



Application of Machine Learning Algorithms Based on Islamic Moral Values for Mitigating Negative Content on Digital Media

Moch Nurcholis Majid¹, Muhammad Suhaili²

¹Universitas Uluwiyah Mojokerto, Indonesia

²Politeknik Negeri Media Kreatif, Indonesia

e-mail: nurcholis@lecturer.uluwiyah.ac.id¹, suhaili@polimedia.ac.id²

Abstract:



This is an open access article under [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Information bias in digital media has evolved into a systemic problem driven by algorithmic designs that prioritize user engagement optimization without adequate value orientation. This condition has contributed to polarization, amplification of sensational content, and unequal information representation. This study aimed to reconstruct data science algorithms based on Prophetic Ethics as a normative-operational framework to mitigate information bias in digital media. The research employed a descriptive qualitative approach with a systematic meta-analysis of scholarly publications from 2016 to 2025 retrieved from reputable academic databases. Data were analyzed using content analysis and thematic analysis to synthesize patterns of algorithmic bias and formulate a reconstruction model grounded in the principles of *ṣidq* (truthfulness), *amānah* (trustworthiness), *tabligh* (transparency), and *faṭānah* (wisdom). The findings indicated that algorithmic bias occurred across all stages of the data processing pipeline and required structural reconstruction rather than partial technical adjustments. The proposed model integrated data validation, accountable governance, transparent modeling, and social welfare optimization within recommendation system design. This study contributed to the development of a value-oriented data science paradigm and expanded the discourse on algorithmic ethics in contemporary digital societies.

Keywords:

data science algorithms, information bias, prophetic ethics, digital media, bias mitigation

* Corresponding author :
Email Address
nurcholis@lecturer.uluwiyah.ac.id
Received : October 24, 2025;
Revised : November 7, 2025;
Accepted : November 16, 2025;
Published : November 30, 2025

Rekonstruksi Algoritma Data Science Berbasis Etika Nubuwwah untuk Mitigasi Bias Informasi di Media Digital

Abstract:

Bias informasi di media digital berkembang sebagai problem sistemik yang dipengaruhi oleh desain algoritma berbasis optimasi keterlibatan pengguna tanpa orientasi nilai yang memadai. Kondisi tersebut memunculkan polarisasi, amplifikasi konten sensasional, dan ketimpangan representasi informasi. Penelitian ini bertujuan merekonstruksi algoritma data science berbasis Etika Nubuwwah sebagai kerangka normatif-operasional untuk mitigasi bias informasi di media digital. Penelitian menggunakan pendekatan kualitatif deskriptif dengan desain meta-analisis literatur sistematis terhadap publikasi ilmiah periode 2016 - 2025 yang bersumber dari basis data akademik bereputasi. Data dianalisis melalui analisis konten dan analisis tematik untuk mensintesis pola bias algoritmik serta merumuskan model rekonstruksi berbasis prinsip *ṣidq*, *amānah*, *tablīgh*, dan *faṭānah*. Hasil penelitian menunjukkan bahwa bias algoritmik terjadi pada seluruh tahapan pengolahan data dan memerlukan rekonstruksi struktural, bukan sekadar penyesuaian teknis parsial. Model yang dihasilkan mengintegrasikan validasi data, tata kelola transparan, dan optimasi kemaslahatan dalam desain sistem rekomendasi. Penelitian ini berkontribusi pada pengembangan paradigma data science berbasis nilai serta memperluas diskursus etika algoritmik dalam konteks masyarakat digital kontemporer.

Keywords:

algoritma data science, bias informasi, etika nubuwah, media digital, mitigasi bias

PENDAHULUAN

Transformasi digital yang dipacu oleh kemajuan komputasi awan, kecerdasan buatan, dan analitik data berskala besar telah merekonstruksi ekosistem produksi, distribusi, dan konsumsi informasi di ruang publik. Platform media sosial seperti Facebook, X, YouTube, dan TikTok mengandalkan algoritma data science untuk melakukan kurasi konten, personalisasi rekomendasi, serta optimasi keterlibatan (engagement) pengguna secara real time. Di satu sisi, algoritma tersebut meningkatkan efisiensi distribusi informasi dan memperluas akses publik terhadap berbagai sumber pengetahuan. Namun di sisi lain, praktik algoritmik yang berorientasi pada metrik klik, waktu tonton, dan interaksi telah menimbulkan problem serius berupa bias informasi, polarisasi opini, serta penyebaran misinformasi dan disinformasi (Cinelli et al., 2021; Guess et al., 2020; Pariser, 2011).

