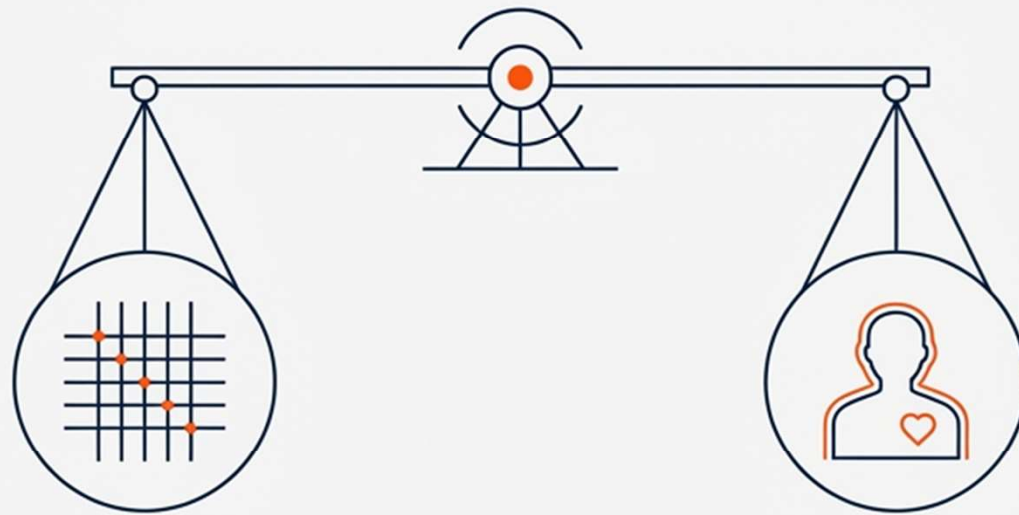


# Bias dan Keadilan dalam Kecerdasan Buatan

Menavigasi Tantangan Etis di Era Algoritma



# Ketegangan Antara Efisiensi dan Keadilan

## Definisi Inti

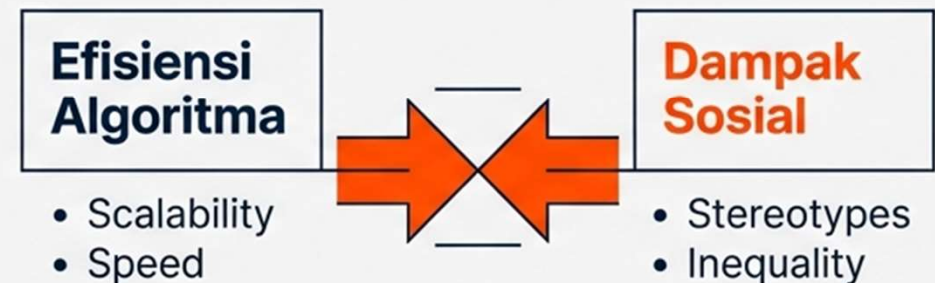
### Bias

Diskriminasi sistematis yang merugikan individu atau kelompok tertentu.

