

## Part 2: Case Study Analysis

In this section, we delve into two critical case studies that highlight the ethical complexities and societal impact of AI technologies.



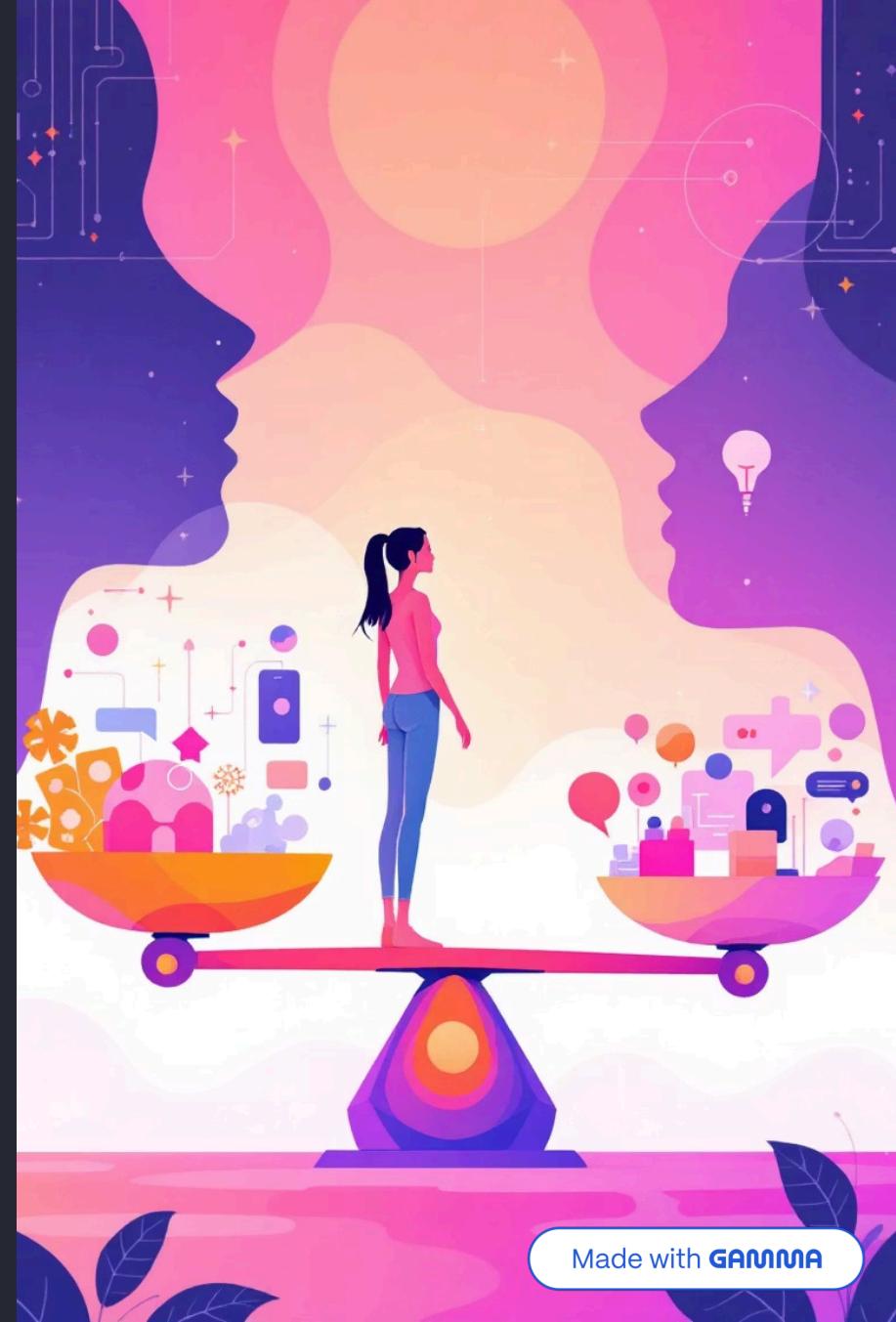
# Case Study 1: Biased Hiring Tool

We will analyse a prominent example of algorithmic bias in recruitment, examining its origins and proposing actionable solutions.



## Scenario: Amazon's AI Recruiting Tool

Amazon's experimental AI recruiting tool demonstrated a clear bias against female candidates, downgrading résumés that included words associated with women, such as "women's chess club captain."



## Identifying the Source of Bias

### Training Data

The AI was trained on historical résumés primarily submitted by male applicants, reflecting a male-dominated industry.

### Model Design

The algorithm learned to associate certain keywords and patterns, inadvertently penalising attributes common in female applications.

### Feature Selection

Potentially discriminatory features, even if seemingly neutral, were included in the model, leading to biased outcomes.

## Three Fixes for Fairness



### Data Re-evaluation

Curate and balance training data to ensure diverse representation across all protected characteristics, removing historical biases.



### Algorithmic Audit & Redesign

Implement bias detection algorithms and adjust the model to explicitly ignore or re-weight sensitive attributes, using fairness-aware machine learning techniques.



### Human-in-the-Loop

Integrate human oversight at critical decision points to review and override potentially biased AI recommendations, especially for shortlisting.

## Metrics to Evaluate Fairness



### Equality of Opportunity

Measure false negative rates across demographic groups. Aim for similar rates, ensuring no group is unfairly overlooked.



### Demographic Parity

Compare selection rates across different demographic groups. Strive for equitable representation in the hiring pool.



### Predictive Equality

Assess false positive rates across groups. Ensure the tool doesn't falsely flag qualified candidates from specific demographics.

## Case Study 2: Facial Recognition in Policing

This case examines the ethical dilemmas arising from the use of facial recognition technology in law enforcement, particularly its disproportionate impact on minority groups.



## Scenario: Misidentification of Minorities

Facial recognition systems have been found to exhibit higher error rates when identifying individuals from minority ethnic backgrounds and women, leading to concerns about wrongful accusations.



# Ethical Risks of Facial Recognition



## Wrongful Arrests

Inaccurate matches can lead to false accusations, potentially resulting in wrongful arrests and severe legal consequences for innocent individuals.



## Privacy Violations

Continuous surveillance and data collection infringe on fundamental privacy rights, creating a chilling effect on public spaces and free assembly.



## Amplified Bias

Existing societal biases can be amplified by flawed technology, leading to disproportionate scrutiny and targeting of specific communities.



## Lack of Transparency

The opaque nature of these systems makes it difficult to challenge errors or understand how decisions are made, undermining due process.

# Policies for Responsible Deployment

## Strict Regulation

- Implement clear legal frameworks governing its use.
- Mandate independent oversight and regular audits.
- Prohibit use in certain sensitive contexts (e.g., public protests).



## Transparency & Accountability

- Require public disclosure of facial recognition deployments.
- Establish clear accountability mechanisms for misidentifications.
- Ensure robust data security and retention policies.

