Proceeding of the 4rd FUAD's International Conference on Strengthening Islamic Studies (FICOSIS)
Vol 4 (2025)
Faculty of Ushuluddin, Adab and Dakwah UIN Ponorogo

ANALISIS BAHASA AL-QUR'AN BERBASIS BIG DATA: Kajian Korpus Dan Semiotika

Aina Shokhiful Asma'
Universitas Islam Negeri KIAI Ageng Muhammad Besari Ponorogo
Email: aaaina067@gmail.com

Moh. Alwy Amru Ghozali
Universitas Islam Negeri KIAI Ageng Muhammad Besari Ponorogo,
Email: alwy@uinponorogo.ac.id

Abstrak

Penelitian ini menelaah cara menggabungkan teknik big data dengan linguistik korpus untuk meneliti bahasa Al-Qur'an, dengan mengandalkan dua sumber digital utama: proyek Tanzil dan Quranic Arabic Corpus (QAC). Kedua basis data tersebut menyediakan teks yang telah terdigitalisasi lengkap dengan tokenisasi, anotasi morfologis, dan penanda semantik, sehingga memungkinkan pemeriksaan yang lebih teliti terhadap struktur morfologi, sebaran leksikal, serta pola-pola semantik yang sebelumnya sulit dijangkau lewat metode tradisional. Fokus kajian diarahkan pada analisis semantik leksem nashara beserta turunannya, dengan tujuan menunjukkan bagaimana data digital dapat mengungkap frekuensi kemunculan, variasi bentuk morfologis, hubungan kolokatif, dan prosodi semantik secara terukur. Hasil penelitian mengindikasikan bahwa pendekatan big data berhasil memetakan pola-pola bahasa Al-Qur'an secara empiris, terutama dalam mengungkap relasi sintagmatik-paradigmatik serta kecenderungan makna kontekstual yang muncul dalam berbagai ayat. Namun, analisis juga menyoroti keterbatasan platform digital seperti QAC, yang tidak memuat konteks historis, sosial, maupun hubungan intertekstual yang menjadi fondasi penting dalam kajian 'ulūm al-Qur'ān. Oleh karena itu, riset berbasis data perlu dipadukan dengan pendekatan tafsir dan hermeneutik agar interpretasi tidak menjadi reduksionistik. Penelitian ini berkontribusi memperkuat metodologi studi Al-Qur'an yang berbasis data digital dan menawarkan model analisis linguistik yang mengintegrasikan teknologi modern dengan tradisi keilmuan Islam klasik, sehingga memperkaya pemahaman tentang struktur dan makna teks suci secara lebih komprehensif.

Kata kunci: ***Big Data, Al-Qur'an,*** **Quranic Arabic Corpus ,** ***Tanzil***

PENDAHULUAN

Linguistik korpus adalah pendekatan empiris yang menggunakan data bahasa alami untuk mendeskripsikan pola bahasa yang digunakan oleh penutur.[1] Dengan memanfaatkan teknologi digital, Al-Qur'an dapat diolah menjadi korpus bahasa Arab klasik yang terstruktur, sehingga memungkinkan analisis linguistik yang lebih mendalam.[2] Proyek seperti Quranic Arabic Corpus dan Tanzil telah memungkinkan Al-Qur'an dianalisis sebagai korpus bahasa Arab klasik, membuka peluang untuk memahami struktur dan pola bahasa Al-Qur'an secara lebih sistematis. Melalui pendekatan ini, peneliti dapat mengidentifikasi pola-pola linguistik yang unik dalam Al-Qur'an, seperti struktur kalimat, penggunaan kata, dan gaya bahasa. Hal ini dapat membantu memahami bagaimana Al-Qur'an menggunakan bahasa untuk menyampaikan pesan dan makna yang terkandung di dalamnya. Sementara itu, semiotika Al-Qur'an menawarkan kerangka untuk memahami makna dan tanda-tanda dalam Al-Qur'an, serta bagaimana Al-Qur'an berinteraksi dengan masyarakat. Semiotika Al-Qur'an melihat Al-Qur'an sebagai kumpulan ujaran dalam bahasa Arab yang terbatas namun terbuka, yang memuat potensi penandaan yang diusulkan Tuhan kepada manusia.[3]

