Journal of Citizenship **Volume 5, Issue 3, September 2026**
E-ISSN 2829-6028

Transparansi Algoritmik dan Akuntabilitas Pelayanan Publik Digital: Studi Eksperimental dalam Konteks Kota Bandung

Rizki Mulyana[1]
[1] **Universitas Serang Raya**
Korespondensi Email: mhmdrizst13@gmail.com

Abstract

The growing use of algorithmic systems in public services raises concerns about transparency, accountability, procedural fairness, and public trust. This study develops a scenario-based experimental framework to examine how algorithmic accessibility and explainability shape trust in algorithms, trust in public officials, perceived accountability, and procedural fairness in digital public services. Two 2 × 2 factorial scenarios distinguish low- and high-discretion decision contexts. The analysis uses a reconstructed/simulated experimental dataset and should therefore be interpreted as a methodological demonstration rather than an empirical representation of Bandung residents. The results indicate that explainability consistently strengthens trust in algorithms and perceived accountability, while accessibility becomes more salient in high-discretion contexts. Both dimensions also improve procedural fairness, which partially mediates the relationship between transparency and accountability. The findings suggest that accountable algorithmic governance requires more than technical disclosure. Public organizations need accessible audit mechanisms, understandable explanations, contestability, human oversight, data correction, and effective redress. The study contributes by distinguishing transparency-as-disclosure from transparency-as-understanding and showing how their effects vary across levels of administrative discretion.

Keywords: *Algorithmic Transparency; Algorithmic Accountability; Digital Public Services; Explainability; Accessibility; Procedural Fairness; Public Trust; Bandung City.*

Abstrak

Pemanfaatan sistem algoritmik dalam pelayanan publik menimbulkan persoalan mengenai transparansi, akuntabilitas, keadilan prosedural, dan kepercayaan warga. Penelitian ini mengembangkan kerangka eksperimen berbasis skenario untuk menganalisis bagaimana accessibility dan explainability memengaruhi kepercayaan terhadap algoritma, kepercayaan terhadap aparatur, persepsi akuntabilitas, dan keadilan prosedural dalam pelayanan publik digital. Dua desain faktorial 2 × 2 digunakan untuk membedakan konteks keputusan berdiskresi rendah dan tinggi. Analisis menggunakan reconstructed/simulated experimental dataset sehingga hasil diposisikan sebagai demonstrasi metodologis, bukan representasi empiris warga Kota Bandung. Hasil menunjukkan bahwa explainability secara konsisten memperkuat kepercayaan terhadap algoritma dan akuntabilitas, sedangkan accessibility menjadi lebih penting pada keputusan berdiskresi tinggi. Kedua dimensi transparansi juga meningkatkan keadilan prosedural, yang memediasi sebagian hubungan antara transparansi dan akuntabilitas. Temuan menunjukkan bahwa tata kelola algoritmik yang akuntabel membutuhkan mekanisme audit, penjelasan yang dapat dipahami, contestability, human oversight, koreksi data, dan redress. Studi ini membedakan transparency-as-disclosure dan transparency-as-understanding serta menunjukkan bahwa pengaruh keduanya bergantung pada tingkat diskresi administratif.

Kata kunci: Transparansi Algoritmik; Akuntabilitas Algoritmik; Pelayanan Publik Digital; Explainability; Accessibility; Keadilan Prosedural; Kepercayaan Publik; Kota Bandung.

Pendahuluan

Transformasi digital telah mengubah cara organisasi publik merancang, menyediakan, dan mengevaluasi pelayanan kepada masyarakat. Perubahan tersebut tidak lagi terbatas pada digitalisasi dokumen atau pemindahan layanan ke platform

daring, tetapi semakin melibatkan sistem komputasional untuk verifikasi, klasifikasi, pemeringkatan, prediksi, rekomendasi, dan dukungan pengambilan keputusan. Dalam administrasi publik, perkembangan ini menandai pergeseran dari street-level bureaucracy menuju system-level dan algorithmic bureaucracy ketika sebagian diskresi administratif dialihkan, dibatasi, atau diarahkan oleh sistem digital (Bovens & & Zouridis, 2002; Bullock, 2019; Vogl et al., 2020; Young et al., 2019). Persoalan utamanya bukan hanya apakah algoritma mampu menghasilkan keputusan secara cepat dan konsisten, tetapi apakah warga dapat memahami alasan keputusan, mengetahui siapa yang bertanggung jawab, dan memperoleh jalur koreksi ketika terjadi kesalahan. Karena keputusan administratif dapat memengaruhi akses terhadap hak, manfaat, dan layanan, penggunaan algoritma perlu dinilai dari efisiensi sekaligus legitimasi, fairness, dan akuntabilitas administratif.

Persoalan menjadi semakin penting ketika algoritma digunakan dalam keputusan yang mempunyai konsekuensi langsung bagi warga. Berbeda dengan sistem administratif sederhana yang hanya mengotomatisasi aturan *if–then* yang relatif mudah ditelusuri, perkembangan kecerdasan artifisial dan *machine learning* memungkinkan sistem menemukan pola dari data dalam skala besar melalui mekanisme yang tidak selalu mudah dijelaskan kepada pengguna maupun aparatur. Kondisi tersebut melahirkan persoalan *algorithmic opacity* atau *black box*, yaitu situasi ketika dasar terbentuknya suatu hasil sulit dipahami atau bahkan tidak dapat diperiksa secara memadai oleh pihak yang terkena dampak keputusan (Burrell, 2016; Guidotti et al., 2018; B. D. Mittelstadt et al., 2016). Dalam administrasi publik, masalah ini mempunyai konsekuensi yang lebih serius dibandingkan banyak aplikasi privat karena warga tidak selalu memiliki pilihan untuk menghindari institusi pemerintah yang menggunakan sistem tersebut.

Perdebatan mengenai algoritma dalam pemerintahan karena itu tidak hanya berkaitan dengan apakah teknologi dapat menghasilkan keputusan yang lebih cepat atau efisien, tetapi juga apakah keputusan tersebut dapat dipertanggungjawabkan. Penggunaan model yang tidak transparan berpotensi menyulitkan warga mengetahui alasan keputusan, menyulitkan aparatur menjelaskan hasil sistem, mengaburkan distribusi tanggung jawab, serta membatasi kemungkinan untuk mengoreksi kesalahan atau mengajukan keberatan. Risiko tersebut dapat semakin besar ketika algoritma menggunakan data yang tidak lengkap, tidak mutakhir, tidak representatif, atau mengandung bias historis sehingga ketidaksetaraan yang sudah ada dapat direproduksi secara otomatis (Eubanks, 2018; Lepri et al., 2018; O'Neil, 2016). Dengan demikian, digitalisasi pelayanan publik tidak secara otomatis menghasilkan tata kelola yang lebih akuntabel.

Salah satu konsep utama yang ditawarkan untuk mengatasi persoalan tersebut adalah **transparansi algoritmik**. Namun, transparansi algoritmik bukan konsep tunggal. Grimmelikhuijsen (2023) membedakan sedikitnya dua dimensi penting, yaitu *accessibility* dan *explainability*. *Accessibility* berkaitan dengan sejauh mana algoritma, dokumentasi, data, atau prosesnya dapat diperiksa oleh pihak yang memiliki kompetensi untuk melakukan pengawasan dan audit. Sementara itu, *explainability* berkaitan dengan kemampuan sistem atau organisasi memberikan alasan mengenai mengapa keputusan tertentu dihasilkan dalam bentuk yang dapat dipahami oleh pihak yang terkena dampaknya. Pembedaan tersebut penting karena membuka kode atau dokumentasi teknis tidak dengan sendirinya menyebabkan warga memahami mengapa

permohonannya ditolak. Sebaliknya, memberikan penjelasan sederhana kepada warga tidak selalu berarti bahwa model yang mendasari keputusan tersebut dapat diaudit secara independen.

