BAB II

TINJAUAN PUSTAKA

2.1 Studi Sebelumnya

Bagian ini menguraikan penelitian-penelitian terdahulu yang memiliki keterkaitan dengan topik penelitian, khususnya dalam konteks privasi digital, keamanan data pribadi, serta perkembangan teknologi *Artificial Intelligence* (AI) di media sosial. Studi-studi terdahulu dipilih berdasarkan relevansi tematik, metodologis, dan konteks penggunaan media sosial yang serupa dengan penelitian ini. Penelitian-penelitian tersebut juga mencakup pembahasan mengenai penyalahgunaan teknologi berbasis AI, seperti *deepfake*, serta respons perilaku pengguna terhadap ancaman digital. Dengan mengkaji penelitian terdahulu, penelitian ini memperoleh landasan teoritis dan empiris yang kuat dalam memahami fenomena yang diteliti.

Penelitian oleh Fauzia dkk. menggunakan metode kuantitatif dengan pendekatan Structural Equation Modeling (SEM) terhadap 391 narasumber pengguna media sosial di Indonesia. Hasil penelitian menunjukkan bahwa variabel kepercayaan, kesadaran, kekhawatiran privasi, dan niat perilaku memiliki pengaruh positif dan signifikan terhadap perilaku berbagi informasi pribadi. Hal ini dibuktikan melalui nilai Critical Ratio (CR) yang melebihi 1,96 serta p-value di bawah 0,05 pada seluruh hipotesis yang diuji. Selain itu, model penelitian menunjukkan kecocokan yang sangat baik dengan nilai CFI sebesar 0,996 dan RMSEA sebesar 0,018, sehingga memperkuat validitas model dalam menjelaskan perilaku pengguna[11].

Penelitian yang dilakukan oleh Rasyidin dkk. membedah fenomena kesadaran mahasiswa dalam memproteksi kerahasiaan data pribadi mereka saat berinteraksi di berbagai *platform* media sosial. Studi ini menggunakan desain kuantitatif deskriptif dengan mengumpulkan data dari 33 orang narasumber melalui instrumen kuesioner digital [12]. Temuan utama menunjukkan bahwa walaupun narasumber memiliki pemahaman teoretis yang mumpuni terkait klasifikasi data sensitif dan risiko siber, hal tersebut tidak berbanding lurus dengan tindakan nyata mereka di ruang digital. Masih ditemukan kelalaian dalam pemanfaatan fitur proteksi seperti autentikasi dua faktor

serta rendahnya frekuensi pembaruan kata sandi secara berkala. Oleh karena itu, hasil studi ini menekankan pentingnya edukasi yang berfokus pada internalisasi perilaku keamanan praktis, bukan sekadar pemberian informasi teknis semata

Lestari, dkk. mengkaji fenomena penyalahgunaan *deepfake* berbasis AI di Indonesia melalui pendekatan kualitatif berbasis studi literatur. Hasil penelitian menunjukkan bahwa penyalahgunaan *deepfake* didorong oleh berbagai motif, seperti keuntungan finansial, dendam pribadi, dan kepentingan seksual [13]. Selain itu, konten *deepfake* memiliki karakteristik yang sulit dikendalikan setelah tersebar di internet, sehingga berpotensi menimbulkan dampak yang luas. Penelitian ini juga menyoroti bahwa regulasi yang ada di Indonesia masih belum cukup kuat untuk menangani kasus penyalahgunaan teknologi *deepfake* secara efektif.

Nurdin dan Nugraha melakukan penelitian menggunakan pendekatan kualitatif deskriptif untuk menganalisis ancaman *deepfake* terhadap keamanan siber. Hasil penelitian menunjukkan bahwa *deepfake* memiliki dampak yang luas, tidak hanya secara teknis tetapi juga secara psikologis, sosial, politis, bahkan struktural [14]. Rendahnya literasi digital masyarakat serta belum optimalnya regulasi memperbesar potensi risiko yang dihadapi pengguna dalam ruang digital. Kondisi ini menyebabkan meningkatnya kerentanan individu dalam menghadapi ancaman manipulasi konten berbasis AI.

