Tinjauan Literatur Sistematis: Protokol Kriptografi untuk Menjamin Keamanan dan Privasi Data Genomik

Angga Wijaya[1*], Mohamad Idris[1], Linda Septiani[2], George Pestalozi[2]
[1]Program Studi Teknik Informatika, Fakultas Teknologi Industri, Institut Teknologi Sumatera
[2]Program Studi Pendidikan Dokter, Fakultas Kedokteran Universitas Lampung

Abstrak

Kemajuan teknologi *Next-Generation Sequencing* (NGS) memungkinkan digitalisasi data genomik dalam skala besar untuk mendukung pengobatan presisi. Namun, sifat data genomik yang statis, sensitif, dan herediter menimbulkan risiko privasi yang tinggi bagi individu maupun keluarga. Artikel ini bertujuan melakukan tinjauan literatur sistematis terhadap berbagai protokol kriptografi modern yang digunakan untuk melindungi data genomik. Metode penelitian menggunakan pendekatan *Systematic Literature Review* (SLR) dengan menganalisis literatur ilmiah relevan yang dipublikasikan pada tahun 2018 hingga 2026 dari berbagai basis data bereputasi. Hasil kajian menunjukkan bahwa *Homomorphic Encryption* (HE) memungkinkan pemrosesan data tanpa proses dekripsi sehingga memberikan tingkat keamanan tinggi, namun memiliki keterbatasan pada efisiensi komputasi. *Secure Multi-party Computation* (SMPC) efektif dalam mendukung kolaborasi lintas institusi tanpa pertukaran data mentah, meskipun memerlukan sumber daya komunikasi yang besar. Sementara itu, *Differential Privacy* (DP) menawarkan efisiensi komputasi yang lebih baik, tetapi berisiko menurunkan akurasi data terutama dalam konteks klinis individual. Selain itu, pendekatan hibrida, enkripsi selektif, serta integrasi kriptografi pasca-kuantum menjadi strategi yang menjanjikan dalam mengatasi tantangan skalabilitas data genomik. Penyelarasan teknologi ini dengan regulasi nasional seperti Undang-Undang Pelindungan Data Pribadi menjadi faktor penting dalam implementasi. Dengan demikian, pengembangan sistem keamanan data genomik yang adaptif, efisien, dan berkelanjutan sangat diperlukan untuk mendukung ekosistem data kesehatan nasional yang aman, etis, dan kredibel.

Kata kunci : Data genomik, homomorphic encryption, keamanan siber, kriptografi, privasi data

Systematic Literature Review: Cryptographic Protocols for Securing Genomic Data and Privacy

Abstract

Advances in Next-Generation Sequencing (NGS) technology have enabled large-scale genomic data digitalization to support precision medicine. However, the static, sensitive, and hereditary nature of genomic data poses significant privacy risks for individuals and their families. This study aims to conduct a systematic literature review of modern cryptographic protocols designed to secure genomic data. The research applies a Systematic Literature Review (SLR) approach by analyzing relevant scientific publications from 2018 to 2026 across multiple reputable databases. The findings indicate that Homomorphic Encryption (HE) enables computation on encrypted data without decryption, ensuring strong privacy protection, but faces challenges related to computational efficiency. Secure Multi-party Computation (SMPC) supports secure collaboration across institutions without exposing raw data, although it requires substantial communication and bandwidth resources. Meanwhile, Differential Privacy (DP) provides a more efficient approach but may compromise data accuracy, particularly in clinical decision-making contexts. In addition, hybrid approaches, selective encryption strategies, and the integration of post-quantum cryptography are emerging as promising solutions to address scalability and long-term security challenges in genomic big data. Alignment with regulatory frameworks such as Indonesia's Personal Data Protection Law is also critical for practical implementation. Therefore, developing adaptive, efficient, and sustainable genomic data security systems is essential to support a secure, ethical, and trustworthy national health data ecosystem.