Pembedaan antara transparency-as-disclosure dan transparency-as-understanding penting dalam pelayanan publik. Keterbukaan memungkinkan pengawasan terhadap desain sistem, data, dan prosedur, sedangkan pemahaman membantu warga mengetahui alasan di balik keputusan individual. Membuka dokumentasi teknis tidak otomatis membuat keputusan mudah dipahami, sementara penjelasan sederhana kepada warga belum tentu menjamin bahwa model dapat diaudit secara independen (Ananny & & Crawford, 2018; Miller, 2019). Karena kebutuhan informasi auditor, aparatur, pengembang, dan warga berbeda, transparansi perlu dirancang sesuai fungsi dan audiensnya.

Transparansi juga berkaitan dengan kepercayaan karena warga harus menerima berbagai keputusan institusi publik, termasuk hasil yang tidak selalu menguntungkan. Literatur algorithm aversion menunjukkan bahwa kepercayaan dapat menurun setelah sistem melakukan kesalahan, sedangkan penerimaan algoritma cenderung lebih tinggi pada tugas teknis dan terstruktur dibanding keputusan yang membutuhkan pertimbangan manusia serta diskresi tinggi (Dietvorst et al., 2015; Lee, 2018; Nagtegaal, 2021). Karena itu, karakter tugas dan tingkat diskresi perlu dipertimbangkan ketika menilai efek transparansi terhadap trust.

Grimmelikhuijsen (2023) menunjukkan bahwa explainability cenderung lebih konsisten memengaruhi keterpercayaan algoritma dibanding accessibility. Penjelasan membantu warga menilai kewajaran alasan keputusan, sedangkan akses audit lebih dekat dengan pengawasan sistemik. Dalam praktik pelayanan publik, keputusan juga sering bersifat human-algorithmic: sistem menyediakan skor, rekomendasi, atau klasifikasi, sementara aparatur menentukan tindak lanjut. Karena itu, evaluasi warga terhadap teknologi dan aparatur saling berkaitan tetapi tidak identik.

Meskipun demikian, kepercayaan tidak dapat disamakan dengan akuntabilitas. Suatu algoritma dapat dipercaya oleh masyarakat tetapi tetap mempunyai mekanisme pertanggungjawaban yang lemah; sebaliknya, sebuah sistem dapat memiliki mekanisme audit dan keberatan yang kuat meskipun tingkat kepercayaan warga terhadapnya belum tinggi. Akuntabilitas algoritmik menuntut kemampuan untuk mengidentifikasi siapa yang bertanggung jawab, bagaimana suatu keputusan dihasilkan, apakah proses dapat diaudit, bagaimana keputusan dapat dipertanyakan, dan bagaimana kesalahan dapat diperbaiki (Busuioc, 2020; Kroll et al., 2016; Meijer & & Grimmelikhuijsen, 2020). Oleh sebab itu, analisis pemanfaatan algoritma dalam pelayanan publik perlu bergerak melampaui pertanyaan mengenai *trust* menuju persoalan *answerability*, *auditability*, *contestability*, *human oversight*, dan *redress.*

Dimensi tersebut semakin relevan apabila dilihat melalui teori keadilan prosedural. Teori ini berangkat dari argumentasi bahwa penerimaan masyarakat terhadap suatu keputusan tidak hanya ditentukan oleh apakah hasil keputusan menguntungkan, tetapi juga oleh bagaimana keputusan tersebut dibuat. Proses yang konsisten, tidak bias, memberikan alasan yang memadai, menyediakan kesempatan untuk menyampaikan suara, dan memungkinkan koreksi akan cenderung dipersepsikan lebih adil (Lind et al., 1990; Tyler & & Huo, 2002; Tyler, 2006). Dalam konteks keputusan algoritmik, *accessibility* dapat memberikan sinyal bahwa proses terbuka terhadap

pengawasan, sedangkan *explainability* memungkinkan warga memahami dasar keputusan. Keduanya berpotensi memengaruhi keadilan prosedural, yang pada akhirnya berhubungan dengan kepercayaan maupun persepsi akuntabilitas.

Namun, tingkat pentingnya transparansi dapat berbeda menurut karakter keputusan. Keputusan yang bersifat administratif, terstandardisasi, dan mempunyai diskresi rendah memungkinkan algoritma memainkan peranan yang lebih dominan karena kriteria keputusan relatif eksplisit. Sebaliknya, pada keputusan kompleks dengan konsekuensi lebih tinggi, algoritma umumnya berfungsi sebagai sistem rekomendasi yang kemudian ditindaklanjuti oleh aparatur. Dalam kondisi tersebut, tanggung jawab manusia menjadi lebih menonjol dan kebutuhan untuk mengetahui bagaimana rekomendasi algoritmik dihasilkan menjadi semakin penting (Buffat, 2015; Bullock et al., 2020; Zouridis et al., 2020). Oleh karena itu, tingkat diskresi merupakan konteks penting dalam mengevaluasi hubungan antara transparansi, kepercayaan, dan akuntabilitas.

Dalam konteks Indonesia, agenda transformasi digital pemerintahan berkembang melalui perluasan pelayanan publik elektronik dan Sistem Pemerintahan Berbasis Elektronik. Transformasi ini menciptakan peluang bagi pemerintah daerah untuk menggunakan pengolahan data dan sistem pendukung keputusan secara lebih intensif. Akan tetapi, berkembangnya kapasitas digital pemerintah perlu diikuti dengan prinsip tata kelola yang mampu menjamin keterbukaan, perlindungan data, keadilan administratif, dan pertanggungjawaban institusional. Pemerintah daerah tidak cukup hanya memastikan bahwa suatu sistem bekerja secara teknis; sistem juga perlu memungkinkan warga mengetahui bahwa teknologi digunakan, memahami faktor utama yang memengaruhi keputusan, mengetahui aktor institusional yang bertanggung jawab, memperbaiki data yang keliru, serta memperoleh mekanisme keberatan dan peninjauan manusia.

Kota Bandung dipilih sebagai konteks analitis karena pelayanan publik digital pada tingkat pemerintah kota menyediakan ruang yang relevan untuk menguji persoalan transparansi dan akuntabilitas dalam interaksi warga-pemerintah. Penelitian ini tidak mengasumsikan bahwa layanan tertentu milik Pemerintah Kota Bandung telah menggunakan kecerdasan artifisial atau algoritma pengambilan keputusan. Skenario yang digunakan bersifat realistis tetapi hipotetis agar persepsi terhadap keputusan algorithm-assisted dapat diuji tanpa membuat klaim mengenai teknologi aktual yang belum diverifikasi. Dengan demikian, rujukan pada Kota Bandung berfungsi sebagai konteks administratif penelitian, bukan pernyataan mengenai implementasi sistem algoritmik tertentu.

Kajian terdahulu telah membahas transformasi diskresi akibat otomatisasi (Bovens & & Zouridis, 2002; Young et al., 2019), opacity dan etika algoritma (Burrell, 2016; B. D. Mittelstadt et al., 2016), keterpercayaan keputusan otomatis (Araujo et al., 2020; Lee, 2018), serta hubungan transparansi algoritmik dengan kepercayaan warga (Grimmelikhuijsen, 2023). Namun, masih terbatas penelitian yang secara simultan menguji accessibility dan explainability terhadap trust, perceived algorithmic accountability, dan procedural fairness pada keputusan publik dengan tingkat diskresi yang berbeda. Sebagian studi menekankan trust sebagai outcome utama, sementara akuntabilitas dan fairness sering dibahas secara normatif dan terpisah. Padahal, keputusan algoritmik dalam pelayanan publik tidak hanya perlu dipercaya, tetapi juga harus dapat diaudit, dijelaskan, dipertanyakan, dan diperbaiki.