Penelitian oleh Dwi, dkk. menggunakan pendekatan fenomenologi dengan melibatkan empat informan pengguna Twitter (Sekarang X). Hasil penelitian menunjukkan bahwa pengguna menerapkan berbagai strategi perlindungan, seperti penggunaan akun anonim, pembatasan konten, dan pemanfaatan fitur keamanan *platform* [15]. Selain itu, ditemukan adanya tekanan psikologis yang muncul ketika pengguna hendak membagikan konten pribadi di media sosial. Hal ini menunjukkan bahwa ancaman di media sosial dapat memicu perubahan perilaku pengguna secara signifikan, baik secara teknis maupun emosional.

Tabel 2.1 di bawah ini merupakan studi sebelumnya, yaitu sebagai berikut:

Tabel 2. 1 Studi Sebelumnya

No	Penulis	Tahun	Tujuan Penelitian	Metode	Hasil & Relevansi
1	Fauzia, dkk.	2024	Menganalisis faktor kesadaran, kepercayaan, kekhawatiran privasi, dan niat perilaku pengguna dalam berbagi informasi	Kuantitatif; SEM; 391 narasumber	Hasil penelitian menunjukkan bahwa seluruh variabel memiliki pengaruh positif dan signifikan dengan nilai CR > 1,96 dan p-value < 0,05. Model penelitian memiliki tingkat kecocokan yang sangat baik dengan nilai CFI sebesar 0,996 dan RMSEA sebesar 0,018. Hal ini menunjukkan bahwa model mampu menjelaskan hubungan antar variabel secara kuat.
2	Rasyidin, dkk.	2025	Mengkaji tingkat kesadaran mahasiswa terhadap keamanan data pribadi	Deskriptif kuantitatif melalui penyebaran kuesioner kepada 33 mahasiswa.	Ditemukan adanya *knowledge-behavior gap* di mana pengetahuan tinggi tidak menjamin perilaku keamanan yang konsisten.

No	Penulis	Tahun	Tujuan Penelitian	Metode	Hasil & Relevansi
3	Lestari, dkk.	2024	Mengkaji penyalahgunaan *deepfake* dan upaya penanggulangan	Kualitatif; studi literatur	Hasil penelitian menunjukkan bahwa penyalahgunaan *deepfake* dipicu oleh berbagai motif, seperti ekonomi, sosial, dan personal. Konten *deepfake* juga sulit dikendalikan setelah tersebar luas di internet. Selain itu, regulasi yang ada masih belum memadai untuk menangani permasalahan ini secara komprehensif. Penyalahgunaan Grok AI merupakan bentuk nyata dari praktik *deepfake* yang berdampak pada privasi pengguna.
4	Nurdin dan Nugraha	2025	Menganalisis dampak *deepfake* terhadap keamanan siber dan masyarakat	Kualitatif deskriptif	Hasil penelitian menunjukkan bahwa *deepfake* memiliki dampak multidimensi, baik secara teknis, psikologis, maupun sosial. Rendahnya literasi digital masyarakat memperbesar risiko penyalahgunaan teknologi ini. Selain itu, regulasi yang belum optimal

No	Penulis	Tahun	Tujuan Penelitian	Metode	Hasil & Relevansi
					turut memperparah kondisi tersebut. Temuan ini menjelaskan bahwa rasa tidak aman pengguna tidak hanya dipengaruhi teknologi, tetapi juga faktor sosial dan psikologis.
5	Dwi, dkk.	2024	Menganalisis strategi perlindungan diri pengguna Twitter dari *cyberbullying*	Kualitatif; fenomenologi	Hasil penelitian menunjukkan bahwa pengguna melakukan berbagai strategi proteksi seperti akun anonim, pembatasan konten, dan penggunaan fitur keamanan. Selain itu, terdapat tekanan psikologis yang dirasakan pengguna saat akan membagikan informasi pribadi. Hal ini menunjukkan adanya perubahan perilaku akibat rasa tidak aman.

2.2 Dasar Teori

2.2.1 *Artificial Intelligence*

Artificial Intelligence (AI) merupakan cabang ilmu komputer yang berfokus pada pengembangan sistem atau mesin yang mampu meniru kemampuan kognitif manusia, seperti berpikir, belajar, memahami bahasa, dan mengambil keputusan. AI bekerja dengan memanfaatkan algoritma dan data dalam jumlah besar untuk mengenali pola tertentu, sehingga sistem dapat memberikan respons atau hasil yang menyerupai tindakan manusia. Dalam konteks digital saat ini, AI tidak hanya digunakan untuk otomasi proses, tetapi juga untuk menghasilkan konten, menganalisis perilaku pengguna, serta mendukung interaksi manusia dengan sistem teknologi.

