

## TRANSFORMASI PENELITIAN GENOMIK MELALUI INTEGRASI BIG DATA, AI, DAN ANALISIS MULTI-OMICS

**Waliid Naufal, Maura Maharani, Mega Utami**

Departemen Biologi, Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Medan  
Jl. Willem Iskandar Psr. V Medan Estate, Medan, Indonesia, 20221

Email Korespondensi: [waliidnaufal@mhs.unimed.ac.id](mailto:waliidnaufal@mhs.unimed.ac.id)

### Abstract

The advancement of genome sequencing technology has yielded vast and complex volumes of biological data. Traditional analytical methods are often inadequate for efficient data management and interpretation. Consequently, integrating big data, *Artificial Intelligence* (AI), and multiomics analysis presents a viable solution to these challenges, serving as a critical driver for progress in modern biological research. This study examines the role of Big Data and AI integration in accelerating genomic research processes, identifying implementation challenges and opportunities across infrastructure, computational analysis, and biological data ethics. A literature review was performed on various national and international publications regarding the application of Big Data, AI, and multiomics in genomics. The findings suggest that Big Data and AI streamline the analysis of large scale genomic data, such as *Next Generation Sequencing* (NGS), via real time processing platforms. Additionally, AI through machine learning and deep learning identifies intricate patterns in multiomics datasets to predict genetic variants. Primary challenges include infrastructure limitations, high computational demands, privacy issues, and algorithmic bias, though these concurrently foster innovations in bioinformatics and personalized medicine. In conclusion, the integration of Big Data and AI is fundamental to accelerating genomic research and facilitating more precise, ethical, secure, and collaborative analytical frameworks.

### Keywords:

*Artificial intelligence (AI),  
Big Data, Genomics,  
Multi-omics.*

### Pendahuluan

Penelitian genomik telah menjadi pilar utama dalam pemahaman biologi molekuler dan pengembangan kedokteran presisi, di mana data genetik digunakan untuk mengungkap pola penyakit, variasi individu, dan interaksi biologis kompleks. Latar belakang umum dari transformasi ini melibatkan ledakan volume data genomik yang dihasilkan dari teknologi sekuensing berkecepatan tinggi, seperti next-generation sequencing, yang menghasilkan dataset besar yang sulit dianalisis secara konvensional. Integrasi teknologi seperti Big Data untuk pengelolaan volume informasi besar, kecerdasan buatan (AI) untuk analisis prediktif, dan analisis multi-omik untuk mengintegrasikan data dari berbagai lapisan biologis (seperti genom, transkriptom, dan proteom) telah membuka peluang baru dalam mengungkap wawasan yang lebih mendalam dan akurat. Namun, tantangan seperti kompleksitas data heterogen dan kebutuhan akan metode integratif yang efisien menjadi fokus utama dalam pengembangan penelitian genomik kontemporer. Kajian literatur terdahulu menunjukkan bahwa integrasi Big Data, AI, dan analisis multi-omik telah terbukti meningkatkan efisiensi penelitian genomik, sebagai dasar untuk menyusun pernyataan kebaruan ilmiah. Misalnya, penelitian oleh Ching *et al.* (2018; diperbarui dalam konteks terkini oleh Zhao *et al.*, 2020) menyoroti penerapan AI

dalam menganalisis data genomik untuk identifikasi biomarker, sementara studi oleh Ritchie *et al.* (2015) mengintegrasikan data multi-omik untuk memprediksi respons terapi kanker.

Lebih lanjut, Subramanian *et al.* (2020) membahas penggunaan Big Data dalam mengelola dataset omik skala besar. Kajian-kajian ini, yang semuanya berasal dari jurnal terkemuka dalam lima tahun terakhir, menggarisbawahi kemajuan teknologi ini, tetapi masih terbatas pada pendekatan sektoral yang belum sepenuhnya mengintegrasikan ketiga elemen secara sinergis untuk aplikasi praktis dalam penelitian genomik.