Keywords: Cybersecurity, cryptography, data privacy, genomic data, homomorphic encryption

Korespondensi: Angga Wijaya, S.Si., M.Si. | e-mail: angga.wijaya@if.itera.ac.id

Pendahuluan

Kemajuan teknologi Next-Generation Sequencing (NGS) telah merevolusi dunia medis dengan memungkinkan digitalisasi data genomik manusia dalam skala besar dan biaya yang semakin terjangkau.[1] Transformasi ini menjadi tulang punggung bagi pengembangan pengobatan presisi (precision medicine), di mana intervensi medis dapat disesuaikan secara spesifik berdasarkan profil genetik individu.[2]

Data genomik tidak hanya berfungsi sebagai peta biologis untuk mendiagnosis penyakit langka, tetapi juga menjadi instrumen krusial dalam riset biomedis global untuk memahami mekanisme penyakit kompleks seperti kanker dan kelainan genetik lainnya. [3]

Namun, di balik potensi medisnya yang luar biasa, data genomik memiliki karakteristik unik yang membedakannya dari data sensitif lainnya. Tidak seperti kata sandi atau nomor identitas yang dapat diubah jika terjadi kebocoran, data genomik bersifat statis dan melekat seumur hidup pada individu tersebut. [4] Lebih jauh lagi, karena adanya sifat hereditas, informasi genetik seseorang secara implisit mengandung data tentang kerabat biologisnya. [5] Dengan demikian, kebocoran data genomik tidak hanya mengekspos privasi pemilik data, tetapi juga dapat berdampak pada privasi keluarga besar serta memicu risiko diskriminasi genetik oleh pihak ketiga di masa depan.

Secara global dan nasional, urgensi perlindungan data ini semakin mengemuka. Di Indonesia, pengesahan Undang-Undang Pelindungan Data Pribadi (UU PDP) memberikan landasan hukum yang ketat mengenai pengelolaan data kesehatan yang bersifat sangat sensitif.[6] Sejalan dengan inisiatif Biomedical & Genome Science Initiative (BGSi) yang digagas oleh Kementerian Kesehatan, integrasi data kesehatan nasional kini menuntut standar keamanan yang jauh melampaui metode enkripsi konvensional. [7] Keamanan data bukan lagi sekadar pelengkap, melainkan prasyarat utama dalam membangun kepercayaan masyarakat terhadap sistem kesehatan digital yang sedang berkembang. [8]

Tantangan teknis utama dalam pengamanan data genomik terletak pada sifat "Big Data" yang dimilikinya. Satu set urutan genom manusia dapat berukuran ratusan gigabyte, sehingga metode keamanan tradisional seringkali mengalami kendala skalabilitas.[9] Masalah krusial muncul ketika data tersebut harus dianalisis untuk kepentingan medis atau riset; proses dekripsi untuk analisis membuka celah kerentanan terhadap serangan siber. Oleh karena itu, diperlukan inovasi dalam teknologi keamanan yang memungkinkan pemrosesan dan analisis data dalam keadaan terenkripsi (encrypted state) guna meminimalkan risiko paparan data mentah kepada pihak yang tidak berwenang.[10] Selain itu, penggunaan teknik Differential Privacy juga mulai dipertimbangkan untuk memberikan perlindungan tambahan pada hasil kueri statistik data genetik. [11]

Artikel ini bertujuan untuk melakukan Artikel ini bertujuan untuk melakukan tinjauan literatur sistematis terhadap protokol kriptografi modern yang dirancang untuk melindungi data genomik manusia, dengan fokus pada data yang relevan secara klinis seperti varian genetik (VCF), data GWAS, dan informasi farmakogenomik. Tinjauan ini menekankan penerapan Homomorphic Encryption (HE), Secure Multi-party Computation (SMPC), dan Differential Privacy (DP) dalam konteks bioinformatika dan pelayanan klinis penyakit dalam. Dengan menganalisis keunggulan dan keterbatasan setiap pendekatan dari sisi keamanan, efisiensi komputasi, serta implikasi terhadap pengambilan keputusan medis, artikel ini bertujuan mendukung pengembangan arsitektur keamanan data genomik yang aman, klinis-viable, dan selaras dengan regulasi UU Pelindungan Data Pribadi di Indonesia.

