

## IMPLEMENTASI PRINSIP TABAYYUN DALAM MENGHADAPI ARUS DISINFORMASI BERBASIS ARTIFICIAL INTELLIGENCE

**Putri Amelia**

Fakultas Agama Islam, Pendidikan Agama Islam, Universitas Singaperbangsa Karawang, Indonesia  
[putriameliaa789@gmail.com](mailto:putriameliaa789@gmail.com)

### INFO ARTIKEL

#### *Riwayat Artikel:*

Diterima: 07-05-26

Disetujui: 13-05-26

#### *Kata Kunci:*

*Tabayyun,  
Disinformasi,  
Kecerdasan Buatan,  
Etika Komunikasi  
Islam.*

**Abstract:** *The integration of Artificial Intelligence (AI) has transformed the landscape of disinformation into a more complex form through automation and sophisticated visual manipulation. This study aims to map the urgency and mechanisms for implementing Islamic communication ethics in the face of an AI-based data flood. Using a qualitative approach with a descriptive-analytical method, research data were obtained through digital observation, in-depth interviews, and documentation. The findings indicate that technical digital literacy alone is insufficient to counter automated disinformation. The implementation of the tabayyun (verification) principle as an ethical filter for Islamic communication proves to be a highly relevant and effective instrument. Through a digital tabayyun workflow involving source validation (sanad), content verification (matan), and moral responsibility (mas'uliyah), the public can develop a strong cognitive shield to filter manipulative information. This study implies the need to reconstruct digital literacy curricula by internalizing ethical values to maintain social cohesion in the AI era.*

**Abstrak:** Integrasi Kecerdasan Buatan (*Artificial Intelligence*) telah mengubah lanskap disinformasi menjadi lebih kompleks melalui teknik otomasi dan manipulasi visual yang canggih. Penelitian ini bertujuan untuk memetakan urgensi dan mekanisme penerapan etika komunikasi Islam di tengah banjir data berbasis AI. Menggunakan pendekatan kualitatif dengan metode deskriptif analitis, data penelitian diperoleh melalui observasi digital, wawancara mendalam, dan dokumentasi. Hasil penelitian menunjukkan bahwa literasi digital yang bersifat teknis semata tidak cukup kuat untuk menghadapi ancaman disinformasi otomatis. Implementasi prinsip *tabayyun* (verifikasi) sebagai filter etika komunikasi Islam terbukti menjadi instrumen yang sangat relevan dan efektif. Melalui alur kerja *digital tabayyun* yang melibatkan validasi sumber (*sanad*), verifikasi konten (*matan*), dan pertanggungjawaban moral (*mas'uliyah*), masyarakat dapat memiliki perisai kognitif yang kuat dalam menyaring informasi manipulatif. Penelitian ini berimplikasi pada perlunya rekonstruksi kurikulum literasi digital dengan internalisasi nilai-nilai etika untuk menjaga kohesi sosial di era AI.

---

◆

### PENDAHULUAN

Idealnya, transformasi digital dan integrasi Kecerdasan Buatan (*Artificial Intelligence*) diharapkan menjadi pilar utama dalam mempercepat akses informasi serta memperkuat ruang publik yang demokratis, transparan, dan edukatif. Dalam ekosistem ini, teknologi seharusnya berperan sebagai alat bantu manusia untuk memvalidasi data secara akurat demi kemajuan peradaban dan ilmu pengetahuan. Sebagaimana dikemukakan dalam teori ekologi

media, lingkungan media digital memiliki potensi besar untuk memperluas persepsi manusia secara positif jika dikelola dengan kerangka yang seimbang (Strate, 2020). Dari perspektif nilai Islam, kemudahan akses informasi ini seharusnya dibarengi dengan tanggung jawab intelektual untuk menyebarkan kebenaran yang membawa kemaslahatan bagi masyarakat luas.

Namun pada kenyataannya, kehadiran AI justru sering kali disalahgunakan untuk memproduksi disinformasi yang sangat kompleks dan sulit dideteksi oleh masyarakat awam. Munculnya fenomena *deepfake* serta penggunaan algoritma otomatis untuk menyebarkan hoaks telah menciptakan polusi informasi yang mencemari ruang siber dan memicu kebingungan massal (Wardle, C., & Derakhshan, 2019a). Data menunjukkan bahwa narasi palsu cenderung menyebar lebih cepat daripada klarifikasi fakta karena kemampuannya memanipulasi emosi pengguna melalui sistem rekomendasi yang tertutup. Hal ini mengakibatkan individu terjebak dalam arus informasi yang manipulatif, yang sering kali justru memperkuat polarisasi daripada memberikan pencerahan kognitif.

