



DAKWAH ISLAM DAN PENAPISAN KANDUNGAN OLEH AI: SATU ANALISIS KRITIS TERHADAP CABARAN PENDAKWAH MEDIA SOSIAL

ISLAMIC PREACHING AND AI CONTENT MODERATION: A CRITICAL ANALYSIS OF THE CHALLENGES FACED BY SOCIAL MEDIA PREACHERS

Ahmad Hamdani Fakhruddin Mukhtar¹, Ahmad Akmal Wan Muhayudin^{2*}

¹ Jabatan Syariah, Universiti Islam Antarabangsa Tuanku Syed Sirajuddin (UniSIRAJ) Perlis, Malaysia

Email: ahmadhamdani@unisiraj.edu.my

² Jabatan Usuluddin, Universiti Islam Antarabangsa Tuanku Syed Sirajuddin (UniSIRAJ) Perlis, Malaysia

Email: ahmadakmal@unisiraj.edu.my

* Corresponding Author

Article Info:

Article history:

Received date: 27.03.2025

Revised date: 14.04.2025

Accepted date: 15.05.2025

Published date: 05.06.2025

To cite this document:

Mukhtar, A. H. F., & Wan Muhayudin, A. A. (2025). Dakwah Islam Dan Penapisan Kandungan Oleh Ai: Satu Analisis Kritis Terhadap Cabaran Pendakwah Media Sosial. *International Journal of Law, Government and Communication*, 10 (40), 192-202.

DOI: 10.35631/IJLGC.1040014

This work is licensed under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)



Abstrak:

Kepesatan inovasi digital seharusnya dimanfaatkan oleh umat Islam sebagai medium dakwah yang penting pada hari ini. Kajian ini bertujuan untuk meneliti cabaran yang dihadapi oleh pendakwah Islam dalam menyampaikan mesej dakwah melalui media sosial, khususnya apabila kandungan mereka ditapis oleh sistem kecerdasan buatan (AI) yang digunakan oleh platform digital utama seperti YouTube, TikTok dan Facebook. Meskipun system penapisan kandungan melalui AI dibangunkan untuk mengekang penyebaran maklumat berbahaya, ia sering gagal menangani konteks keagamaan dan budaya, sehingga mengakibatkan kandungan dakwah yang sah disekat secara tidak wajar. Kajian ini menggunakan pendekatan kualitatif melalui kaedah analisis kandungan terhadap laporan media, laporan pengguna dan kajian terdahulu yang mendokumentasikan kes penapisan kandungan dakwah. Dapatan kajian menunjukkan bahawa penapisan oleh AI bukan sahaja berpunca daripada algoritma yang dilatih dengan data yang bias, tetapi turut dipengaruhi oleh laporan pengguna serta polisi penyedia platform yang kabur terhadap kandungan berunsur keagamaan. Kajian mencadangkan keperluan penambahbaikan sistem penapisan kandungan yang lebih kontekstual, adil dan telus. Kajian ini turut mencadangkan penglibatan aktif institusi dakwah dan pendakwah secara umumnya dalam memperkasa literasi digital dalam kalangan pendakwah agar mereka mampu menyesuaikan diri dengan cabaran algoritma tanpa mengorbankan nilai asas dakwah.

Katakunci:

Dakwah Islam, Penapisan Kandungan, Kecerdasan Buatan (AI), Algoritma, Ketidakadilan Epistemik, Media Sosial, Literasi Digital

Abstract:

The rapid advancement of digital innovation should be harnessed by Muslims as a vital medium for contemporary da'wah. This study aims to examine the challenges faced by Islamic preachers in conveying their message through social media, particularly when their content is filtered by artificial intelligence (AI) systems employed by major digital platforms such as YouTube, TikTok, and Facebook. Although AI-based content moderation systems are developed to curb the spread of harmful information, they frequently fail to account for religious and cultural contexts, resulting in the unjustified restriction of legitimate da'wah content. This research adopts a qualitative approach, employing content analysis of media reports, user complaints, and previous studies that document cases of da'wah content censorship. The findings indicate that AI-driven censorship stems not only from algorithms trained on biased datasets but is also influenced by user reports and ambiguous platform policies concerning religious content. The study recommends improvements to content moderation systems to make them more contextual, fair, and transparent. It further suggests the active involvement of Islamic institutions and preachers in strengthening digital literacy among da'wah practitioners, enabling them to navigate algorithmic challenges without compromising the fundamental values of da'wah.

Keywords:

Islamic Dakwah, Content Filtering, Artificial Intelligence (AI), Algorithm, Epistemic Injustice, Social Media, Digital Literacy.

Pengenalan

Dakwah Islam merupakan satu usaha suci dalam menyampaikan ajaran Islam kepada masyarakat secara aman dan berhemah, sama ada kepada umat Islam sendiri mahupun kepada bukan Islam. Seiring dengan perkembangan teknologi digital, pendekatan dakwah kini semakin beralih ke platform dalam talian yang lebih terbuka, pantas dan bersifat global. Dalam persekitaran ini, kecerdasan buatan (AI) memainkan peranan penting dalam menyaring dan mengawal kandungan yang tersebar di media sosial. Namun, kemunculan AI sebagai "penjaga gerbang" digital turut membawa cabaran baru terhadap kebebasan bersuara dan penyebaran mesej keagamaan.

Kemajuan pesat dalam teknologi AI tanpa dinafikan telah merevolusikan pendekatan terhadap mutu pengurusan kandungan digital, khususnya di platform seperti Facebook, YouTube dan TikTok yang kini bergantung kepada algoritma penapisan automatik. Sistem ini direka bentuk untuk mengesan kandungan yang dianggap melanggar garis panduan komuniti, termasuk ujaran kebencian, ekstremisme, dan maklumat palsu (Gorwa, 2020). Walaupun pendekatan ini bertujuan untuk memastikan keselamatan dan keharmonian dalam ruang digital, ia turut menimbulkan kebimbangan terhadap ketelusan dan keadilan dalam pelaksanaan proses penapisan tersebut.

Dalam konteks kandungan keagamaan, khususnya dakwah Islam, terdapat peningkatan laporan mengenai insiden penapisan yang berlaku secara automatik tanpa pertimbangan terhadap konteks sebenar kandungan yang disampaikan. Perkataan atau frasa seperti *jihad*, *syariah*, dan *khilafah* sering kali ditandai oleh sistem AI sebagai elemen sensitif, meskipun digunakan dalam kerangka akademik atau dakwah yang sah (Rahman & Samat, 2021). Fenomena ini menimbulkan persoalan sama ada algoritma yang digunakan bersifat neutral, atau sebaliknya memperkukuh bias sedia ada terhadap kandungan Islamik yang dikaitkan secara stereotaip dengan ekstremisme (Noble, 2018).

Permasalahan kajian ini dianggap signifikan kerana tanggungjawab dakwah merupakan kewajipan yang dipikul oleh setiap individu Muslim, tanpa mengira tahap kemampuan atau latar belakang pengetahuan mereka. Selagi seseorang itu memiliki kapasiti untuk menyampaikan kebaikan dan mencegah kemungkaran, maka tanggungjawab tersebut tetap perlu dilaksanakan mengikut kemampuan masing-masing. Allah SWT berfirman dalam Surah Ali 'Imran ayat 110: "*Kamu adalah sebaik-baik umat yang dilahirkan bagi (faedah) umat manusia, (kerana) kamu menyuruh berbuat segala perkara yang baik dan melarang daripada segala perkara yang salah (buruk dan keji), serta kamu pula beriman kepada Allah.*" Ayat ini menegaskan kedudukan dakwah sebagai asas dalam pembinaan umat Islam yang cemerlang. Seiring dengan itu, sabda Rasulullah SAW: "*Sampaikanlah daripadaku walaupun satu ayat.*" (Riwayat al-Bukhari, no. 3461) memperlihatkan bahawa penyebaran ilmu Islam adalah tanggungjawab kolektif yang perlu dilaksanakan mengikut kadar kemampuan.

Namun, pendekatan dalam penyampaian dakwah harus mengambil kira konteks semasa, sebagaimana yang ditunjukkan oleh Rasulullah SAW dalam fasa awal dakwah di Mekah yang dijalankan secara tertutup bagi melindungi kepentingan umat Islam pada ketika itu. Dalam era digital kini, media sosial menjadi platform utama dakwah, namun penyampaian tersebut semakin mencabar apabila kandungan Islam yang sah ditapis oleh sistem AI atas tafsiran yang terdedah kepada maklumat yang tidak tepat. Hal ini menimbulkan keperluan untuk mempertahankan hak dakwah dalam ruang digital, khususnya apabila pencegahan kemungkaran sendiri dianjurkan oleh Nabi SAW mengikut tahap kemampuan, sama ada dengan tangan, lisan atau hati (Sahih Muslim, no. 49). Justeru, memahami kesan penapisan AI terhadap dakwah adalah langkah penting dalam memastikan amanah ini dapat terus dilaksanakan secara adil dan berkesan.