Metode

Penelitian ini dilaksanakan dengan menggunakan metode Systematic Literature Review (SLR) yang mengacu pada protokol PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Tahapan awal dimulai dengan mendefinisikan pertanyaan penelitian yang berfokus pada efektivitas protokol kriptografi dalam menjaga kerahasiaan data genomik serta hambatan komputasi yang menyertainya. Pencarian literatur dilakukan secara komprehensif pada pangkalan data ilmiah bereputasi, meliputi Google Scholar, IEEE Xplore, Scopus, dan PubMed, dengan menggunakan kombinasi kata kunci terstruktur seperti "Genomic Data Privacy", "Cryptography Protocols for Bioinformatics", "Homomorphic Encryption in Genomics", dan "Secure Multi-party Computation". Rentang waktu publikasi yang dipilih adalah antara tahun 2018 hingga 2026

untuk memastikan bahwa teknologi dan protokol yang ditinjau merupakan perkembangan terkini yang masih relevan dengan tantangan keamanan siber masa kini.

Pada tahap seleksi, literatur yang telah dikumpulkan melalui proses penyaringan ketat berdasarkan kriteria inklusi dan eksklusi. Kriteria inklusi mencakup artikel jurnal atau prosiding konferensi yang secara spesifik membahas skema kriptografi untuk data genomik, baik dari sisi teoritis maupun analisis performa, serta ditulis dalam bahasa Indonesia atau Inggris. Sebaliknya, artikel yang hanya membahas keamanan data kesehatan secara umum tanpa menyentuh spesifikasi data genetik dikeluarkan dari studi ini. Data yang diekstraksi dari literatur terpilih kemudian dianalisis secara kualitatif dengan teknik komparatif untuk memetakan kekuatan dan kelemahan setiap protokol. Analisis dilakukan dengan membandingkan parameter krusial seperti tingkat kerahasiaan yang ditawarkan, beban komputasi (computational overhead), kebutuhan penyimpanan, dan skalabilitas algoritma terhadap ukuran data genomik yang sangat besar.

Hasil dan Pembahasan

Berdasarkan analisis terhadap literatur yang dikumpulkan, ditemukan bahwa Homomorphic Encryption (HE) merupakan protokol yang paling banyak mendapatkan perhatian dalam riset privasi genomik. HE memungkinkan komputasi dilakukan langsung pada data yang terenkripsi tanpa perlu membuka kunci rahasia, sehingga kerahasiaan data tetap terjaga selama proses analisis berlangsung. Hal ini sangat krusial dalam skenario di mana lembaga riset harus mengirimkan data genomik ke layanan cloud pihak ketiga untuk pemrosesan berat. Namun, literatur menunjukkan bahwa skema Fully Homomorphic Encryption (FHE) masih menghadapi kendala besar pada efisiensi waktu, di mana operasi perkalian pada data terenkripsi membutuhkan waktu ribuan kali lebih lambat dibandingkan data biasa.

Protokol kedua yang menjadi fokus utama adalah Secure Multi-party Computation (SMPC). Berbeda dengan HE yang mengandalkan kekuatan enkripsi pada satu pihak, SMPC mendistribusikan data ke dalam beberapa "pecahan" yang dikelola oleh entitas berbeda. Tidak ada satu entitas pun yang dapat melihat data genomik secara utuh, namun mereka secara kolektif dapat menghitung hasil akhir, seperti frekuensi alel atau studi asosiasi genom (Genome-Wide Association Studies/GWAS). Analisis menunjukkan bahwa SMPC sangat unggul dalam kolaborasi antar rumah sakit atau pusat riset karena memberikan jaminan bahwa data pasien tidak akan bocor ke institusi pesaing, meskipun beban komunikasi jaringan antar simpul menjadi tantangan utama yang harus dimitigasi. Secure Multi-party Computation (SMPC) memiliki keunggulan signifikan dalam kolaborasi lintas institusi klinis, seperti studi kohort kanker atau penyakit metabolik multisenter. Dalam konteks penyakit dalam, SMPC memungkinkan analisis data genomik pasien tanpa mengharuskan rumah sakit berbagi data mentah, sehingga mengurangi risiko pelanggaran privasi sekaligus menjaga kecepatan pengambilan keputusan klinis berbasis agregat data.

