



Dilema Etika dan Legalitas *Web Scraping*: Analisis Integritas Praktisi TI Berdasarkan Regulasi Digital Indonesia

Evy Nurmiati^{1*}, Luthfi Herianda Putra²

¹⁻²Sistem Informasi, UIN Syarif Hidayatullah Jakarta, Indonesia

Email: evy.nurmiati@uinjkt.ac.id^{1*}, luthfi.herianda24@mhs.uinjkt.ac.id²

*Penulis Korespondensi: evy.nurmiati@uinjkt.ac.id

Abstract. Massive digital transformation across various sectors has positioned data as a vital instrument in accelerating technological innovation, particularly in the development of the Artificial Intelligence (AI) ecosystem. One of the most effective yet controversial data collection methods to date is web scraping. This research aims to deeply analyze the ethical dilemmas and the legal boundaries of web scraping practices in Indonesia, specifically in the context of providing large-scale datasets for AI model training. The research method employed is descriptive qualitative with a literature review approach toward various positive legal instruments, platform privacy policies, and international professional codes of ethics. The results of the analysis indicate that although web scraping offers significant technical efficiency for various industrial sectors, this practice often overlooks the aspect of consent from data owners and potentially violates Law Number 28 of 2014 concerning Copyright and Law Number 27 of 2022 concerning Personal Data Protection (PDP Law). The emphasis on the integrity of Information Technology (IT) practitioners, referring to the ACM (Association for Computing Machinery) code of ethics, becomes highly crucial as a moral compass amidst the void of specific regulations regarding scraping at the national level. This research concludes that the true professionalism of an IT practitioner is measured not only by the mastery of data extraction technology but also by compliance with legal boundaries, transparency, and respect for private property rights to create a fair, secure, and transparent digital ecosystem.

Keywords: Copyright; Data Practitioner Integrity; IT Legality; Professional Ethics; Web Scraping.

Abstrak. Transformasi digital yang masif di berbagai sektor telah menempatkan data sebagai instrumen vital dalam mengakselerasi inovasi teknologi, khususnya pada pengembangan ekosistem Kecerdasan Buatan (*Artificial Intelligence*). Salah satu metode pengumpulan data yang dinilai paling efektif namun tetap kontroversial hingga saat ini adalah *web scraping*. Penelitian ini bertujuan untuk menganalisis secara mendalam dilema etika serta batasan legalitas praktik *web scraping* di Indonesia, terutama dalam konteks penyediaan *dataset* skala besar untuk pelatihan model AI. Metode penelitian yang diterapkan adalah kualitatif deskriptif dengan pendekatan studi pustaka (*literature review*) terhadap berbagai instrumen hukum positif, kebijakan privasi platform, serta kode etik profesi internasional. Hasil analisis menunjukkan bahwa meskipun *web scraping* memberikan efisiensi teknis yang signifikan bagi berbagai sektor industri, praktik ini sering kali mengabaikan aspek persetujuan (*consent*) dari pemilik data dan berpotensi melanggar UU Nomor 28 Tahun 2014 tentang Hak Cipta serta UU Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi (UU PDP). Penekanan pada integritas praktisi Teknologi Informasi (TI) yang mengacu pada kode etik ACM (*Association for Computing Machinery*) menjadi sangat krusial sebagai kompas moral di tengah kekosongan regulasi spesifik mengenai *scraping* di tingkat nasional. Penelitian ini menyimpulkan bahwa profesionalisme sejati seorang praktisi TI tidak hanya diukur dari penguasaan teknologi ekstraksi data, melainkan juga dari kepatuhan terhadap batasan legalitas, transparansi, serta penghormatan terhadap hak milik pribadi guna menciptakan ekosistem digital yang adil, aman, dan transparan.

Kata kunci: Etika Profesi; Hak Cipta; Integritas Praktisi Data; Legalitas IT; Pengambilan Data dari Web.