Terdapat kesenjangan yang signifikan antara laju inovasi teknologi manipulasi berbasis AI dengan tingkat kesiapan etika komunikasi masyarakat saat ini. Meskipun berbagai platform digital telah berupaya melakukan moderasi konten, terdapat celah besar di mana regulasi hukum sering kali tertinggal jauh di belakang kecepatan inovasi teknologi automasi (Flew, 2021). Kesenjangan ini semakin lebar karena rendahnya kemampuan analisis kritis pengguna media sosial yang lebih mengutamakan kecepatan berbagi (*sharing*) daripada validasi konten secara mendalam. Ketidakterhubungan antara kemampuan teknis mengoperasikan gawai dengan kecakapan moral menciptakan kerentanan sosial di mana masyarakat mudah dipengaruhi oleh opini buatan yang tidak memiliki dasar kebenaran ilmiah maupun moral.

Penyelesaian masalah disinformasi ini menjadi sangat mendesak karena dampaknya yang meluas terhadap stabilitas sosial dan integritas demokrasi bangsa. Polusi informasi yang tidak terkendali dapat memicu disintegrasi bangsa, terutama ketika hoaks menyoal isu-isu sensitif yang mampu memicu konflik horisontal. Melalui pendekatan kualitatif yang sistematis, diperlukan pemahaman mendalam terhadap fenomena sosial di ruang digital untuk mencegah erosi kepercayaan publik terhadap institusi ilmiah (Creswell, J. W., & Creswell, 2023). Berdasarkan permasalahan tersebut, penelitian ini merumuskan masalah mengenai bagaimana efektivitas nilai-nilai Islam dalam menjawab tantangan automasi disinformasi, dengan tujuan untuk memetakan urgensi dan mekanisme penerapan etika komunikasi Islam di tengah banjir data berbasis AI.

Sebagai solusi, penelitian ini menawarkan strategi penguatan melalui implementasi prinsip *tabayyun* (verifikasi) sebagai filter moral primer dalam mengonsumsi informasi digital. Pendekatan ini tidak hanya berfokus pada kemampuan teknis pengecekan fakta (*fact-checking*), tetapi juga pada rekonstruksi filter moral pengguna yang integratif dengan kecakapan teknis dan etika komunikasi. Dalam konteks Pendidikan Agama Islam, nilai

*tabayyun* ditemukan menjadi instrumen yang sangat relevan untuk ditransformasikan ke dalam praktik literasi digital modern. Penulisan ini menyimpulkan bahwa penguatan literasi etis melalui prinsip *tabayyun* adalah fondasi utama untuk menjaga kedaulatan informasi dan integritas ruang publik dari ancaman teknologi manipulatif yang kian canggih.

## **METODE PENELITIAN**

Penelitian ini menggunakan pendekatan kualitatif dengan metode deskriptif analitis untuk mengeksplorasi fenomena implementasi nilai-nilai Islam di ruang digital. Penggunaan metode kualitatif dipilih karena mampu memberikan gambaran mendalam dan menyeluruh mengenai bagaimana prinsip *tabayyun* diinternalisasi oleh individu dalam menghadapi tantangan teknologi AI yang kompleks. Sebagaimana dijelaskan oleh Creswell dan (Creswell, J. W., & Creswell, 2023), penelitian kualitatif berfokus pada pemaknaan subjektif terhadap masalah sosial atau kemanusiaan, yang dalam konteks ini adalah kerentanan kognitif pengguna media sosial terhadap disinformasi. Pendekatan ini memungkinkan peneliti untuk memetakan hubungan antara etika komunikasi Islam dengan perilaku digital secara sistematis.