Penelitian ini menekankan pada tiga aspek. Pertama, penelitian membedakan accessibility sebagai bentuk transparency-as-disclosure dan explainability sebagai transparency-as-understanding. Kedua, penelitian menempatkan trust, perceived algorithmic accountability, dan procedural fairness dalam satu model analitis sehingga dapat dibedakan antara penerimaan psikologis dan mekanisme pertanggungjawaban. Ketiga, penelitian membandingkan keputusan berdiskresi rendah dan tinggi untuk menguji apakah kebutuhan warga terhadap transparansi berubah seiring meningkatnya kompleksitas dan konsekuensi keputusan.

Berdasarkan kesenjangan tersebut, penelitian ini menjawab tiga pertanyaan utama: (1) bagaimana accessibility dan explainability memengaruhi kepercayaan terhadap algoritma dan aparatur; (2) bagaimana kedua dimensi transparansi tersebut memengaruhi perceived algorithmic accountability dan procedural fairness; dan (3) apakah pengaruh transparansi berubah antara keputusan berdiskresi rendah dan keputusan berdiskresi tinggi?

Penelitian ini bertujuan menganalisis pengaruh accessibility dan explainability terhadap trust in algorithm, trust in public official, perceived algorithmic accountability, dan procedural fairness dalam pelayanan publik digital. Penelitian juga membandingkan pola pengaruh tersebut antara keputusan berdiskresi rendah dan tinggi serta mengevaluasi secara eksploratif peran procedural fairness sebagai mekanisme yang menghubungkan transparansi dengan akuntabilitas. Secara teoretis, studi ini diarahkan untuk memperkuat behavioral public administration dan algorithmic governance; secara metodologis, studi mendemonstrasikan desain eksperimen skenario yang dapat diuji kembali menggunakan responden aktual.

Metode

Penelitian ini menggunakan pendekatan kuantitatif dalam tradisi behavioral public administration melalui scenario-based survey experiment. Struktur penelitian mengadaptasi logika Grimmelikhuijsen (2023)dengan dua eksperimen faktorial 2 × 2 yang memanipulasi accessibility (rendah/tinggi) dan explainability (rendah/tinggi), sementara konteks keputusan dibedakan menjadi diskresi rendah dan diskresi tinggi. Perbedaan utama penelitian ini adalah penambahan perceived algorithmic accountability dan procedural fairness sebagai outcome, selain trust in algorithm dan trust in public official.

Penelitian ini menggunakan reconstructed pilot/simulated experimental dataset yang dikembangkan untuk mendemonstrasikan penerapan desain eksperimen berbasis skenario. Dataset terdiri atas 960 profil partisipan dewasa sintetis dengan dua observasi per profil, yaitu satu pada skenario diskresi rendah dan satu pada skenario diskresi tinggi, sehingga total terdapat 1.920 observasi skenario. Dengan empat kondisi eksperimen, masing-masing sel memuat 240 observasi pada setiap skenario. Seluruh estimasi kuantitatif diperlakukan sebagai hasil simulasi metodologis dan tidak ditafsirkan sebagai representasi empiris masyarakat Kota Bandung. Ukuran dataset dirancang untuk mendemonstrasikan pengujian main effects dan interaction effects secara stabil; justifikasi ukuran sampel penelitian lapangan tetap perlu dihitung ulang berdasarkan desain, power, dan karakteristik populasi aktual (Lakens & & Caldwell, 2021).

Skenario 1 menggambarkan keputusan pelayanan administratif digital berdiskresi rendah. Partisipan diminta membayangkan mengajukan permohonan layanan administratif kota secara daring, tetapi sistem menolak permohonan karena dinilai tidak memenuhi salah satu kriteria verifikasi. Aparatur memiliki ruang diskresi terbatas karena keputusan terutama didasarkan pada aturan terstruktur. Skenario 2 menggambarkan keputusan berdiskresi tinggi: sistem digital menghasilkan rekomendasi prioritas verifikasi untuk suatu program layanan kota berdasarkan beberapa informasi administrasi, kemudian aparatur memutuskan tindak lanjut yang berdampak pada pemohon. Kedua skenario sengaja bersifat realistis tetapi hipotetis dan tidak memakai nama aplikasi atau program aktual milik Pemerintah Kota Bandung.

Dalam kondisi low accessibility, partisipan diberi informasi bahwa dokumentasi model dan cara kerja sistem tidak tersedia untuk pemeriksaan pihak eksternal dan tidak ada mekanisme audit independen yang dapat menguji penggunaan informasi. Dalam kondisi high accessibility, partisipan diberi informasi bahwa dokumentasi sistem tersedia bagi auditor/pakar independen, sumber data dan aturan utama dapat diperiksa sesuai batas perlindungan data, serta terdapat jejak audit. Perlakuan ini tidak menyamakan accessibility dengan kewajiban membuka seluruh source code kepada publik.

Dalam kondisi low explainability, sistem hanya menyampaikan bahwa permohonan tidak memenuhi kriteria atau direkomendasikan untuk pemeriksaan lebih lanjut, tanpa alasan substantif. Dalam kondisi high explainability, sistem memberikan alasan utama, jenis informasi yang berpengaruh, hubungan informasi tersebut dengan keputusan, serta penjelasan mengenai langkah yang dapat dilakukan warga untuk memperbaiki data atau mengajukan keberatan. Penjelasan dirancang sebagai citizen-facing explanation, bukan uraian teknis model machine learning.

Empat kondisi eksperimen adalah: (1) low accessibility–low explainability; (2) high accessibility–low explainability; (3) low accessibility–high explainability; dan (4) high accessibility–high explainability. Assignment dilakukan secara acak dan diblok agar setiap kondisi memiliki jumlah observasi yang seimbang. Urutan skenario diskresi rendah dan tinggi juga dikonseptualisasikan sebagai randomized order dengan distraction task singkat di antaranya, mengikuti desain sumber untuk mengurangi spill-over effect.

Trust in algorithm diukur melalui empat item Likert 1–7 yang mengadaptasi dimensi kompetensi dan integritas pada artikel sumber: penggunaan informasi yang benar, ketepatan keputusan, kejujuran penilaian, dan penggunaan informasi yang relevan. Skala trust in public official menggunakan empat indikator paralel sehingga perbandingan antarkedua objek kepercayaan tetap konsisten. Perceived algorithmic accountability terdiri atas enam item yang merepresentasikan answerability, auditability, identifiable responsibility, contestability, human oversight, dan possibility of redress. Procedural fairness diukur dengan empat item mengenai konsistensi, ketidakberpihakan, kesempatan memperoleh alasan, dan kesempatan memperbaiki keputusan. Reliabilitas seluruh skala berada pada tingkat tinggi dalam dataset simulasi:

alpha trust algoritma = 0,909 dan 0,912; trust aparatur = 0,888 dan 0,896; accountability = 0,932 dan 0,934; procedural fairness = 0,885 dan 0,892 untuk Study 1 dan Study 2 secara berurutan.

