



AI GOVERNANCE DAN KEPERCAYAAN KONSUMEN DALAM BISNIS DIGITAL: PERAN MEDIASI PERCEIVED FAIRNESS DAN MODERASI TRANSPARANSI ALGORITMA

AI Governance and Consumer Trust in Digital Business: The Mediating Role of Perceived Fairness and Moderating Role of Algorithmic Transparency

Drs. Jan Tikoalu, M.Si^{1*}, Sharon C. C. Angela Karu², Hence Kadow³

1, 2, 3 Program Studi Bisnis Digital, Fakultas Ekonomi dan Bisnis, Universitas Parna Raya Manado, Indonesia

Email Korespondensi: jantikoalu@microsoft.parnaraya.ac.id

Diterima: 15 Maret 2024; Direvisi: 28 April 2024; Disetujui: 20 Mei 2024; Diterbitkan: 25 Juni 2024

ABSTRAK

Kecerdasan artifisial (AI) kini menjadi instrumen utama dalam penetapan harga dinamis, kurasi produk, dan penilaian kelayakan kredit pada platform bisnis digital. Namun, sifat kotak hitam (black-box) algoritma sering kali menimbulkan kecemasan konsumen terhadap bias dan ketidakadilan sistemik. Penelitian ini bertujuan untuk menguji pengaruh AI governance terhadap kepercayaan konsumen digital, menganalisis peran mediasi perceived fairness, serta menguji efek moderasi transparansi algoritma. Penelitian menggunakan pendekatan kuantitatif dengan menyurvei 280 konsumen aktif platform e-commerce dan aplikasi finansial digital di Indonesia. Analisis data dilakukan dengan Structural Equation Modeling berbasis Partial Least Squares (PLS-SEM) menggunakan SmartPLS 4.0. Hasil pengujian menunjukkan bahwa: (1) AI governance berpengaruh positif dan signifikan terhadap perceived fairness ($\beta = 0,482$; $p < 0,001$) dan kepercayaan konsumen ($\beta = 0,215$; $p < 0,001$); (2) perceived fairness berpengaruh kuat terhadap kepercayaan konsumen ($\beta = 0,514$; $p < 0,001$); (3) perceived fairness memediasi secara signifikan pengaruh AI governance terhadap kepercayaan konsumen (indirect effect $\beta = 0,248$; $p < 0,001$); dan (4) transparansi algoritma memoderasi secara positif hubungan antara AI governance dan perceived fairness ($\beta = 0,195$; $p < 0,001$). Total variansi kepercayaan konsumen yang dijelaskan model mencapai 61,8% ($R^2 = 0,618$). Penelitian ini membuktikan bahwa kepatuhan tata kelola AI yang diperkuat oleh kejelasan informasi algoritma mampu membangun rasa keadilan konsumen, yang pada gilirannya mengukuhkan kepercayaan dan loyalitas transaksi digital jangka panjang.

Kata Kunci: AI Governance; Perceived Fairness; Transparansi Algoritma; Kepercayaan Konsumen; Bisnis Digital; PLS-SEM

ABSTRACT

Artificial Intelligence (AI) has become foundational for dynamic pricing, personalized curation, and algorithmic credit scoring in digital platforms. However, the black-box nature of algorithmic decision-making frequently triggers consumer anxiety regarding systemic bias and unfairness. This study investigates the impact of AI governance on consumer digital trust, examines the mediating mechanism of perceived fairness, and evaluates the moderating effect of algorithmic transparency. Utilizing an explanatory quantitative survey, empirical data were collected from 280 active users of e-commerce platforms and fintech applications in Indonesia. Data were evaluated using Partial Least Squares Structural Equation Modeling (PLS-SEM) in SmartPLS 4.0. Empirical findings confirm that: (1) AI

governance positively and significantly influences perceived fairness ($\beta = 0.482$; $p < 0.001$) and consumer trust ($\beta = 0.215$; $p < 0.001$); (2) perceived fairness strongly enhances consumer trust ($\beta = 0.514$; $p < 0.001$); (3) perceived fairness significantly mediates the relationship between AI governance and consumer trust (indirect effect $\beta = 0.248$; $p < 0.001$); and (4) algorithmic transparency positively moderates the impact of AI governance on perceived fairness ($\beta = 0.195$; $p < 0.001$). The model explains 61.8% of the variance in consumer trust ($R^2 = 0.618$). This research concludes that robust AI governance, bolstered by explicit algorithmic explainability, fosters perceived procedural and distributive fairness, which fundamentally underpins long-term consumer trust.