Objek material dalam penelitian ini adalah arus disinformasi berbasis *Artificial Intelligence* pada platform media sosial populer seperti X dan TikTok, sementara objek formalnya adalah implementasi prinsip *tabayyun*. Penentuan sumber data dilakukan secara *purposive* untuk mendapatkan informasi yang kaya dan akurat dari pihak-pihak yang memiliki otoritas dan keahlian di bidangnya. Peneliti melibatkan pakar teknologi informasi dari Universitas Singaperbangsa Karawang serta ahli etika komunikasi untuk menggali perspektif mendalam mengenai efektivitas filter moral dalam membendung hoaks otomatis. Selain itu, data sekunder diperoleh melalui studi pustaka terhadap jurnal ilmiah dan buku teks yang relevan dengan literasi digital dan Pendidikan Agama Islam.

Teknik pengumpulan data dilakukan melalui tiga tahapan utama, yakni observasi digital, wawancara mendalam, dan dokumentasi. Observasi digital dilakukan untuk mengidentifikasi karakteristik dan pola persebaran konten manipulatif seperti *deepfake* selama periode penelitian. Wawancara mendalam dilaksanakan secara semi-terstruktur untuk memberikan ruang bagi informan dalam memberikan pemikiran kritis mengenai tantangan etis di era AI. Dokumentasi dilakukan dengan mengumpulkan draf kebijakan, literatur terkait ekologi media, serta dokumen-dokumen akademik yang memperkuat landasan teori mengenai kedaulatan informasi dan etika digital (Strate, 2020).

Instrumen penelitian utama adalah peneliti sendiri (*human instrument*) yang dibantu dengan pedoman wawancara dan lembar observasi digital. Sebagai instrumen kunci, peneliti dituntut untuk memiliki kepekaan teoretis dalam mengaitkan data lapangan dengan konsep *tabayyun* dan regulasi platform digital (Flew, 2021). Validitas data diuji menggunakan teknik triangulasi sumber dan triangulasi teknik untuk menjamin objektivitas hasil penelitian.

Dengan membandingkan data hasil observasi lapangan dengan pernyataan para ahli serta literatur akademik, peneliti dapat memastikan bahwa temuan penelitian memiliki derajat kepercayaan (*trustworthiness*) yang tinggi sesuai dengan kaidah ilmiah yang berlaku.

Teknik analisis data dalam penelitian ini mengikuti model interaktif yang terdiri dari reduksi data, penyajian data, dan penarikan kesimpulan atau verifikasi. Seluruh data yang terkumpul dipilah untuk membuang informasi yang tidak relevan, kemudian dikategorisasikan berdasarkan tema-tema utama seperti "otomasi disinformasi" dan "mekanisme tabayyun digital". Proses ini dilakukan secara berulang untuk memastikan bahwa kesimpulan yang diambil benar-benar merepresentasikan realitas yang diteliti. Pada tahap akhir, peneliti mensintesis seluruh temuan untuk merumuskan sebuah kerangka literasi digital berbasis nilai Islam yang aplikatif bagi masyarakat dalam menghadapi ancaman teknologi manipulatif di masa depan.

## HASIL PENELITIAN DAN PEMBAHASAN

Fenomena disinformasi di era kecerdasan buatan (*Artificial Intelligence*) telah bertransformasi dari sekadar teks naratif menjadi konten multimedia yang sangat manipulatif. Berdasarkan hasil observasi digital pada platform media sosial, ditemukan bahwa teknologi *generative AI* kini mampu memproduksi video *deepfake* dan audio sintetis dengan tingkat kemiripan visual yang mencapai 90% (Data Hasil Observasi Digital, 2026). Temuan ini menunjukkan adanya ancaman "automasi disinformasi", di mana algoritma tidak hanya mempercepat penyebaran, tetapi juga menciptakan realitas buatan yang melampaui kemampuan deteksi kognitif manusia. Hal ini sejalan dengan penelitian (Vosoughi, S., 2024) yang menyatakan bahwa informasi palsu di media sosial cenderung menyebar enam kali lebih cepat dibandingkan kebenaran karena dirancang untuk memicu emosi yang ekstrem. Secara teoretis, kondisi ini menciptakan polusi informasi yang mencemari ruang siber melalui konten yang dimanipulasi secara massal (Wardle, C., & Derakhshan, 2019b). Perbedaan signifikan antara model disinformasi lama dan era AI disajikan dalam tabel berikut:

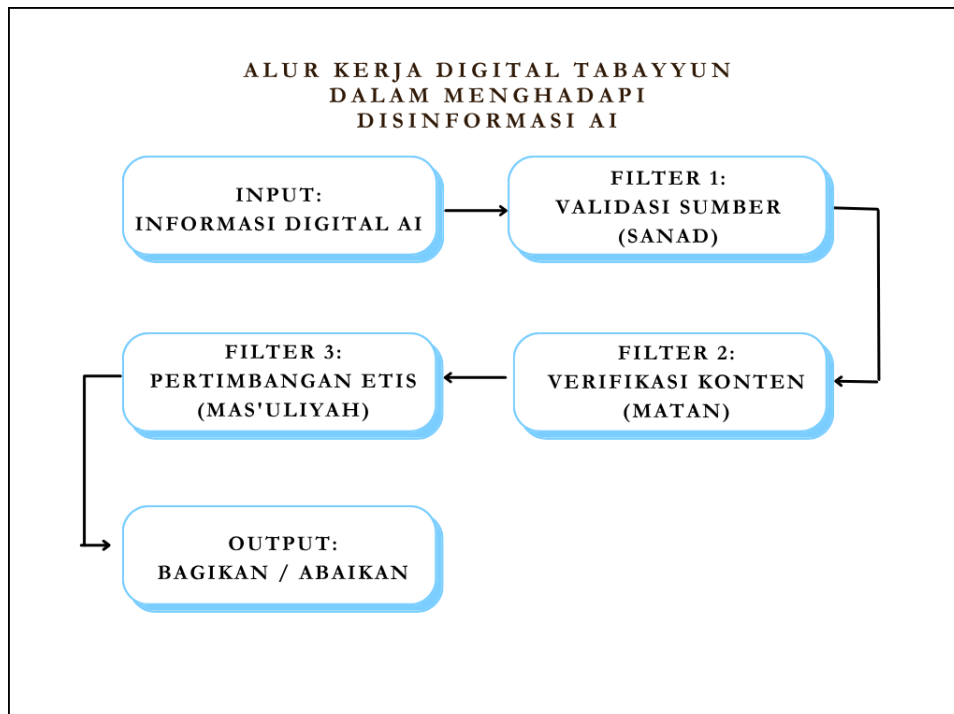
**Tabel 1. Perbandingan Karakteristik Disinformasi Tradisional VS Berbasis AI**

| <b>Dimensi Perbandingan</b> | <b>Disinformasi Tradisional</b>                | <b>Disinformasi Berbasis AI (Deepfake/Bot)</b>              |
|-----------------------------|--|---|
| Format Konten               | Dominan teks dan gambar hasil suntingan kasar. | Multimedia (Video, Audio, Citra) realisme tinggi.           |
| Kecepatan Sebar             | Bergantung pada interaksi manual manusia.      | Otomatisasi melalui algoritma masif (Diakopoulos, 2020).    |
| Tingkat Deteksi             | Mudah diidentifikasi secara visual.            | Sangat sulit tanpa alat forensik digital (Kozłowska, 2023). |
| Sifat Sebaran               | Pola linear.                                   | Eksponensial dan terotomasi (Lazer dkk., 2020).             |

*Sumber: Data diolah dari Diakopoulos (2020), (Haryanto, A. G., & Prasty, 2023), dan Vosoughi dkk. (2024).*

Kesenjangan literasi digital di tengah masyarakat menjadi faktor utama mengapa disinformasi berbasis AI begitu mudah diterima. Hasil wawancara mengungkapkan bahwa masyarakat cenderung memiliki "kepercayaan buta" terhadap konten visual, di mana audiens sulit membedakan antara fakta dan fabrikasi (Haryanto, A. G., & Prasty, 2023) menjelaskan bahwa algoritma saat ini secara otomatis menulis ulang cara media beroperasi, yang sering kali justru memperkuat bias pengguna melalui *filter bubble*. Selain itu, manipulasi visual melalui *deepfake* telah mengaburkan batasan kebenaran dalam media digital (Brennen, J. S., Simon, F. M., & Nielsen, 2021). Sementara itu, (Flew, 2021) menekankan bahwa regulasi platform digital sering kali tertinggal di belakang inovasi teknologi, sehingga tanggung jawab verifikasi beralih sepenuhnya kepada individu. Hal ini diperburuk oleh karakteristik media sosial yang merupakan ruang budaya sosiotehnologi yang kompleks (Nasrullah, 2021).