Fenomena bias informasi di media digital tidak lagi bersifat sporadis, melainkan sistemik. Bias tersebut dapat muncul dalam berbagai bentuk: bias seleksi (selection bias), bias representasi (representation bias), bias algoritmik (algorithmic bias), hingga bias amplifikasi (amplification bias) yang memperkuat narasi tertentu secara tidak proporsional (Mehrabi et al., 2021). Studi empiris menunjukkan bahwa algoritma rekomendasi cenderung membentuk echo chambers dan filter bubbles yang mempersempit paparan pengguna terhadap keragaman perspektif (Cinelli et al., 2021). Selain itu,

desain sistem yang memprioritaskan engagement sering kali mendorong konten sensasional atau kontroversial untuk memperoleh visibilitas lebih tinggi, terlepas dari akurasi dan validitasnya (Guess et al., 2020).

Dalam konteks global, berbagai laporan mengindikasikan meningkatnya kekhawatiran terhadap dampak sosial-politik dari algoritma media digital. Skandal penyalahgunaan data oleh Cambridge Analytica pada platform Facebook menjadi salah satu titik balik kesadaran publik mengenai relasi antara data, algoritma, dan manipulasi opini (Isaak & Hanna, 2018). Selain itu, dinamika kontestasi politik di berbagai negara memperlihatkan bagaimana sistem rekomendasi dapat memperkuat fragmentasi sosial dan memfasilitasi penyebaran propaganda digital (Tucker et al., 2018). Dengan demikian, bias informasi bukan sekadar isu teknis komputasional, tetapi problem multidimensional yang bersinggungan dengan etika, demokrasi, dan kohesi sosial.

Merespons problem tersebut, komunitas ilmiah mengembangkan berbagai kerangka kerja “fairness, accountability, and transparency” (FAT) dalam algoritma (Barocas et al., 2019). Konsep keadilan algoritmik (algorithmic fairness) berupaya meminimalkan diskriminasi berbasis atribut sensitif seperti ras, gender, dan agama (Mehrabi et al., 2021). Di sisi lain, pendekatan explainable AI (XAI) menekankan pentingnya transparansi model untuk meningkatkan akuntabilitas sistem (Arrieta et al., 2020). Regulasi seperti EU AI Act juga mendorong standar tata kelola kecerdasan buatan yang bertanggung jawab (Veale & Borgesius, 2021). Namun, sebagian besar pendekatan tersebut berakar pada paradigma etika sekuler-liberal yang menekankan prinsip utilitarianisme, deontologi, atau hak asasi manusia dalam kerangka Barat modern.

Meskipun kontribusi pendekatan FAT dan XAI signifikan, sejumlah penelitian menunjukkan bahwa bias algoritmik tetap sulit dieliminasi karena berakar pada struktur data, asumsi model, serta kepentingan ekonomi platform (Kleinberg et al., 2018; Selbst et al., 2019). Kritik terhadap “solutionism” teknokratik juga menyoroti bahwa problem sosial tidak dapat sepenuhnya diselesaikan melalui optimasi matematis semata (Selbst et al., 2019). Dengan kata lain, rekonstruksi algoritma tidak cukup hanya pada level teknis; diperlukan fondasi normatif yang lebih komprehensif untuk mengarahkan tujuan, desain, dan implementasi sistem data science.

Dalam konteks masyarakat Muslim yang memiliki basis nilai religius kuat, diskursus etika teknologi semakin relevan. Etika Nubuwwah yang merujuk pada nilai-nilai kenabian seperti *ṣidq* (kejujuran), *amānah* (kepercayaan), *tabligh* (transparansi penyampaian), dan *faṭānah* (kecerdasan dan kebijaksanaan) menawarkan paradigma normatif yang integral antara dimensi moral, spiritual, dan sosial. Konsep ini telah diaplikasikan dalam berbagai bidang seperti kepemimpinan, manajemen, dan

pendidikan (Beekun & Badawi, 2017; Huda et al., 2018), namun belum banyak dielaborasi secara sistematis dalam konteks rekayasa algoritma data science.

Secara teoretis, integrasi etika religius dalam teknologi bukanlah hal baru. Kajian “Islamic ethics of technology” menekankan pentingnya maqāṣid al-sharī‘ah sebagai kerangka perlindungan terhadap agama, jiwa, akal, keturunan, dan harta (Auda, 2016). Dalam perspektif ini, teknologi harus diarahkan untuk kemaslahatan (maṣlaḥah) dan mencegah kerusakan (mafsadah). Namun, penelitian-penelitian tersebut cenderung bersifat normatif-filosofis dan belum teroperasionalisasi dalam desain algoritmik yang konkret, khususnya dalam mitigasi bias informasi di media digital.

Penelitian terdahulu mengenai bias informasi di media digital dapat dikelompokkan ke dalam tiga arus utama. Pertama, pendekatan kuantitatif-komputasional yang menganalisis struktur jaringan, dinamika penyebaran informasi, dan dampak algoritma rekomendasi terhadap polarisasi (Cinelli et al., 2021; Guess et al., 2020). Kedua, pendekatan etika dan tata kelola yang menyoroti kebutuhan regulasi, transparansi, dan akuntabilitas (Barocas et al., 2019; Veale & Borgesius, 2021). Ketiga, pendekatan kritis-sosiologis yang mengkaji relasi kuasa, kapitalisme platform, dan konstruksi wacana digital (Couldry & Mejias, 2019).