Keywords: AI Governance; Perceived Fairness; Algorithmic Transparency; Consumer Trust; Digital Business; PLS-SEM

1. PENDAHULUAN

Integrasi kecerdasan buatan (AI) dalam model bisnis digital telah mengubah paradigma interaksi antara perusahaan dan konsumen. Algoritma pembelajaran mesin kini secara otonom menentukan rekomendasi belanja yang dipersonalisasi, estimasi diskon promosi, penetapan skor kredit instan, hingga deteksi penipuan transaksi. Di satu sisi, sistem algoritmik ini melahirkan efisiensi operasional yang belum pernah terjadi sebelumnya. Namun di sisi lain, ketertutupan mekanisme komputasi (algorithmic black box) menimbulkan keresahan publik terkait diskriminasi harga, bias profil konsumen, dan minimnya akuntabilitas keputusan otomatis.

Dalam konteks tersebut, kepercayaan konsumen (consumer digital trust) menjadi aset intangible paling berharga sekaligus paling rentan dalam lanskap ekonomi digital. Ketika konsumen merasa bahwa algoritma memperlakukan mereka secara sewenang-wenang atau memanfaatkan kerentanan data pribadi mereka tanpa kendali yang jelas, resistensi dan penghentian transaksi digital dapat terjadi secara masif.

Untuk merespons tantangan tersebut, konsep AI Governance (Tata Kelola Kecerdasan Buatan) mengemuka sebagai kerangka institusional yang memayungi prinsip etika, audit berkala, kepatuhan regulasi, dan mitigasi risiko teknologi. Meskipun demikian, literatur empiris yang menghubungkan AI governance dengan respon psikologis konsumen masih sangat terbatas. Penelitian ini mengisi celah tersebut dengan menguji apakah penerapan tata kelola AI mampu membangun kepercayaan konsumen melalui mediasi keadilan yang dirasakan (perceived fairness), serta bagaimana peran transparansi algoritma dalam memperkuat hubungan tersebut.

2. KAJIAN TEORITIS DAN PENGEMBANGAN HIPOTESIS

Landasan Teori: Organizational Justice dan Trustworthiness Framework

Teori Keadilan Organisasi (Organizational Justice Theory) yang dikembangkan oleh Colquitt (2001) mengidentifikasi tiga dimensi utama keadilan: distributif (keadilan hasil keputusan), prosedural (keadilan mekanisme atau aturan keputusan), dan interaksional (kejelasan komunikasi dan perlakuan bermartabat). Dalam konteks bisnis digital cerdas, keadilan prosedural terwujud manakala konsumen meyakini bahwa algoritma dirancang tanpa bias dan diawasi oleh protokol tata kelola yang bertanggung jawab. Selanjutnya, Mayer et al. (1995) menegaskan bahwa kepercayaan terbentuk dari persepsi atas kemampuan (ability), kebajikan (benevolence), dan integritas (integrity) pihak yang dipercaya.