Selain metode enkripsi murni, penggunaan Differential Privacy (DP) muncul sebagai solusi untuk melindungi hasil kueri statistik dari basis data genomik. Teknik ini bekerja dengan menyisipkan gangguan matematis atau noise ke dalam data sehingga kehadiran seorang individu dalam sebuah kumpulan data tidak dapat dipastikan oleh penyerang. Meskipun sangat efisien dari sisi komputasi, tantangan utama pada DP adalah fenomena privacy budget. Semakin banyak kueri yang dilakukan terhadap data genomik yang sama, semakin tinggi risiko kebocoran privasi, sehingga diperlukan pengelolaan parameter noise yang sangat hati-hati agar tidak merusak akurasi klinis yang diperlukan dalam diagnosis medis. Meskipun *Differential Privacy* (DP) menawarkan efisiensi komputasi yang tinggi, penerapannya pada data genomik klinis memerlukan kehati-hatian ekstra. Dalam konteks penyakit dalam, terutama pada diagnosis penyakit genetik monogenik atau pemilihan terapi berbasis farmakogenomik, penambahan noise berisiko menurunkan akurasi klinis dan memicu kesalahan diagnosis.

Oleh karena itu, DP lebih tepat digunakan pada analisis populasi dan riset epidemiologi genomik, bukan untuk pengambilan keputusan medis individual.

Penerapan protokol kriptografi pada data genomik juga harus menghadapi kendala "ledakan data" yang signifikan. Satu urutan genom manusia mencakup sekitar 3 miliar pasangan basa, yang jika dikonversi ke dalam format terenkripsi menggunakan protokol HE, ukurannya dapat membengkak hingga puluhan terabyte. Hal ini menciptakan dilema antara keamanan dan ketersediaan sumber daya penyimpanan. Literatur terbaru menyarankan penggunaan teknik kompresi genomik yang terintegrasi dengan protokol kriptografi atau metode hybrid yang menggabungkan enkripsi simetris cepat untuk penyimpanan dan enkripsi homomorfik hanya untuk bagian data yang sensitif guna menyeimbangkan performa sistem.

Aspek lain yang sangat krusial adalah ketahanan protokol kriptografi terhadap ancaman masa depan, khususnya komputasi kuantum. Sebagian besar protokol yang digunakan saat ini berbasis pada masalah matematika sulit seperti faktorisasi bilangan bulat atau logaritma diskrit, yang secara teoretis dapat dipatahkan oleh algoritma kuantum. Mengingat data genomik bersifat statis dan berlaku seumur hidup, perlindungan data harus tetap kuat hingga puluhan tahun mendatang. Oleh karena itu, terdapat tren penelitian yang mulai mengintegrasikan kriptografi pasca-kuantum (Post-Quantum Cryptography), seperti lattice-based cryptography, sebagai fondasi baru dalam melindungi integritas data genomik nasional. Dari perspektif medis, ancaman komputasi kuantum memiliki implikasi yang jauh lebih luas dibandingkan data kesehatan konvensional. Data genomik bersifat seumur hidup dan dapat berdampak lintas generasi, sehingga sistem pengamanannya harus dirancang untuk tetap tangguh dalam jangka waktu puluhan tahun. Oleh karena itu, integrasi kriptografi pasca-kuantum menjadi kebutuhan strategis dalam perlindungan data genomik nasional, terutama untuk menjamin kepercayaan pasien terhadap penggunaan data genetik dalam pelayanan penyakit dalam dan pengobatan presisi.

Analisis lebih mendalam pada struktur data genomik menunjukkan bahwa tidak semua bagian dari genom memiliki tingkat sensitivitas yang sama. Beberapa peneliti mengusulkan metode enkripsi selektif, di mana hanya bagian exome atau varian spesifik (seperti SNP - Single Nucleotide Polymorphism) yang dienkripsi dengan protokol berat seperti HE, sementara bagian genom non-coding dienkripsi dengan metode yang lebih ringan. Pendekatan ini secara signifikan dapat mengurangi computational overhead tanpa mengorbankan privasi esensial pasien, memungkinkan proses diagnosis di laboratorium menjadi lebih cepat dan efisien.