Perkembangan AI yang semakin pesat membawa dampak signifikan terhadap ekosistem media digital, khususnya media sosial. Kehadiran AI generatif memungkinkan terciptanya teks, gambar, audio, dan video secara otomatis dengan tingkat realisme yang tinggi. Di satu sisi, teknologi ini memberikan kemudahan dan efisiensi, namun di sisi lain menimbulkan risiko penyalahgunaan, terutama ketika AI digunakan untuk memanipulasi informasi. Penelitian di Indonesia menunjukkan bahwa peningkatan pemanfaatan AI memerlukan perhatian serius terhadap aspek etika, akurasi informasi, serta dampaknya terhadap kepercayaan publik di ruang digital [16].

2.2.2 *Deepfake*

Deepfake adalah teknologi manipulasi media yang menggunakan teknik pembelajaran mesin untuk menciptakan atau memodifikasi konten visual dan audio sehingga menyerupai individu nyata. Teknologi ini mampu meniru wajah, suara, serta ekspresi seseorang dengan tingkat kemiripan yang tinggi, sehingga sulit dibedakan dari konten asli. Seiring perkembangan AI generatif, pembuatan *deepfake* menjadi semakin mudah dan dapat dilakukan oleh masyarakat umum dengan perangkat lunak yang tersedia secara luas.

Keberadaan *deepfake* menimbulkan tantangan serius dalam dunia komunikasi digital, terutama terkait keaslian informasi dan perlindungan identitas

individu. Penelitian di Indonesia menyoroti bahwa *deepfake* berpotensi digunakan untuk penipuan, pencemaran nama baik, dan penyebaran disinformasi. Selain merugikan individu yang menjadi objek manipulasi, *deepfake* juga berdampak pada menurunnya kepercayaan publik terhadap konten digital secara keseluruhan, karena pengguna menjadi semakin sulit membedakan antara fakta dan rekayasa [17].

2.2.3 Rasa Aman

Rasa aman dalam konteks digital merujuk pada persepsi individu terhadap tingkat keamanan dan perlindungan data pribadi saat menggunakan teknologi atau *platform* online. Konsep ini berkaitan erat dengan *information privacy concern*, yaitu tingkat kekhawatiran individu terhadap potensi penyalahgunaan informasi pribadinya. Semakin tinggi tingkat kekhawatiran (*perceived risk*) tersebut, maka semakin besar kecenderungan individu untuk melakukan tindakan perlindungan terhadap privasinya (*coping behavior*) [10]. Hal ini menunjukkan bahwa rasa aman bukan hanya kondisi objektif, tetapi juga dipengaruhi oleh bagaimana individu memaknai risiko yang ada di lingkungan digital.

Dalam kajian perilaku pengguna, rasa aman sering dijelaskan melalui pendekatan *perceived* atau persepsi individu terhadap suatu ancaman. Beberapa konsep penting yang sering digunakan adalah risiko yang dirasakan (*perceived risk*), kerentanan yang dirasakan (*perceived vulnerability*), dan tingkat keparahan yang dirasakan terhadap suatu ancaman (*perceived severity*). Ketika pengguna merasa dirinya rentan terhadap penyalahgunaan data atau melihat dampak yang serius dari suatu ancaman, maka tingkat *privacy concern* akan meningkat, sehingga mendorong perilaku perlindungan seperti membatasi unggahan atau mengatur privasi akun. Hal ini juga didukung oleh temuan penelitian bahwa persepsi ancaman memiliki pengaruh signifikan terhadap kepedulian privasi pengguna [10].

Dalam konteks penelitian ini, fenomena penyalahgunaan teknologi AI seperti *deepfake* atau Grok AI dapat meningkatkan *perceived risk* dan *perceived vulnerability* pengguna terhadap unggahan foto pribadi. Ketika pengguna menyadari bahwa foto yang mereka unggah berpotensi disalahgunakan, maka

persepsi rasa aman akan menurun dan mendorong perubahan perilaku dalam menggunakan media sosial. Oleh karena itu, rasa aman dalam penelitian ini tidak hanya dilihat sebagai kondisi teknis, tetapi sebagai hasil dari persepsi pengguna terhadap risiko digital yang berkembang.