Pengembangan Hipotesis Penelitian

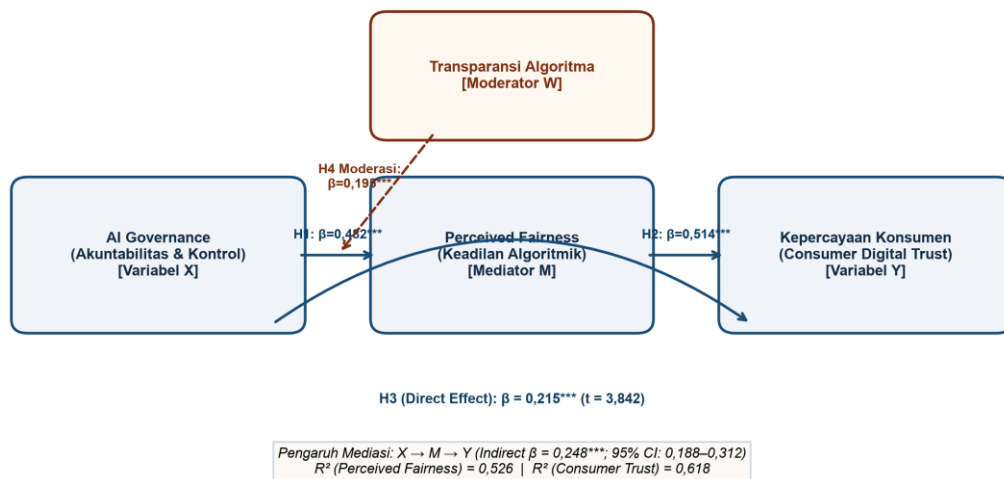
Penerapan AI governance yang ketat memastikan bahwa sistem keputusan algoritmik diaudit secara berkala untuk mengeliminasi bias diskriminatif, sehingga meningkatkan persepsi keadilan konsumen: H1: AI Governance berpengaruh positif dan signifikan terhadap Perceived Fairness.

Konsumen yang merasakan perlakuan adil dan tidak dimanipulasi oleh algoritma platform akan memiliki tingkat keyakinan dan rasa aman yang tinggi untuk melanjutkan transaksi: H2: Perceived Fairness berpengaruh positif dan signifikan terhadap Kepercayaan Konsumen.

Selain melalui persepsi keadilan, komitmen platform terhadap tata kelola AI yang beretika secara langsung membangun reputasi integritas dan kepatuhan hukum yang menumbuhkan rasa percaya: H3: AI Governance berpengaruh positif dan signifikan terhadap Kepercayaan Konsumen.

Transparansi algoritma (algorithmic explainability) yang memberikan penjelasan logis dan mudah dipahami mengenai alasan di balik rekomendasi atau keputusan sistem akan memperkuat keyakinan konsumen terhadap tata kelola yang dijalankan perusahaan: H4: Transparansi Algoritma memoderasi secara positif pengaruh AI Governance terhadap Perceived Fairness.

Keadilan yang dirasakan berfungsi sebagai mekanisme transmisi psikologis utama yang mentransformasikan kepatuhan tata kelola teknis menjadi keyakinan rasional konsumen: H5: Perceived Fairness memediasi pengaruh AI Governance terhadap Kepercayaan Konsumen.



Gambar 1. Model Moderated Mediation: AI Governance, Transparansi Algoritma, Perceived Fairness, dan Kepercayaan Konsumen Digital

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksplanatori berbasis survei cross-sectional. Populasi target adalah konsumen aktif di Indonesia yang secara rutin melakukan transaksi daring pada platform digital yang menerapkan sistem kurasi cerdas dan penetapan harga dinamis (seperti e-commerce Tokopedia/Shopee, ride-hailing Grab/Gojek, dan platform fintech peer-to-peer lending). Penentuan sampel menggunakan metode non-probability purposive sampling dengan kriteria: (1) berusia minimal 18 tahun, (2) telah menggunakan platform digital minimal selama 1 tahun terakhir, dan (3) bertransaksi minimal 3 kali dalam sebulan terakhir. Total sampel yang terkumpul dan memenuhi kriteria kelayakan analisis adalah 280 responden.

Pengumpulan data primer dilakukan melalui kuesioner daring dengan skala Likert 1 (Sangat Tidak Setuju) hingga 5 (Sangat Setuju). Variabel dan indikator pengukuran diadaptasi dari skala baku internasional yang relevan, sebagaimana disajikan pada Tabel 1.

