Ethics and Law for Data Science - Assignment 2
By Athika Fatima - 101502209

1) **What are some of the challenges that AI will pose in the judicial system based on these two cases?**

Answer: AI in the judicial system presents several challenges based on the cases of Loomis v. Wisconsin and Robert McDaniel:

- **Lack of Transparency & Due Process Concerns**
  - In Loomis v. Wisconsin, the COMPAS risk assessment tool used in sentencing was a closed-source algorithm, meaning its internal workings were not disclosed.
  - Defendants and their lawyers could not challenge the accuracy, biases, or methodology behind its risk assessment, violating due process rights.
  - AI-based decisions should be explainable, but current proprietary algorithms create a black-box problem in the justice system.
- **Automation Bias & Overreliance on AI**
  - The Supreme Court of Wisconsin justified the use of COMPAS by stating that human judges could ignore its recommendations.
  - However, studies show that human decision-makers tend to trust AI outputs, even when they might be flawed.
  - This overreliance on AI could lead to biased and unfair sentencing.
- **Presumption of Innocence vs. Predictive Policing**
  - In the case of Robert McDaniel, predictive AI flagged him as a potential criminal not because of his actions, but due to his social network and neighborhood.
  - This form of profiling resembles redlining, where AI disproportionately targets marginalized communities.
- **Bias & Discrimination**
  - AI tools like COMPAS have shown to disproportionately label Black individuals as high risk while misidentifying White individuals as low risk.
  - Predictive policing, as seen in McDaniel's case, often reinforces existing systemic biases by targeting poor and racialized communities.
- **Lack of Accountability & Appeals Process**
  - If an AI system makes an error, who is responsible? The software developers? The judge? The state?
  - Individuals affected by AI-driven decisions lack clear avenues to appeal or challenge AI-generated conclusions.
- **Ethical & Human Rights Violations**

- AI is being used to determine who deserves freedom and who should be monitored, raising major human rights concerns.
- The use of AI in criminal justice must align with fairness, accountability, and transparency, or it risks amplifying injustices rather than solving them.

2) What is the impact of algorithms on fundamental liberties?

Answer: Algorithms significantly impact fundamental liberties, often in ways that challenge fairness, privacy, and due process. Here are some key areas where they influence human rights and freedoms:

1. Due Process & Fair Trial Rights:

- AI-driven risk assessment tools (like COMPAS in Loomis v. Wisconsin) influence judicial decisions without transparency, making it difficult for defendants to challenge outcomes.
- The lack of explainability in AI decisions can violate the right to a fair trial and due process if individuals do not understand or cannot contest how a decision was made.

2. Presumption of Innocence & Predictive Policing

- In Robert McDaniel's case, AI flagged him as a potential criminal based on his environment and social connections, not his actions.
- Predictive policing shifts law enforcement focus onto individuals before they commit any crime, reversing the burden of proof and threatening the presumption of innocence.

3. Privacy Violations & Mass Surveillance

- AI-powered surveillance systems collect vast amounts of personal data, often without consent, leading to mass government surveillance and potential violations of privacy rights.
- Examples include facial recognition technology, which has been criticized for racial bias and misuse in authoritarian regimes.

4. Discrimination & Inequality

- AI often reflects and reinforces societal biases, leading to discriminatory decisions in criminal justice, employment, and finance.
- In judicial settings, algorithms like COMPAS have been found to disproportionately classify Black individuals as high-risk, leading to harsher sentencing.
- AI-based hiring tools have also shown bias against women and minority groups.

5. Freedom of Expression & Information Control

- Social media algorithms shape what people see and can suppress certain viewpoints based on biased training data.

- Automated moderation systems may censor dissenting opinions, threatening freedom of speech in political discourse.

6. Right to Freedom & Liberty

- AI is being used to predict and prevent crime, but when misused, it can wrongfully restrict people's movements, freedoms, or even lead to wrongful arrests.
- The use of automated decision-making in parole and sentencing can unfairly deprive individuals of their freedom based on opaque and potentially flawed predictions.

7. Lack of Accountability & Redress Mechanisms

- AI-driven decisions often lack clear accountability—if an algorithm makes an unfair or incorrect decision, who is responsible?
- The absence of clear appeal processes for AI-generated decisions can deny individuals justice when errors occur.

3) How do you think we can make 'algorithms designers and developers (People) more accountable'?

Answer: As AI becomes increasingly integrated into critical decision-making systems, ensuring accountability among algorithm designers and developers is essential. The risks of biased, opaque, and unregulated AI—such as those seen in Loomis v. Wisconsin and the case of Robert McDaniel—demand stronger legal frameworks, ethical standards, and technical safeguards.

