/Virtue Ethics
      Character: virtue, vices, integrity

Agents?

- Humans: users, software & hardware engineers, salespeople
- Organizations: ISPs, app development cooperatives, regulatory agencies

Increasingly,

- Products themselves: ArcGIS, ChatGPT

WICHITA STATE UNIVERSITY

APPLICATION OF ETHICAL CONSIDERATIONS

2. Outcome-centered ethical theories

a. Consequentialism

   Health, environmental impact, economic impact, etc.

b. Utilitarianism

   What's useful for happiness

What matters?

- Scale
- Conformity/Diversity
- Distribution, Equity
- Flexibility, Adaptibility

APPLICATION OF ETHICAL CONSIDERATIONS

3. Relationship-centered

a. Care ethics

Trust, honesty, understanding, communication, authenticity, power asymmetries

Relationships

- Parenting
- Nursing
- Financial Management
- Corrections

Add 3rd party: AI assistants, mediation of relationships

APPLICATION OF ETHICAL CONSIDERATIONS

4. Public/justice-centered

a. Human rights

   Right to privacy, right to be forgotten, right to not be stalked

b. Social contract

   Prudence-derived obligations between individuals & the state and state-like parties, e.g. multinationals

c. Social justice

Fairness, authentic inclusion, equity, freedoms, empowerment

Ripped from the Headlines…

PUBLIC/JUSTICE-CENTERED HEADLINES

- Does your rewards card know if you're pregnant? Privacy experts sound the alarm
- Oregon is dropping an artificial intelligence tool used in child welfare system
- Oh, the Irongy: After a discrediting campaign, DHS pauses a board created to combat disinformation
- U.S. warns of discrimination in using artificial intelligence to screen job candidates
- CRISPR's 'ancestry problem' misses cancer targets in those of African descent