Di samping itu, penggunaan algoritma dalam menilai dan menapis kandungan dalam talian yang dikatakan melanggar garis panduan komuniti turut menjadi antara permasalahan utama yang diketengahkan dalam kajian ini, khususnya apabila sistem ini gagal memahami konteks keagamaan dan budaya yang kompleks. Akibatnya, berlaku diskriminasi dalam proses penilaian kandungan, di mana mesej dakwah yang sah dan berlandaskan prinsip keamanan disalah tafsir sebagai ujaran kebencian atau kandungan ekstremis. Ketidakkampuan AI untuk membezakan antara teguran agama yang membina dengan seruan kebencian telah mengakibatkan penyekatan kandungan yang sepatutnya dibenarkan untuk tujuan pendidikan dan kerohanian.

Lebih membimbangkan, diskriminasi ini bukan sahaja menghalang kebebasan bersuara pendakwah Muslim, tetapi turut menjejaskan persepsi masyarakat terhadap kefahaman Islam itu sendiri. Penapisan yang berat sebelah dan tidak telus memberi gambaran seolah-olah ajaran Islam tidak sesuai dengan nilai masyarakat digital global, walhal kandungan dakwah sering kali bertujuan untuk memperkukuh moral, keharmonian sosial dan nilai-nilai sejagat. Oleh itu,

diskriminasi dalam penilaian kandungan bukan sekadar isu teknikal, tetapi juga menyentuh aspek hak asasi, integriti agama, dan keadilan maklumat dalam persekitaran digital.

Justeru kajian ini dijalankan bertujuan untuk meneroka cabaran yang dihadapi oleh para pendakwah Islam dalam menyampaikan mesej agama melalui platform digital yang dikawal oleh sistem penapisan automatik. Tumpuan diberikan kepada bagaimana bias dalam data latihan dan mekanisme pelaporan pengguna boleh memberi kesan terhadap kebebasan bersuara serta penyebaran maklumat Islam yang sah dan damai. Dalam era yang menyaksikan perbincangan global tentang kerangka etika dan hak asasi dalam pembangunan AI, isu ini menjadi semakin penting untuk diteliti secara kritis daripada sudut perspektif keagamaan dan teknologi semasa.

Selain itu, kajian ini turut menumpukan analisis terhadap sistem kecerdasan buatan (AI) menapis kandungan dakwah Islam di media sosial, khususnya apabila kandungan keagamaan disalah tafsir sebagai ujaran kebencian atau ekstremisme. Penelitian ini akan mengenal pasti kes-kes semasa yang dialami oleh pendakwah serta meneliti pola penapisan melalui sumber sekunder seperti laporan media dan kajian terdahulu. Analisis ini penting bagi memahami bagaimana AI, yang direka bentuk berdasarkan data dan input tertentu, boleh gagal menangani konteks agama dan budaya dengan tepat.

Dalam masa yang sama, kajian ini meneliti bentuk bias algoritma yang menyumbang kepada penapisan tidak adil serta membuka ruang kepada cadangan pendekatan yang lebih telus dan kontekstual untuk memastikan kebebasan dakwah Islam terus terpelihara dalam ruangan digital. Antara cadangan termasuk pembangunan garis panduan AI yang lebih peka terhadap konteks keagamaan, penyertaan pakar Islam dalam proses penilaian kandungan, serta penekanan terhadap peranan manusia dalam penapisan automatik. Dalam pada itu, kajian ini turut menggesa para pendakwah kontemporari untuk meningkatkan kecekapan dalam kemahiran digital, selaras dengan perkembangan teknologi yang semakin meresap ke dalam kehidupan masyarakat pelbagai lapisan. Dalam era digital ini, kebergantungan terhadap teknologi seperti kecerdasan buatan (AI) dalam mendapatkan maklumat keagamaan semakin meningkat, terutamanya dalam kalangan generasi muda. Fenomena ini memerlukan perhatian serius kerana ia boleh menjejaskan kedudukan dan kepercayaan terhadap autoriti keagamaan sekiranya kandungan yang diakses tidak ditapis dengan konteks dan kefahaman yang betul.

Bil	Domain Isu	Contoh Masalah
1	Kelemahan AI memahami konteks agama	AI gagal bezakan dakwah mendidik dengan ucapan kebencian, menyebabkan kandungan Islamik positif turut dipadam.
2	Sekatan ruang digital komuniti beragama	Pendakwah boleh kehilangan akses akaun (“de-platformed”), menjejaskan motivasi dan capaian dakwah mereka.
3	Bias dalam sistem laporan dan moderasi	Sistem flagging disalah guna oleh pengguna prejudis; AI menerima laporan tanpa semakan kontekstual terhadap kandungan Islam.
4	Kurangnya ketelusan dan garis panduan	Kandungan "borderline" tiada kriteria jelas; pendakwah Islam sukar kenal

		pasti batas, menyebabkan penapisan diri secara berlebihan.
5	Tekanan politik terhadap platform digital	Syarikat tunduk kepada tekanan kerajaan dengan undang-undang menghina agama, menyekat kandungan dakwah tanpa semakan adil.

Jadual 1: Isu Berkaitan Kebebasan Bersuara dan Dakwah Digital Menurut USCIRF

Sumber: United States Commission on International Religious Freedom. (2021).

Sorotan Literatur

Bagi menjalankan kajian ini, beberapa kajian lepas yang penting telah disoroti secara literatur bagi mendapatkan gambaran awal mengenai tajuk yang akan dikaji.

Dakwah Islam dalam Era Digital

Dalam era digital ini, dakwah Islam semakin berkembang melalui penggunaan media sosial sebagai medium utama untuk menyampaikan mesej keagamaan, khususnya di kalangan generasi muda. Media sosial, seperti YouTube, Facebook, dan TikTok, menyediakan platform yang mudah diakses dan memungkinkan penyebaran maklumat secara cepat dan meluas. Salim et al., (2024) dalam kajian bertajuk “*Inovasi Dakwah di Era Digital Melalui Media Sosial*” memfokuskan kajian pada menjelaskan bagaimana media sosial dan teknologi seperti kecerdasan buatan (AI) dan Internet of Things (IoT) dapat menjadi sarana yang efektif dalam menyampaikan pesan dakwah kepada masyarakat luas. Artikel tersebut menggunakan pendekatan kualitatif melalui kajian literatur, di mana penulis mengumpulkan data dari buku-buku, jurnal ilmiah, serta artikel terkait dakwah dan teknologi.

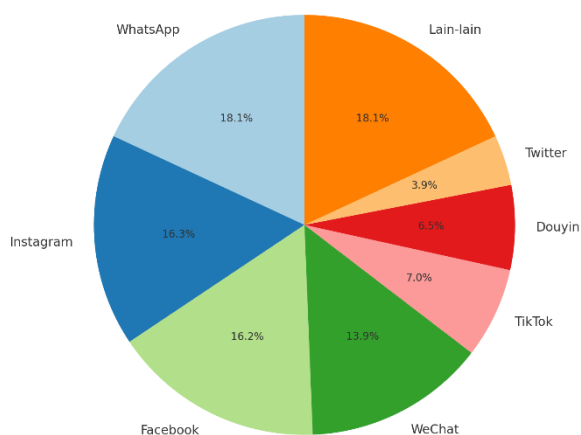
Hasil kajian mendapati bahawa inovasi dakwah digital mencakup penggunaan media sosial seperti YouTube, Instagram, dan TikTok untuk menyampaikan mesej keagamaan secara kreatif dan menarik, serta pemanfaatan teknologi AI untuk membantu perencanaan dan penyusunan strategi dakwah. Di samping itu, pendekatan dakwah digital juga membantu memperluas jangkauan dakwah, menjangkau masyarakat yang sebelumnya sulit diakses melalui metod konvensional. Kajian ini menegaskan pentingnya reformulasi metode dakwah agar relevan dengan karakteristik masyarakat digital yang haus informasi dan terbiasa dengan media cepat serta interaktif (Salim et al., 2024).