Manipulation check menggunakan dua ukuran: persepsi bahwa sistem dapat diperiksa/audit oleh pihak eksternal dan persepsi bahwa alasan keputusan cukup dapat dipahami. Variabel kontrol mencakup usia, jenis kelamin, tingkat pendidikan, literasi digital, familiaritas dengan AI, dan kepercayaan umum pada pemerintah. Analisis utama menggunakan ordinary least squares dengan heteroskedasticity-consistent HC3 standard errors pada masing-masing studi. Model lintas konteks menggunakan clustered standard errors pada level partisipan untuk mengakomodasi dua observasi per profil. Mediasi eksploratif procedural fairness terhadap accountability dihitung dengan bootstrap 2.000 resampling. Seluruh koefisien dibaca pada skala outcome 1–7.

Karena penelitian ini menggunakan dataset simulatif dan tidak melibatkan responden manusia aktual, tidak ada data pribadi yang dikumpulkan maupun ethical approval number yang diklaim. Apabila desain ini diterapkan pada penelitian lapangan, informed consent, persetujuan etik, minimisasi data, debriefing, dan perlindungan data pribadi wajib diterapkan sesuai standar penelitian yang berlaku.

Tabel 1. Kondisi Eksperimen Faktorial 2 × 2

Kondisi	Accessibility	Explainability	Deskripsi
1	Rendah	Rendah	Model tidak dapat diaudit eksternal; keputusan tanpa alasan individual.
2	Tinggi	Rendah	Dokumentasi dapat diperiksa pihak independen; keputusan tanpa alasan individual.
3	Rendah	Tinggi	Model tidak dapat diaudit eksternal; keputusan disertai alasan dan faktor utama.
4	Tinggi	Tinggi	Dokumentasi dapat diaudit; keputusan disertai alasan, faktor utama, dan jalur koreksi.

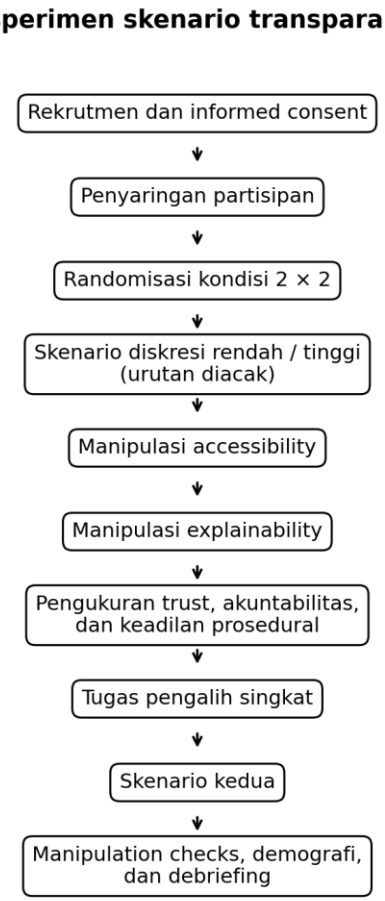
Sumber: Desain eksperimen peneliti, 2026.

Tabel 2. Karakteristik Reconstructed Pilot Dataset dan Uji Keseimbangan Randomisasi

Karakteristik	Rerata/%	SD	Uji keseimbangan empat kondisi
Jumlah profil partisipan sintetis	960		

Karakteristik	Rerata/%	SD	Uji keseimbangan empat kondisi
Usia	35.45	10.97	p balance Study 1 = 0.217; Study 2 = 0.106
Perempuan (%)	50.7	-	p = 0.758; 0.121
Pendidikan (1–5)	3.15	1.04	p = 0.617; 0.540
Literasi digital (1–7)	4.68	1.04	p = 0.706; 0.945
Familiaritas AI (1–7)	3.85	1.25	p = 0.990; 0.686
Kepercayaan umum pemerintah (1–7)	4.04	0.97	p = 0.696; 0.174

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.



Gambar 1. Prosedur Eksperimen Berbasis Vignette.

Sumber: Desain penelitian peneliti, diadaptasi dari Grimmelikhuijsen (2023).

Hasil dan Diskusi

Karakteristik Data, Keseimbangan Randomisasi, dan Reliabilitas

Reconstructed pilot dataset terdiri atas 960 profil partisipan sintetis dengan rerata usia 35,45 tahun (SD = 10,97), dan 50,7% berjenis kelamin perempuan. Rerata literasi digital adalah 4,68 dari 7, familiaritas AI 3,85, dan kepercayaan umum pada pemerintah

4,04. Uji keseimbangan menunjukkan tidak terdapat perbedaan berarti antar-empat kondisi eksperimen pada usia, jenis kelamin, pendidikan, literasi digital, familiaritas AI, maupun kepercayaan umum pada pemerintah pada kedua skenario. Pola keseimbangan tersebut menunjukkan bahwa struktur simulasi menghasilkan distribusi karakteristik yang relatif seimbang antar-kondisi eksperimen.

Konsistensi internal skala juga memadai. Nilai alpha untuk trust in algorithm sebesar 0,909 pada diskresi rendah dan 0,912 pada diskresi tinggi. Trust in public official memiliki alpha 0,888 dan 0,896; perceived accountability 0,932 dan 0,934; sedangkan procedural fairness 0,885 dan 0,892. Dengan demikian, perbedaan hasil antarkondisi tidak terutama didorong oleh inkonsistensi internal ukuran dalam dataset simulasi.

Uji Manipulasi

Tabel 3. Hasil *Manipulation Checks*

Konteks	Outcome check	Perlakuan utama	B	SE	p
Diskresi rendah	Accessibility check	Accessibility	2.247	0.078	<0,001
Diskresi rendah	Explainability check	Explainability	2.495	0.079	<0,001
Diskresi tinggi	Accessibility check	Accessibility	2.355	0.079	<0,001
Diskresi tinggi	Explainability check	Explainability	2.320	0.078	<0,001

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

Manipulasi bekerja secara kuat pada arah yang diharapkan. Pada skenario diskresi rendah, high accessibility meningkatkan skor accessibility check sebesar B = 2.247 ($p < 0,001$), sedangkan high explainability meningkatkan explainability check sebesar B = 2.495 ($p < 0,001$). Pada skenario diskresi tinggi, koefisien masing-masing adalah B = 2.355 dan B = 2.320 (keduanya $p < 0,001$). Interaksi antara accessibility dan explainability tidak signifikan dalam manipulation checks, menunjukkan bahwa kedua perlakuan relatif berhasil memanipulasi konstruk yang berbeda.

Tabel 4. Rerata Outcome Per Kondisi Eksperimen

Konteks	Kondisi akses–penjelasan	Trust algoritma	Trust aparatur	Akuntabilitas	Keadilan prosedural
Diskresi rendah	Rendah–Rendah	4.14	4.41	4.28	3.96
Diskresi rendah	Rendah–Tinggi	4.69	4.54	5.05	4.57
Diskresi rendah	Tinggi–Rendah	4.20	4.35	4.58	4.18
Diskresi rendah	Tinggi–Tinggi	4.66	4.43	5.37	4.66
Diskresi tinggi	Rendah–Rendah	3.74	4.37	3.88	3.54

Konteks	Kondisi akses–penjelasan	Trust algoritma	Trust aparatur	Akuntabilitas	Keadilan prosedural
Diskresi tinggi	Rendah–Tinggi	4.16	4.60	4.79	4.20
Diskresi tinggi	Tinggi–Rendah	4.16	4.49	4.54	3.90
Diskresi tinggi	Tinggi–Tinggi	4.68	4.81	5.26	4.41