1. Legal & Regulatory Measures

Governments must establish AI transparency laws, requiring companies to disclose how their algorithms function and what data they use. The EU AI Act and the U.S. Algorithmic Accountability Act set a precedent for such regulations. Additionally, developers should be held legally responsible when AI systems cause harm, similar to liability laws in medical and consumer industries. Bias audits should be mandatory to identify and mitigate discriminatory AI decisions. Individuals affected by AI-driven judgments, such as parole denials or predictive policing, should have the right to appeal based on clear explanations of the AI's reasoning.

2. Industry Standards & Ethical Development

Developers should undergo AI ethics certification similar to professional licensing in other industries. Ethical frameworks like the IEEE Ethically Aligned Design guidelines should be adopted across organizations. AI teams must also be diverse, as homogeneous

groups often reinforce biases in data and decision-making. Algorithmic Impact Assessments (AIA) should be required before deployment to evaluate risks related to fairness and discrimination.

3. Technical Safeguards for Fairness

To prevent AI from acting as an unchallengeable "black box," Explainable AI (XAI) should be a standard practice, ensuring that decisions can be reviewed and understood. AI must not make final decisions in high-stakes areas like criminal sentencing—human oversight must be integrated into AI workflows. Open-source AI development and peer reviews can help identify flaws before they cause harm. Additionally, developers must use unbiased datasets, as poor data quality often leads to discriminatory outcomes, as seen in facial recognition AI failures.

4. Public Awareness & Oversight

Algorithmic literacy should be promoted so individuals understand their rights when AI affects their lives. Governments should establish AI oversight boards to regulate high-risk AI applications. Protecting whistleblowers—such as Timnit Gebru, who exposed racial bias in AI—can also help maintain ethical AI development.

5. Corporate Responsibility & Incentives

Companies that deploy harmful AI should face fines and penalties, similar to how GDPR enforces data privacy laws. Ethical AI development should be incentivized through government funding and recognition programs. AI testing in algorithmic sandboxes before full-scale deployment can prevent unintended consequences.

4) Draw Ethical concerns from the literature in your case study analysis

Answer: The cases of Loomis v. Wisconsin and Robert McDaniel highlight significant ethical concerns regarding the use of AI in criminal justice. These concerns align with key themes in the literature—

1. Fairness and Validity of Algorithms

One of the major ethical concerns in Loomis v. Wisconsin is the fairness and validity of AI-based risk assessments. The COMPAS algorithm classified Loomis as high-risk based on historical data patterns rather than individualized factors, raising concerns about bias and accuracy. AI models often rely on historical crime data, which may reflect systemic inequalities in policing and sentencing. Similarly, in Robert McDaniel's case, AI flagged him as a risk not based on his actions but his social connections, demonstrating how flawed assumptions can lead to unjust outcomes.

2. Presumption of Innocence

The presumption of innocence is a fundamental legal principle, yet AI-driven risk assessments challenge this notion. In McDaniel's case, he became the subject of police surveillance without committing a crime, simply due to his environment and associations. This predictive policing approach undermines due process, as individuals are treated as potential criminals based on statistical probabilities rather than actual behavior.

3. Right to a Fair Trial

In Loomis v. Wisconsin, the court relied on COMPAS scores for sentencing, raising concerns about whether Loomis received a fair trial. AI models often provide risk assessments without explaining their reasoning, limiting the defendant's ability to challenge them. Without transparency, defendants cannot question the validity of AI-generated evidence, which compromises their right to a fair defense.

4. Principles of Legality

The principle of legality requires that laws and legal decisions be clear and based on established legal norms. The use of proprietary AI algorithms in criminal sentencing conflicts with this principle because defendants cannot access or understand the logic behind their own risk assessments. In Loomis v. Wisconsin, the court admitted that COMPAS used a closed-source algorithm, making it impossible for the defense to scrutinize its fairness or accuracy. This raises concerns about whether AI decisions align with established legal frameworks.

5. Principles of Non-Discrimination and Equality

Both case studies highlight how AI can reinforce systemic discrimination. The COMPAS algorithm has been shown to disproportionately label Black defendants as high-risk, reflecting historical racial biases in criminal justice. Similarly, Robert McDaniel was flagged as a risk due to his residence in a high-crime area, reinforcing socioeconomic and racial profiling. When AI replicates existing inequalities, it violates the principle of non-discrimination and fails to uphold equal treatment under the law.

6. Opacity of AI / Explainability / Transparency

A critical ethical issue is the opacity of AI systems, making it difficult for defendants, lawyers, and judges to understand or challenge AI-based decisions. In Loomis v. Wisconsin, the court justified its ruling by claiming COMPAS was one of several factors in sentencing, but in reality, judges may place undue weight on AI-generated risk scores due to automation bias. The inability to audit or question the AI's decision-making process undermines accountability and erodes trust in the justice system.