Menerusi pemanfaatan tersebut, pendakwah bebas menggunakan pelbagai bentuk kandungan seperti infografik, video pendek, dan kapsyen ringkas untuk menyampaikan mesej dakwah dengan cara yang bersahaja namun berimpak. Pendekatan ini memanfaatkan prinsip *hikmah* dan *mau'izah hasanah*, sesuai dengan kaedah dakwah Islam yang menekankan penyampaian yang lembut dan mengajak kepada kebaikan.

Menurut Abd Bakar et al. (2025) dalam kajian “*Media Sosial dan Isu 3R: Platform Perbincangan atau Pemangkin Ketegangan?*” membincangkan secara menyeluruh peranan media sosial dalam menyuarakan pandangan berkaitan isu-isu sensitif 3R (agama, kaum dan institusi diraja) khusus di Malaysia, dengan menggunakan pendekatan kualitatif menerusi analisis dokumen dan kajian literatur. Meskipun fokus utama adalah pada kebebasan bersuara, ketegangan sosial dan dasar penapisan, dapatan kajian ini memberikan nilai tambah penting kepada wacana dakwah Islam secara digital. Ini kerana dalam konteks masyarakat Malaysia

yang majmuk, penyampaian dakwah melalui media sosial tidak dapat dipisahkan daripada sensitiviti keagamaan dan kawalan kandungan.

Kajian ini menunjukkan bagaimana media sosial telah mengubah landskap komunikasi rakyat Malaysia — termasuk aktivis dan pengengaruh — untuk berbicara mengenai isu agama secara terbuka, walaupun berhadapan risiko undang-undang atau penapisan. Walaupun tidak memberi fokus langsung kepada dakwah Islam, artikel ini menyingkap satu aspek penting iaitu bagaimana dakwah dalam era digital kini beroperasi dalam ruang yang kompleks dan dikawal. Media sosial diiktiraf sebagai medium penting untuk menyampaikan mesej agama kepada khalayak pelbagai latar, namun ia juga menjadi medan konflik apabila dakwah menyentuh isu-isu yang dianggap sensitif seperti status Islam sebagai agama Persekutuan, pensijilan halal atau kedudukan institusi agama (Abd Bakar et al., 2025).



Rajah 1 : Statistik Penggunaan Media Sosial Pilihan di Malaysia

Sumber : DateReportal Global Digital Insight (2023)

Selain itu, artikel ini turut mengetengahkan bagaimana teknologi digital seperti algoritma media sosial dan kawalan kandungan boleh menyekat suara tertentu yang membawa mesej berbentuk keagamaan. Namun kajian tersebut hanya menyentuh penapisan yang dilakukan ke atas mesej yang berunsur hasutan atau kebencian, penyebaran maklumat tidak sah dan berbentuk provokasi (Abd Bakar et al., 2025). Adapun bagi penapisan yang dilakukan secara bias terhadap kandungan dakwah yang benar dan tidak provokatif tidak disentuh oleh penulis.

Penggunaan Kecerdasan Buatan (AI) dalam Islam

Islam adalah agama yang menerima semua bentuk kemajuan teknologi, termasuk perkembangan dalam kecerdasan buatan. Pendekatan Islam terhadap teknologi secara umumnya didasarkan pada konsep bahawa pengetahuan adalah anugerah dari Allah yang harus dimanfaatkan untuk kemaslahatan umat manusia.

Kajian oleh Khoirunnisa et al. (2023) mempersembahkan pendekatan menyeluruh terhadap isu kecerdasan buatan (AI) dari sudut pandang Islam dengan memberi fokus kepada potensi dan ancaman yang wujud dalam penggunaannya. Metodologi kajian ini menggunakan pendekatan kualitatif deskriptif dengan kaedah kajian perpustakaan sebagai sumber utama pengumpulan data terhadap data sekunder pelbagai sumber seperti artikel jurnal, buku akademik, laporan media dan tesis terdahulu berkaitan kecerdasan buatan (AI) dan perspektif Islam. Analisis data

dijalankan secara tematik dan naratif bagi mengenal pasti peluang, ancaman serta garis panduan etika penggunaan AI menurut perspektif Islam.

Dapatan kajian menjelaskan pemanfaatan AI dengan cara yang tepat amat membantu Muslim dalam menyelesaikan permasalahan atau memenuhi keperluan mereka seperti pendidikan al-Quran, pengeluaran fatwa, penyebaran dakwah serta mempertingkatkan keberkesanan rantaian industri halal. Penggunaan teknologi seperti algoritma Decision Tree dan aplikasi mudah alih turut mempercepat akses kepada ilmu dan perkhidmatan Islam. Namun begitu, AI turut membawa ancaman jika tidak dikawal, termasuk risiko penyebaran nilai-nilai murtad, radikalisme dan keganasan secara halus melalui kandungan digital. Justeru, penggunaan AI mestilah berteraskan prinsip etika Islam, menjaga kemaslahatan manusia, tidak melanggar syariah serta menegakkan nilai kemanusiaan dan keadilan sejajar dengan tuntutan maqasid syariah (Khoirunnisa et al., 2023).

Selain itu, kajian yang hampir sama oleh El-Hady dan Zenrif (2024) bertajuk “*Pandangan Islam Terhadap Etika Kecerdasan Buatan (Artificial Intelligence) Dalam Kehidupan Sehari-Hari*” membincangkan etika penggunaan kecerdasan buatan (AI) dalam kehidupan seharian menurut perspektif Islam, dengan menggunakan pendekatan kualitatif melalui kajian literatur dan wawancara mendalam bersama pakar etika Islam dan teknologi. Melalui analisis terhadap sumber klasik dan kontemporari, kajian ini menekankan bahawa Islam menerima kemajuan teknologi termasuk AI selagi ia mendatangkan kemaslahatan dan tidak melanggar prinsip syariah. Antara prinsip etika yang diketengahkan ialah keadilan (*‘adl*), amanah (tanggungjawab), maslahah (kemanfaatan), dan perlindungan maqasid syariah, khususnya *hifz al-nafs*, *hifz al-‘aql*, dan *hifz al-mal*. AI dilihat mampu menyumbang kepada bidang kesihatan, pendidikan, dan ekonomi, namun juga berisiko menjejaskan privasi, mencetuskan diskriminasi algoritma, menghakis nilai kemanusiaan dan mewujudkan ketimpangan sosial jika tidak dikawal secara beretika.

Meskipun kajian-kajian meneliti perspektif Islam terhadap yang menyentuh kelebihan dan kekurangannya terhadap dakwah, namun ia tidak memberi fokus khusus kepada isu penapisan kandungan dakwah oleh AI mahupun pengalaman pendakwah Islam dalam berhadapan dengan algoritma media sosial.

Cabaran Kecerdasan Buatan (AI) Terhadap Umat Islam

Meskipun kajian yang secara khusus meneliti cabaran AI terhadap para pendakwah di media sosial masih sangat terbatas, namun terdapat sejumlah kajian yang telah mengupas implikasi teknologi ini terhadap umat Islam secara umum. Justeru, penelusuran terhadap kajian-kajian tersebut amat signifikan bagi mengenal pasti cabaran-cabaran yang berkemungkinan turut mempengaruhi perkembangan dakwah, walaupun dalam bentuk yang tidak dinyatakan secara langsung. Penelitian ini membolehkan kita memahami secara lebih mendalam kesan tersirat AI terhadap konteks penyampaian dakwah dalam era digital.

Antara kajian tersebut ialah penulisan oleh Hosni et al., (2023) dengan tajuk “*Isu dan Cabaran Chatgpt terhadap Pengajian Islam*” memfokuskan analisis cabaran dan isu yang timbul daripada penggunaan aplikasi kecerdasan buatan khususnya ChatGPT dalam bidang pendidikan agama iaitu Pengajian Islam. Objektif utama kajian adalah untuk mengenal pasti implikasi penggunaan AI terhadap kredibiliti ilmu Islam, masalah etika dan tanggungjawab, serta risiko plagiarisme dalam kalangan pelajar. Metodologi kajian ini bersifat kualitatif dengan pendekatan analisis kandungan. Data dikumpul melalui kajian kepustakaan yang merangkumi

artikel, jurnal, laporan berita dan dokumen berkaitan. Kajian ini menelusuri bagaimana kemunculan ChatGPT yang menyerupai komunikasi manusia memberi kesan terhadap cara penyampaian, rujukan dan penyelidikan dalam konteks keilmuan Islam.