### Keadilan (Fairness)

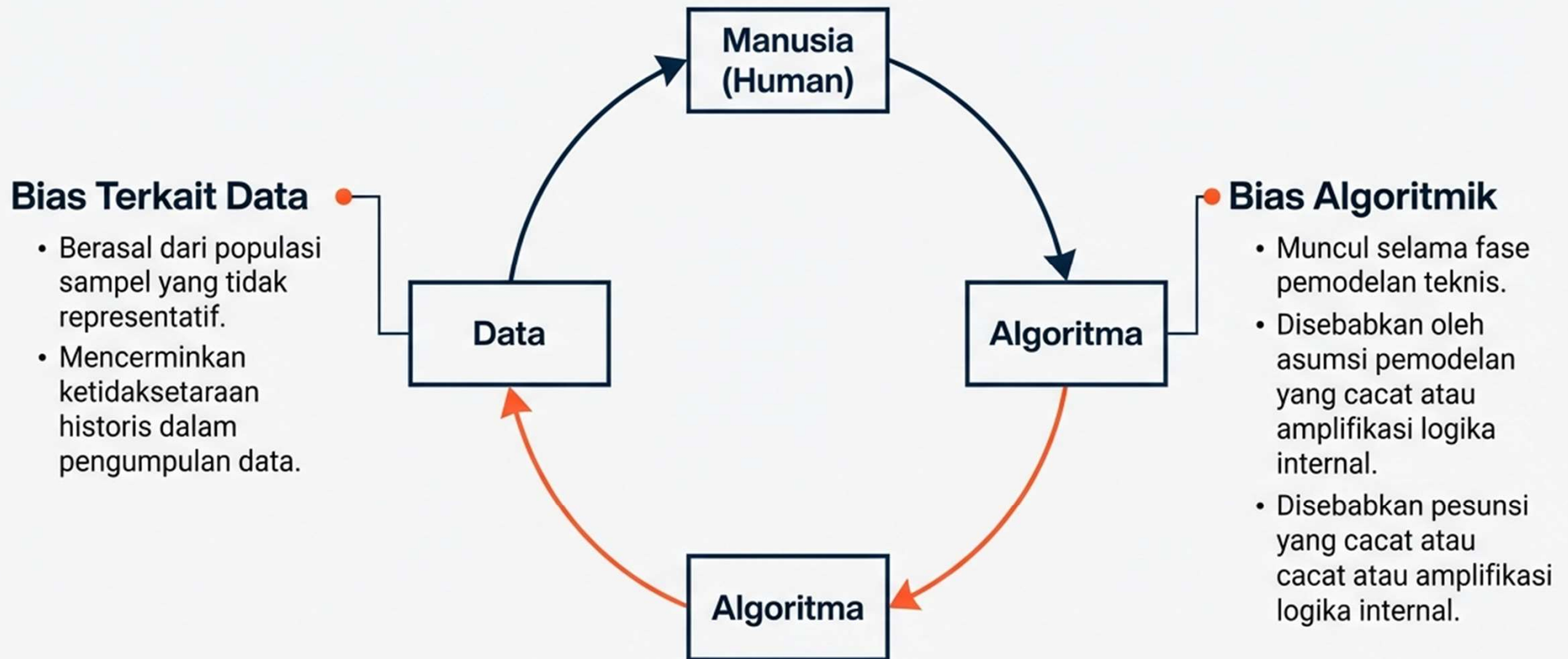
Pengembangan dan penerapan algoritma yang setara, memastikan tidak ada demografi yang dirugikan.

## Konflik Utama

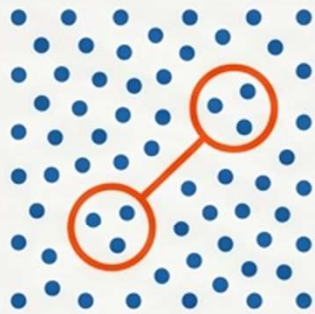


**Konsekuensi bias bukan sekadar kesalahan teknis, melainkan manifestasi ketidakadilan nyata.**

# Asal Mula Masalah: Lingkaran Bias



# Taksonom I: Pengumpulan dan Seleksi Data



## Sampling Bias

Data tidak mewakili populasi yang dituju.

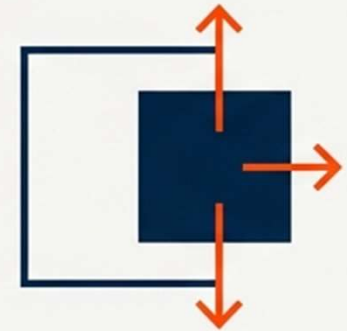
Contoh: Studi kesehatan yang hanya mencakup masyarakat perkotaan.



## Selection Bias

Kesalahan sistematis di mana seleksi partisipan tidak acak.

Termasuk: Self-selection & Attrition bias.



## Exclusion Bias

Penghilangan sistematis kelompok tertentu.

Contoh: Survei digital yang mengabaikan lansia tanpa akses internet.



# Taksonomi II: Interpretasi Sejarah dan Budaya

## Historical Bias

Distorsi akibat perspektif subjektif masa lalu.

**Kasus:** Alat perekrutan AI yang dilatih dengan data puluhan tahun yang didominasi pria.



## Cultural Bias

Menilai fenomena melalui lensa budaya sendiri.



**Krisis WEIRD: Riset AI terlalu berpusat pada masyarakat Barat, Terdidik, Industrialisasi, Kaya, dan Demokratis.**



## Confirmation Bias

Mengutamakan informasi yang mendukung keyakinan awal.

