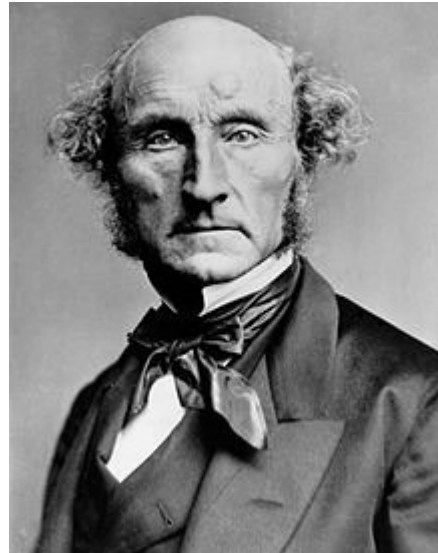# MSBA 5507.1   Ethics, Risk Management and Data Security

Accuracy, Fairness, Autonomy and Transparency
July 29, 2022

John Stuart Mills

Ken DeJarnette
ken@kdejarnette.com
ken.dejarnette@dominican.edu
213-399-8706
Office hours: Wed 1-3 pm by appointment

Non-users/non-subscribers

Anti-Autonomy
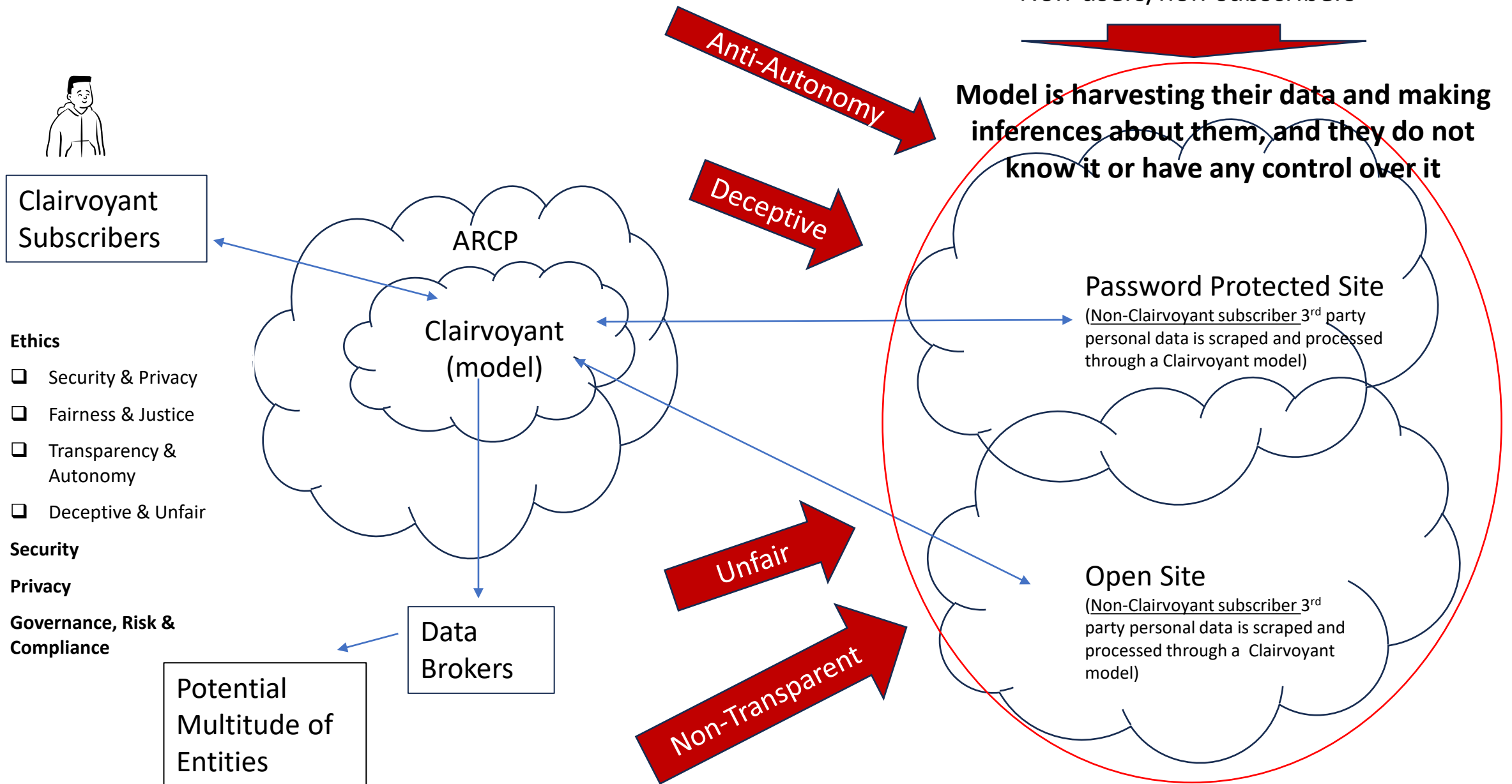
Model is harvesting their data and making inferences about them, and they do not know it or have any control over it

Clairvoyant Subscribers

ARCP

Clairvoyant (model)

Deceptive

Password Protected Site
(Non-Clairvoyant subscriber 3rd party personal data is scraped and processed through a Clairvoyant model)

Ethics

❑ Security & Privacy

❑ Fairness & Justice

❑ Transparency & Autonomy

❑ Deceptive & Unfair

Security

Privacy

Governance, Risk & Compliance

Unfair

Open Site
(Non-Clairvoyant subscriber 3rd party personal data is scraped and processed through a Clairvoyant model)