APPLICATION OF ETHICAL CONSIDERATIONS

5. Community/culture-centered

a. Confucian ethics

Lessons from the Analects, social roles

b. African ethics

Ubuntu, pre/post-Colonial, North vs. sub-Saharan traditions

Considerations

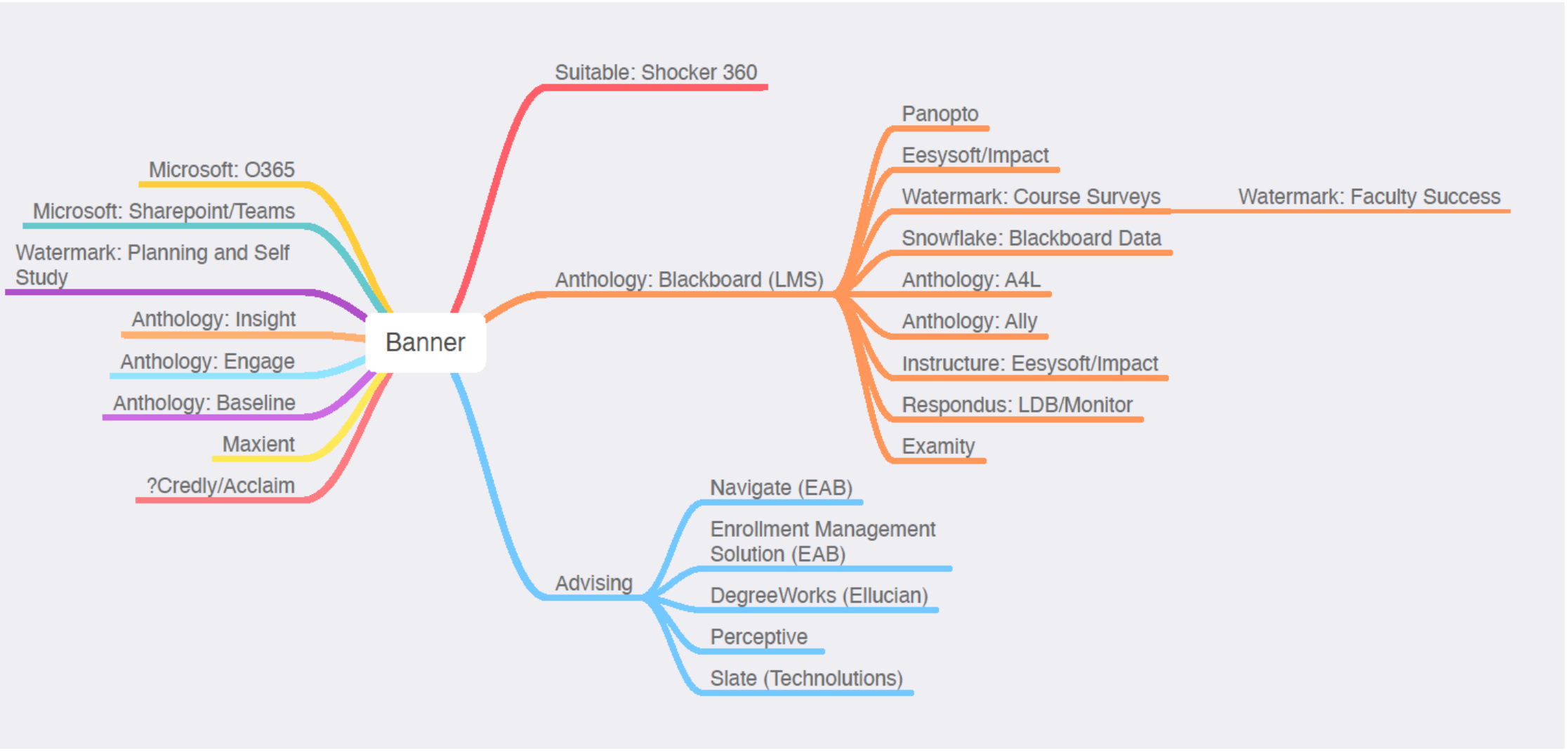
How are communities more than interpersonal relationships?

How should traditions and social roles inform what's good?

Is privacy *different* for different communities?

What we can learn from an Indigenous approach to AI

WHAT'S OUR DATA PROFILE AT WSU?



RESPECTING STUDENT DATA PRIVACY: BEYOND FERPA

Empathy Strategy: Put yourself in the student's shoes. Upon reading A, what's your reaction? How might that inform what you do, knowing B?

From ARC23Jan: Data Reporting in Bb and Panopto

A. "The Office of Instructional Resources now has full access to the entire Bb database (read only!) That means we can track every single click inside of Bb"
B. "Impact Course Reports are available to all instructors using Blackboard"

(Hint: trust, right to not be stalked, communication…)

WSU WICHITA STATE UNIVERSITY

TEACHING RESPONSIBLE USE OF GENERATIVE AI

Socratic Strategy: Play the *what if* game, vary the context, and see what turns out to be salient or decisive. Keep playing!

Background Lens: The Black Box Problem and Explainable AI, The Salmon of Doubt (from Imran Musaji in ARC23Jan: ARC Book Club: Robot Proof)

Q: What responsibilities do professionals have in using GenAI?

1. What if professionals in your field use GenAI, which is proprietary DL?
2. What if professionals in your field use GenAI, which is an open source NN product?
3. What if GenAI is very new and professionals aren't using it?

RESPECTING STUDENT INTELLECTUAL PROPERTY

Socratic Strategy 2.0: Pose a scenario, vary the context, and ask different *kinds* of questions.

Suppose: You have access to a nice bank of well-graded student work AND we could train BbAI on that data set to grade student work.

1. Is it ethical to use our students' work this way?
2. Suppose we inform students when they enroll that their work may be or will be used to train an AI-grader. Is it ethical now?
3. Does it matter whether the instructor or the institution is the agent?
4. Is there a danger of algorithmic drift over time?