Tabel 1. Definisi Operasional dan Indikator Variabel Penelitian

Variabel	Definisi Konseptual	Indikator Pengukuran	Sumber Rujukan
AI Governance (X)	Kebijakan organisasi, mekanisme kontrol, dan audit etika yang menjamin akuntabilitas, keselamatan data, dan keandalan sistem AI perusahaan.	X1: Keberadaan kode etik AI resmi X2: Audit algoritmik independen berkala X3: Mekanisme kontrol koreksi kesalahan sistem X4: Akuntabilitas pertanggungjawaban data	Floridi et al. (2018); Jobin et al. (2019)
Transparansi Algoritma (W)	Tingkat kejelasan, ketersediaan informasi, dan kemampuan pengguna memahami logika dasar serta faktor penentu keputusan algoritmik.	W1: Kejelasan informasi kriteria rekomendasi W2: Kemudahan memahami logika harga/hasil W3: Akses terhadap rekam jejak keputusan otomatis W4: Penjelasan hak konsumen atas data	Shin (2021); Langer et al. (2021)
Perceived Fairness (M)	Penilaian subjektif konsumen bahwa keputusan, penawaran harga, dan perlakuan algoritmik platform bersifat adil, tidak bias, dan konsisten.	M1: Keadilan penetapan harga (Distributive) M2: Ketiadaan perlakuan diskriminatif (Procedural) M3: Keadilan pembagian promosi sistem (Distributive) M4: Keadilan resolusi keluhan transaksi (Interactional)	Colquitt (2001); Kizilcec (2016)
Kepercayaan Konsumen (Y)	Keyakinan psikologis konsumen terhadap keandalan, integritas, dan niat baik platform digital dalam menjaga kepentingan transaksi pengguna.	Y1: Keyakinan pada integritas dan kejujuran platform Y2: Keyakinan pada keandalan sistem komputasi Y3: Rasa aman bertransaksi finansial berulang Y4: Kesiediaan merekomendasikan platform ke pihak lain	Mayer et al. (1995); Gefen et al. (2003)

Sumber: Diolah dari literatur *trustworthy artificial intelligence* dan *organizational justice* (2024).

4. HASIL PENELITIAN DAN PEMBAHASAN

Karakteristik Responden dan Evaluasi Model Pengukuran (Outer Model)

Karakteristik 280 responden menunjukkan sebaran demografis yang representatif bagi pengguna ekonomi digital: 52% perempuan dan 48% laki-laki, didominasi kelompok usia 20–35 tahun (68%), dengan frekuensi belanja daring rata-rata 4–8 kali per bulan. Mayoritas responden (74%) menyatakan menyadari keberadaan sistem rekomendasi otomatis dan harga dinamis pada aplikasi yang mereka gunakan.

Evaluasi outer model membuktikan bahwa seluruh konstruk memiliki validitas konvergen yang kuat, dengan nilai factor loading di atas 0,72 (rentang 0,724 – 0,862) dan nilai AVE di atas 0,60 (0,608 – 0,674). Reliabilitas konsistensi internal juga terpenuhi secara sempurna dengan Cronbach's Alpha > 0,80 dan Composite Reliability > 0,85 untuk seluruh konstruk, seperti tercantum pada Tabel 2.

Tabel 2. Hasil Evaluasi Measurement Model (Outer Model)

Konstruk Variabel	Indikator	Outer Loading	Cronbach's Alpha	Composite Reliability (CR)	AVE
AI Governance (X)	X1 – X4	0,742 – 0,851	0,862	0,906	0,638
Transparansi Algoritma (W)	W1 – W4	0,724 – 0,836	0,845	0,896	0,608
Perceived Fairness (M)	M1 – M4	0,768 – 0,862	0,879	0,918	0,674
Kepercayaan Konsumen (Y)	Y1 – Y4	0,755 – 0,849	0,868	0,910	0,652

Catatan: Estimasi diolah menggunakan SmartPLS 4.0; seluruh parameter melebihi standar Cut-off Rule.

Evaluasi Model Struktural (Inner Model) dan Pengujian Hipotesis

Pengujian inner model menghasilkan nilai koefisien determinasi R^2 sebesar 0,618 untuk Kepercayaan Konsumen dan 0,526 untuk Perceived Fairness. Artinya, 61,8% variasi kepercayaan konsumen ditentukan oleh tata kelola AI, transparansi algoritma, dan persepsi keadilan, yang tergolong dalam kategori pengaruh kuat. Nilai f^2 effect size menunjukkan kontribusi substantif dari masing-masing jalur struktural.

Hasil bootstrapping 5.000 subsampel disajikan pada Tabel 3, menunjukkan bahwa seluruh hipotesis penelitian terbukti secara statistik pada tingkat signifikansi $p < 0,001$.