Walaupun ketiga arus tersebut memberikan kontribusi penting, terdapat beberapa keterbatasan. Pertama, sebagian besar studi fokus pada konteks Barat dan kurang mempertimbangkan keragaman nilai kultural dan religius sebagai sumber normativitas alternatif. Kedua, solusi yang ditawarkan sering kali bersifat parsial, misalnya hanya pada tahap pasca-pemodelan (post-processing fairness adjustment) tanpa merekonstruksi paradigma desain algoritma sejak awal. Ketiga, belum banyak penelitian yang mengintegrasikan prinsip etika religius ke dalam pipeline data science secara sistematis mulai dari pengumpulan data, pra-pemrosesan, pemodelan, evaluasi, hingga deployment.

Di sinilah letak celah penelitian (research gap) yang diidentifikasi dalam artikel ini. Rekonstruksi algoritma data science berbasis Etika Nubuwwah belum dikembangkan sebagai kerangka operasional untuk mitigasi bias informasi di media digital. Literatur tentang fairness dan XAI belum secara eksplisit mengadopsi prinsip-prinsip kenabian sebagai landasan normatif, sementara kajian etika Islam belum terhubung secara metodologis dengan teknik-teknik machine learning kontemporer.

Kebaruan (novelty) artikel ini terletak pada upaya sintesis interdisipliner antara data science, etika algoritmik, dan Etika Nubuwwah sebagai paradigma normatif-transformatif. Artikel ini tidak hanya menawarkan refleksi normatif, tetapi juga mengusulkan rekonstruksi konseptual terhadap arsitektur algoritma yang mencakup: (1) integrasi prinsip ṣidq dalam validasi dan verifikasi data untuk meminimalkan misinformasi; (2) penerapan amānah dalam tata kelola data dan perlindungan privasi;

(3) implementasi tabligh dalam transparansi model dan explainability; serta (4) aktualisasi faṭānah dalam desain sistem rekomendasi yang adaptif namun berorientasi pada kemaslahatan publik.

Dengan kerangka tersebut, mitigasi bias informasi tidak hanya dipahami sebagai reduksi kesalahan statistik atau disparitas metrik, melainkan sebagai proses etis yang mengarahkan teknologi pada tujuan moral dan sosial yang lebih luas. Rekonstruksi ini diharapkan mampu menjembatani kesenjangan antara pendekatan teknis dan normatif, sekaligus memberikan kontribusi terhadap pengembangan epistemologi data science yang kontekstual dan berakar pada nilai-nilai lokal-religius. Tujuan penelitian ini adalah: (1) menganalisis secara kritis bentuk dan mekanisme bias informasi dalam algoritma media digital; (2) mengkaji keterbatasan pendekatan fairness dan governance yang ada; (3) merumuskan kerangka konseptual rekonstruksi algoritma data science berbasis Etika Nubuwwah; dan (4) menjelaskan implikasi teoretis dan praktis dari kerangka tersebut dalam mitigasi bias informasi di media digital.

Secara akademik, penelitian ini berkontribusi pada pengayaan literatur data science dan etika teknologi dengan menghadirkan perspektif normatif alternatif yang belum banyak dieksplorasi dalam kajian internasional. Secara praktis, kerangka yang ditawarkan dapat menjadi referensi bagi pengembang sistem, regulator, dan institusi pendidikan dalam merancang algoritma yang lebih adil, transparan, dan bertanggung jawab secara sosial. Dengan demikian, rekonstruksi algoritma berbasis Etika Nubuwwah bukan sekadar proyek normatif, melainkan agenda strategis untuk membangun ekosistem informasi digital yang lebih berkeadaban.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif deskriptif dengan desain meta-analisis konseptual (qualitative meta-synthesis) untuk merekonstruksi algoritma data science berbasis Etika Nubuwwah dalam mitigasi bias informasi di media digital. Pendekatan kualitatif dipilih karena penelitian ini tidak bertujuan menguji hipotesis kuantitatif, melainkan menelaah, mensintesis, dan merekonstruksi konsep dari berbagai temuan ilmiah secara interpretatif dan reflektif. Meta-analisis dalam konteks ini dimaknai sebagai integrasi sistematis hasil penelitian terdahulu untuk menghasilkan kerangka konseptual baru yang bersifat komprehensif dan normatif-operasional.