Dapatan kajian mendapati bahawa penggunaan AI, khususnya ChatGPT, membawa cabaran besar terhadap integriti ilmu Islam. Antaranya termasuk penyebaran maklumat yang tidak tepat, potensi manipulasi terhadap teks al-Quran dan hadith, serta kehilangan elemen sanad dalam proses penyampaian ilmu. Selain itu, ChatGPT juga berpotensi mencetuskan isu berat sebelah (bias) terhadap Islam, seperti pengaitan Islam dengan keganasan. Terdapat juga kebimbangan terhadap etika penggunaan AI iaitu pada siapa yang bertanggungjawab jika AI menyampaikan maklumat salah atau menyesatkan. Justeru, kajian ini menekankan bahawa umat Islam perlu bersikap berhati-hati dan menggunakan AI dengan berlandaskan prinsip syarak, demi menjaga ketulenan ajaran Islam dan mengelakkan kemudaratan (Hosni et al., 2023). Kajian terdahulu ini hanya memfokuskan interaksi dengan aplikasi AI Chatgpt yang bukan merupakan platform media sosial namun kajian baharu ini akan memfokuskan bias yang berlaku daripada AI dalam platform media sosial.

Selain itu, berdasarkan kajian bertajuk *"Pendapatan Islam terhadap Etika Kecerdasan Buatan (Artificial Intelligence) dalam Kehidupan Sehari-hari"* oleh El-Hady dan Zenrif (2024), objektif utama penyelidikan ini adalah untuk meneliti perspektif Islam terhadap penggunaan kecerdasan buatan (AI) dalam kehidupan seharian serta merumuskan prinsip-prinsip etika Islam yang boleh dijadikan panduan dalam menangani cabaran kemajuan teknologi ini. Fokus kajian tertumpu kepada aspek etika, moral dan syariah yang menjadi asas dalam penilaian terhadap penggunaan AI, sejauh mana ia selari dengan prinsip-prinsip maqasid syariah seperti pemeliharaan jiwa (ḥifẓ al-nafs), akal (ḥifẓ al-'aql), harta (ḥifẓ al-māl) dan agama (ḥifẓ al-dīn). Kajian ini juga bertujuan mengenalpasti cabaran dan risiko etika yang timbul daripada penggunaan AI serta bagaimana umat Islam boleh menanggapi teknologi ini secara bertanggungjawab dan beretika.

Dari segi metodologi, kajian ini menggunakan pendekatan kualitatif dengan kaedah studi literatur serta wawancara mendalam. Studi literatur merangkumi analisis terhadap kitab-kitab klasik, jurnal ilmiah, dan buku kontemporari yang membincangkan etika Islam dan teknologi. Kaedah hermeneutik turut digunakan untuk menafsirkan teks-teks agama secara kontekstual, agar sesuai dengan realiti moden khususnya dalam isu-isu berkaitan kecerdasan buatan. Di samping itu, wawancara bersama para ulama, ahli teknologi dan cendekiawan Muslim turut dijalankan bagi memperoleh pandangan langsung mengenai penerapan nilai-nilai Islam dalam pembangunan serta penggunaan AI.

Dapatan kajian menunjukkan bahawa Islam secara prinsip menyokong penggunaan teknologi yang memberikan manfaat kepada umat manusia dengan menekankan kepentingan etika dan tanggungjawab dalam penggunaannya. AI berpotensi memberi kemaslahatan dalam bidang kesihatan, pendidikan, dan ekonomi, tetapi ia juga membawa risiko yang besar seperti penyalahgunaan data peribadi, diskriminasi algoritma, dehumanisasi, serta kehilangan nilai-nilai spiritual dan sosial. Kajian ini menyarankan agar pembangunan dan penerapan AI harus dilaksanakan dengan mengutamakan prinsip keadilan, amanah, niat yang baik, serta perlindungan terhadap hak asasi manusia, selaras dengan tuntutan syariah Islam. Maka, satu kerangka etika Islam yang kukuh perlu dibentuk bagi memastikan AI tidak menjadi alat yang merosakkan, sebaliknya menjadi pemangkin kepada kesejahteraan sejagat (El-Hady & Zenrif, 2024).

Tema	Kajian	Fokus Kajian	Kaedah	Penemuan Utama	Kelebihan	Kelemahan / Had
Dakwah Islam dalam Era Digital	Salim et al. (2024) – <i>Inovasi Dakwah di Era Digital Melalui Media Sosial</i>	Penggunaan media sosial dan teknologi AI/IoT dalam dakwah	Kualitatif (kajian literatur)	Media sosial sebagai saluran efektif dakwah; AI bantu strategi dakwah; kandungan perlu kreatif & interaktif	Menyediakan cadangan praktikal untuk pendakwah	Tidak fokus kepada penapisan atau isu bias
	Abd Bakar et al. (2025) – <i>Media Sosial dan Isu 3R</i>	Media sosial dan kebebasan bersuara dalam isu agama, kaum, raja	Kualitatif (analisis dokumen & literatur)	Media sosial sebagai medan dakwah & konflik; sensitiviti keagamaan dikawal; penapisan wujud dalam isu 3R	Kaitkan realiti sosial Malaysia dengan cabaran dakwah	Tidak fokus kepada bias AI secara langsung
Penggunaan Kecerdasan Buatan (AI) dalam Islam	Khoirunnisa et al. (2023) – <i>AI dalam Perspektif Islam</i>	Peluang, ancaman & garis panduan etika AI dalam Islam	Kualitatif (kajian perpustakaan, analisis tematik)	AI bantu pendidikan al-Quran, dakwah & industri halal; risiko nilai sesat & radikalisme; perlu ikut maqasid syariah	Seimbang antara manfaat dan ancaman AI	Tidak fokus kepada media sosial atau penapisan
	El-Hady & Zenrif (2024) – <i>Etika AI dalam Kehidupan Sehari-hari</i>	Pandangan Islam terhadap etika penggunaan AI	Kualitatif (literatur, wawancara, hermeneutik)	Islam sokong AI beretika; tekankan maqasid syariah; risiko diskriminasi & dehumanisasi	Gunakan sumber klasik & kontemporari; libatkan pakar	Tidak fokus kepada konteks dakwah atau penapisan
Cabaran Kecerdasan Buatan (AI) Terhadap Umat Islam	Hosni et al. (2023) – <i>Isu dan Cabaran ChatGPT terhadap Pengajian Islam</i>	Ancaman AI (ChatGPT) terhadap ilmu Islam: bias, etika, ketulenan ilmu	Kualitatif (analisis kandungan)	Risiko maklumat salah; tiada sanad; potensi bias terhadap Islam; persoalan tanggungjawab	Fokus praktikal kepada realiti pengguna Muslim	Tidak menyentuh media sosial secara langsung
	El-Hady & Zenrif (2024) – <i>Etika AI dalam</i>	Etika AI dalam konteks risiko sosial & agama	Kualitatif (literatur, wawancara)	AI boleh bawa manfaat & bahaya; perlukan kerangka etika Islam kukuh		

Jadual 2: Perbandingan Sorotan Kajian Lepas Berdasarkan Tema

Metodologi Kajian

Kajian ini dijalankan menggunakan pendekatan kualitatif dengan kaedah analisis kandungan yang tertumpu kepada sumber-sumber sekunder. Pendekatan ini dipilih kerana ia memberi ruang untuk memahami secara mendalam bagaimana kandungan dakwah Islam ditapis atau dihadkan oleh sistem kecerdasan buatan (AI) dalam ruang digital. Ia juga sesuai untuk menjelajahi corak dan naratif penapisan yang berlaku secara halus, tanpa perlu melibatkan temu bual atau kajian lapangan secara langsung (Given, 2020; Neuendorf, 2017).

Data dikumpulkan secara purposif, iaitu dengan memilih dokumen dan sumber yang berkait rapat dengan isu dan objektif kajian. Empat kategori utama bahan rujukan digunakan:

1. **Laporan media digital**, yang melaporkan kejadian penapisan terhadap kandungan keagamaan di platform seperti Facebook, YouTube, TikTok dan Instagram.
2. **Artikel akademik dan jurnal ilmiah**, yang mengupas isu berkaitan algoritma, penapisan kandungan digital, AI dan kebebasan bersuara.
3. **Dokumen rasmi platform sosial**, termasuk garis panduan komuniti, polisi penyederhanaan kandungan, dan laporan ketelusan tahunan.
4. **Catatan pengguna media sosial**, seperti pendakwah, pempengaruh dan aktivis Muslim yang berkongsi pengalaman mereka tentang sekatan kandungan atau isu seperti *shadow banning*.

Semua data yang digunakan adalah sumber terbuka dan boleh diakses secara sah. Ini menjadikan pendekatan ini tidak hanya praktikal, tetapi juga relevan dalam konteks kajian digital semasa.