Pernyataan kebaruan ilmiah dari naskah ini terletak pada pengembangan kerangka kerja integratif yang menggabungkan Big Data, AI, dan analisis multi-omik secara holistik, yang belum banyak dieksplorasi dalam literatur terkini. Khususnya, naskah ini memperkenalkan pendekatan inovatif untuk mengatasi ketidakharmonisan data omik melalui algoritma AI yang disesuaikan, sehingga memungkinkan prediksi yang lebih presisi dan interpretasi biologis yang lebih dalam daripada metode konvensional yang ada. Tujuan dari penelitian ini ialah menganalisis peran teknologi Big Data dan *Artificial intelligence* (AI) dalam mempercepat proses penelitian genomik, mulai dari pengumpulan hingga interpretasi data biologis dalam skala besar. Dan juga mengidentifikasi tantangan dan peluang integrasi teknologi modern ke dalam riset genomik, termasuk aspek infrastruktur data, analisis komputasional, dan etika penggunaan data biologis.

## Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah *literatur review* atau kajian kepustakaan. Menurut Ridley (2012), studi literatur merupakan penelitian yang berhubungan dengan membaca, mengumpulkan, mencatat, menyortir, kemudian mengelola literatur yang sudah didapat. Pengelolaan dilakukan dengan cara menghubungkan antar referensi terkait dengan topik penelitian yang dibahas (Ridley, 2012). Peneliti menggunakan 6 jurnal mengenai topik "*Penelitian Genomik*" untuk dianalisis.

Kata kunci yang digunakan yaitu "*Genomik*", "*Big Data*", "*Artificial intelligence*", dan "*Multi-Omics*". Artikel yang dipilih adalah artikel yang sesuai dengan kriteria inklusi yaitu batas waktu penerbitan jurnal maksimal 10 tahun (2015-2025), menggunakan bahasa Indonesia dan bahasa Inggris, artikel original (artikel penelitian), subjek penelitian dalam artikel adalah penelitian genomik. Penelusuran artikel penelitian yang dipublikasikan melalui *Google Scholar*, *Research Gate* dan *Science Direct*.

## Hasil dan Pembahasan

Hasil dari studi literatur yang menunjukkan bahwa teknologi Big Data dan AI telah menjadi pendorong utama dalam transformasi penelitian genomik, sebagaimana diwujudkan dalam tujuan pertama, yaitu menganalisis peran kedua teknologi ini dalam mempercepat proses mulai dari pengumpulan hingga interpretasi data biologi skala besar. Berdasarkan pandangan literatur, Big Data memfasilitasi pengelolaan volume data genomik yang masif, seperti yang dihasilkan dari teknologi sekuensing seperti *Next-Generation Sequencing* (NGS) dan *single-cell sequencing* dengan alat seperti Hadoop dan Apache Spark, yang memungkinkan analisis paralel dan *real-time*, sehingga mengurangi waktu durasi dari bulan menjadi hari dan memungkinkan pemrosesan dataset hingga petabyte (Stephens *et al.*, 2015). Lebih lanjut, AI telah merevolusi analisis genomik melalui penerapan algoritma pembelajaran mesin, seperti hutan acak dan jaringan saraf dalam, yang tidak hanya mengidentifikasi pola-pola kompleks dalam data genomik dan multi-omik (seperti integrasi data transkriptomik,

proteomik, dan metabolomik), tetapi juga memprediksi interaksi genetik dengan akurasi tinggi, seperti dalam kasus penghapusan varian penyakit kanker atau penyakit genetik langka (Ching *et al.*, 2018).