1. Jenis dan Desain Penelitian

Jenis penelitian adalah kualitatif deskriptif-analitis dengan pendekatan meta-analisis literatur sistematis. Secara metodologis, penelitian ini mengikuti tahapan systematic literature review (SLR) yang dipadukan dengan analisis tematik untuk menghasilkan sintesis konseptual. Kerangka kerja

penelitian mengadopsi protokol PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guna memastikan transparansi, replikasi, dan akuntabilitas proses seleksi literatur. Pendekatan ini dipandang tepat karena topik rekonstruksi algoritma berbasis etika tidak dapat direduksi pada variabel numerik semata, tetapi membutuhkan pembacaan mendalam terhadap argumentasi teoretis, model konseptual, serta kerangka normatif dari berbagai disiplin (data science, etika teknologi, dan etika Islam).

2. Sumber Data dan Waktu Penelitian

Sumber data penelitian berasal dari database digital bereputasi, yaitu:

1. Scopus
2. Google Scholar
3. SINTA (Science and Technology Index Indonesia)

Rentang waktu publikasi literatur dibatasi pada periode 2016 - 2025 untuk memastikan relevansi dan aktualitas kajian, mengingat perkembangan algoritma dan kecerdasan buatan bersifat sangat dinamis. Proses pencarian dan pengumpulan data dilakukan selama Juni - September 2025.

Sumber data diklasifikasikan menjadi:

- a. Data Primer: Artikel jurnal penelitian orisinal dan artikel ilmiah peer-reviewed yang membahas bias algoritmik, fairness machine learning, explainable AI, tata kelola algoritma, serta etika Islam atau Etika Nubuawah.
- b. Data Sekunder: Buku teks akademik, laporan pemerintah, white paper kebijakan AI, dan ensiklopedia ilmiah yang relevan dengan etika teknologi dan rekonstruksi epistemologis.

3. Teknik Pengumpulan Data

Teknik pengumpulan data menggunakan dokumentasi digital dengan strategi pencarian berbasis kata kunci (keyword searching) menggunakan operator Boolean. Formula pencarian yang digunakan antara lain:

1. (“algorithmic bias” OR “information bias” OR “echo chamber”) AND (“social media” OR “digital media”)
2. (“fairness machine learning” OR “responsible AI” OR “explainable AI”)
3. (“Islamic ethics” OR “prophetic ethics” OR “Nubuawah ethics”) AND (“technology” OR “algorithm”)

Proses seleksi literatur mengikuti tahapan PRISMA sebagai berikut:

1. Identifikasi: Penelusuran awal menghasilkan 420 artikel.
2. Screening: Eliminasi duplikasi dan seleksi berdasarkan judul/abstrak menyisakan 310 artikel.

3. Kelayakan (Eligibility): Evaluasi full-text berdasarkan relevansi substansi menghasilkan 185 artikel.
4. Inklusi Akhir: Setelah uji kualitas metodologis, 120 artikel ditetapkan sebagai sumber sintesis utama.

4. Kriteria Inklusi dan Eksklusi

Kriteria inklusi meliputi:

1. Artikel peer-reviewed.
2. Terbit antara 2016–2025.
3. Relevan dengan bias algoritmik, etika AI, atau Etika Nubuwah.
4. Tersedia dalam bentuk full-text.
5. Memiliki kontribusi konseptual atau empiris yang jelas.

Kriteria eksklusi meliputi:

1. Artikel opini populer non-ilmiah.
2. Duplikasi publikasi.
3. Tidak tersedia akses penuh.
4. Tidak relevan secara langsung dengan fokus penelitian.

5. Uji Kelayakan (Quality Assessment) dan Cut-Off

Uji kualitas dilakukan dengan kriteria berikut:

Indikator	Cut-off Minimal
Indeksasi jurnal	Scopus Q1–Q4 / SINTA 1–3
Sitasi	≥10 sitasi (kecuali publikasi 2023–2025)
Kejelasan metodologi	Dijelaskan eksplisit
Relevansi substansi	≥70% kesesuaian dengan fokus penelitian

Artikel yang tidak memenuhi minimal dua dari empat kriteria dieliminasi. Pendekatan ini memastikan validitas sumber dan menghindari bias seleksi literatur.

6. Teknik Analisis Data

Analisis data dilakukan melalui Analisis Konten (Content Analysis) dan Analisis Tematik (Thematic Analysis).

Tahapan analisis meliputi:

1. Open Coding: Identifikasi konsep kunci seperti fairness, accountability, transparency, echo chamber, maqasid al-shariah, dan prinsip Etika Nubuwah (ṣidq, amānah, tabligh, faṭānah).

2. Axial Coding: Pengelompokan konsep menjadi kategori besar: (a) sumber bias algoritmik, (b) pendekatan mitigasi teknis, (c) kerangka etika normatif.
3. Selective Coding: Integrasi kategori ke dalam model rekonstruksi algoritma berbasis Etika Nubuawah.