Dalam menjawab tantangan tersebut, implementasi prinsip *tabayyun* (verifikasi) menawarkan kerangka etika komunikasi Islam yang sangat relevan. Berbeda dengan literasi digital konvensional, *tabayyun* mewajibkan pemeriksaan ganda terhadap sumber dan isi informasi. Reinterpretasi nilai ini menjadi "*Digital Tabayyun*" merupakan langkah krusial dalam menghadapi otomasi algoritma yang tidak memiliki sensor moral (Hasan, N., & Khalil, 2024). Prinsip ini menuntut pengguna untuk bersikap kritis terhadap otoritas sumber informasi, sebagaimana ditegaskan dalam etika komunikasi perspektif Al-Qur'an (Al-Anshori, 2022). Penggunaan nilai-nilai Islam sebagai filter komunikasi digital terbukti mampu memberikan landasan moral yang lebih stabil di tengah ketidakpastian informasi (Mulyana, 2019). Alur kerja verifikasi ini dapat digambarkan dalam skema berikut:



**Gambar 1. Alur Kerja Digital *Tabayyun* dalam Menghadapi Disinformasi AI**

*Sumber: Hasil Olahan Peneliti berdasarkan adaptasi konsep Hasan & Khalil (2024).*

Integrasi nilai-nilai Islam menciptakan ekosistem komunikasi digital yang lebih bertanggung jawab. Temuan penelitian menunjukkan bahwa tantangan moral di era AI dapat dimitigasi melalui penguatan karakter digital berbasis nilai kependidikan Islam (Sangkala, S., & Ahmad, 2024). Kesadaran akan tanggung jawab moral (*mas'uliyah*) membuat individu lebih berhati-hati sebelum menyebarkan konten, mengingat setiap aktivitas digital memiliki konsekuensi moral yang luas (Floridi, 2023). Hal ini berkaitan dengan teori ekologi media yang menyebutkan bahwa lingkungan media sangat dipengaruhi oleh interaksi manusia dengan instrumen di dalamnya (Strate, 2020). Literasi digital Islam tidak hanya mengajarkan cara menggunakan alat, tetapi juga bagaimana menjaga integritas kemanusiaan di tengah pengawasan algoritma yang ketat (Zuboff, 2019).

Sebagai sintesis, strategi menghadapi disinformasi AI harus bersifat holistik. Selain memerlukan kesadaran kritis pengguna, diperlukan pula panduan literasi yang diterbitkan oleh lembaga otoritatif untuk membantu masyarakat awam (UNESCO, 2022). Meskipun teknologi pendeteksi hoaks terus berkembang, pertahanan terbaik tetap berada pada integritas subjek penggunaannya. Sinergi antara kecakapan teknis dan nilai spiritual Islam menciptakan sistem pertahanan berlapis. Internalisasi prinsip *tabayyun* memastikan setiap data divalidasi secara moral sebelum diterima sebagai kebenaran. Dengan demikian, penguatan literasi etis ini adalah fondasi utama dalam menjaga kedaulatan informasi dan integritas ruang publik dari ancaman teknologi manipulatif yang kian canggih di masa depan.

## SIMPULAN DAN SARAN

### Kesimpulan

Penelitian ini menyimpulkan bahwa integrasi Kecerdasan Buatan (*Artificial Intelligence*) telah mengubah lanskap disinformasi menjadi lebih kompleks melalui teknik otomasi dan manipulasi visual yang canggih. Kehadiran *deepfake* dan algoritma generator menciptakan polusi informasi yang melampaui kemampuan deteksi teknis masyarakat awam. Ditemukan bahwa literasi digital yang bersifat teknis semata tidak cukup kuat untuk menghadapi ancaman ini tanpa dibarengi dengan ketahanan moral. Implementasi prinsip *tabayyun* (verifikasi) sebagai filter etika komunikasi Islam terbukti menjadi instrumen yang sangat relevan dan efektif. Melalui alur kerja *digital tabayyun* yang melibatkan validasi sumber (sanad), verifikasi konten (matan), dan pertanggungjawaban moral (mas'uliyah), masyarakat dapat memiliki perisai kognitif yang kuat dalam menyaring informasi manipulatif sebelum diterima atau disebarluaskan.