Kajian ini berlangsung selama empat hingga enam bulan, merangkumi fasa pengumpulan data, analisis kandungan serta penulisan dapatan akhir.

Bagi menganalisis data, penyelidik menggunakan kaedah analisis tematik, seperti yang dicadangkan oleh Braun dan Clarke (2006). Prosesnya melibatkan beberapa peringkat:

1. Membaca dan menyaring bahan-bahan yang relevan.
2. Mengenal pasti corak dan istilah yang berulang contohnya, istilah agama yang sering dianggap 'sensitif' oleh sistem AI.
3. Mengkategorikan corak-corak tersebut kepada beberapa tema utama seperti sebab penapisan, kesan kepada kebebasan bersuara, serta potensi bias algoritma.
4. Menilai secara kritikal bagaimana sistem AI berfungsi dalam menapis kandungan dan apakah kesannya terhadap penyampaian dakwah Islam.

Melalui pendekatan ini, kajian bukan sahaja dapat mengenal pasti corak penapisan kandungan, tetapi juga membantu memahami bagaimana teknologi canggih seperti AI boleh memberi kesan besar terhadap dakwah dalam masyarakat moden.

Dapatan Kajian dan Perbincangan

Berdasarkan metodologi kajian yang digunakan, kajian ini menghasilkan beberapa dapatan.

Realiti Penapisan Kandungan Dakwah Islam di Platform Digital

Cabaran utama dalam gerakan dakwah kontemporari adalah penapisan kandungan yang dikenakan oleh platform digital, yang memberi kesan terhadap kadar pemahaman dan

penyampaian mesej dakwah. Walaupun keyakinan umat Islam terhadap kebenaran tetap terpelihara, meskipun berhadapan dengan cabaran dan halangan, seperti yang dinyatakan dalam firman Allah: "*Mereka hendak memadamkan cahaya (agama) Allah dengan mulut mereka, tetapi Allah tetap menyempurnakan cahaya-Nya, sekalipun orang-orang kafir tidak menyukainya*" (Surah As-Saff, 61:8), namun tidak dapat dinafikan bahawa impak terhadap saluran dakwah pada aspek praktikalnya tetap membimbangkan.

Kajian ini mendapati bahawa terdapat beberapa kes yang menunjukkan bagaimana kandungan dakwah Islam seringkali menjadi sasaran penapisan oleh platform digital, sama ada melalui intervensi automatik sistem AI, dasar kerajaan, atau tekanan daripada pengguna. Salah satu contoh signifikan yang memperlihatkan cabaran terhadap dakwah Islam dalam ruang digital ialah kes yang melibatkan pendakwah mualaf terkenal Malaysia, Firdaus Wong Wai Hung. Beliau merupakan tokoh awam dalam bidang dakwah yang sangat aktif di media sosial seperti Facebook, YouTube dan TikTok. Beliau juga mengasaskan organisasi dakwah Multiracial Reverted Muslims (MRM) dan dikenali menerusi pendekatan interaktif serta responsif terhadap isu-isu semasa berkaitan akidah, akhlak dan kefahaman Islam khususnya dalam kalangan golongan mualaf (Malay Mail, 2024).

Kes ini tercetus apabila satu video yang dimuat naik oleh Firdaus Wong di platform TikTok sekitar pertengahan tahun 2023 telah menimbulkan kontroversi dalam kalangan masyarakat pelbagai kaum dan agama di Malaysia. Kandungan video tersebut dilaporkan memaparkan penjelasan tentang hak seseorang individu yang telah baligh untuk memeluk agama Islam tanpa keperluan mendapatkan keizinan ibu bapa, berdasarkan prinsip hukum syarak. Walaupun penerangan tersebut bersifat ilmiah dan berdasarkan hukum Islam, beberapa pihak mendakwa bahawa ia berpotensi menggalakkan penukaran agama secara rahsia oleh kanak-kanak bawah umur.

Kontroversi ini semakin memuncak apabila aduan rasmi dan laporan polis dikemukakan oleh badan bukan kerajaan (NGO) dan aktivis yang menuntut agar kandungan video itu dipadamkan. Mereka mendakwa bahawa penyebaran video tersebut bertentangan dengan prinsip keharmonian antara agama serta undang-undang keluarga dan perlindungan kanak-kanak di Malaysia. Akibat tekanan awam ini, Mahkamah Tinggi Kuala Lumpur pada 3 Oktober 2024 telah mengeluarkan perintah interim supaya Firdaus Wong memadamkan video tersebut daripada TikTok serta semua platform media sosial, sementara menunggu keputusan perbicaraan penuh (Malay Mail, 2024).

Firdaus Wong dalam kenyataan balasnya menegaskan bahawa beliau tidak pernah menggalakkan sesiapa memeluk Islam secara rahsia atau secara tidak sah menurut undang-undang. Sebaliknya, beliau menjelaskan bahawa kenyataannya bersandarkan hukum Islam yang mengiktiraf kebebasan beragama bagi individu yang telah mencapai usia baligh. Walaubagaimanapun, tindakan mahkamah dan penurunan kandungan tersebut oleh TikTok menunjukkan wujudnya kecenderungan terhadap penapisan kandungan dakwah yang berasaskan kepada tafsiran luar terhadap sensitiviti masyarakat, bukan semata-mata kepada pelanggaran eksplisit terhadap garis panduan platform digital.

Apa yang menjadi persoalan penting ialah sejauh mana penapisan ini dilakukan berdasarkan pertimbangan yang adil dan kontekstual. Dalam konteks ini, platform TikTok telah mengambil langkah mematuhi perintah mahkamah tanpa mempertimbangkan konteks sebenar penyampaian dakwah tersebut. Tindakan ini memperlihatkan bagaimana kandungan dakwah

Islam boleh menjadi sasaran penapisan apabila berhadapan dengan tekanan undang-undang dan persepsi sosial, meskipun kandungan tersebut tidak melanggar mana-mana prinsip eksplisit dalam garis panduan platform.

Selain itu dalam kes yang lain di luar negara, Ahmad Musa Jibril, seorang pendakwah Palestin-Amerika, kembali aktif di media sosial selepas dibebaskan dari penjara pada 2012. Walaupun ceramah beliau tidak secara langsung menghasut keganasan, kandungannya dikaitkan dengan inspirasi kepada individu yang terlibat dalam aktiviti ekstremis, termasuk dalam serangan London Bridge pada 2017 (ClickOnDetroit, 2017).

Pada mulanya, YouTube enggan memadamkan kandungan beliau kerana tidak melanggar garis panduan komuniti. Namun, tekanan awam dan media memaksa YouTube memperkenalkan dasar baharu pada 2017 untuk menghadkan capaian kepada kandungan ekstremis, termasuk menambah label amaran dan menyekat fungsi komen serta pengiklanan. Kes ini menimbulkan dilema antara kebebasan bersuara dan keselamatan awam. Walaupun kandungan Ahmad Musa Jibril tidak melanggar undang-undang secara langsung, persepsi awam dan tekanan politik mempengaruhi tindakan platform digital. Ini menekankan keperluan untuk dasar penapisan yang telus dan melibatkan pakar agama dalam menilai kesesuaian kandungan dakwah.

Dari sudut yang lebih luas, kes-kes ini menggambarkan cabaran utama terhadap kebebasan bersuara bagi pendakwah Islam dalam era digital, terutama apabila kandungan berbentuk keagamaan mudah disalah tafsir sebagai bertentangan dengan norma majoriti atau perundangan sivil. Ia turut membuka perbincangan mengenai kedudukan media sosial sebagai ruang awam yang kononnya neutral tetapi hakikatnya dikawal oleh algoritma serta terikat dengan kuasa undang-undang dan tekanan politik tertentu.

Selain itu, penapisan kandungan dakwah seperti ini juga menunjukkan ketidakseragaman dalam pendekatan platform digital terhadap isu agama. Tidak ada kejelasan tentang bagaimana kandungan Islam dikategorikan sebagai sensitif, bahkan tidak wujud garis panduan yang konsisten untuk menilai sama ada sesuatu kandungan itu perlu diturunkan. Implikasi ini membawa kepada satu bentuk ketidakadilan epistemik di mana suara keagamaan, khususnya dari sudut pandang Islam, menjadi semakin terpinggir dalam ekosistem digital.