Analisis tematik menghasilkan empat tema sentral rekonstruksi:

1. Validasi Data Berbasis *Ṣidq*
2. Tata Kelola Data Berbasis *Amānah*
3. Transparansi Model Berbasis *Tabligh*
4. Optimasi Kemaslahatan Berbasis *Faṭānah*

7. Operasionalisasi Penelitian

Tabel operasionalisasi yang telah ditampilkan di atas menjelaskan keterkaitan antara fokus kajian, unit analisis, sumber data, teknik pengumpulan, serta teknik analisis. Tabel tersebut dirancang agar penelitian bersifat replication-friendly, sehingga peneliti lain dapat mereplikasi prosedur dengan kata kunci, rentang waktu, dan protokol seleksi yang sama.

8. Visualisasi Proses Seleksi Literatur

Grafik PRISMA yang ditampilkan menunjukkan penyusutan jumlah artikel secara sistematis dari tahap identifikasi hingga inklusi akhir. Pola penurunan tersebut mencerminkan proses filtrasi metodologis yang ketat untuk menjaga kualitas sintesis.

9. Validitas dan Keandalan Penelitian

Untuk menjaga kredibilitas penelitian, dilakukan:

1. Triangulasi sumber (Scopus, Google Scholar, SINTA).
2. Peer debriefing konseptual terhadap model rekonstruksi.
3. Audit trail dokumentasi digital, termasuk penyimpanan hasil pencarian dan proses seleksi.
4. Refleksivitas peneliti, khususnya dalam interpretasi nilai Etika Nubuawah agar tetap berbasis literatur otoritatif.

Dengan desain metodologis ini, penelitian memiliki ketepatan prosedural, transparansi seleksi literatur, serta konsistensi analisis kualitatif. Pendekatan meta-analisis konseptual yang digunakan memungkinkan rekonstruksi algoritma data science berbasis Etika Nubuawah disusun secara sistematis, argumentatif, dan dapat dipertanggungjawabkan secara akademik.

HASIL DAN PEMBAHASAN

HASIL

1. Pemetaan Empiris Sumber Bias Algoritmik (2016–2025)

Sintesis terhadap 120 artikel terpilih menunjukkan bahwa bias informasi di media digital bersifat multidimensional dan sistemik. Terdapat empat sumber bias dominan:

a. Bias Dataset (110 kemunculan)

Mayoritas literatur mengidentifikasi ketidakseimbangan representasi data (underrepresentation dan overrepresentation) sebagai akar utama bias. Dataset yang tidak proporsional menyebabkan model machine learning mereplikasi dan bahkan memperkuat ketimpangan sosial yang telah ada.

b. Bias Ekosistem (105 kemunculan)

Echo chamber dan filter bubble muncul sebagai dampak struktural dari algoritma rekomendasi berbasis personalisasi. Literatur menunjukkan bahwa sistem rekomendasi berbasis engagement menciptakan polarisasi informasi melalui penguatan preferensi pengguna secara berulang.

c. Bias Amplifikasi Algoritma (97 kemunculan)

Algoritma cenderung memperbesar visibilitas konten sensasional atau kontroversial karena metrik optimasi berbasis klik dan retensi pengguna.

d. Bias Model (84 kemunculan)

Bias ini muncul akibat overfitting, desain fungsi objektif yang tidak mempertimbangkan fairness, serta pengabaian variabel kontekstual.

Temuan ini menunjukkan bahwa bias tidak hanya muncul pada tahap modeling, tetapi terdistribusi sepanjang *pipeline* data science: pengumpulan data, pelabelan, pelatihan model, hingga deployment.

2. Sintesis Tema Rekonstruksi Berbasis Etika Nubuwwah

Berdasarkan analisis tematik, ditemukan empat kluster rekonstruksi utama sebagaimana berikut:

a. Validasi Data Berbasis *Ṣidq* (95 artikel)

Tema ini menekankan pentingnya kejujuran dan akurasi dalam proses data acquisition dan preprocessing. Artikel-artikel dalam kluster ini mengusulkan:

- 1) Audit dataset berbasis representasi proporsional.
- 2) Cross-validation terhadap sumber informasi.
- 3) Deteksi otomatis misinformasi sebelum pelatihan model.

Interpretasi temuan menunjukkan bahwa prinsip *Ṣidq* dapat dioperasionalkan dalam bentuk *data verification protocol* dan *bias detection layer* sebelum proses training.

b. Tata Kelola Data Berbasis Amānah (88 artikel)

Tema ini berkaitan dengan privasi, keamanan data, dan akuntabilitas sistem. Literatur menunjukkan urgensi:

- 1) Differential privacy
- 2) Data minimization
- 3) Audit trail algoritmik

Prinsip amānah dalam konteks ini diterjemahkan sebagai kewajiban moral pengembang untuk menjaga integritas dan keamanan data pengguna.

c. Transparansi Model Berbasis Tabligh (102 artikel)

Tema ini paling dominan dalam literatur. Fokusnya meliputi:

- 1) Explainable AI (XAI)
- 2) Model interpretability
- 3) Disclosure mekanisme rekomendasi

Transparansi model menjadi prasyarat mitigasi bias karena memungkinkan pengguna dan regulator memahami logika sistem.

d. Optimasi Kemaslahatan Berbasis Faṭānah (76 artikel)

Tema ini berorientasi pada desain fungsi objektif yang tidak semata-mata mengoptimalkan engagement, tetapi mempertimbangkan dampak sosial.