1. LATAR BELAKANG

Di era transformasi digital saat ini, data telah menjadi komoditas yang sangat berharga, sering kali dianalogikan sebagai "emas baru" atau "tambang baru" yang menggerakkan inovasi dan efisiensi operasional bisnis sebagaimana dinyatakan oleh Rizquina & Ratnasari (2023). Fenomena ini didorong oleh perkembangan pesa. Kecerdasan Buatan (AI) yang membutuhkan dataset dalam skala masif untuk proses pelatihan algoritma (*training models*).

Salah satu metode utama yang digunakan untuk pengumpulan data tersebut adalah *web scraping*, yakni teknik ekstraksi data secara otomatis dari halaman web melalui struktur HTML menurut penjelasan Al Hafizh dkk. (2025). Implementasi *web scraping* telah meluas di berbagai sektor di Indonesia. Mulyono & Hindrayani (2024) mencatat penggunaannya dalam pemantauan berita terpadu di instansi pemerintah, sedangkan Javier dkk. (2025) menyoroti pengumpulan data laporan keuangan di Bursa Efek Indonesia untuk analisis investasi. Selain itu, Yuda (2025) menjelaskan pemanfaatannya dalam ekstraksi data pasar *e-commerce* untuk analisis tren harga dan kompetitor. Secara teknis, penggunaan pustaka seperti *BeautifulSoup* dan *Selenium* dalam bahasa pemrograman *Python* telah terbukti efektif dalam menangani konten dinamis dan meningkatkan efisiensi pengumpulan data dibandingkan metode manual, seperti yang diuraikan oleh Arsianto & Gautama (2025) serta Hidayati dkk. (2024).

Namun, di balik efisiensi teknisnya, praktik *web scraping* memicu dilema etika dan legalitas yang kompleks. Pengumpulan data secara otomatis sering kali dilakukan tanpa persetujuan eksplisit dari pemilik data, yang menurut Fatimah dkk. (2025) dalam perspektif etika komunikasi, merupakan kesenjangan serius dalam hal *consent* atau persetujuan. Selain itu, Abdullah (2024) mengingatkan bahwa AI generatif yang melatih modelnya berdasarkan rekaman percakapan atau data yang dikumpulkan dari internet berisiko melanggar hak privasi jika tidak didasarkan pada persetujuan sah dari subjek data.

Secara hukum, Rezon dkk. (2025) menyatakan bahwa Indonesia belum memiliki regulasi spesifik yang mengatur batasan *web scraping* dalam konteks pelatihan AI sehingga menciptakan kerentanan bagi karya kreatif nasional. Praktik ini berpotensi bersinggungan dengan berbagai instrumen hukum yang ada, seperti UU Nomor 28 Tahun 2014 tentang Hak Cipta yang melindungi hak ekonomi dan moral pencipta sebagaimana dianalisis oleh Tanujaya (2024) serta Putra & Nurmiati (2025). Selain itu, perlindungan data pribadi juga menjadi perhatian utama sebagaimana diatur dalam UU. Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi (UU PDP), yang menurut Syarifuddin dkk. (2025) menekankan pentingnya dasar pemrosesan data yang sah.

Kesenjangan antara inovasi teknologi dan kerangka hukum ini menuntut integritas tinggi dari para praktisi TI. Kode Etik dan Perilaku Profesional ACM (2018) menekankan bahwa setiap tindakan profesional komputer harus mempertimbangkan kepentingan publik sebagai pertimbangan utama. Praktisi diharapkan tidak hanya mengejar kemajuan teknologi, tetapi juga menghormati batasan legalitas seperti dokumen *robots.txt* dan kebijakan privasi platform guna menghindari praktik "penjarahan data" secara masif, sebagaimana ditekankan oleh Solove & Hartzog (2024) serta Putra & Nurmiati (2025).

Oleh karena itu, diperlukan pemahaman yang komprehensif mengenai batasan etika dan hukum dalam pemanfaatan AI dan *web scraping* untuk memastikan bahwa perkembangan teknologi tetap selaras dengan nilai-nilai kemanusiaan, perlindungan hak individu, dan kepastian hukum bagi semua pemangku kepentingan.