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

Tabel 5. Model Regresi Skenario Diskresi Rendah

Outcome	Prediktor	B	SE	95% CI	p
Trust algoritma	Accessibility	0.068	0.095	[-0.119; 0.255]	0,476
Trust algoritma	Explainability	0.559	0.091	[0.380; 0.738]	<0,001
Trust algoritma	Interaksi A×E	-0.104	0.134	[-0.366; 0.158]	0,436
Trust algoritma	R^2	0.061			
Trust aparatur	Accessibility	-0.063	0.093	[-0.244; 0.119]	0,499
Trust aparatur	Explainability	0.123	0.093	[-0.060; 0.306]	0,188
Trust aparatur	Interaksi A×E	-0.039	0.132	[-0.299; 0.220]	0,767
Trust aparatur	R^2	0.011			
Akuntabilitas	Accessibility	0.299	0.091	[0.122; 0.476]	<0,001
Akuntabilitas	Explainability	0.761	0.088	[0.588; 0.934]	<0,001
Akuntabilitas	Interaksi A×E	0.023	0.123	[-0.218; 0.264]	0,852
Akuntabilitas	R^2	0.172			

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

Studi 1: Keputusan Digital Berdiskresi Rendah

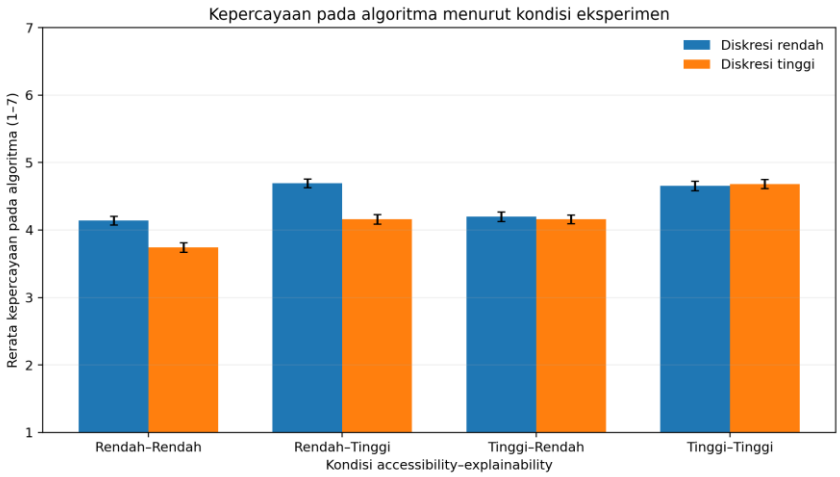
Pada kondisi baseline low accessibility-low explainability, rerata trust in algorithm adalah 4,14. Menambahkan explainability menaikkan rerata menjadi 4,69 ketika accessibility tetap rendah, sedangkan accessibility saja hanya mengubah rerata menjadi 4,20. Regresi yang mengendalikan usia, jenis kelamin, pendidikan, literasi digital, familiaritas AI, dan kepercayaan umum pada pemerintah mengonfirmasi pola tersebut: explainability berpengaruh positif terhadap trust in algorithm (B = 0,559; SE = 0,091; 95% CI [0,380; 0,738]; p < 0,001), sedangkan accessibility tidak signifikan (B = 0,068; p = 0,476). Interaksi accessibility × explainability juga tidak signifikan. Pola ini

menunjukkan bahwa pada keputusan yang kriterianya relatif terstruktur, kebutuhan transparansi terutama bersifat interpretatif: warga lebih membutuhkan alasan yang dapat dipahami dibandingkan akses teknis terhadap dokumentasi atau mekanisme audit. Dengan kata lain, transparency-as-understanding lebih relevan bagi pembentukan trust pada konteks diskresi rendah.

Kepercayaan pada aparatur menunjukkan pola berbeda. Accessibility (B = -0,063; p = 0,499) dan explainability (B = 0,123; p = 0,188) tidak memberikan dukungan statistik yang memadai pada konteks diskresi rendah. Ketika proses administratif dipersepsikan sangat terstruktur, warga dapat memisahkan evaluasi sistem dari evaluasi aparatur yang terutama menjalankan aturan. Dalam konteks tersebut, aparatur lebih berfungsi sebagai mekanisme pengawasan terakhir daripada pengambil keputusan substantif, sehingga evaluasi warga terhadap aparatur tidak banyak berubah akibat variasi transparansi sistem.

Untuk accountability, kedua dimensi transparansi justru signifikan. Accessibility meningkatkan perceived accountability sebesar B = 0,299 (p = 0,001), sedangkan explainability mempunyai pengaruh lebih besar, B = 0,761 (p < 0,001). Rerata accountability meningkat dari 4,28 pada kondisi rendah–rendah menjadi 5,37 pada kondisi tinggi–tinggi. Temuan ini menunjukkan bahwa trust dan accountability adalah konstruk yang berbeda: accessibility mungkin belum cukup untuk membuat algoritma lebih dipercaya oleh warga, tetapi akses audit eksternal tetap meningkatkan persepsi bahwa sistem dapat dimintai pertanggungjawaban. Dengan demikian, transparency-as-disclosure dapat berfungsi pada level institusional meskipun efek psikologisnya pada trust lebih terbatas.

Procedural fairness juga meningkat melalui accessibility (B = 0,224; p = 0,008) dan terutama explainability (B = 0,596; p < 0,001). Pola ini sejalan dengan procedural justice theory: alasan keputusan memberikan sinyal bahwa otoritas tidak bertindak secara arbitrer, sementara kemungkinan audit memberi sinyal bahwa prosedur terbuka untuk pengawasan (Lind et al., 1990; Tyler, 2006).



Gambar 2. Rerata Kepercayaan Pada Algoritma Menurut Kondisi Accessibility–Explainability. Error Bar = Standard Error.

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

Tabel 6. Model Regresi Skenario Diskresi Tinggi

Outcome	Prediktor	B	SE	95% CI	p
Trust algoritma	Accessibility	0.422	0.099	[0.228; 0.616]	<0,001
Trust algoritma	Explainability	0.418	0.099	[0.224; 0.613]	<0,001
Trust algoritma	Interaksi A×E	0.103	0.138	[-0.168; 0.374]	0,455
Trust algoritma	R^2	0.094			
Trust aparatur	Accessibility	0.113	0.096	[-0.075; 0.300]	0,239
Trust aparatur	Explainability	0.219	0.091	[0.040; 0.397]	0,016
Trust aparatur	Interaksi A×E	0.091	0.133	[-0.170; 0.351]	0,495
Trust aparatur	R^2	0.032			
Akuntabilitas	Accessibility	0.662	0.089	[0.487; 0.837]	<0,001
Akuntabilitas	Explainability	0.919	0.090	[0.743; 1.095]	<0,001
Akuntabilitas	Interaksi A×E	-0.189	0.123	[-0.430; 0.051]	0,122
Akuntabilitas	R^2	0.231			