Data Brokers

Potential Multitude of Entities

Non-Transparent

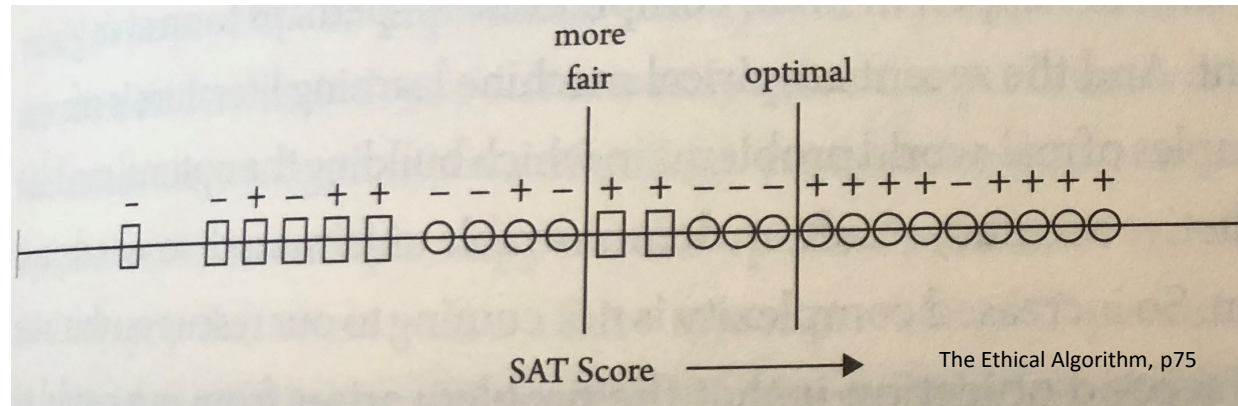| TOPIC	DESCRIPTION/SOURCES
Accuracy	S. Andrew Schroeder, *How Many Have Died*, Issues in Science and Technology (2020)

Discussion	❏ How is counting and exercise in interests and values?
	  ❏ A man suffers from depression, gets drunk, drives and fails to see a poorly lit and marked turn in the road and crashes the car and dies? What is the cause of death?
	  ❏ What does it mean to attribute a death to Covid?
	  ❏ How do you count lives saved by Covid?
	  ❏ How do you account for the lives lost (but unrelated to having Covid) because of Covid?
	  ❏ What do you think about using excess mortality as a means of determining Covid deaths?

Accuracy and fairness   Michael Kearns and Aaron Roth, **The Ethical Algorithm,** Oxford Univ. Press (2020).

What does fairness mean?

- Scientific/ technical notions of fairness
  - Statistical parity
  - Equality of false negatives
  - Equality of positive predictive value

The Ethical Algorithm, p75

- Societal Notions of fairness
  - Equal opportunity
  - Equality of outcomes
  - Socially just
  - Procedurally fair

- Are accuracy and fairness compatible?

- Is reduction in accuracy worth the gain in fairness?

- How do data, design and use impact accuracy and fairness?

Fairness/Bias   Kate Crawford, *The trouble with Bias* (NIPS 2017 Keynote)

Judgements based on preconceived notions or prejudices, usually in a way considered to be unfair

- ❏ Is bias a technical problem or a societal problem?
- ❏ Why does Kate Crawford think classification is so significant an issue?
- ❏ If we try to improve accuracy or scrub for bias/unfairness
  - ❏ Who decides?
  - ❏ Whose idea of fair/neutrality?
- ❏ Fundamental questions
  - ❏ Who is benefited?
  - ❏ Who is harmed?
- ❏ What is the importance of redress (the ability to appeal)?

Discussion
(Relationship between accuracy and fairness/bias)

Bias Harms

Allocation (resources)

- ❏ Immediate
- ❏ Quantifiable
- ❏ Discrete
- ❏ Transactional

Withholds an opportunity or resource

Representation (identity)

- ❏ Long-term
- ❏ Formularization difficult
- ❏ Diffuse
- ❏ Cultural

Stereotypes, denigrates, underrepresents

TOPIC	DESCRIPTION/SOURCES
Transparency/ Autonomy	Michael Kearns and Aaron Roth, **The Ethical Algorithm,** Oxford Univ. Press (2020).

Discussion

(The relationship between transparency and autonomy)

- ❏ What is transparency?
    - ❏ Data?
    - ❏ Algorithm?
    - ❏ Output?
- ❏ Is transparency possible?
- ❏ What is the difference between transparency and interpretability?
- ❏ What is autonomy?
- ❏ How is autonomy impacted by algorithms?
- ❏ What is the relationship between transparency and autonomy?

TOPIC	DESCRIPTION/SOURCES
Five Cs	Mike Loukides, Hillary Mason, and DJ Patil, **Ethics and Data Science,** O'Reilly Media (2018), pp. 18-20.

	❑ Consent	"At each step ask whether appropriate and necessary consent has been provided"
	❑ Clarity	"Users must have clarity about what data they are providing, what is going to be done with the data, and any downstream consequences of how their data is used"
Discussion	❑ Consistency	"Trust requires consistency over time. You can't trust someone who is unpredictable."
	❑ Control (and transparency)	"Once you have given your data to a service, you must be able to understand what is happening to your data."
	❑ Consequences (and harm)	"It is essential to ask whether the data that is being collected could cause harm to an individual or a group."

Case Study   *Hiring by Machine*, Princeton Dialogues on AI and Ethics

Discussion
- ❑ How did the Paris system "learn" to disqualify an applicant like Hara?
- ❑ How do you operationalize fairness technically?
- ❑ Should companies have to disclose automated decisioning systems in hiring systems?
- ❑ What are the implications of artificial intelligence/deep learning with respect to consent?