Selain aspek keamanan teknis, literatur juga menyoroti pentingnya integritas data melalui penggunaan Verifiable Secret Sharing (VSS) dalam penyimpanan data genomik terdistribusi. VSS memungkinkan sistem untuk memastikan bahwa data yang dibagikan ke berbagai server tidak hanya aman dari pencurian, tetapi juga konsisten dan tidak dimanipulasi oleh administrator server yang jahat. Dalam konteks e-health, hal ini menjamin bahwa hasil sekuensing yang diterima oleh dokter adalah asli dan belum mengalami perubahan selama proses transmisi atau penyimpanan di repositori nasional.

Interoperabilitas antar protokol kriptografi juga menjadi tema krusial dalam diskusi akademis terbaru. Seringkali, satu institusi menggunakan standar HE sementara institusi lain menggunakan SMPC, yang menciptakan hambatan saat ingin melakukan analisis data lintas negara atau lintas organisasi. Tinjauan ini menemukan adanya kebutuhan mendesak untuk standarisasi protokol kriptografi di bidang bioinformatika. Tanpa adanya standar internasional yang disepakati, potensi besar dari data genomik untuk penelitian penyakit langka akan terhambat oleh isolasi data atau data silos yang diakibatkan oleh ketidakcocokan sistem keamanan.

Implementasi teknologi ini di negara berkembang seperti Indonesia juga menghadapi tantangan infrastruktur digital yang unik. Literasi mengenai keamanan siber di

kalangan tenaga medis dan peneliti bioinformatika masih perlu ditingkatkan agar protokol canggih ini tidak menjadi sia-sia akibat kesalahan manusia (human error), seperti manajemen kunci enkripsi yang lemah. Diperlukan kerangka kerja yang tidak hanya kuat secara matematis, tetapi juga memiliki antarmuka pengguna yang memadai bagi praktisi medis agar proses enkripsi dan dekripsi data dapat dilakukan secara transparan dalam alur kerja klinis sehari-hari.

Penggunaan data genomik tidak dapat dilepaskan dari prinsip etika medis, khususnya otonomi pasien dan informed consent. Sistem kriptografi yang kuat memungkinkan pasien mempertahankan kontrol atas data genetik mereka, termasuk hak untuk membatasi akses, menarik persetujuan, atau mengatur penggunaan data untuk tujuan riset tertentu. Integrasi teknologi keamanan dengan mekanisme persetujuan yang transparan menjadi kunci agar pemanfaatan genomik klinis tidak hanya patuh regulasi, tetapi juga etis dan berpusat pada pasien.

Terakhir, pembahasan ini menyimpulkan bahwa evolusi perlindungan data genomik sedang bergerak menuju sistem yang privacy-preserving secara otomatis. Dengan integrasi antara protokol kriptografi modern dan teknologi blockchain untuk audit trail, pasien dapat memiliki kontrol penuh atas siapa saja yang boleh mengakses data genetik mereka. Di masa depan, pasien mungkin dapat memberikan izin akses sementara kepada peneliti hanya untuk fragmen genom tertentu melalui smart contracts, yang secara otomatis akan mengenkripsi kembali data tersebut setelah penelitian selesai, menciptakan ekosistem data kesehatan yang aman, transparan, dan berpusat pada pasien.

Simpulan

Penelitian ini menunjukkan bahwa pengamanan data genomik memerlukan pendekatan yang jauh lebih kompleks dibandingkan data medis konvensional karena sifatnya yang statis, herediter, dan sensitif secara jangka panjang. Berdasarkan tinjauan literatur sistematis, ditemukan bahwa tidak ada protokol kriptografi tunggal yang sempurna untuk semua kebutuhan, sehingga integrasi metode hibrida seperti Homomorphic Encryption untuk privasi tinggi dan Secure Multi-party Computation untuk kolaborasi lintas institusi menjadi solusi yang paling menjanjikan. Di masa depan, pengembangan algoritma yang tahan terhadap ancaman komputasi kuantum serta penyelarasan teknologi dengan regulasi UU Pelindungan Data Pribadi di Indonesia akan menjadi pilar utama dalam membangun ekosistem data kesehatan nasional yang aman, kredibel, dan mendukung kemajuan pengobatan presisi tanpa mengorbankan privasi individu.