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

Studi 2: Keputusan Digital Berdiskresi Tinggi

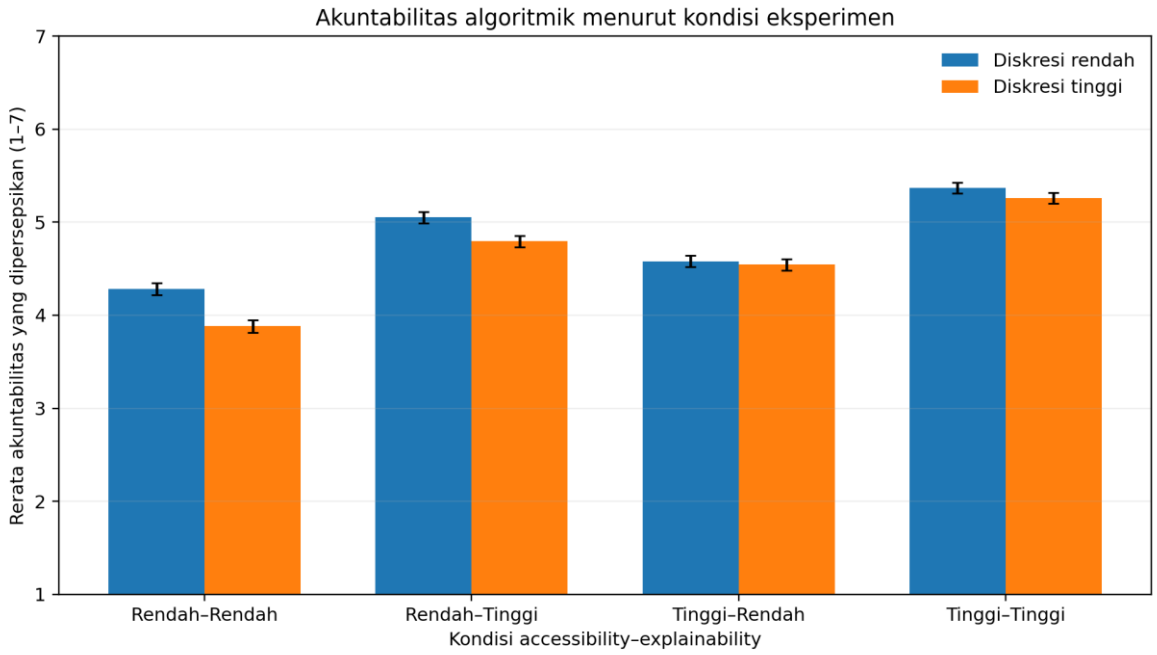
Pada skenario diskresi tinggi, baseline trust in algorithm lebih rendah (3,74) dibanding skenario diskresi rendah, konsisten dengan literatur yang menunjukkan bahwa penerimaan algoritma menjadi lebih problematik ketika keputusan kompleks dan membutuhkan judgment manusia (Lee, 2018; Nagtegaal, 2021). Berbeda dari Studi 1, accessibility meningkatkan trust in algorithm secara signifikan (B = 0,422; SE = 0,099; p < 0,001) dan explainability juga signifikan (B = 0,418; SE = 0,099; p < 0,001). Kondisi tinggi-tinggi menghasilkan rerata trust algoritma 4,68. Perbedaan ini menunjukkan bahwa semakin besar konsekuensi dan ruang diskresi, semakin besar kebutuhan warga untuk mengetahui bukan hanya alasan individual, tetapi juga apakah sistem dapat diaudit secara independen.

Untuk trust in public official, accessibility tidak signifikan (B = 0,113; p = 0,239), tetapi explainability berpengaruh positif (B = 0,219; p = 0,016). Artinya, ketika aparatur memiliki tanggung jawab lebih besar untuk menindaklanjuti rekomendasi algoritma, kualitas penjelasan sistem dapat memengaruhi evaluasi warga terhadap manusia yang

menggunakan sistem tersebut. Penjelasan yang memadai membuat tanggung jawab aparatur lebih terlihat karena warga dapat menilai apakah tindak lanjut manusia masuk akal atau sekadar berlindung di balik rekomendasi mesin.

Perceived accountability menunjukkan efek terkuat. Accessibility menghasilkan B = 0,662 ($p < 0,001$) dan explainability B = 0,919 ($p < 0,001$). Rerata accountability naik dari 3,88 pada kondisi rendah–rendah menjadi 5,26 pada kondisi tinggi–tinggi. Efek ini memperkuat argumen bahwa dalam keputusan berkonsekuensi tinggi, warga tidak hanya membutuhkan alasan individual tetapi juga jaminan bahwa model, data, dan proses dapat diperiksa oleh pihak yang kompeten. Di sinilah accessibility lebih dekat dengan systemic accountability, sedangkan explainability berfungsi sebagai citizen-facing accountability.

Interaksi accessibility × explainability tidak signifikan pada trust maupun accountability. Hal ini menunjukkan bahwa accessibility dan explainability tidak bekerja secara substitutif maupun sinergis secara statistik dalam dataset simulasi. Efek keduanya lebih tepat dipahami sebagai kontribusi yang relatif independen dan aditif: explainability tidak menghapus kebutuhan auditability, sedangkan accessibility tidak menggantikan reason-giving. Keduanya merupakan lapisan tata kelola yang berbeda.



Gambar 3. ***Perceived Algorithmic Accountability*** **Menurut Kondisi Eksperimen. Error Bar = Standard Error.**

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

Analisis Lintas Konteks, Keadilan Prosedural, dan Evaluasi Hipotesis

Tabel 7 Interaksi Transparansi × Konteks Diskresi

Outcome	Interaksi	B	SE	95% CI	p
Trust algoritma	Accessibility × diskresi tinggi	0.355	0.138	[0.084; 0.626]	0,010

Outcome	Interaksi	B	SE	95% CI	p
Trust algoritma	Explainability × diskresi tinggi	-0.135	0.138	[-0.405; 0.136]	0,330
Trust aparatur	Accessibility × diskresi tinggi	0.171	0.132	[-0.088; 0.430]	0,195
Trust aparatur	Explainability × diskresi tinggi	0.104	0.129	[-0.149; 0.357]	0,419
Akuntabilitas	Accessibility × diskresi tinggi	0.365	0.129	[0.113; 0.617]	0,005
Akuntabilitas	Explainability × diskresi tinggi	0.158	0.125	[-0.087; 0.402]	0,206
Keadilan prosedural	Accessibility × diskresi tinggi	0.126	0.115	[-0.100; 0.352]	0,275
Keadilan prosedural	Explainability × diskresi tinggi	0.062	0.117	[-0.167; 0.291]	0,594

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

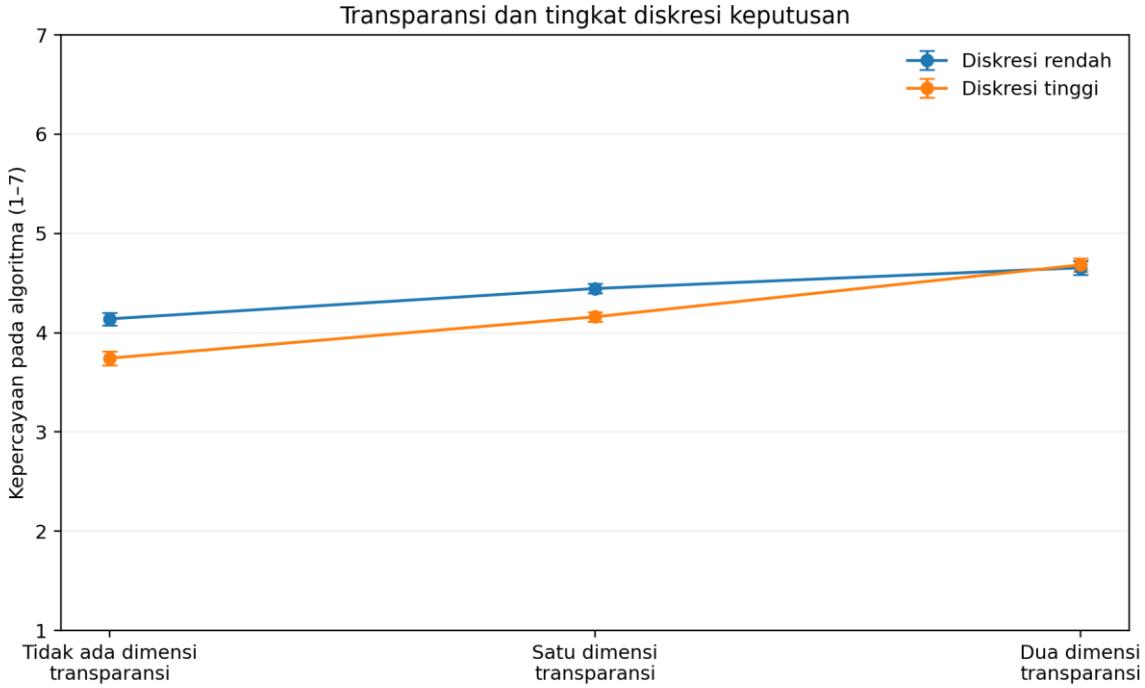
Model lintas konteks menunjukkan bahwa interaksi accessibility × diskresi tinggi signifikan untuk trust in algorithm (B = 0,355; p = 0,010) dan accountability (B = 0,365; p = 0,005). Sebaliknya, interaksi explainability × diskresi tinggi tidak signifikan untuk trust algoritma, trust aparatur, accountability, maupun procedural fairness. Dengan demikian, efek explainability relatif stabil pada dua konteks, sedangkan nilai tambah accessibility menjadi lebih nyata ketika keputusan memerlukan diskresi dan berdampak lebih tinggi. Temuan ini merupakan salah satu kontribusi utama penelitian karena menunjukkan bahwa kebutuhan terhadap transparansi tidak seragam: semakin kompleks keputusan, semakin penting assurance bahwa model, data, dan proses dapat diaudit secara institusional.