# Taksonomi III: Pengukuran dan Pelaporan



## Measurement Bias (Bias Pengukuran)

Kesalahan sistematis dalam pencatatan data atau kalibrasi instrumen.

- ✦ Observer Bias: Predisposisi peneliti mempengaruhi data (misal: kesalahan pelabelan gender medis).

## Reporting Bias (Bias Pelaporan)

Pelaporan temuan secara selektif, seringkali hanya mengunggulkan hasil positif.

**Solusi:** Pengungkapan semua hasil studi, termasuk yang tidak konklusif.

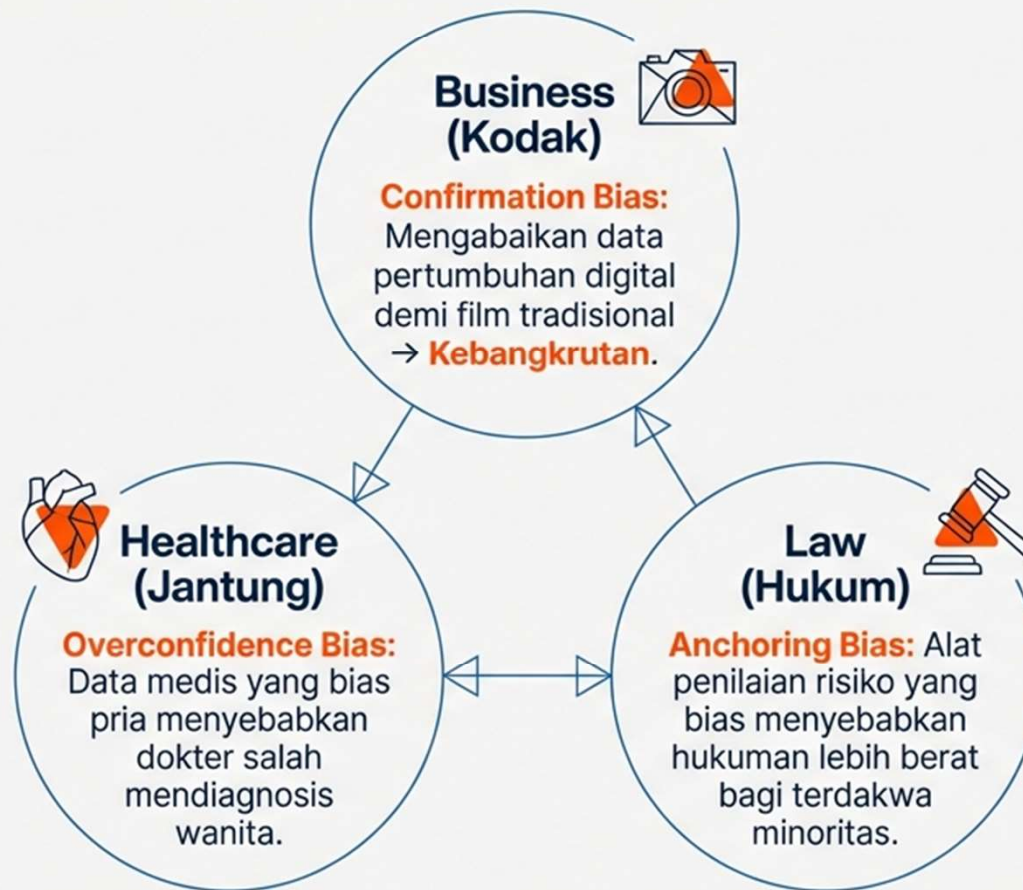
# Strategi Mitigasi untuk Dataset

<b>Stratified Sampling</b>	<p>Membagi populasi ke dalam strata dan mengambil sampel acak dari masing-masing.</p> <p><b>Dampak: Mencegah penghapusan kelompok minoritas.</b></p>
<b>Randomization</b>	<p>Memastikan setiap individu memiliki peluang seleksi yang sama.</p> <p><b>Dampak: Menetralkan predisposisi peneliti.</b></p>
<b>Propensity Score Matching</b>	<p>Memasangkan partisipan berdasarkan sifat serupa.</p> <p><b>Dampak: Menyimulasikan keadilan dalam studi observasional.</b></p>
<b>Oversampling &amp; Undersampling</b>	<p>Menyesuaikan ukuran kelompok untuk menyeimbangkan representasi.</p> <p><b>Dampak: Memberikan bobot setara pada suara minoritas.</b></p>



# Biaya dari Keputusan yang Bias

Output AI bertemu heuristik manusia.