Bias Algoritma dan Ketidakadilan Epistemik dalam Dakwah Islam

Salah satu cabaran etika utama dalam penggunaan kecerdasan buatan (AI), khususnya sistem yang berasaskan algoritma pembelajaran mesin, ialah risiko berlakunya bias atau ketidakberkecualian dalam keputusan yang dihasilkan. Ketidakadilan ini lazimnya berpunca daripada data latihan yang tidak bersifat representatif atau mengandungi kecenderungan diskriminatif, sama ada secara tersurat atau tersirat. Dalam konteks Islam, ketidakadilan seperti ini adalah satu bentuk pelanggaran terhadap prinsip asas keadilan (*al-'adl*), yang merupakan nilai teras dalam ajaran Islam (El-Hady & Zenrif, 2024).

Keadilan menurut perspektif Islam bukan sahaja dituntut dalam hubungan antara individu, malah meliputi sistem sosial, ekonomi dan teknologi. Sebarang keputusan atau tindakan yang membawa kepada penindasan, diskriminasi kaum, gender atau latar belakang sosioekonomi adalah bertentangan dengan semangat keadilan yang dituntut oleh al-Qur'an dan Sunnah. Firman Allah SWT: "*Sesungguhnya Allah menyuruh (kamu) berlaku adil dan berbuat kebajikan...*" (Surah al-Nahl, 16:90) menjadi asas kepada keperluan memastikan setiap aspek kehidupan, termasuk pembangunan teknologi, menepati piawaian keadilan yang hakiki.

Sebagai contoh, dalam sektor kewangan, terdapat laporan yang menunjukkan bahawa sesetengah algoritma pemarkahan kredit memberikan skor yang lebih rendah secara tidak adil kepada kumpulan etnik atau jantina tertentu. Ketidakeimbangan ini berpunca daripada data sejarah yang memaparkan pola ketidakadilan sedia ada, yang kemudiannya dipelajari oleh sistem AI tanpa disaring secara kritis. Kesan daripada sistem sedemikian bukan sahaja menjejaskan peluang ekonomi kelompok tertentu, malah menyumbang kepada jurang sosial yang semakin melebar. Dalam konteks Islam, sistem seperti ini bukan sahaja tidak beretika, malah perlu diperbetulkan agar selari dengan maqasid al-shari'ah yang menekankan keadilan sosial dan pemeliharaan hak asasi manusia (El-Hady & Zenrif, 2024).

Dalam perkembangan dakwah Islam sendiri, penggunaan kecerdasan buatan (AI) seharusnya menawarkan potensi besar untuk memperluas jangkauan dan keberkesanan penyampaian mesej keagamaan. Namun, penggunaan AI juga menimbulkan cabaran serius berkaitan dengan akurasi maklumat, autoriti keagamaan, dan ketidakadilan epistemik yang mungkin berlaku akibat bias algoritma (Batubara, 2024).

AI mempunyai kemampuan untuk mengakses, memproses, dan menyebarkan maklumat dalam skala besar dengan kelajuan tinggi. Keupayaan ini memungkinkan penyebaran mesej dakwah kepada audiens yang lebih luas berbanding dengan kaedah tradisional. Namun, jika maklumat yang disampaikan tidak akurat, ia boleh membawa kepada kesalahpahaman yang meluas dalam kalangan umat Islam. Kesalahan dalam interpretasi teks-teks keagamaan, seperti ayat Al-Qur'an atau hadith, oleh AI yang tidak dilatih dengan baik atau tidak diawasi, boleh menyebabkan penyebaran tafsiran yang menyimpang daripada ajaran Islam yang sebenar.

Salah satu isu kritikal yang timbul adalah berkaitan dengan autoriti dalam mengesahkan kebenaran maklumat keagamaan yang disampaikan oleh AI. Sekiranya berlaku dakwaan bahawa penyampaian dakwah oleh AI adalah menyimpang dan bercanggah dengan nilai sebenar Islam, persoalan utama yang timbul adalah: siapakah yang mempunyai autoriti untuk mengesahkan atau menolak dakwaan tersebut? Adakah pihak yang mempunyai kuasa terhadap platform AI tersebut, atau para ulama dan cendekiawan Islam yang sepatutnya menjadi rujukan utama dalam menentukan kesahihan maklumat keagamaan?

Justeru, bias algoritma dalam AI boleh menyebabkan ketidakadilan epistemik, di mana sistem AI mungkin mengutamakan sumber atau tafsiran tertentu yang tidak mewakili kepelbagaian pandangan dalam Islam. Ini boleh mengakibatkan marginalisasi terhadap perspektif minoriti atau tafsiran alternatif yang sah dalam tradisi Islam (Batubara, 2024). Ketidakadilan epistemik juga boleh berlaku apabila AI menghasilkan maklumat yang memperkukuh stereotaip atau prasangka tertentu, yang boleh menjejaskan pemahaman dan amalan keagamaan individu.

Dari sudut falsafah pengetahuan, fenomena ini dapat dianalisis menerusi lensa teori *epistemic injustice* oleh Fricker (2007). Dalam konteks ini, pendakwah Islam sering kali mengalami dua bentuk ketidakadilan epistemik — *testimonial injustice* apabila suara mereka tidak dianggap sahih oleh sistem, dan *hermeneutical injustice* apabila kandungan mereka tidak dapat difahami kerana ketiadaan kerangka penafsiran yang sesuai dalam algoritma. Ini menunjukkan bahawa sistem AI secara tidak langsung mengekalkan ketidaksaksamaan struktur pengetahuan yang mengutamakan naratif dominan, sekali gus meminggirkan wacana Islam yang sah dan intelektual.

Antara bentuk tapisan yang dilakukan oleh media sosial adalah pada mencadangkan kandungan-kandungan yang bersesuaian kepada pengguna. Berdasarkan kajian Albadi et al. (2022), bias algoritma terhadap kandungan keagamaan turut berlaku pada content recommendation yang juga melibatkan agama Islam. Mereka mendedahkan bahawa sistem saranan video YouTube mempunyai kecenderungan mengekalkan dan menyebarkan kandungan berunsur kebencian terhadap agama tertentu, termasuk terhadap sesama mazhab Islam sendiri. Melalui analisis lebih 350,000 video berbahasa Arab, mereka mendapati bahawa sebahagian besar kandungan yang disarankan oleh algoritma adalah bersifat intoleransi agama dan diskriminatif. Ini memberi kesan langsung kepada aktiviti dakwah Islam yang sah apabila kandungan dakwah yang moderat berisiko terbenam di sebalik limpahan kandungan yang lebih ekstrem atau prejudis.

Kajian tersebut menunjukkan bahawa hampir 30% daripada video yang diperoleh melalui hasil carian YouTube mengandungi kandungan kebencian, manakala lebih 20% daripada saranan video pertama selepas tontonan juga tergolong dalam kategori yang sama. Kandungan kebencian ini bukan sahaja menasarkan agama lain seperti Kristian dan Atheisme, tetapi turut menasarkan sesama Muslim, seperti mazhab Syiah dan juga Sunni, yang sering dijadikan bahan tuduhan atau fitnah (Albadi et al., 2022). Keadaan ini mencerminkan bagaimana algoritma tidak neutral terhadap kandungan agama, sebaliknya mampu menguatkan polarisasi dengan mengitar semula kandungan yang serupa secara berterusan — satu fenomena yang disebut sebagai *'hate loop'*. Situasi ini bukan sahaja memberi kesan terhadap persepsi umum terhadap Islam tetapi juga menyukarkan usaha dakwah yang berpaksikan ilmu dan nilai perdamaian.

Lebih membimbangkan, kajian ini turut membuktikan bahawa profil personalisasi pengguna memainkan peranan besar dalam menentukan kandungan yang disarankan. Pengguna yang mempunyai sejarah tontonan video berunsur radikal atau identiti agama tertentu (seperti Syiah radikal) lebih cenderung menerima saranan kandungan yang ekstrem. Sebagai contoh, profil pengguna Syiah yang bersifat radikal menerima 29% lebih banyak kandungan kebencian berbanding Syiah moderat, manakala pengguna lelaki Sunni menerima lebih banyak kandungan kebencian berbanding wanita Sunni. Secara keseluruhan, personalisasi meningkatkan risiko pendedahan kepada kandungan kebencian sebanyak 46% berbanding pengguna neutral tanpa sejarah tontonan. Ini secara tidak langsung memberi cabaran besar kepada para pendakwah Islam yang ingin menyampaikan mesej dakwah secara digital tetapi tersekat atau dipinggirkan oleh keutamaan algoritma yang berat sebelah terhadap interaksi dan kontroversi (Albadi et al., 2022).