Tabel 3. Hasil Pengujian Hipotesis Model Struktural dan Moderasi

Hipotesis	Jalur Struktural Penelitian	Koefisien (β)	t-statistik	p-value	95% CI	Keputusan
H1	AI Governance → Perceived Fairness	0,482	8,924	< 0,001	[0,378; 0,584]	Diterima
H2	Perceived Fairness → Kepercayaan Konsumen	0,514	9,412	< 0,001	[0,408; 0,618]	Diterima
H3	AI Governance → Kepercayaan Konsumen (Direct)	0,215	3,842	< 0,001	[0,108; 0,324]	Diterima
H4	AI Governance * Transparansi → Perceived Fairness	0,195	4,118	< 0,001	[0,102; 0,286]	Diterima
H5	AI Governance → Fairness → Kepercayaan (Mediasi)	0,248	6,782	< 0,001	[0,188; 0,312]	Diterima

Catatan: Pengujian moderasi menggunakan pendekatan two-stage calculation; taraf signifikansi $p < 0,001$.

Pembahasan dan Implikasi Praktis

Hasil penelitian menegaskan bahwa tata kelola AI yang komprehensif terbukti secara empiris meningkatkan persepsi keadilan konsumen ($\beta = 0,482$; $p < 0,001$) dan memicu lahirnya kepercayaan digital yang kokoh ($\beta = 0,215$; $p < 0,001$). Temuan mediasi (H5: indirect effect $\beta = 0,248$; $p < 0,001$) memperlihatkan bahwa jalur psikologis keadilan memegang peranan krusial: konsumen tidak semata-mata membutuhkan regulasi tertulis perusahaan, melainkan merasakan langsung bahwa sistem tidak melakukan diskriminasi atau manipulasi harga secara tidak wajar.

Selain itu, konfirmasi atas hipotesis moderasi (H4: $\beta = 0,195$; $p < 0,001$) memberikan wawasan baru dalam literatur trustworthy AI. Efek positif tata kelola terhadap rasa keadilan akan menjadi berlipat ganda apabila platform secara transparan memberikan fitur penjelasan algoritmik (misalnya tombol "Mengapa produk/harga ini ditampilkan untuk Anda?"). Keterbukaan informasi ini mereduksi asimetri informasi dan mengubah persepsi algoritma dari ancaman manipulatif menjadi asisten belanja yang dapat diandalkan.

5. KESIMPULAN, KETERBATASAN, DAN SARAN

Penelitian ini menyimpulkan bahwa AI governance dan transparansi algoritma merupakan pilar determinan dalam memupuk perceived fairness dan kepercayaan konsumen pada ekosistem ekonomi digital Indonesia. Keadilan algoritmik terbukti menjadi jembatan mediasi kunci yang menghubungkan tata kelola teknologi dengan loyalitas transaksi pelanggan.

Keterbatasan penelitian ini bersumber dari desain survei berbasis self-report pada konsumen umum yang memiliki tingkat pemahaman teknis AI yang beragam. Riset selanjutnya disarankan untuk menggunakan pendekatan eksperimen faktorial (misalnya mensimulasikan antarmuka belanja dengan berbagai variasi level penjelasan algoritmik) guna menguji respon psikologis dan perilaku transaksi riil secara lebih mendalam.

DAFTAR PUSTAKA

- Colquitt, J. A. (2001). On the dimensionality of organizational justice: A construct validation of a measure. *Journal of Applied Psychology*, 86(3), 386–400. <https://doi.org/10.1037/0021-9010.86.3.386>
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in online shopping: An integrated model. *MIS Quarterly*, 27(1), 51–90. <https://doi.org/10.2307/30036519>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Kizilcec, R. F. (2016). How much information? Effects of transparency on trust in an algorithmic interface. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2390–2395. <https://doi.org/10.1145/2858036.2858402>
- Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sesing, A., & Baum, K. (2021). What do we want from Explainable Artificial Intelligence (XAI)? A stakeholder perspective on XAI and a conceptual model guiding its application. *Artificial Intelligence*, 296, 103473. <https://doi.org/10.1016/j.artint.2021.103473>
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734. <https://doi.org/10.5465/amr.1995.9508080332>

Shin, D. (2021). The perception of humanness in conversational journalism: An algorithmic information-processing approach. *New Media & Society*, 23(3), 560–584.
<https://doi.org/10.1177/1461444820903851>