2. KAJIAN TEORITIS

Analisis terhadap fenomena *web scraping* memerlukan landasan konseptual yang kuat, mengingat praktiknya yang berada di persimpangan antara kemajuan teknis, kepatuhan hukum, dan integritas moral. Pemahaman mengenai bagaimana data diekstraksi secara otomatis menjadi dasar untuk melihat sejauh mana teknologi ini dapat dioptimalkan tanpa mencederai hak pihak lain. Oleh karena itu, bagian ini akan memaparkan kerangka teori yang dimulai dari aspek teknis operasional, tinjauan terhadap instrumen hukum positif yang berlaku di Indonesia, hingga landasan etika profesi yang menjadi standar perilaku bagi para praktisi Teknologi Informasi (TI). Integrasi dari berbagai landasan teoritis ini diharapkan dapat memberikan perspektif yang utuh dalam memetakan batasan-batasan yang harus dipatuhi dalam ekosistem digital yang dinamis.

Konsep *Web Scraping* dan Ekstraksi Data

Web scraping merupakan teknik otomatisasi untuk mengekstraksi data dari halaman situs web dengan memetakan struktur HTML dokumen tersebut. Al Hafizh dkk. (2025) menjelaskan bahwa metode ini krusial dalam menyediakan dataset masif bagi pengembangan Kecerdasan Buatan (AI). Efisiensi pengumpulan data ini didukung oleh penggunaan pustaka pemrograman dinamis seperti *BeautifulSoup* dan *Selenium* yang mampu menggantikan metode pengambilan data manual secara signifikan (Arsianto & Gautama, 2025). Landasan Hukum Digital di Indonesia. Regulasi digital di Indonesia memberikan batasan ketat terhadap aktivitas pemrosesan data otomatis. Putra & Nurmiati (2025) menggarisbawahi bahwa legalitas *scraping* bersinggungan langsung dengan dua instrumen utama: UU No. 27 Tahun 2022 (UU PDP): Mengatur kewajiban dasar pemrosesan data yang sah melalui persetujuan eksplisit (*explicit consent*) dari subjek data (Abdullah, 2024). UU No. 28 Tahun 2014 (Hak Cipta): Melindungi konten orisinal dari ekstraksi tanpa izin yang dapat merugikan hak ekonomi dan moral pencipta (Rezon dkk., 2025). Teori Etika Profesi TI. Dalam membedah perilaku praktisi data, penelitian ini menggunakan kerangka etika komunikasi dan profesi sebagai pisau analisis. Menurut Fatimah dkk. (2025), terdapat tiga perspektif etika utama yang menjadi standar evaluasi tindakan profesional, yaitu Deontologi (berorientasi pada kewajiban dan aturan), Teleologi (berorientasi pada konsekuensi dan manfaat), serta Etika Kebajikan (berorientasi pada

integritas karakter individu). Ketiga teori ini digunakan untuk mengukur sejauh mana praktisi TI mempertahankan integritasnya di tengah ambiguitas regulasi nasional (Syarifuddin dkk., 2025).

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif deskriptif untuk mengeksplorasi fenomena dilema etika dan legalitas dalam praktik *web scraping*. Sebagaimana dijelaskan oleh Fatimah dkk. (2025), pendekatan kualitatif sangat efektif untuk mengidentifikasi kesenjangan antara etika komunikasi dengan praktik nyata di lapangan, terutama terkait masalah persetujuan (*consent*). Metode utama yang diterapkan adalah studi pustaka (*library research*). Menurut Syarifuddin dkk. (2025), analisis terhadap literatur yang ada sangat penting untuk memetakan implikasi etis dari penerapan AI dalam konteks TI, khususnya mengenai bagaimana praktisi menavigasi dilema moral saat mengelola data dalam skala besar. Data sekunder yang digunakan dalam penelitian ini mencakup: Dokumen Hukum: UU Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi (UU PDP), UU Nomor 28 Tahun 2014 tentang Hak Cipta, dan UU ITE. Literatur Ilmiah: Jurnal-jurnal nasional dan internasional mengenai implementasi teknis *web scraping* (seperti penelitian oleh Arsianto & Gautama (2025) dan Javier dkk. (2025)) serta artikel mengenai etika profesi. Kebijakan Platform: Dokumen kebijakan privasi dan ketentuan penggunaan dari berbagai sektor seperti *e-commerce*, perbankan, dan bursa efek.