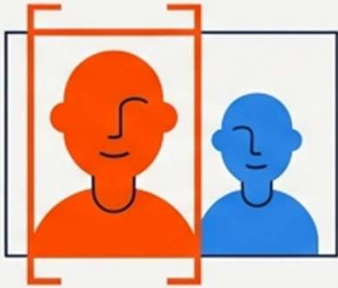


# Studi Kasus Kegagalan Teknologi

## Twitter (2020)

### Masalah

Algoritma cropping gambar memprioritaskan wajah kulit putih.



### Penyebab

Model 'saliency' meniru atensi manusia yang mewarisi **bias rasial**.

## Amazon (2018)

### Masalah

Alat perekrutan AI mempenalti resume yang mengandung kata 'wanita'.



### Penyebab

Data historis 10 tahun mencerminkan **dominasi pria** di industri teknologi.

## Google (2019)

### Masalah

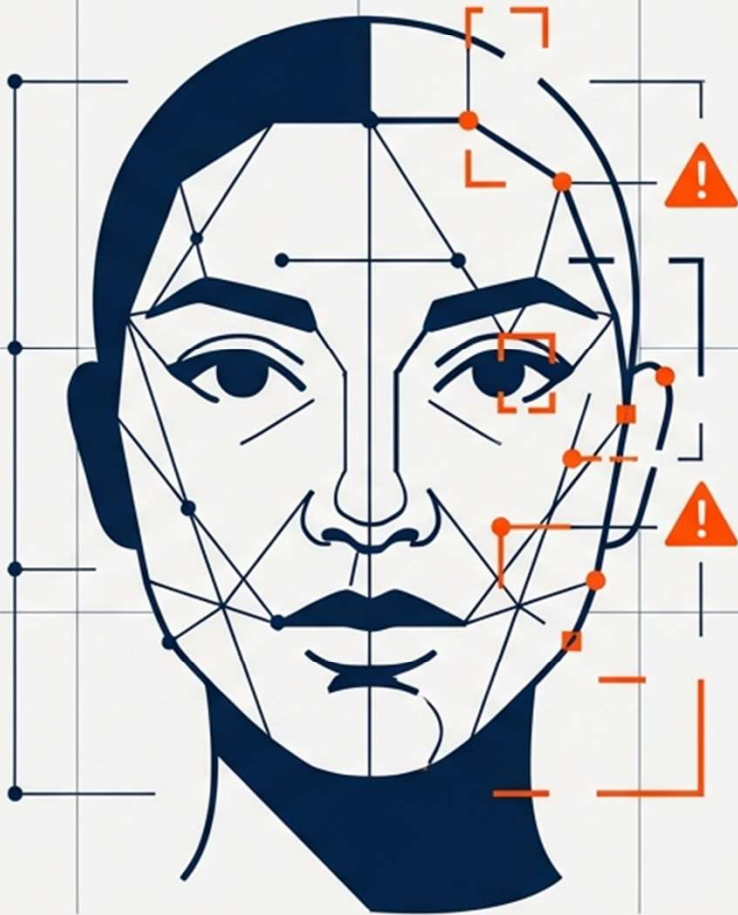
Dewan Etika (ATEAC) **dibubarkan dalam dua minggu**.