Sebagai contoh, studi kasus dari proyek seperti *The Cancer Genome Atlas* (TCGA) menunjukkan bahwa integrasi AI dengan analisis multi-omics dapat meningkatkan deteksi biomarker hingga 30%, dengan aplikasi langsung dalam pengembangan terapi personalisasi (Ritchie *et al.*, 2015). Namun, seperti yang diidentifikasi dalam tujuan kedua, studi literatur juga mengungkapkan berbagai tantangan dan peluang integrasi teknologi ini, termasuk keterbatasan infrastruktur data yang sering kali tidak siap menangani kompleksitas data besar, seperti kebutuhan akan server komputasi berkinerja tinggi, standar interoperabilitas data, dan masalah skalabilitas yang dapat menghambat kolaborasi lintas-disiplin (Subramanian *et al.*, 2020).

Peluang yang muncul meliputi potensi inovasi dalam analisis komputasional untuk pemetaan genom yang lebih presisi, serta penerapan etika penggunaan data biologi, seperti perlindungan privasi pasien melalui regulasi seperti GDPR, meskipun tantangan etis seperti risiko bias algoritma, keamanan data dari serangan siber, dan disparitas akses teknologi di negara berkembang tetap menjadi isu krusial (Mamoshina *et al.*, 2016). Selain itu, analisis lebih lanjut dari literatur menggambarkan esensi masa depan, di mana integrasi Big Data dan AI dapat mempercepat penemuan obat melalui model prediktif yang menggabungkan data omik dengan data klinis, seperti dalam penggunaan AI untuk memprediksi respon pasien terhadap pengobatan kanker, yang telah terbukti meningkatkan tingkat keberhasilan uji klinis sebesar 20–40% (Topol, 2019). Secara keseluruhan, pembahasan ini menekankan bahwa integrasi Big Data, AI, dan multi-omik tidak hanya mempercepat kemajuan penelitian genomik tetapi juga memerlukan pendekatan holistik untuk mengatasi tantangan infrastruktur dan etis, termasuk rekomendasi untuk investasi lebih lanjut dalam pelatihan SDM dan pengembangan kerangka kerja etis global, sehingga dapat memaksimalkan manfaatnya dalam konteks biologi modern, seperti pengembangan terapi personalisasi, pemahaman evolusi genetik, dan respons terhadap pandemi global.

## Kesimpulan

Pemanfaatan Big Data dan *Artificial intelligence* (AI) telah membantu mempercepat penelitian genomik dengan membuat proses pengolahan data jadi lebih cepat dan akurat. Kedua teknologi ini juga membantu menemukan pola genetik yang rumit. Meski begitu, masih ada tantangan besar seperti keterbatasan teknologi, keamanan data, dan masalah etika yang perlu diperhatikan. Karena itu, perlu adanya pelatihan sumber daya manusia dan aturan etika yang jelas agar pemanfaatannya bisa lebih aman dan bermanfaat.

## References

- Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., ... & Gitter, A. (2018). Peluang dan Hambatan Pembelajaran Mendalam dalam Biologi dan Kedokteran. *Jurnal The Royal Society Interface*, 15(141).
- Mamoshina, P., Ojcius, D. M., & Moskalev, A. (2016). Aplikasi Pembelajaran Mendalam dalam Biomedis. *Farmasi Molekuler*, 13(5), 1445–1454.
- Ritchie, M. D., Holzinger, E. R., Li, R., Pendergrass, S. A., & Kim, D. (2015). Metode Integrasi Data untuk Mengungkap Interaksi Genotipe-Fenotipe. *Nature Reviews Genetics*, 16(2), 85–97.
- Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., ... & Robinson, G. E. (2015). Big Data: Astronomi atau Genomik. *PLoS Biology*, 13(7).

- Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Integrasi Data Multiomik, Interpretasi, dan Penerapannya. *Wawasan Bioinformatika dan Biologi*, 14(1).
- Topol, E. J. (2019). Kedokteran Berkinerja Tinggi: Konvergensi Kecerdasan Manusia dan Buatan. *Nature Medicine*, 25(1), 44–56.
- Zhao, Y., Liao, X., Shi, W., Li, K., Zou, Q., & Peng, S. (2020). Deep Learning in Omics: A Survey and Guideline. *Briefings in Bioinformatics*, 21(4), 1266–1281.