Teknik analisis data dilakukan dengan analisis isi tematik (*thematic content analysis*). Proses ini bertujuan untuk menyintesis informasi dari berbagai sumber guna menemukan pola perilaku praktisi data dan mengukur tingkat kesesuaian tindakan mereka terhadap regulasi yang berlaku di Indonesia. Putra & Nurmiati (2025) menekankan bahwa penggunaan analisis dokumen hukum dan kebijakan privasi sangat krusial untuk menjaga integritas profesi di tengah ambiguitas aturan yang ada. Melalui metode ini, penelitian diharapkan dapat memberikan gambaran komprehensif mengenai batasan yang harus dipatuhi oleh praktisi TI.

4. HASIL DAN PEMBAHASAN

Bagian ini memaparkan hasil analisis mendalam mengenai kompleksitas praktik *web scraping* di Indonesia melalui kacamata hukum, etika, dan integritas profesional. Temuan penelitian menunjukkan bahwa meskipun secara teknis instrumen ekstraksi data telah mencapai efisiensi yang tinggi, tetapi pemanfaatannya masih menghadapi tantangan besar dalam menyelaraskan inovasi dengan regulasi yang ada. Pembahasan akan dimulai dengan membedah aspek legalitas melalui instrumen hukum positif nasional, yang kemudian akan

dikaitkan dengan dilema etika komunikasi yang sering dihadapi oleh para praktisi. Lebih lanjut, analisis akan difokuskan pada peran integritas praktisi Teknologi Informasi (TI) sebagai fondasi utama dalam menjaga keseimbangan antara kepentingan bisnis dan perlindungan hak individu. Seluruh rangkaian pembahasan ini akan ditutup dengan perumusan strategi mitigasi yang mengintegrasikan solusi teknis dan komitmen moral guna menciptakan ekosistem data yang lebih bertanggung jawab di era kecerdasan buatan.

Analisis Legalitas Berdasarkan Regulasi Indonesia

Implementasi *web scraping* di Indonesia saat ini berada dalam ruang transisi hukum yang ketat. Berdasarkan analisis terhadap UU Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi (UU PDP), setiap aktivitas pengumpulan data yang melibatkan data spesifik maupun umum wajib memiliki dasar pemrosesan yang sah. Syarifuddin dkk. (2025) menekankan bahwa salah satu dasar utama tersebut adalah persetujuan eksplisit (*explicit consent*). Dalam konteks pelatihan AI, Abdullah (2024) mengingatkan bahwa pengumpulan data dari percakapan atau interaksi pengguna tanpa izin yang jelas merupakan pelanggaran serius terhadap privasi. Hal ini diperkuat oleh ketentuan dalam Pasal 28G ayat (1) UUD 1945 yang menjamin perlindungan atas harta benda dan martabat individu di ruang digital.

Dari perspektif kekayaan intelektual, UU Nomor 28 Tahun 2014 tentang Hak Cipta menjadi batasan legalitas lainnya. Praktik ekstraksi data massal sering kali mengabaikan hak eksklusif pencipta konten. Rezon dkk. (2025) memberikan studi komparasi yang menunjukkan bahwa tanpa regulasi spesifik mengenai *scraping* di Indonesia, karya kreatif nasional menjadi sangat rentan dieksploitasi oleh pengembang AI. Sebagai contoh, Arifin dkk. (2025) menyoroti dampak AI pada gaya visual spesifik seperti Ghibli, di mana ekstraksi data tanpa izin dapat mengikis nilai ekonomi dan moral dari pencipta aslinya. Selain itu, berdasarkan UU ITE, sistem AI diklasifikasikan sebagai agen elektronik sehingga segala kerugian hukum yang timbul dari aktivitas otomatisasi ini menjadi tanggung jawab penyelenggara sistem tersebut (Putra & Nurmiati, 2025).