Temuan tersebut selaras dengan perdebatan dalam literatur. Explainability terutama menjawab kebutuhan warga untuk mengetahui mengapa keputusan tertentu terjadi, sehingga kegunaannya muncul baik pada prosedur sederhana maupun kompleks. Accessibility, sebaliknya, menjadi lebih penting ketika warga memiliki alasan lebih kuat untuk mempertanyakan bagaimana model dibangun, apakah data tepat, dan apakah keputusan dapat diaudit. Ini menjelaskan mengapa accessibility yang relatif lemah pada trust dalam skenario diskresi rendah menjadi signifikan pada skenario diskresi tinggi.

Analisis mediasi eksploratif menunjukkan bahwa procedural fairness berpotensi menjadi salah satu mekanisme yang menghubungkan transparansi dengan accountability. Pada Studi 1, indirect effect accessibility terhadap accountability melalui

fairness sebesar 0,056 (95% bootstrap CI [0,016; 0,098]), sedangkan indirect effect explainability sebesar 0,183 [0,132; 0,241]. Pada Studi 2, indirect effect accessibility sebesar 0,100 [0,057; 0,144] dan explainability sebesar 0,209 [0,158; 0,266]. Seluruh interval tidak melintasi nol. Hasil ini mendukung H9 secara eksploratif, tetapi tidak menunjukkan bahwa fairness merupakan satu-satunya mekanisme yang bekerja.

Namun mediasi tersebut tidak boleh dibaca sebagai bukti bahwa fairness adalah satu-satunya mekanisme. Accessibility juga dapat bekerja melalui persepsi auditability dan institutional control, sedangkan explainability dapat bekerja melalui pemahaman, voice, dan kemampuan menggugat keputusan. Studi eksperimental lanjutan perlu mengukur mekanisme tersebut secara terpisah.



Gambar 4. Hubungan Tingkat Transparansi dan Kepercayaan pada Algoritma pada Dua Konteks Diskresi.

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

Tabel 9. Ringkasan Evaluasi Hipotesis

Hipotesis	Prediksi	Diskresi rendah	Diskresi tinggi	Evaluasi umum
H1	Accessibility → trust algoritma	Tidak didukung	Didukung	Campuran
H2	Explainability → trust algoritma	Didukung	Didukung	Didukung
H3	Accessibility → trust aparatur	Tidak didukung	Tidak didukung	Tidak didukung

Hipotesis	Prediksi	Diskresi rendah	Diskresi tinggi	Evaluasi umum
H4	Explainability → trust aparatur	Tidak didukung	Didukung	Campuran
H5	Accessibility → accountability	Didukung	Didukung	Didukung
H6	Explainability → accountability	Didukung	Didukung	Didukung
H7	Explainability lebih konsisten daripada accessibility	Didukung	Sebagian	Didukung
H8	Efek lebih kuat pada diskresi tinggi	Sebagian	Sebagian	Campuran/sebagian
H9	Fairness memediasi transparency → accountability	Didukung eksploratif	Didukung eksploratif	Didukung eksploratif

Sumber: Diolah peneliti dari reconstructed pilot dataset (simulasi), 2026.

Transparansi sebagai Keterbukaan dan Transparansi sebagai Pemahaman

Temuan utama penelitian bukan bahwa satu bentuk transparansi harus dipilih dan bentuk lainnya ditinggalkan, melainkan bahwa keduanya mengatasi masalah yang berbeda. Accessibility mengurangi organizational opacity dengan membuka ruang audit, pemeriksaan model, dan penilaian kualitas data. Explainability mengurangi citizen opacity dengan memberi alasan yang dapat dipahami oleh individu yang terkena keputusan. Pemerintah yang hanya membuka dokumentasi teknis berisiko memenuhi transparency-as-disclosure tetapi gagal menghasilkan transparency-as-understanding. Sebaliknya, pemerintah yang hanya memberi penjelasan singkat tanpa mekanisme audit dapat menciptakan ilusi akuntabilitas karena warga mendapat alasan tetapi tidak ada pihak independen yang mampu memeriksa apakah alasan tersebut benar-benar mencerminkan cara sistem bekerja.

Literatur explainable AI mendukung pembedaan tersebut. Penjelasan yang baik harus menyesuaikan kebutuhan audiens; warga tidak membutuhkan detail teknis yang sama dengan auditor atau pengembang (Friedrich & & Zanker, 2011; Miller, 2019; B. Mittelstadt et al., 2019). Pada saat yang sama, model high-stakes sebaiknya tidak bergantung pada post-hoc explanation saja jika model yang lebih interpretable dapat digunakan (Rudin, 2019). Dalam pelayanan publik, desain penjelasan karena itu harus dipadukan dengan dokumentasi model, log keputusan, pengujian bias, dan mekanisme audit.

Secara teoretis, penelitian ini memperluas behavioral public administration dengan menunjukkan bahwa transparansi algoritmik memiliki sedikitnya dua fungsi. Explainability terutama bekerja pada level citizen-facing legitimacy dengan membantu warga memahami alasan keputusan, sedangkan accessibility bekerja pada level systemic accountability dengan membuka kemungkinan audit dan pengawasan. Pembedaan ini penting karena sebuah sistem dapat mudah dijelaskan tetapi sulit diaudit, atau dapat diaudit tetapi tidak cukup dipahami oleh warga.

Akuntabilitas Human–Algorithm dan Implikasi bagi Pelayanan Publik Digital

Algoritma bukan aktor hukum yang dapat memikul tanggung jawab secara mandiri. Ketika sistem memberikan rekomendasi, pertanggungjawaban tetap harus dapat ditelusuri pada organisasi publik, pejabat berwenang, unit pengelola data, pengembang, dan—jika ada—vendor teknologi. Penggunaan vendor tidak boleh menghasilkan accountability outsourcing. Kontrak pengadaan perlu memastikan akses audit, dokumentasi data dan model, kewajiban melaporkan perubahan versi, mekanisme pengujian bias, serta kemampuan organisasi publik menjelaskan keputusan kepada warga. Intellectual property dapat dilindungi, tetapi perlindungan tersebut tidak boleh menghapus kewajiban pemerintah untuk memenuhi standar pelayanan publik dan memberikan alasan keputusan.