WSU Wichita State University

(Question was raised by Darren in ARC23Jan: Academic Honesty in the Age of AI)

ACADEMIC AND PROFESSIONAL INTEGRITY

Leadership/Context Control Strategy: To the extent that you have decision-making authority, shape the context to be conducive to ethics.

- Transparency: Define what's allowed, what's not, and explain why.
- Shared governance: Listen to people! Have a little epistemic humility.
- Reduce temptation. Good oversight helps a lot.

2 cases (next slide)

ACADEMIC AND PROFESSIONAL INTEGRITY

Case: ChatGPT can write code. What should faculty do?

- Detecting it: Is it really to spec? Is it efficient? Is it well commented?
- Using it: Can the student modify it to new specs? Can they construct good prompts?

Is this more than a natural progression from using software libraries?

Case: RoseTTAFold can design new drugs. What should faculty do?

- Are these good targets? Can we synthesize them? Can we predict safety or efficacy?
- How should we prioritize validation?

What constitutes research skill in your specialty?

WSU | WICHITA STATE UNIVERSITY

STEP BACK AND RECENTER

Pluralistic Strategy: Apply all those ethical theories. (Keep it Socratic by using questions and what ifs from all angles to reach conclusions.)

- What would Kant say? What would Aristotle say? What would Confucius say?
- What are the downstream effects, unintended consequences, unintended uses, bugs, benefits, etc. that might come out of this?
- Will this promote happiness, health, well-being? (For whom?)
- Is this honest? Is it fair?
- Will it change us in ways that matter (development or corruption)?

WSU Wichita State University

COOPERATIVE STRATEGIES

1. Do your homework. Find out what other people have said/done, and why, instead of completely reinventing the wheel.
2. Work as a team.
    a. Thinking through something together is often more effective than working alone.
    b. Feeling alone is disempowering and rather unpleasant.
    c. If/when you need to advocate, there is strength in numbers and well-vetted reasons.

RATIONALIST & SENTIMENTALIST APPROACHES

Rationalist

- Intellect, reason, understanding are what make us moral agents.
  - Critical reasoning, explanation are the best means of ethical improvement.
  - Communication and education should be lecture-style, assertoric, essay form.

Sentimentalist

- Empathy, shame, and the moral sentiments or reactive attitudes are what make us capable of morality.
  - Engaging and training our affects are the means to ethical improvement.
  - Storytelling, narrative forms, imagery, lived experience are best for communication and education.