Dilema Etika Komunikasi dan Profesi

Kesenjangan antara kemajuan teknologi dan kematangan etika menjadi inti dari dilema praktisi TI. Melalui kacamata etika komunikasi, Fatimah dkk. (2025) mengidentifikasi tiga pola kesenjangan utama: Deontologi: Banyak praktisi mengabaikan kewajiban moral untuk menghormati aturan platform demi mengejar efisiensi pengumpulan data. Teleologi: Terjadi ketidakseimbangan di mana manfaat ekonomi dari penggunaan data hasil *scraping* sering kali dianggap lebih penting daripada risiko kebocoran privasi pengguna. Etika Kebajikan: Terdapat penurunan integritas karakter di mana praktisi TI cenderung memanfaatkan "celah hukum"

daripada mengedepankan transparansi. Syarifuddin dkk. (2025) menambahkan bahwa dalam era AI, tantangan etis ini semakin kompleks karena data yang diproses sangat masif sehingga pengawasan terhadap penggunaan data yang bertanggung jawab sering kali terabaikan oleh para profesional TI.

Integritas Praktisi TI di Era Kecerdasan Buatan (AI)

Integritas praktisi data diukur dari kemampuan mereka menyeimbangkan kecanggihanteknologi dengan tanggung jawab kemanusiaan. Prinsip *High-Tech-High-Touch* menjadi relevan di sini; semakin tinggi teknologi yang digunakan, semakin tinggi pula tuntutan moral yang melekat pada penggunaannya. Putra & Nurmiati (2025) menegaskan bahwa kepatuhan terhadap Kode Etik ACM adalah standar mutlak bagi praktisi TI untuk tetap profesional. Praktisi diharapkan tidak hanya terpaku pada keberhasilan teknis seperti yang diuraikan oleh Arsianto & Gautama (2025) dalam pengembangan aplikasi *crawling*, tetapi juga harus memastikan bahwa metode pengumpulan tersebut tidak merugikan pihak lain secara ilegal. Menjaga integritas berarti memilih untuk menggunakan jalur legal, seperti API resmi, meskipun teknik *web scraping* menawarkan jalan pintas yang lebih murah, tetapi berisiko secara etis dan hukum.

Strategi Mitigasi: Integrasi Solusi Teknis dan *Ethical by Design*

Sebagai bentuk nyata dari integritas praktisi TI dalam menghadapi ambiguitas regulasi, diperlukan langkah-langkah mitigasi yang menggabungkan kecanggihan sistem dengan tanggung jawab moral. Strategi mitigasi ini tidak hanya bertujuan untuk memenuhi standar hukum nasional seperti UU PDP, tetapi juga untuk membangun kepercayaan publik terhadap ekosistem data.

Secara teknis, penguatan keamanan data menjadi pilar utama. Menurut Syarifuddin dkk. (2025), praktisi harus mengimplementasikan protokol keamanan tingkat lanjut, termasuk enkripsi data dan otentikasi multi-faktor guna mencegah akses tidak sah terhadap data hasil ekstraksi. Lebih lanjut, transparansi dalam tata kelola data dapat ditingkatkan melalui pemanfaatan teknologi *blockchain*. Sebagaimana dijelaskan oleh Wira & Suryawijaya (2023), implementasi *blockchain* memungkinkan terciptanya sistem rekam jejak yang akuntabel (*audit trail*), sehingga asal-usul dan penggunaan data dapat diverifikasi secara terbuka. Selain itu, penghormatan terhadap batasan teknis platform, seperti penggunaan API resmi dan kepatuhan terhadap protokol *robots.txt*, merupakan manifestasi teknis dari etika profesional yang harus dijunjung tinggi (Putra & Nurmiati, 2025).