Dengan demikian, kombinasi linguistik korpus dan semiotika Al-Qur'an dapat membantu memahami Al-Qur'an sebagai teks yang kaya makna dan relevan dengan kehidupan manusia. Pendekatan ini dapat membantu memahami bagaimana Al-Qur'an menggunakan bahasa untuk menyampaikan pesan dan makna, serta bagaimana Al-Qur'an berinteraksi dengan masyarakat dan budaya. Linguistik korpus dan semiotika Al-Qur'an dapat membantu memahami Al-Qur'an dari perspektif yang berbeda, yaitu sebagai teks bahasa dan sebagai tanda-tanda yang membawa makna. Dengan demikian, kita dapat memahami Al-Qur'an secara lebih komprehensif dan mendalam. Pendekatan ini juga dapat membantu memahami bagaimana Al-Qur'an relevan dengan kehidupan manusia dan bagaimana Al-Qur'an dapat diaplikasikan dalam konteks modern.

---

[1] Fuzi Fauziah Nur Hizbullah, Fazlurrahman, "Penyusunan Model Korpus Al-Qur'an Digital," *AL-AZHAR INDONESIA SERI HUMANIORA* 3, no. 3 (2016). 217

[2] Herpindo Herpindo et al., "Pembelajaran Dan Pengajaran Tata Bahasa Berdasarkan Korpus," *Risenologi* 8, no. 2 (2023): 25–37. 4

[3] Ahmad Sihabul Millah, *Semiotika Al-Qur'an Mohammad Arkoun* (Yogyakarta: lintang hayuning buwana, 2022). 34

Dalam konteksLinguistik korpus memberikan cara yang terukur dan akurat untuk meneliti bahasa Al-Qur'an, namun kajian tentang kitab suci masih sangat dipengaruhi oleh tradisi tafsir yang bersifat interpretatif. Pendekatan tradisional menekankan pentingnya latar belakang sejarah, sebab-sebab turunnya ayat (asbāb al-nuzūl), serta otoritas para mufasir dalam menafsirkan makna. Kedua pendekatan ini memiliki peran yang berbeda namun saling melengkapi. Dengan memanfaatkan analisis korpus, peneliti dapat mengungkap pola struktural dan semantik secara objektif, sementara tafsir interpretatif membantu menyingkap dimensi historis, kontekstual, dan nilai-nilai spiritual yang tersirat. Penggunaan kedua metode secara bersamaan memungkinkan pemahaman yang lebih lengkap dan mendalam tentang Al-Qur'an. Korpus linguistik menyajikan data kuantitatif yang dapat memvalidasi atau memperkaya interpretasi tradisional, sedangkan tafsir klasik menambahkan nuansa kontekstual yang tidak dapat ditangkap oleh angka semata. Dengan mengintegrasikan keduanya, studi Al-Qur'an dapat menghasilkan gambaran yang lebih holistik, menggabungkan keakuratan analitis dengan kekayaan makna yang bersumber dari tradisi keilmuan Islam.

Analisis big data yang fokus pada kuantifikasi leksikal dan pola statistik seringkali dianggap tidak mampu menangkap kedalaman makna wahyu.[4] Analisis big data dapat membantu memahami struktur dan pola bahasa Al-Qur'an, namun tidak dapat menangkap makna yang lebih dalam dan spiritual dari wahyu. Wahyu Al-Qur'an memiliki makna yang kompleks dan multidimensional, yang tidak dapat direduksi menjadi hanya angka-angka dan statistik. Oleh karena itu, analisis big data harus digunakan sebagai alat bantu, bukan sebagai satu-satunya metode untuk memahami Al-Qur'an. Penting untuk mempertimbangkan konteks, nuansa bahasa, dan makna spiritual dalam memahami Al-Qur'an, sehingga kita dapat memperoleh pemahaman yang lebih komprehensif dan mendalam tentang wahyu.

Ketegangan konseptual muncul antara paradigma tekstual-empiris berbasis data dan paradigma hermeneutik-normatif berbasis otoritas penafsiran.[5] Paradigma tekstual-empiris berbasis data fokus pada analisis objektif dan presisi struktur dan pola bahasa Al-Qur'an, sedangkan paradigma hermeneutik-normatif berbasis otoritas penafsiran fokus

[4] Hossein Hassani et al., "Text Mining in Big Data Analytics" (n.d.): 1–34. 6

[5] Ali Mahfuz Munawar et al., "Epistemological Distinctions between the Living Qur ' an Methodology and Hermeneutics : A Theoretical Study in Contemporary Tafsir Studies" 1, no. 4 (2025): 461–476. 462

pada pemahaman makna dan konteks Al-Qur'an dalam kerangka tradisi dan otoritas keagamaan. Hal ini menimbulkan ketegangan antara kedua paradigma ini karena perbedaan pendekatan dan fokus dalam memahami Al-Qur'an. Paradigma tekstual-empiris berbasis data dapat dianggap tidak mampu menangkap makna yang lebih dalam dan spiritual dari Al-Qur'an, sehingga paradigma hermeneutik-normatif berbasis otoritas penafsiran menjadi sangat penting untuk memahami konteks dan makna Al-Qur'an. Dengan demikian, mempertimbangkan kedua paradigma ini sangat penting dalam memahami Al-Qur'an, sehingga kita dapat memperoleh pemahaman yang lebih komprehensif dan mendalam tentang wahyu.