RATIONALIST REFERENCES

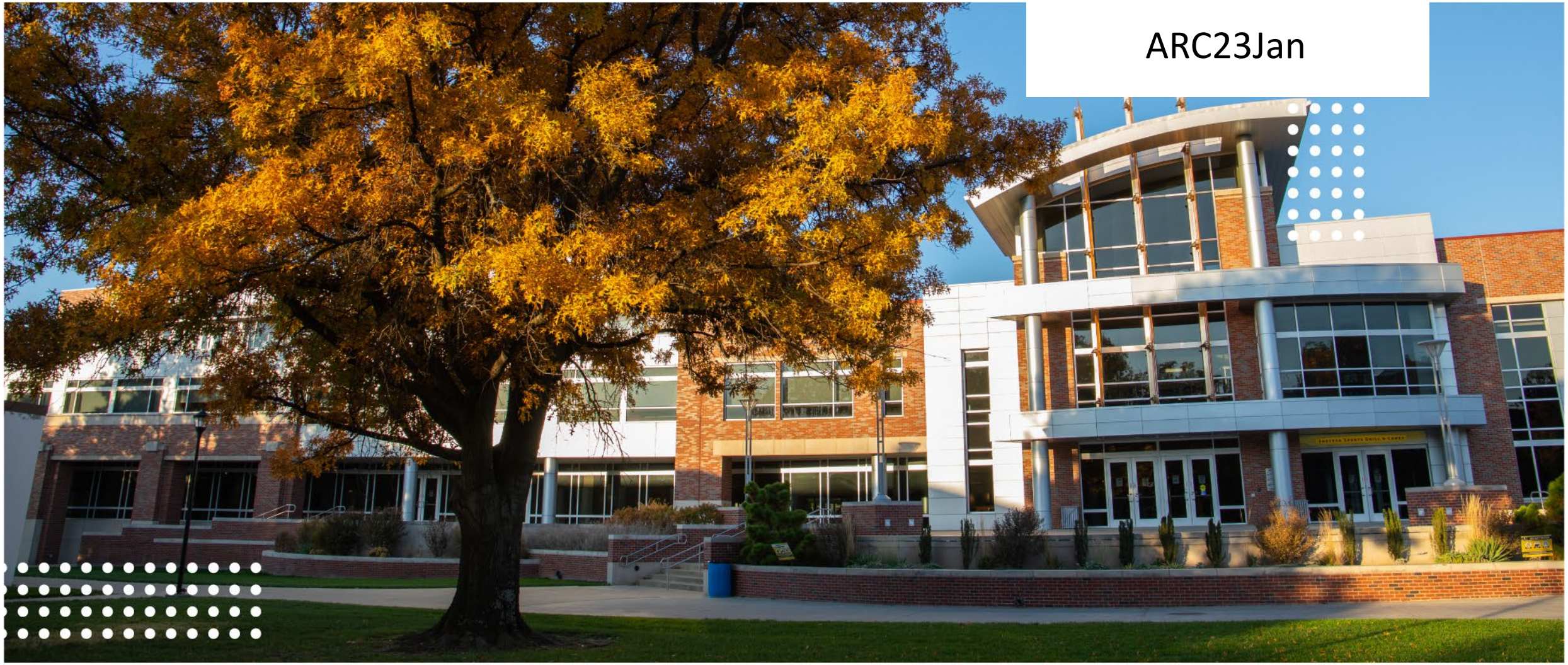
1. Aragon *et al* Human-Centered Data Science: An Introduction.
2. Boylan & Teays Ethics in the AI, Technology, and Information Age.
3. Collmann & Matei Ethical Reasoning in Big Data: An Exploratory Analysis.
4. D'Ignazio & Klein Data Feminism.
5. Fleming & Bruce Responsible Data Science: Transparency and Fairness in Algorithms.
6. Goldstein & Nost The Nature of Data: Infrastructures, Environments, Politics.
7. Katz *et al* Legal Informatics.
8. Kearns & Roth The Ethical Algorithm.
9. Martens Data Science Ethics: Concepts, Techniques and Cautionary Tales.
10. Martin Ethics of Data and Analytics: Concepts and Cases.
11. Shukla *et al* Data Ethics and Challenges.
12. Williamson, Hugh F. and Leonelli, Sabina (eds) Towards Responsible Plant Data Linkage: Data Challenges for Agricultural Research and Development.

WSU WICHITA STATE UNIVERSITY

SENTIMENTALIST REFERENCES

1. Benjamin, Ruha Race After Technology: Abolitionist Tools for the New Jim Code.
2. Brayne, Sarah Predict and Surveil: Data, Discretion, and the Future of Policing.
3. Broussard, Meredith Artificial Unintelligence: How Computers Misunderstand the World.
4. Christian, Brian The Alignment Problem: Machine Learning and Human Values.
5. Crawford, Kate Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.
6. Eubanks, Virginia Automating Inequality: How High-Tech tools Profile, Police, and Punish the Poor.
7. Fisher, Max The Chaos Machine: The Inside Story of How Social Media Rewired our Minds and Our World.
8. Hicks, Mar Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing.
9. Noble, Safiya Umoja Algorithms of Oppression: How Search Engines Reinforce Racism.
10. O'Neil, Cathy Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.
11. Pasquale, Frank New Laws of Robotics: Defending Human Expertise in the Age of AI. Harvard.
12. Wachter-Boettcher, Sara Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech.
13. Zuboff, Shoshana The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.

WICHITA STATE UNIVERSITY



ARC23Jan

Thank you for your *good will*! ;)

WICHITA STATE UNIVERSITY