ABSTRAK

TANGGUNG JAWAB PLATFORM MEDIA SOSIAL INSTAGRAM YANG MENGALAMI KEGAGALAN PADA PROSES PENYARINGAN KONTEN MENGGUNAKAN KECERDASAN BUATAN (*ARTIFICIAL INTELLIGENCE)*

Oleh:

MUHAMMAD YAFI JAWAD RIADI

Perkembangan kecerdasan buatan *Artificial Intelligence* (AI) telah memengaruhi praktik penyaringan konten di media sosial. AI digunakan untuk mendeteksi dan memblokir konten sensitif, namun sering kali mengalami kegagalan yang menimbulkan persoalan hukum dan sosial. Berdasarkan hal tersebut, penelitian ini membahas mengenai tanggung jawab hukum platform media sosial Instagram atas kegagalan penyaringan konten oleh kecerdasan buatan serta sanksi hukum yang dapat dikenakan apabila sistem tersebut tidak berfungsi sebagaimana mestinya dalam menyaring konten sensitif. Kegagalan tersebut kerap disebabkan oleh kesalahan klasifikasi, bias algoritmik, serta ketidakmampuan AI memahami konteks sosial dan bahasa, yang berpotensi merugikan pengguna serta menimbulkan pertanggungjawaban hukum bagi penyelenggara platform.

Penelitian ini merupakan penelitian hukum normatif dengan tipe penelitian deskriptif. Pendekatan yang digunakan meliputi pendekatan perundang-undangan dan pendekatan konseptual. Data penelitian diperoleh melalui studi kepustakaan yang kemudian dianalisis secara kualitatif untuk menarik kesimpulan yang relevan dengan permasalahan penelitian.

Hasil penelitian menunjukkan bahwa tanggung jawab hukum platform media sosial Instagram atas kegagalan penyaringan konten oleh kecerdasan buatan bersifat administratif, regulatif, dan perdata sebagaimana diatur dalam PP Nomor 71 Tahun 2019 dan Permenkominfo Nomor 5 Tahun 2020. Kegagalan moderasi konten dapat berupa *false negative*, yaitu tidak terdeteksinya konten sensitif atau ilegal, serta *false positive*, yaitu penghapusan terhadap konten yang sah. Atas kelalaian tersebut, platform dapat dikenai sanksi administratif berupa teguran tertulis, denda, pembatasan akses, hingga pencabutan izin operasional. Selain itu, tanggung jawab regulatif menuntut kepatuhan terhadap ketentuan UU ITE dan prinsip penyelenggaraan sistem elektronik yang andal dan bertanggung jawab.

Kata kunci: Penyaringan Konten, Kecerdasan Buatan, Tanggung Jawab Hukum.

ABSTRACT

THE RESPONSIBILITY OF THE INSTAGRAM SOCIAL MEDIA PLATFORM REGARDING FAILURES IN CONTENT FILTERING PROCESSES USING ARTIFICIAL INTELLIGENCEE.

By

MUHAMMAD YAFI JAWAD RIADI

The development of Artificial Intelligence (AI) has significantly influenced the practice of content moderation on social media platforms. AI is utilized to detect and block sensitive content; however, it often fails, resulting in various legal and social issues. This study discusses the legal responsibility of the social media platform Instagram for the failure of AI-based content moderation, as well as the legal sanctions that may be imposed when such systems do not function properly in filtering sensitive content. These failures are frequently caused by misclassification, algorithmic bias, and the inability of AI to understand social and linguistic contexts, which may harm users and give rise to legal liability for the platform provider.

This research is a normative legal study with a descriptive research type. The approaches used include the statutory approach and the conceptual approach. The data were obtained through library research and analyzed qualitatively.

The results of this study indicate that the legal liability of the social media platform Instagram for failures in content moderation using artificial intelligence is administrative, regulatory, and civil in nature, as stipulated in Government Regulation Number 71 of 2019 and Regulation of the Minister of Communication and Informatics Number 5 of 2020. Failures in content moderation may take the form of false negatives, namely the failure to detect sensitive or illegal content, as well as false positives, namely the removal of lawful content. As a result of such negligence, the platform may be subject to administrative sanctions in the form of written warnings, administrative fines, access restrictions, or revocation of operational licenses. Furthermore, regulatory liability requires compliance with the provisions of the Electronic Information and Transactions Law as well as the principles of reliable and responsible electronic system operation.

Keywords: Content Moderation, Artificial Intelligence, Legal Responsibility.