Selain langkah teknis, internalisasi prinsip *Ethical by Design* menjadi krusial dalam setiap tahapan pengembangan sistem ekstraksi data. Hal ini berarti pertimbangan etika harus

diintegrasikan sejak tahap perancangan, bukan sekadar menjadi pelengkap di akhir proses. Fatimah dkk. (2025) menekankan bahwa pengembangan kerangka etika komunikasi yang kuat, yang mengedepankan prinsip keadilan dan transparansi, adalah solusi jangka panjang untuk menutupi kesenjangan regulasi. Dengan menerapkan pendekatan ini, organisasi dan praktisi TI dapat memastikan bahwa inovasi teknologi yang mereka kembangkan tetap selaras dengan perlindungan privasi dan hak kekayaan intelektual (Syarifuddin dkk., 2025). Integrasi antara solusi teknis yang transparan dan komitmen etis inilah yang pada akhirnya membentuk standar profesionalisme baru di era kecerdasan buatan.

5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa praktik *web scraping* di Indonesia masih berada dalam zona abu-abu hukum yang menuntut integritas tinggi dari praktisi TI. Meskipun memberikan efisiensi teknis yang signifikan dalam berbagai sektor seperti properti dan *e-commerce*, aktivitas ini memiliki risiko pelanggaran terhadap UU PDP dan UU Hak Cipta, terutama jika dilakukan tanpa persetujuan eksplisit. Profesionalisme seorang praktisi TI tidak hanya diukur dari kemampuannya menavigasi ambiguitas hukum, tetapi juga dari komitmennya untuk menerapkan standar etika global seperti Kode Etik ACM dan prinsip *Ethical by Design*. Akhirnya, sinergi antara langkah teknis yang transparan dan internalisasi nilai-nilai moral menjadi kunci utama dalam melindungi hak individu di tengah masifnya inovasi teknologi AI. Berdasarkan temuan tersebut, praktisi Teknologi Informasi diharapkan untuk tidak hanya mengutamakan efisiensi teknis, tetapi juga mulai mengadopsi prinsip *privacy by design* melalui audit etis mandiri serta memprioritaskan penggunaan API resmi guna menghormati batasan legalitas pemilik platform. Bagi penyelenggara sistem elektronik atau organisasi, diperlukan penyusunan panduan etika internal yang spesifik mengatur tata kelola data untuk pelatihan AI agar profesional TI dapat bekerja tanpa tekanan ekonomi yang mengabaikan hak asasi pengguna. Di sisi regulasi, pemerintah disarankan segera merumuskan aturan turunan UU PDP yang secara spesifik mengatur *automated data collection* guna memberikan kepastian hukum dan melindungi karya kreatif nasional. Terakhir, bagi peneliti selanjutnya, disarankan untuk melakukan penelitian empiris kuantitatif mengenai tingkat kepatuhan praktisi TI terhadap UU PDP atau studi komparasi mendalam mengenai implementasi *EU AI Act* sebagai referensi penguatan kerangka regulasi di Indonesia.

UCAPAN TERIMA KASIH

Penulis memanjatkan puji syukur kepada Allah SWT atas selesainya artikel penelitian ini. Ucapan terima kasih yang tulus penulis sampaikan kepada Ibu Evy Nurmiati, M.Kom., selaku dosen pengampu mata kuliah Etika Profesi Teknologi Informasi (EPTI) di Program Studi Sistem Informasi UIN Syarif Hidayatullah Jakarta, atas bimbingan, ilmu, dan arahan yang telah diberikan selama proses perkuliahan. Penulis juga berterima kasih kepada seluruh pihak yang telah memberikan dukungan serta referensi yang berharga sehingga artikel mengenai integritas praktisi data ini dapat disusun dengan baik. Semoga penelitian ini dapat memberikan kontribusi positif bagi pemahaman etika profesi di lingkungan akademik maupun praktisi TI.