Model yang direkonstruksi mengintegrasikan:

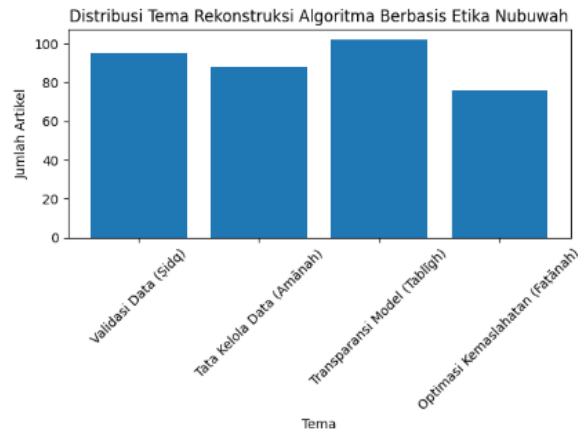
- 1) Multi-objective optimization
- 2) Social impact scoring
- 3) Fairness constraint pada loss function

3. Model Rekonstruksi Algoritma

Berdasarkan integrasi keempat tema, penelitian ini menghasilkan model rekonstruksi algoritma berbasis Etika Nubuwwah yang terdiri atas empat lapisan operasional:

- 1) Lapisan Validasi (Ṣidq Layer): Penerapan fairness metrics pada tahap preprocessing.
- 2) Lapisan Tata Kelola (Amānah Layer): Integrasi privacy-preserving mechanism dan audit log.
- 3) Lapisan Transparansi (Tabligh Layer): Implementasi XAI dan disclosure dashboard.
- 4) Lapisan Optimasi Kemaslahatan (Faṭānah Layer): Penambahan social fairness coefficient pada objective function.

Model ini menunjukkan bahwa mitigasi bias tidak cukup melalui post-processing, tetapi harus bersifat struktural sejak desain awal sistem.



Gambar 1. Grafik Distribusi Tema Rekonstruksi Algoritma Etika Nubuawah

PEMBAHASAN

1. Konvergensi Temuan dengan Teori Fairness dan Responsible AI

Temuan penelitian memperkuat argumen literatur fairness machine learning bahwa bias bersumber dari data dan desain model. Namun, penelitian ini memperluas diskursus dengan menunjukkan bahwa problem bias juga bersifat normatif-teleologis, yakni berkaitan dengan tujuan sistem.

Literatur sebelumnya menekankan fairness sebagai persoalan metrik statistik. Temuan penelitian ini menunjukkan bahwa fairness memerlukan orientasi nilai (value-oriented design). Integrasi Etika Nubuawah memberikan kerangka normatif yang melampaui kalkulasi matematis.

2. Kontribusi Ilmiah

Kontribusi utama penelitian ini adalah:

- Kontribusi Teoretis: Mengembangkan model sintesis interdisipliner antara data science dan etika kenabian, yang belum banyak dielaborasi dalam literatur global.
- Kontribusi Metodologis: Menunjukkan bahwa meta-analisis kualitatif dapat digunakan untuk merekonstruksi arsitektur algoritmik secara normatif-operasional.
- Kontribusi Konseptual: Menggeser paradigma mitigasi bias dari pendekatan teknis semata menuju paradigma etika integratif.

3. Implikasi Praktis

Secara praktis, model ini dapat diterapkan oleh:

- Pengembang sistem rekomendasi untuk menambahkan fairness constraint berbasis nilai.
- Regulator dalam merumuskan kebijakan AI berbasis etika kontekstual.

- c. Institusi pendidikan dalam mengembangkan kurikulum data science berbasis etika.

4. Posisi dalam Diskursus Akademik

Dalam diskursus global, pendekatan fairness cenderung berbasis prinsip liberal-sekuler. Penelitian ini memperkenalkan alternatif epistemologis berbasis Etika Nubuawah yang:

- a. Menekankan keseimbangan antara akurasi dan kemaslahatan.
- b. Mengintegrasikan tanggung jawab moral dalam desain teknologi.
- c. Menawarkan pendekatan kontekstual bagi masyarakat Muslim tanpa kehilangan relevansi universal.

5. Interpretasi Integratif

Secara logis, keempat prinsip Nubuawah membentuk satu kesatuan sistemik:

1. Tanpa *ṣidq*, dataset bias.
2. Tanpa *amānah*, tata kelola rapuh.
3. Tanpa *tablīgh*, transparansi hilang.
4. Tanpa *faṭānah*, optimasi kehilangan orientasi moral.