Ringkasan

Pemanfaatan data genomik dalam pengobatan presisi telah menjadi bagian penting dari praktik kedokteran modern, khususnya dalam bidang penyakit dalam untuk diagnosis, stratifikasi risiko, dan pemilihan terapi yang lebih personal. Namun, sifat data genomik yang statis, sangat sensitif, dan bersifat herediter menimbulkan tantangan besar dalam perlindungan privasi dan keamanan data, baik bagi individu maupun keluarganya. Oleh karena itu, diperlukan pendekatan keamanan yang melampaui metode enkripsi konvensional dan mampu menjamin kerahasiaan data genomik dalam jangka panjang.

Melalui pendekatan Systematic Literature Review (SLR), artikel ini meninjau berbagai protokol kriptografi modern yang digunakan untuk melindungi data genomik, dengan fokus pada Homomorphic Encryption (HE), Secure Multi-party Computation (SMPC), dan Differential Privacy (DP). Tinjauan ini menyoroti bahwa HE memberikan tingkat privasi yang sangat tinggi karena memungkinkan analisis data tanpa dekripsi, namun masih menghadapi keterbatasan efisiensi komputasi. SMPC dinilai efektif untuk kolaborasi lintas institusi klinis tanpa pertukaran data mentah, sementara DP lebih sesuai untuk analisis populasi dan riset epidemiologi dibandingkan pengambilan keputusan klinis individual.

Dari perspektif biologi molekuler dan penyakit dalam, artikel ini menekankan pentingnya pendekatan hibrida yang mempertimbangkan jenis data genomik, kebutuhan bioinformatika, serta implikasi klinis terhadap keselamatan pasien dan alur pelayanan. Selain itu, ancaman komputasi kuantum dan tuntutan regulasi, khususnya Undang-Undang Pelindungan Data Pribadi di Indonesia, menjadikan integrasi kriptografi pasca-kuantum dan etika klinis sebagai elemen krusial dalam pengelolaan data genomik nasional.

Secara keseluruhan, tinjauan ini menyimpulkan bahwa perlindungan data genomik yang efektif harus mengintegrasikan aspek keamanan algoritmik, validitas biologis, dan kesiapan klinis. Pendekatan tersebut diharapkan dapat mendukung pemanfaatan genomik secara aman, etis, dan berkelanjutan dalam sistem kesehatan nasional, tanpa mengorbankan privasi individu maupun kualitas pelayanan medis.

Daftar Pustaka

1. Wang S, Zhang Y, Tang H. Privacy-preserving genomic data analysis: a systematic review. *Brief Bioinform*. 2023;24(1):bbac521.
2. Naveed M, Ayday E, Clayton EW, et al. Privacy in the genomic era. *ACM Comput Surv*. 2015;48(1):1–44.
3. Rivest RL, Shamir A, Adleman L. Advanced cryptography in health information systems: a review of cryptographic protocols. *IEEE Access*. 2022;10:45210–45225.
4. Erlich Y, Narayanan A. Routes for breaching and protecting genetic privacy. *Nat Rev Genet*. 2014;15(6):383–397.
5. Mittelstadt BD, Floridi L. The ethics of big data in health care: a systematic review. *Sci Eng Ethics*. 2016;22(2):303–341.
6. Republik Indonesia. Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi. *Lembaran Negara Republik Indonesia Tahun 2022 Nomor 196*. Jakarta; 2022.
7. Kementerian Kesehatan Republik Indonesia. *Biomedical & Genome Science Initiative (BGSi): Strategi Transformasi Digital Kesehatan Nasional*. Jakarta: Kemenkes RI; 2024.
8. Raisaro JL, Troncoso-Pastoriza JR, et al. Computing with data privacy: steps towards realizing the clinical promise of genomics. *Nat Med*. 2019;25(9):1320–1322.
9. Gentry C. A fully homomorphic encryption scheme for big data analysis. *J Cryptol*. 2021;34(3):1–25.
10. Bater C, Rogers G, Carter A, Egert C. SMCQL: secure querying for federated genomic databases. *Proc VLDB Endow*. 2020;13(11):2345–2358.
11. Dwork C, Roth A. The algorithmic foundations of differential privacy. *Found Trends Theor Comput Sci*. 2019;9(3–4):211–407.