DAFTAR REFERENSI

- Abdullah, A. M. (2024). Pelindungan hak privasi terhadap pengumpulan data pribadi oleh AI generatif berdasarkan percakapan dengan pengguna. *Padjadjaran Law Review*, 12(2), 156–172.
- ACM. (2018). *ACM code of ethics and professional conduct*. New York, NY: Association for Computing Machinery.
- Al Hafizh, M. R., Fauzan, A. D., Rahayu, W. I., Mustaqim, K., & Hanum, R. (2025). Web scraping data ulasan pelanggan untuk kemajuan bisnis e-commerce pada official store dan non-official store dengan pendekatan Natural Language Processing. *DSI: Jurnal Data Science Indonesia*, 5(1), 1–9.
- Arifin, M. N., Mahmud, M. A., Haris, F. M., & Ramadhan, M. H. (2025). Analysing the impact of AI on Ghibli style visual effects in the creative industry. *Brilliance: Research of Artificial Intelligence*, 5(1).
- Arsianto, J. L., & Gautama, T. K. (2025). Pengembangan aplikasi web scraping untuk crawling web data dari situs e-commerce properti. *Jurnal Strategi*, 7(1), 163–172.
- Fatimah, F., Yasin, A., Amelia, N., & Hasibuan, A. Z. N. (2025). Etika komunikasi dalam pengumpulan dan pemanfaatan data pribadi di era digital. *Jurnal Bisnis Mahasiswa (JBME)*, 4(2), 3069–3079.
- Hidayati, L., Kusuma, L. P., Agustini, D., & Ardhana, V. Y. P. (2024). Implementasi web scraping untuk pengumpulan data media sosial lingkup pemerintah provinsi NTB. *Jurnal Sistem Informasi dan Informatika (Simika)*, 7(1), 62–71.
- Javier, N., Satoto, B. D., & Negara, Y. D. P. (2025). Implementasi teknik web scraping untuk pengumpulan data laporan keuangan perusahaan di Bursa Efek Indonesia (IDX). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(2).
- Mulyono, F. N. H., & Hindrayani, K. M. (2024). Penerapan teknik web scraping untuk membangun website pemantauan berita terpadu KLHK. *SENASIK: Seminar Nasional Sistem Informasi dan Komputer*, 4(1), 56–61.
- Putra, L. H., & Nurmiati, E. (2025). Integritas praktisi data: Dilema etika dan batasan legalitas web scraping bisnis dalam perspektif UU hak cipta Indonesia. *Jurnal EPTI*.
- Republik Indonesia. (1945). *Undang-Undang Dasar Negara Republik Indonesia Tahun 1945*. Jakarta: Sekretariat Jenderal MPR RI.

- Republik Indonesia. (2014). *Undang-Undang Nomor 28 Tahun 2014 tentang Hak Cipta*. Jakarta: Sekretariat Negara.
- Republik Indonesia. (2022). *Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi*. Jakarta: Sekretariat Negara.
- Rezon, A., Montolalu, Y., Kenuwiarja, H., & Sugianto, F. (2025). Studi komparasi hak cipta atas proses data scrapping AI di Indonesia, Uni Eropa, dan Amerika. *Anthology: Inside Intellectual Property Rights*, 3(1).
- Rizquina, A. Z., & Ratnasari, C. I. (2023). Implementasi web scraping untuk pengambilan data pada website e-commerce. *Jurnal Teknologi dan Sistem Informasi Bisnis (JTEKSIS)*, 5(4), 377–383.
- Solove, D. J., & Hartzog, W. (2024). The great scrape: The clash between scraping and privacy. *California Law Review*, 113, 1521.
- Syarifuddin, Saprizal, A. M., Pebriana, R., Margaret, T., & Swastina, L. (2025). Etika profesi TI di era AI: Tantangan perlindungan data dan privasi pengguna. Universitas Sari Mulia Banjarmasin.
- Tanujaya, O. (2024, Agustus 1). Gaya Ghibli dan tantangan hak cipta di era AI. Diakses dari <https://id.wikipedia.org/wiki/Artikel>
- Wira, T., & Suryawijaya, E. (2023). Memperkuat keamanan data melalui teknologi blockchain: Mengeksplorasi implementasi sukses dalam transformasi digital di Indonesia. *Jurnal Sains dan Teknologi*, 2(1), 55–67.
- Yuda, M. A. D. (2025). Implementasi web scraping untuk ekstraksi data penjual dan produk panel surya di e-commerce. *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, 5(1).