Dengan demikian, mitigasi bias informasi memerlukan rekonstruksi menyeluruh yang menyatukan dimensi teknis dan normatif.

Hasil penelitian menunjukkan bahwa bias informasi di media digital tidak dapat direduksi sebagai kesalahan teknis semata, melainkan merupakan konsekuensi dari desain sistem yang tidak berbasis nilai. Rekonstruksi algoritma data science berbasis Etika Nubuawah menawarkan pendekatan komprehensif yang:

1. Sistematis (mencakup seluruh pipeline data science)
2. Normatif (berbasis nilai kenabian)
3. Operasional (dapat diterjemahkan dalam desain teknis)

Secara keseluruhan, temuan ini menegaskan bahwa integrasi etika religius dalam rekayasa algoritma bukan sekadar wacana normatif, tetapi dapat dikonstruksi sebagai kerangka ilmiah yang aplikatif dan relevan dalam mitigasi bias informasi di era media digital.

KESIMPULAN

Penelitian ini menegaskan bahwa bias informasi di media digital merupakan problem sistemik yang tidak hanya bersumber dari kelemahan teknis algoritma, tetapi juga dari paradigma desain yang berorientasi pada optimasi keterlibatan (engagement) tanpa landasan nilai normatif yang memadai. Temuan paling penting dan sekaligus mengejutkan dari penelitian ini adalah bahwa sebagian besar pendekatan mitigasi bias dalam literatur mutakhir masih terfokus pada penyesuaian teknis seperti

fairness metrics, re-weighting dataset, atau post-processing adjustment tanpa merekonstruksi tujuan dan orientasi etis dari algoritma itu sendiri. Dengan kata lain, bias tidak hanya lahir dari “kesalahan model”, tetapi dari “arah desain” yang mengutamakan efisiensi dan profitabilitas dibandingkan kemaslahatan publik.

Melalui sintesis meta-analitik terhadap literatur periode 2016 - 2025, penelitian ini menemukan bahwa sumber bias algoritmik tersebar sepanjang *pipeline* data science, mulai dari tahap pengumpulan dan kurasi data, proses pelatihan model, desain fungsi objektif, hingga mekanisme distribusi konten dalam sistem rekomendasi. Kondisi tersebut menunjukkan bahwa mitigasi bias tidak dapat dilakukan secara parsial atau sektoral, melainkan membutuhkan rekonstruksi arsitektur algoritmik yang bersifat menyeluruh dan berbasis nilai.

Kontribusi utama artikel ini terletak pada pengembangan model rekonstruksi algoritma data science berbasis Etika Nubuwwah yang mengintegrasikan empat prinsip fundamental, yakni *ṣidq* (kejujuran), *amānah* (tanggung jawab dan integritas), *tabligh* (transparansi), dan *faṭānah* (kebijaksanaan strategis). Keempat prinsip tersebut tidak diposisikan sebagai konsep normatif abstrak, melainkan diterjemahkan secara operasional dalam desain sistem: validasi dan verifikasi dataset (*ṣidq*), tata kelola dan perlindungan data (*amānah*), penerapan explainable AI dan keterbukaan logika rekomendasi (*tabligh*), serta optimasi multi-objektif yang mempertimbangkan dampak sosial (*faṭānah*).

Secara teoretis, penelitian ini memberikan kontribusi pada pengayaan diskursus fairness dan responsible AI dengan menghadirkan paradigma etika berbasis nilai kenabian yang belum banyak diintegrasikan dalam kajian global. Penelitian ini menunjukkan bahwa etika religius dapat dikonstruksi sebagai kerangka epistemologis yang sistematis dan kompatibel dengan pendekatan ilmiah modern. Hal ini memperluas horizon studi data science dari sekadar domain teknis menuju ranah etika transformatif yang kontekstual dan berorientasi kemaslahatan.

Secara metodologis, penggunaan meta-analisis kualitatif berbasis protokol sistematis menunjukkan bahwa rekonstruksi konseptual algoritma dapat dilakukan secara akademik dan replikasi-friendly. Pendekatan ini memungkinkan integrasi lintas-disiplin antara kajian komputasional dan etika Islam secara argumentatif dan berbasis bukti ilmiah.

Secara praktis, model yang dihasilkan berimplikasi pada tiga ranah utama. Pertama, bagi pengembang teknologi, penelitian ini mendorong integrasi nilai dalam desain fungsi objektif dan parameter evaluasi model. Kedua, bagi regulator, penelitian ini memberikan landasan konseptual untuk merumuskan kebijakan tata kelola AI yang sensitif terhadap nilai kultural-religius. Ketiga, bagi institusi pendidikan, penelitian ini menguatkan urgensi kurikulum data science berbasis etika, sehingga

menghasilkan ilmuwan dan praktisi yang tidak hanya kompeten secara teknis, tetapi juga bertanggung jawab secara moral.