### Pelajaran

**Pengawasan etis** membutuhkan **penyelarasan ideologis**, bukan sekadar formalitas komersial.

# Sorotan: Proyek Gender Shades



**Penemuan:** Proyek ini mengungkap disparitas ras dan gender yang signifikan dalam alat analisis wajah komersial.

**Dampak:** Tingkat kesalahan substansial lebih tinggi pada wanita berkulit gelap.

**Deteksi yang ketat mengarah pada mitigasi yang bermakna.**



# Deteksi dan Manajemen Algoritma



## Teknik Deteksi (Detection)

**Audit Studies:** Pengujian pihak ketiga yang independen.

**Disparate Impact Analysis:** Uji statistik untuk hasil yang berbeda berdasarkan ras/gender.

**Explainability Tools:** Perangkat lunak untuk mendekonstruksi 'kotak hitam' AI.



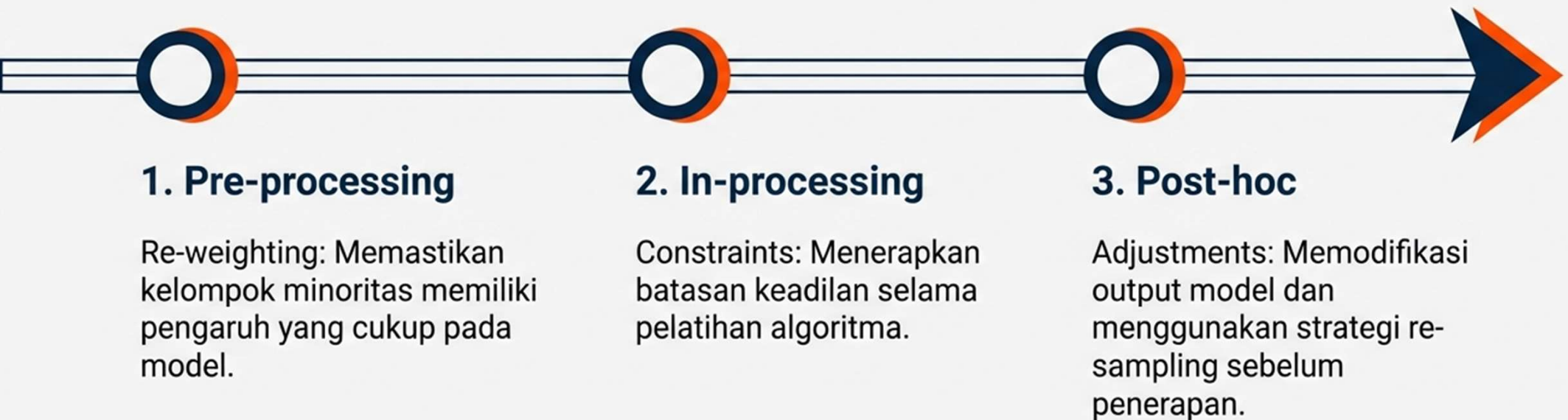
## Strategi Manajemen (Management)

**Diverse Training Data:** Refleksi heterogenitas dunia nyata.

**Fairness Constraints:** Pemrograman kriteria keadilan secara eksplisit.

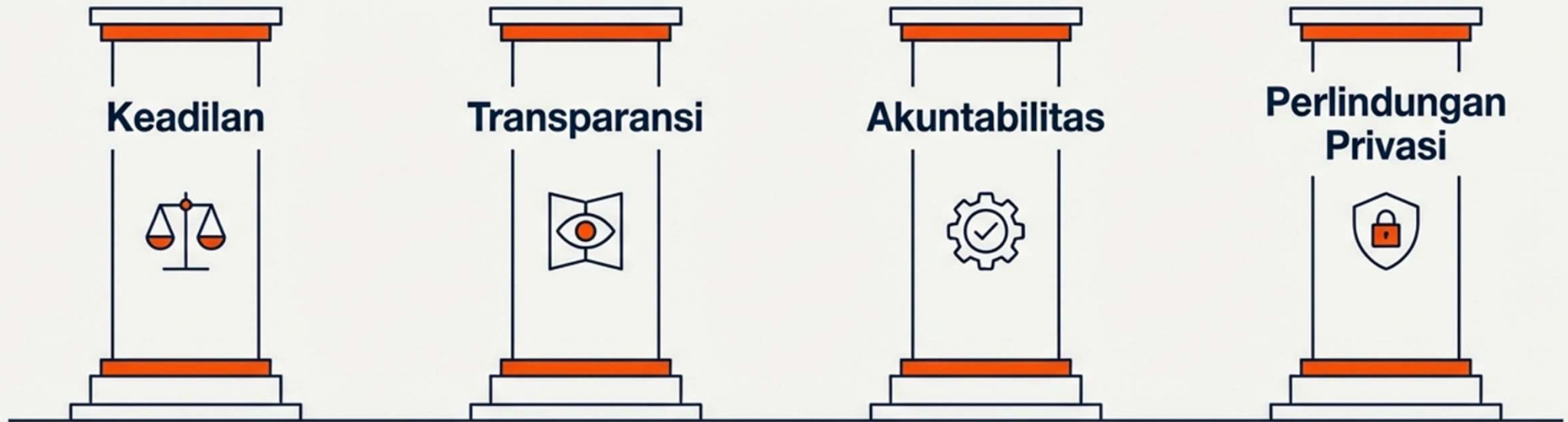
**Fairness-Aware Modeling:** Perlindungan teknis yang menyesuaikan bias saat memproses data baru.