Dapatan ini menunjukkan bahawa sistem saranan automatik berasaskan AI, seperti yang digunakan oleh YouTube tidaklah bebas nilai, sebaliknya dibentuk oleh logik keterlibatan (*engagement*) dan sejarah penggunaan, yang akhirnya boleh mencipta ruang yang tidak adil kepada kandungan agama. Dalam konteks dakwah Islam, bias algoritma seperti ini bukan sahaja menyempitkan ruang penyebaran mesej Islam yang bersifat dakwah dan tarbiah, tetapi juga boleh mencetuskan salah faham dan ketegangan antara agama dan intra-agama.

Implikasi terhadap Kebebasan Bersuara dan Strategi Dakwah Digital

Implikasi penapisan kandungan terhadap kebebasan bersuara dalam dakwah Islam adalah sangat signifikan, khususnya dalam kalangan generasi muda yang menjadikan media sosial sebagai medan utama penyampaian ilmu dan mesej keagamaan masa kini. Pendakwah Islam berdepan dengan tekanan untuk menyesuaikan strategi penyampaian mereka agar tidak

melanggar garis panduan algoritma, termasuk mengelakkan istilah-istilah tertentu yang berpotensi mencetuskan penapisan. Ini bukan sahaja membataskan kebebasan bersuara dan ekspresi agama, malah boleh mengubah bentuk penyampaian dakwah sehingga lari daripada keasliannya. Pendakwah terpaksa berpindah ke platform yang kurang popular atau menggunakan bahasa yang lebih 'selamat' demi mengelakkan sekatan, sekali gus mengurangkan capaian dan impak dakwah mereka.

Justeru, pelaksanaan dasar penapisan kandungan oleh platform media sosial yang bergantung kepada sistem kecerdasan buatan (AI) menimbulkan implikasi serius terhadap kebebasan menyampaikan dakwah Islam di ruang digital. Laporan oleh United States Commission on International Religious Freedom (USCIRF, 2021) telah mengenal pasti beberapa kelemahan utama dalam pendekatan sedia ada, yang secara tidak langsung memberi kesan negatif terhadap komuniti agama, termasuk pendakwah Muslim.

Pertama, sistem AI yang digunakan untuk mengesan kandungan bermasalah sering kali gagal memahami konteks agama dan budaya dalam kandungan yang dimuat naik. USCIRF mengkritik bahawa algoritma penapisan secara automatik sukar membezakan antara ucapan kebencian sebenar dan kandungan yang bertujuan menolak atau melawan kebencian itu sendiri. Keadaan ini menyebabkan kandungan dakwah yang positif dan mendidik turut berisiko untuk disekat atau dipadam secara tidak wajar (USCIRF, 2021). Implikasi ini sangat signifikan bagi pendakwah Islam yang cuba melibatkan diri dalam wacana dakwah kritikal atau menyampaikan pandangan terhadap ketidakadilan agama yang berlaku secara global.

Kedua, penapisan automatik berasaskan AI juga telah menyebabkan penghapusan ruang digital bagi komuniti agama yang sepatutnya mempunyai hak untuk bersuara dan menyampaikan ajaran agama. USCIRF menyatakan bahawa dasar ini bukan sahaja menghalang usaha komuniti agama untuk melawan diskriminasi, tetapi juga berisiko menyebabkan mereka "di-de-platform", iaitu kehilangan akses sepenuhnya terhadap platform digital (USCIRF, 2021). Bagi pendakwah muda atau bebas yang menjadikan media sosial sebagai saluran utama dakwah, kehilangan akses ini bukan sahaja menyekat capaian mereka, malah memberi tekanan psikologi yang menjejaskan motivasi untuk terus berdakwah secara aktif.

Ketiga, proses penapisan kandungan oleh syarikat media sosial juga terdedah kepada bias yang lahir daripada sistem laporan pengguna (*flagging*) dan keputusan moderator manusia yang tidak telus. USCIRF menegaskan bahawa sistem ini boleh menyuburkan prejudis sedia ada terhadap komuniti agama minoriti (USCIRF, 2021). Sebagai contoh, dalam suasana masyarakat yang tidak ramah terhadap Islam, pengguna lain boleh melaporkan kandungan dakwah sebagai bermasalah walaupun kandungan tersebut bersifat ilmiah atau neutral. Sistem AI pula cenderung mempercayai laporan tersebut tanpa semakan kontekstual, lalu bertindak secara tidak adil terhadap pendakwah Islam.

Keempat, laporan ini turut mengkritik kekurangan ketelusan dan akauntabiliti dalam proses penapisan kandungan. USCIRF menyatakan bahawa walaupun berlaku peningkatan dari segi kawalan terhadap kandungan ekstrem, masih wujud jumlah besar kandungan yang disebut sebagai "*borderline content*" — iaitu kandungan yang sukar diklasifikasikan sebagai salah atau tidak. Namun, proses membuat keputusan terhadap kandungan seperti ini kekal tidak telus, termasuk latihan dan garis panduan yang digunakan oleh moderator. Ini menjadikan pendakwah Islam tidak mempunyai rujukan jelas untuk mengetahui batasan kandungan yang

dibenarkan, seterusnya mendorong mereka untuk menapis sendiri ucapan mereka secara berlebihan demi mengelakkan sekatan (USCIRF, 2021).

Akhir sekali, USCIRF membangkitkan isu bahawa banyak syarikat media sosial tunduk kepada tekanan politik dari sesetengah negara yang menggunakan undang-undang menghina agama untuk menyekat kandungan digital. Ini termasuk negara-negara yang mempunyai undang-undang kabur dan berat sebelah terhadap Islam itu sendiri. Syarikat seperti Facebook, menurut laporan tersebut, telah beberapa kali menyekat kandungan atas permintaan kerajaan tanpa semakan menyeluruh terhadap kesahihan aduan tersebut (USCIRF, 2021). Hal ini jelas menunjukkan bahawa kandungan dakwah Islam bukan sahaja berisiko dipadamkan oleh sistem automatik, tetapi juga oleh tekanan luar yang tidak bersifat adil atau objektif.

Kesemua dapatan ini menunjukkan bahawa pelaksanaan dasar penapisan kandungan oleh AI dalam media sosial mempunyai implikasi mendalam terhadap kebebasan dakwah Islam, bukan hanya dari sudut teknikal, tetapi juga dari sudut hak asasi dan psikososial. Oleh itu, usaha reformasi sistem penapisan kandungan perlu dilakukan secara menyeluruh, termasuk dengan melibatkan komuniti agama dan pakar syariah dalam proses pembangunan dasar digital, agar dakwah Islam dapat terus disampaikan dalam suasana yang adil dan kondusif.

Dalam masa yang sama, institusi keagamaan dan badan-badan dakwah juga perlu memperkukuh literasi digital dalam kalangan pendakwah agar mereka lebih bersedia menghadapi cabaran dalam ruang maya yang tidak sentiasa ramah kepada agama. Pendedahan kepada kemahiran digital terhadap pengguna dan pendakwah juga perlu dilatih agar mereka memahami cara bekerja algoritma dan memanfaatkan ruang digital dengan lebih strategik.

Kesimpulan

Dakwah Islam merupakan suatu usaha berterusan yang berhadapan dengan pelbagai bentuk cabaran merentas masa, tempat dan suasana. Ia bukan sekadar tanggungjawab para ulama atau pendakwah rasmi, tetapi satu kewajipan kolektif umat Islam yang perlu dilaksanakan mengikut kemampuan dan tahap ilmu masing-masing. Hal ini bertepatan dengan firman Allah SWT: *"Wahai orang yang berselimut! Bangunlah, lalu berilah peringatan!"* (Surah al-Muddaththir, 74:1–2). Ayat ini menunjukkan perintah langsung kepada Nabi Muhammad SAW untuk memulakan misi dakwahnya setelah menerima wahyu, sekaligus menzahirkan bahawa dakwah adalah tugas besar yang perlu diangkat walaupun berdepan tentangan.

Dalam konteks semasa, kajian ini memberi tumpuan kepada cabaran terkini yang dihadapi oleh pendakwah Islam dalam era kecerdasan buatan (AI), terutamanya berkaitan dengan penapisan kandungan di media sosial. Meskipun teknologi AI membawa banyak manfaat dalam mengekang penyebaran kandungan berbahaya, penerapannya dalam menilai kandungan dakwah perlu diteliti agar tidak berlaku ketidakadilan atau bias algoritma. Proses penapisan haruslah berpaksikan prinsip keadilan, mengambil kira sensitiviti budaya dan agama, serta menjamin hak kebebasan bersuara dalam penyampaian ajaran Islam yang sah.

Disamping itu, beberapa cadangan boleh dikemukakan. Pertama, syarikat teknologi perlu membangunkan sistem AI yang lebih inklusif dan kontekstual dengan melibatkan pakar agama dan budaya dalam proses pembangunan algoritma. Kedua, proses semakan kandungan perlu melibatkan lebih banyak campur tangan manusia yang memahami konteks tempatan sebelum tindakan sekatan diambil.