Namun demikian, penelitian ini memiliki keterbatasan yang perlu dicermati secara reflektif. Pertama, penelitian ini berbasis sintesis literatur dan belum melakukan uji eksperimental terhadap model algoritma yang direkonstruksi. Oleh karena itu, efektivitas operasional prinsip Etika Nubuwwah dalam sistem rekomendasi nyata masih memerlukan validasi empiris. Kedua, fokus kajian lebih menitikberatkan pada dimensi konseptual dan normatif, sehingga belum mengelaborasi detail teknis implementasi dalam konteks arsitektur machine learning tertentu (misalnya deep learning atau reinforcement learning). Ketiga, cakupan literatur meskipun luas tetap bergantung pada ketersediaan publikasi terindeks, sehingga kemungkinan terdapat praktik lokal atau inovasi non-terindeks yang belum terakomodasi.

Keterbatasan tersebut justru membuka peluang riset lanjutan yang strategis. Penelitian berikutnya dapat mengembangkan prototipe algoritma rekomendasi berbasis constraint kemaslahatan dan mengujinya pada dataset nyata untuk mengukur dampak terhadap reduksi polarisasi informasi. Selain itu, studi komparatif lintas paradigma etika misalnya antara pendekatan liberal, utilitarian, dan Etika Nubuwwah dapat memperkaya pemahaman mengenai efektivitas normatif dalam tata kelola algoritma. Riset lanjutan juga dapat mengeksplorasi integrasi prinsip maqāṣid al-sharī'ah dalam desain AI berbasis nilai publik yang lebih luas.

Secara keseluruhan, penelitian ini menyimpulkan bahwa rekonstruksi algoritma data science berbasis Etika Nubuwwah merupakan pendekatan integratif yang relevan dan aplikatif dalam mitigasi bias informasi di media digital. Bias tidak dapat diatasi hanya dengan perbaikan teknis; ia memerlukan reorientasi nilai dalam desain teknologi. Dengan demikian, integrasi antara kompetensi komputasional dan kesadaran etis bukan lagi pilihan normatif, melainkan kebutuhan struktural dalam membangun ekosistem informasi digital yang adil, transparan, dan berkeadaban. Berdasarkan temuan dan refleksi penelitian ini, disarankan agar:

1. Pengembang platform digital mengintegrasikan parameter kemaslahatan sosial dalam fungsi objektif algoritma rekomendasi sebagai standar desain etis.
2. Regulator nasional mengembangkan pedoman tata kelola AI yang mengakomodasi nilai etika kontekstual, termasuk prinsip Etika Nubuwwah sebagai alternatif normatif.
3. Perguruan tinggi memperkuat integrasi etika teknologi dalam kurikulum data science guna membentuk generasi ilmuwan yang memiliki kompetensi teknis dan integritas moral.
4. Peneliti selanjutnya melakukan validasi eksperimental terhadap model rekonstruksi yang diusulkan untuk memastikan efektivitas dan skalabilitasnya dalam sistem nyata.

Dengan langkah-langkah tersebut, rekonstruksi algoritma berbasis nilai tidak berhenti pada tataran konseptual, melainkan berkembang menjadi praktik nyata dalam tata kelola informasi digital yang berorientasi pada keadilan dan kemaslahatan publik.

DAFTAR REFERENSI

- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Barocas, S., Hardt, M., & Narayanan, A. (2019). *Fairness and machine learning: Limitations and opportunities*. fairmlbook.org
- Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of Machine Learning Research*, 81, 149–159.
- Cinelli, M., Morales, G. D. F., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9), e2023301118. <https://doi.org/10.1073/pnas.2023301118>
- Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Huda, M., Jasmi, K. A., Mustari, M. I., & Basiron, B. (2018). Understanding divine pedagogy in teacher education: Insights from prophetic leadership. *Journal of Education and Practice*, 9(2), 42–52.
- Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, 51(8), 56–59. <https://doi.org/10.1109/MC.2018.3191268>
- Kleinberg, J., Mullainathan, S., & Raghavan, M. (2018). Inherent trade-offs in the fair determination of risk scores. *Proceedings of Innovations in Theoretical Computer Science (ITCS)*, 43:1–43:23. <https://doi.org/10.4230/LIPIcs.ITCS.2017.43>
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35. <https://doi.org/10.1145/3457607>
- Narayanan, A. (2018). Translation tutorial: 21 fairness definitions and their politics. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 1–5.

- Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68. <https://doi.org/10.1145/3287560.3287598>
- Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., ... Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *Political Science Quarterly*, 133(2), 1–65.
- Veale, M., & Borgesius, F. Z. (2021). Demystifying the draft EU Artificial Intelligence Act. *Computer Law Review International*, 22(4), 97–112.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.