Implikasi berikut harus dipahami sebagai arah desain tata kelola yang perlu diuji melalui penelitian empiris, bukan rekomendasi kebijakan yang langsung diturunkan dari kondisi aktual Pemerintah Kota Bandung. Jika pemerintah daerah menggunakan atau akan menggunakan algoritma untuk memverifikasi, memprioritaskan, menyaring, atau merekomendasikan keputusan pelayanan, desain tata kelola dapat mempertimbangkan algorithm register internal, documentation standard, algorithmic impact assessment, independent audit, human oversight pada keputusan berdampak tinggi, explanation notice untuk warga, complaint and appeal mechanism, hak koreksi data, audit trail, model monitoring, bias testing, vendor accountability, perlindungan privasi, dan komunikasi publik. Prinsip tersebut perlu disesuaikan dengan dasar hukum, kapasitas institusional, serta karakter layanan yang benar-benar digunakan.

Contestability harus mendapat perhatian khusus. Warga perlu mengetahui bahwa algoritma digunakan, bagian keputusan mana yang dipengaruhi algoritma, alasan utama hasil, siapa aktor institusional yang bertanggung jawab, bagaimana mengajukan keberatan, bagaimana memperbaiki data, dan apakah keputusan dapat ditinjau ulang oleh manusia. Tanpa contestability, explainability dapat berhenti pada pemberian informasi tanpa daya koreksi; tanpa redress, accountability menjadi simbolik.

Data governance juga tidak dapat dipisahkan dari algorithmic accountability. Model yang transparan tetap dapat menghasilkan keputusan tidak adil jika data tidak akurat, tidak mutakhir, atau tidak representatif. Audit harus mencakup provenance data, kelengkapan, kualitas, fairness, dasar hukum pemrosesan, retention, serta hak koreksi. Karena itu, akuntabilitas algoritmik lebih tepat dipahami sebagai rantai tata kelola dari data input, model, rekomendasi, tindakan aparatur, hingga mekanisme keberatan.

Kesimpulan

Penelitian ini menunjukkan bahwa transparansi algoritmik terdiri atas dua dimensi yang menjalankan fungsi berbeda dalam tata kelola pelayanan publik digital. Explainability terutama memperkuat pemahaman warga dan kepercayaan terhadap keputusan algoritmik, sedangkan accessibility lebih berkaitan dengan auditability dan perceived accountability. Nilai accessibility menjadi lebih penting pada keputusan dengan tingkat diskresi yang lebih tinggi, sementara pengaruh explainability relatif stabil pada dua konteks keputusan. Analisis eksploratif juga menunjukkan bahwa procedural fairness berpotensi menjadi salah satu mekanisme yang menghubungkan transparansi dengan perceived accountability. Kontribusi utama penelitian terletak pada pembedaan transparency-as-disclosure dan transparency-as-understanding serta integrasinya dengan trust, accountability, fairness, dan administrative discretion. Namun, karena seluruh hasil berasal dari reconstructed/simulated experimental dataset, temuan harus dipahami sebagai demonstrasi metodologis dan memerlukan validasi melalui penelitian lapangan sebelum digunakan sebagai dasar kebijakan.

Ucapan Terima Kasih

Penulis menyampaikan penghargaan kepada para peneliti yang mengembangkan literatur transparansi algoritmik, akuntabilitas AI, keadilan prosedural, dan behavioral public administration. Kerangka eksperimental Grimmelikhuijsen (2023) menjadi salah satu acuan metodologis utama. Naskah menggunakan reconstructed pilot dataset yang bersifat simulatif untuk demonstrasi metodologis dan tidak mengklaim adanya pengumpulan data lapangan maupun pendanaan eksternal untuk survei responden.

Referensi

Ananny, M., & & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, *20*(3), 973–989.

Araujo, T., Helberger, N., Kruikemeier, S., & & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & Society*, *35*, 611–623.

Bovens, M., & & Zouridis, S. (2002). From street-level to system-level bureaucracies: How information and communication technology is transforming administrative discretion and constitutional control. *Public Administration Review*, *62*(2), 174–184.

Buffat, A. (2015). Street-level bureaucracy and e-government. *Public Management Review*, *17*(1), 149–161.

Bullock, J. B. (2019). Artificial intelligence, discretion, and bureaucracy. *The American Review of Public Administration*, *49*(7), 751–761.

Bullock, J. B., Young, M. M., & & Fang, Y.-F. (2020). Artificial intelligence, bureaucratic form, and discretion in public service. *Information Polity*, *25*(4), 491–506.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, *3*(1), 1–12.

Busuioc, M. (2020). Accountable artificial intelligence: Holding algorithms to account.

Public Administration Review, *81*, 825–836. https://doi.org/10.1111/puar.13293

Dietvorst, B. J., Simmons, J. P., & & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, *144*(1), 114–126.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Friedrich, G., & & Zanker, M. (2011). A taxonomy for generating explanations in recommender systems. *AI Magazine*, *32*(3), 90–98.

Grimmelikhuijsen, S. (2023). Explaining why the computer says no: Algorithmic transparency affects the perceived trustworthiness of automated decision-making. *Public Administration Review*, *83*(2), 241–262. https://doi.org/10.1111/puar.13483

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, *51*(5), 93.

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & & Yu, H. (2016). Accountable algorithms. *University of Pennsylvania Law Review*, *165*, 633–705.

Lakens, D., & & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, *4*(1), 2515245920951503.

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1), 2053951718756684.

Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology*, *31*(4), 611–627.

Lind, E. A., Kanfer, R., & & Earley, P. C. (1990). Voice, control, and procedural justice: Instrumental and noninstrumental concerns in fairness judgments. *Journal of Personality and Social Psychology*, *59*(5), 952–959.

Meijer, A., & & Grimmelikhuijsen, S. (2020). *Responsible and accountable algorithmization: How to generate citizen trust in governmental usage of algorithms*. In M. Schuilenburg & R. Peeters (Eds.), The algorithmic society: Technology, power, and knowledge. Routledge.

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, *267*, 1–38.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2), 205395171667967.

Mittelstadt, B., Russell, C., & & Wachter, S. (2019). Explaining explanations in AI. In *In Proceedings of the Conference on Fairness, Accountability, and Transparency* (hal. 279–288).

Nagtegaal, R. (2021). The impact of using algorithms for managerial decisions on public employees' procedural justice. *Government Information Quarterly*, *38*(1), 101536.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215.

Tyler, T. R. (2006). *Why people obey the law*. Yale University Press.

Tyler, T. R., & & Huo, Y. J. (2002). *Trust in the law: Encouraging public cooperation with the police and courts*. Russell Sage Foundation.

Vogl, T. M., Seidelin, C., Ganesh, B., & & Bright, J. (2020). Smart technology and the emergence of algorithmic bureaucracy: Artificial intelligence in UK local authorities. *Public Administration Review*, *80*(6), 946–961.

Young, M. M., Bullock, J. B., & & Lecy, J. D. (2019). Artificial discretion as a tool of governance: A framework for understanding the impact of artificial intelligence on public administration. *Perspectives on Public Management and Governance*, *2*(4), 301–313.

Zouridis, S., van Eck, M., & & Bovens, M. (2020). *Automated discretion*. In T. Evans & P. Hupe (Eds.), Discretion and the quest for controlled freedom (pp. 313–329). Palgrave Macmillan.