



Prinsip AI

Tujuan membangun AI
yang bermanfaat

Tanggung Jawab: Prinsip Kami

Terlepas dari optimisme kami terhadap potensi AI, kami menyadari bahwa teknologi canggih dapat menimbulkan berbagai tantangan penting yang harus ditangani dengan jelas, cermat, dan tegas. Prinsip AI ini menjelaskan komitmen kami untuk mengembangkan teknologi secara bertanggung jawab serta menetapkan area aplikasi tertentu yang tidak akan kami kembangkan.

Tujuan aplikasi AI

1. Bermanfaat untuk kehidupan sosial.

Jangkauan luas teknologi baru semakin menyentuh masyarakat secara keseluruhan. Kemajuan dalam AI akan memiliki dampak transformatif dalam berbagai bidang, termasuk kesehatan, keamanan, energi, transportasi, manufaktur, dan hiburan. Karena kami mempertimbangkan potensi pengembangan dan penggunaan teknologi AI, kami akan memperhitungkan berbagai faktor sosial dan ekonomi, dan akan melanjutkan jika kami yakin bahwa keseluruhan manfaatnya secara substansial lebih besar dibandingkan risiko dan kerugian yang dapat diperkirakan.

AI juga meningkatkan kemampuan kami untuk memahami makna konten dalam skala besar. Kami akan berupaya menyediakan informasi yang berkualitas tinggi dan akurat menggunakan AI, sambil tetap menghormati norma-norma budaya, sosial, dan hukum di semua negara tempat kami beroperasi. Selain itu, kami akan terus mengevaluasi secara cermat kapan teknologi kami akan tersedia untuk keperluan non-komersial.

2. Menghindari timbulnya atau diterapkannya bias yang tidak adil.

Algoritme dan kumpulan data AI dapat mencerminkan, menerapkan, atau mengurangi bias yang tidak adil. Kami mengakui bahwa membedakan bias yang adil dari yang tidak adil tidak selalu mudah, dan berbeda di setiap budaya dan masyarakat. Kami akan berupaya untuk menghindari dampak yang tidak adil pada banyak orang, terutama yang terkait dengan karakteristik sensitif seperti ras, etnis, gender, kebangsaan, pendapatan, orientasi seksual, kemampuan, dan keyakinan agama atau politik.

3. Dibuat dan diuji keamanannya.

Kami akan terus mengembangkan dan menerapkan praktik keselamatan dan keamanan yang kuat untuk menghindari hasil yang tidak diinginkan yang menimbulkan risiko bahaya. Kami akan mendesain sistem AI agar selalu Berhati-hati, dan berupaya mengembangkannya sesuai dengan praktik terbaik dalam riset keamanan AI. Dalam kasus yang sesuai, kami akan menguji teknologi AI dalam lingkungan yang terbatas dan memantau pengoperasiannya setelah deployment.

4. Dapat dipertanggungjawabkan kepada orang lain.

Kami akan mendesain sistem AI yang menyediakan peluang yang tepat untuk menerima masukan, penjelasan yang relevan, dan pengajuan banding. Teknologi AI kami akan tunduk pada kontrol dan kendali manusia.

5. Menerapkan prinsip-prinsip desain privasi.

Kami akan menerapkan prinsip-prinsip privasi kami dalam pengembangan dan penggunaan teknologi AI. Kami akan memberikan kesempatan untuk mendapatkan pemberitahuan dan izin, mendorong arsitektur dengan perlindungan privasi, serta menyediakan transparansi dan kontrol yang sesuai atas penggunaan data.

6. Menjunjung tinggi standar keunggulan ilmiah.

Inovasi teknologi berakar pada metode ilmiah dan komitmen guna membuka jalan untuk riset, ketelitian intelektual, integritas, dan kolaborasi. Alat AI memiliki potensi untuk mendorong hadirnya cakupan baru dari riset ilmiah dan pengetahuan di bidang-bidang penting seperti biologi, kimia, kedokteran, dan ilmu lingkungan. Kami menginginkan standar keunggulan ilmiah yang tinggi seiring kami berupaya memajukan pengembangan AI.

7. Tersedia untuk penggunaan yang sesuai dengan prinsip-prinsip berikut.

Banyak teknologi memiliki beberapa kegunaan. Kami akan bekerja untuk membatasi aplikasi yang berpotensi berbahaya atau merusak. Saat kami mengembangkan dan men-deploy teknologi AI, kami akan mengevaluasi kemungkinan penggunaan dengan mempertimbangkan faktor-faktor berikut:

- Tujuan dan penggunaan utama: tujuan utama dan kemungkinan penggunaan teknologi dan aplikasi, termasuk seberapa relevan solusi tersebut dengan atau dapat diadaptasikan untuk penggunaan yang berbahaya
- Sifat dasar dan keunikan: apakah kami menyediakan teknologi yang unik atau yang lebih umum tersedia
- Skala: apakah penggunaan teknologi ini akan memiliki pengaruh yang signifikan
- Sifat keterlibatan Google: apakah kami menyediakan alat dengan tujuan umum, mengintegrasikan alat untuk pelanggan, atau mengembangkan solusi kustom

Aplikasi AI yang tidak akan kami kembangkan

Selain tujuan di atas, kami tidak akan mendesain atau men-deploy AI dalam area aplikasi berikut:

1. Teknologi yang menyebabkan atau cenderung membahayakan secara keseluruhan. Jika ada risiko bahaya yang besar, kami hanya akan melanjutkan ketika kami yakin bahwa manfaatnya secara substansial jauh lebih besar daripada risikonya, dan akan menerapkan batasan keamanan yang sesuai.
2. Senjata atau teknologi lain yang penerapan atau tujuan utamanya adalah untuk menyebabkan atau secara langsung menimbulkan cedera pada banyak orang.
3. Teknologi yang mengumpulkan atau menggunakan informasi untuk pengawasan yang melanggar norma-norma yang diterima secara internasional.
4. Teknologi yang tujuannya bertentangan dengan prinsip-prinsip hukum internasional dan hak asasi manusia yang diterima secara luas.

Dengan semakin banyaknya pengalaman kami dalam bidang ini, daftar ini dapat berubah.