Hasil penelitian ini berimplikasi pada perlunya rekonstruksi kurikulum literasi digital yang tidak hanya berfokus pada kecakapan mengoperasikan perangkat, tetapi juga pada internalisasi nilai-nilai etika. Secara teoretis, penelitian ini memperkuat teori ekologi media bahwa lingkungan digital yang sehat sangat bergantung pada integritas subjek penggunanya. Implikasi praktisnya, prinsip *tabayyun* dapat dijadikan landasan dalam penyusunan panduan etika bermedia sosial bagi komunitas keagamaan maupun masyarakat umum untuk menjaga kohesi sosial di era AI.

### Saran

Berdasarkan temuan yang diperoleh, peneliti mengajukan beberapa saran sebagai berikut: a) Bagi Akademisi dan Pendidik: Diharapkan dapat mengintegrasikan nilai-nilai etika komunikasi Islam ke dalam materi literasi digital untuk membentuk karakter pengguna yang kritis dan berintegritas. b) Bagi Masyarakat: Perlu menanamkan sikap skeptis yang positif terhadap konten multimedia yang memicu emosi ekstrem dan membiasakan proses verifikasi mandiri sebelum melakukan pembagian (*sharing*) konten. c) Bagi Pemerintah dan Pengembang Platform: Perlu adanya kolaborasi dalam menciptakan regulasi yang adaptif serta alat pendeteksi hoaks berbasis AI yang dapat diakses dengan mudah oleh publik. d) Bagi Peneliti Selanjutnya: Mengingat keterbatasan penelitian ini yang berfokus pada aspek kualitatif-etika, peneliti selanjutnya disarankan untuk melakukan penelitian kuantitatif guna mengukur tingkat efektivitas pelatihan *digital tabayyun* terhadap penurunan persebaran hoaks di kelompok masyarakat tertentu.

## DAFTAR REFERENSI

- Al-Anshori, M. Z. (2022). *Etika Komunikasi di Era Digital: Perspektif Al-Qur'an dan Hadis*. Pustaka Al-Kautsar.
- Brennen, J. S., Simon, F. M., & Nielsen, R. K. (2021). Beyond (Mis)representation: Visuals in COVID-19 Misinformation. *The International Journal of Press/Politics*, 26(1).
- Creswell, J. W., & Creswell, J. D. (2023). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* ((6th ed.)). SAGE Publications.
- Flew, T. (2021). *Regulating Platforms*. Polity Press.
- Floridi, L. (2023). *The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities*. Oxford University Press.
- Haryanto, A. G., & Prastya, N. M. (2023). Deepfake and the Future of Truth in Digital Media. *Jurnal Komunikasi Ikatan Sarjana Komunikasi Indonesia*, 8(2).
- Hasan, N., & Khalil, M. (2024). Digital Tabayyun: Reinterpreting Islamic Ethics in the Age of Social Media Algorithms. *Journal of Islamic Studies and Culture*, 12(1).
- Mulyana, D. (2019). *Ilmu Komunikasi: Suatu Pengantar*. Remaja Rosdakarya.
- Nasrullah, R. (2021). *Media Sosial: Perspektif Komunikasi, Budaya, dan Sosioteknologi*. Simbiosis Rekatama Media.
- Sangkala, S., & Ahmad, H. (2024). Artificial Intelligence dalam Pendidikan Islam: Peluang dan Tantangan Moral. *Jurnal Pendidikan Islam Indonesia*, 10(1).
- Strate, L. (2020). *Media Ecology: An Approach to Understanding the Human Condition*. Peter Lang Inc.
- UNESCO. (2022). *Journalism, Fake News & Disinformation: A Handbook for Journalism Education and Training*. UNESCO Publishing.
- Vosoughi, S., dkk. (2024). *The Spread of True and False News Online*. MIT Press Collections.
- Wardle, C., & Derakhshan, H. (2019a). *Information Disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe.
- Wardle, C., & Derakhshan, H. (2019b). *Information Disorder: Toward an interdisciplinary framework for research and policymaking*. Council of Europe.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. PublicAffairs.