# Intervensi Teknis di Sepanjang Siklus Hidup





# Ekosistem Regulasi dan Etika



## Lansekap Regulasi



**GDPR (EU): “Hak untuk Penjelasan”**  
bagi warga.



**HIPAA (Healthcare):**  
Standar non-diskriminasi ketat untuk data pasien.



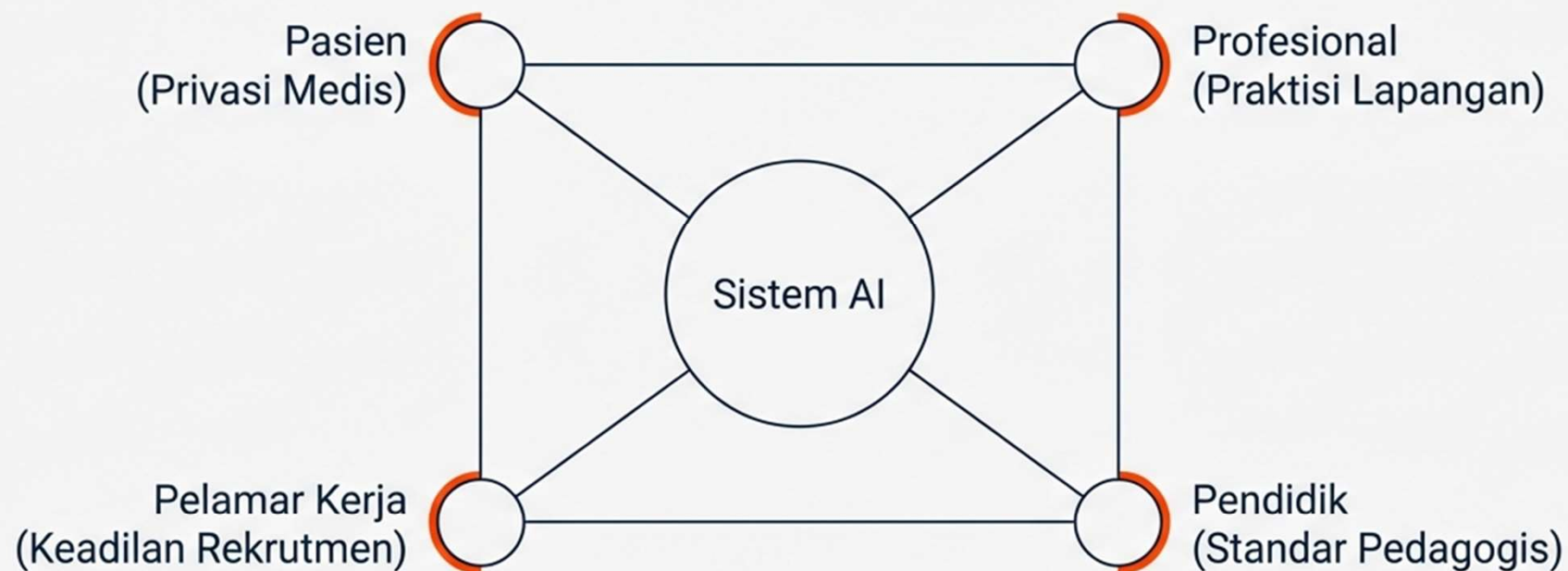
**Kasus HUD vs Facebook:**  
Preseden hukum untuk akuntabilitas iklan.



**Inovasi Pemerintah:**  
AI Registries di Amsterdam & Helsinki.

# Keterlibatan Pemangku Kepentingan

Keterlibatan adalah keharusan moral dan praktis.



**Strategi:** Pelibatan sejak fase desain & Komunikasi transparan tentang batasan sistem

# **Menuju Masa Depan yang Bertanggung Jawab**

**Kita harus bergerak  
melampaui efisiensi semata  
menuju AI yang bertanggung  
jawab, etis, adil.**

- ◆ Gunakan pendekatan multidisiplin (Sosiolog + Insinyur).
- ◆ Lakukan audit independen secara berkala.
- ◆ Prioritaskan Data yang Beragam atau gunakan data sintetik.