Ketiga, pendakwah dan institusi Islam disaran memperkuat penggunaan platform alternatif yang memberi lebih ruang kepada kebebasan bersuara secara bertanggungjawab. Keempat, kerjasama strategik antara NGO Islam dan badan hak asasi digital perlu diwujudkan bagi mendesak ketelusan dalam dasar penapisan kandungan oleh platform global.

Akhir sekali, kesedaran dalam kalangan pendakwah dan pengguna media sosial perlu dipertingkatkan tentang bagaimana algoritma berfungsi, serta bagaimana untuk merangka kandungan yang mesra AI tanpa menjejaskan kebenaran mesej Islam. Hanya melalui pendekatan yang inklusif dan telus, sistem penapisan kandungan digital dapat menyokong pembangunan masyarakat global yang adil, terbuka dan menghormati kepelbagaian.

Berdasarkan dapatan dan perbincangan yang telah dikemukakan, dapat dirumuskan bahawa objektif utama kajian ini ialah untuk menganalisis cabaran yang dihadapi oleh pendakwah Islam dalam menyampaikan dakwah melalui media sosial apabila kandungan mereka ditapis oleh sistem kecerdasan buatan (AI), serta mencadangkan pendekatan penyelesaian yang lebih adil dan kontekstual.

Penemuan kajian menunjukkan bahawa sistem AI semasa masih belum mampu memahami konteks kandungan keagamaan secara menyeluruh, sekali gus menyebabkan penapisan tidak wajar terhadap mesej dakwah yang sah. Hal ini membuktikan bahawa kebimbangan yang dibangkitkan dalam objektif kajian adalah berasas. Tambahan pula, kajian ini berjaya mengenal pasti bentuk bias dalam algoritma, termasuk ketidakjelasan garis panduan platform dan pengaruh laporan pengguna yang bersifat subjektif. Justeru, objektif kedua kajian iaitu mengenal pasti punca sebenar penapisan kandungan dakwah boleh dikatakan telah dicapai secara umum.

Begitu juga antara cadangan yang dikemukakan dalam kajian ini seperti pembangunan sistem AI yang lebih sensitif kepada konteks, penyertaan pakar agama dalam proses pembangunan teknologi, serta literasi digital dalam kalangan pendakwah secara langsung menjawab objektif kajian untuk mencadangkan pendekatan alternatif yang lebih telus. Kesemua elemen ini mengukuhkan justifikasi bahawa objektif kajian bukan sahaja tercapai, tetapi turut menyediakan asas yang kukuh untuk tindakan susulan dalam menangani isu penapisan kandungan dakwah Islam dalam era digital.

Pengakuan dan Penghargaan

Pengkaji ingin merakamkan penghargaan terhadap beberapa dapatan penting kajian ini yang menyumbang kepada perkembangan wacana dakwah Islam dalam era kecerdasan buatan (AI):

Pengesanan Realiti Penapisan Kandungan Dakwah

Kajian ini berjaya mengenal pasti bentuk penapisan terhadap kandungan dakwah Islam yang berlaku melalui sistem AI di media sosial, termasuk insiden nyata yang melibatkan pendakwah kontemporari. Dapatan ini memberi gambaran jelas bahawa kandungan yang sah dari segi keagamaan juga berisiko disekat akibat kelemahan sistem penapisan automatik.

Pendedahan Terhadap Bias Algoritma dan Ketidakadilan Epistemik

Analisis menunjukkan bahawa bias dalam algoritma AI bukan sahaja menjejaskan penyebaran maklumat Islam, tetapi turut menimbulkan bentuk ketidakadilan epistemik yang meminggirkan suara keagamaan yang sah. Isu ini jarang disentuh secara langsung dalam kajian sedia ada dan memberi sumbangan baru kepada bidang etika teknologi dan dakwah digital.

Cadangan Strategik untuk Reformasi Penapisan Kandungan

Kajian ini turut menyumbang cadangan praktikal dalam bentuk pembentukan sistem penapisan yang lebih telus dan kontekstual. Antaranya termasuk penglibatan pakar agama dalam pembangunan algoritma, serta keperluan membina literasi digital dalam kalangan pendakwah untuk menyesuaikan strategi dakwah mereka dalam landskap digital yang semakin kompleks.

Penekanan terhadap Kewajipan Kolektif Dakwah

Kajian ini menegaskan bahawa dakwah bukan hanya tugas institusi agama, tetapi merupakan tanggungjawab kolektif umat Islam mengikut kemampuan masing-masing. Penemuan ini memperkukuh kepentingan untuk mempertahankan ruang digital sebagai medan dakwah yang inklusif dan terbuka.

Rujukan

- Abd Bakar, S. N., Abdul Halim, A. H., Ismail, N., Mohd Sharif, S., Ismail, N., & Rameli, M. F. P. (2025). Media sosial dan isu 3R: Platform perbincangan atau pemangkin ketegangan? *Journal of Islamic Philanthropy & Social Finance (JIPSF)*, 7(1), 112–124. https://doi.org/10.24191/JIPSF/v7n12025_112-124
- al-Bukhari, M. I. (n.d.). *Ṣaḥīḥ al-Bukhārī* (ḥadīth no. 3461). Beirut: Dār Ibn Kathīr.
- Albadi, N., Kurdi, M., & Mishra, S. (2022). Deradicalizing YouTube: Characterization, detection, and personalization of religiously intolerant Arabic videos. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), Article 505. <https://doi.org/10.1145/3555618>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Brenner, D. M., Chaudhary, Y., & Geraci, R. M. (2024). Technical and religious perspectives on AI misinformation and disinformation. *SSRN*. <https://ssrn.com/abstract=4360413>
- ClickOnDetroit. (2017, June 6). London Bridge attack: Michigan cleric reportedly involved in radicalization of attacker. <https://www.clickondetroit.com/news/2017/06/06/london-bridge-attack-michigan-cleric-reportedly-involved-in-radicalization-of-attacker/>
- El-Hady, E. H. F., & Zenrif, M. F. (2024). Pandangan Islam terhadap etika kecerdasan buatan (Artificial Intelligence) dalam kehidupan sehari-hari. *Nuansa: Jurnal Penelitian Ilmu Sosial dan Keagamaan Islam*, 21(2), 84–98. <http://ejournal.iainmadura.ac.id/index.php/nuansa/index>
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Given, L. M. (Ed.). (2020). *The SAGE encyclopedia of qualitative research methods*. SAGE Publications.
- Hosni, A., Md Ariffin, M. F., & Ishak, H. (2023). Isu dan cabaran ChatGPT terhadap pengajian Islam. *Al-Turath: Journal of Al-Quran and Al-Sunnah*, 8(1), i–xvii. <https://spaj.ukm.my/jalturath/index.php/jalturath/article/view/205>
- Khoirunnisa, A., Rohman, F., Azizah, H. A., Ardianti, D., Maghfiroh, A. L., & Noor, A. M. (2023). Islam in the midst of AI (Artificial Intelligence) struggles: Between opportunities and threats. *SUHUF: International Journal of Islamic Studies*, 35(1), 1–5. <https://doi.org/10.23917/suhuf.v35i1.22365>
- Malay Mail. (2024, October 3). High Court orders controversial Muslim preacher Firdaus Wong to remove conversion video for minors from TikTok pending trial. <https://www.malaymail.com/news/malaysia/2024/10/03/high-court-orders-controversial-muslim-preacher-firdaus-wong-to-remove-conversion-video-for-minors-from-tiktok-pending-trial/152435>

- Muslim ibn al-Hajjaj. (n.d.). *Sahih Muslim* (Hadis no. 49). Beirut: Dar al-Fikr.
- Neuendorf, K. A. (2017). *The content analysis guidebook* (2nd ed.). SAGE.
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
- Rahman, F., & Samat, N. (2021). Islamic preaching in the digital age: Challenges of content moderation on social media platforms. *Journal of Islamic Media Studies*, 3(1), 45–60.
- Salim, N., Amanda, N. C., & Rohman. (2024). Inovasi dakwah di era digital melalui media sosial. *Qaulan Baligha: Jurnal Ilmu Manajemen Dakwah*, 1(2).
- United States Commission on International Religious Freedom (USCIRF). (2021). *Protecting religious freedom online*. <https://www.uscifr.gov/sites/default/files/2021-12/2021%20Factsheet%20-%20Protecting%20Religious%20Freedom%20Online.pdf>