Mengintegrasikan metode digital dengan kajian tradisional membuka peluang baru untuk menafsirkan Al-Qur'an secara lebih utuh.[6] Analisis komputasional mengungkap pola bahasa dan struktur teks secara objektif, sementara tradisi tafsir menyediakan konteks historis, sebab-sebab turunnya ayat, serta otoritas mufasir yang penting bagi pemahaman makna. Dengan memadukan kedua perspektif, Al-Qur'an tidak lagi dilihat sekadar sebagai kumpulan kata, melainkan sebagai karya yang sarat makna dan relevan dengan kehidupan manusia. Karena teksnya bersifat kompleks dan berlapis, diperlukan pendekatan holistik yang menggabungkan analisis kuantitatif dan kualitatif untuk memperoleh pemahaman yang lebih dalam.

Namun, integrasi ini tidak tanpa tantangan. Perbedaan metodologi, epistemologi, dan ontologi antara pendekatan digital dan tradisional dapat menimbulkan kesulitan.[7] Penelitian lebih lanjut diperlukan untuk menemukan cara mengharmoniskan kedua cara tersebut secara efektif dan efisien. Jika berhasil, integrasi ini akan memperdalam pemahaman tentang Al-Qur'an, agama Islam, dan budaya yang terkait, serta memperkuat dialog antar-agama dan kebudayaan dalam konteks zaman kini.

Artikel ini menawarkan pendekatan sintetik yang mengintegrasikan data digital Al-Qur'an dengan perangkat linguistik korpus dan teori semiotika Qur'an. Pendekatan ini bertujuan untuk memperluas basis data yang dapat digunakan dalam menganalisis teks Al-Qur'an secara empiris, bukan untuk menggantikan metode tafsir tradisional. Melalui

[6] U I N Sunan, Ampel Surabaya, and Indonesia Email, "THE ROLE OF DIGITAL TOOLS IN QUR ' ANIC RESEARCH AND TAFSIR : Improving Methodology and Interpretation" 12, no. 1 (2025): 124–136. 129

[7] Kais Dukes and Nizar Habash, "Morphological Annotation of Quranic Arabic," *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010* (2010): 2530–2536. 2530

analisis frekuensi kata, kolokasi, dan jaringan semantik dalam Quranic Arabic Corpus, penelitian ini berupaya menunjukkan bagaimana big data dapat memperkaya studi linguistik Al-Qur'an. Teori semiotika Qur'an digunakan untuk menafsirkan pola kebahasaan tersebut sebagai sistem tanda yang merepresentasikan makna teologis dan konseptual di dalam teks. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan metode analisis linguistik Qur'an yang lebih efektif dan efisien.

Pendekatan sintetik yang digunakan dalam penelitian ini memungkinkan untuk memahami struktur dan makna teks Al-Qur'an secara lebih komprehensif. Analisis frekuensi kata, kolokasi, dan jaringan semantik dapat membantu mengidentifikasi pola-pola linguistik yang unik dalam Al-Qur'an, sedangkan teori semiotika Qur'an dapat membantu menafsirkan makna teologis dan konseptual yang terkandung dalam pola-pola tersebut. Dengan demikian, penelitian ini dapat memberikan wawasan baru tentang struktur dan makna teks Al-Qur'an.

Penelitian sebelumnya telah menerapkan analisis berbasis korpus terhadap Al-Qur'an, seperti yang dilakukan oleh Arfan Ahmed (2022) yang menegaskan posisi korpus Qur'an sebagai salah satu sumber linguistik Arab paling lengkap dan terbuka di dunia.[8] Studi semantik prosodi kata oleh Ghuzayyin (2025) juga menunjukkan bahwa prosodi semantik ditentukan oleh niat, konteks, fungsi pragmatis, kolokat, dan kontrusi gramatikal. Namun, penelitian yang secara eksplisit menautkan pendekatan big data dengan teori semiotika Qur'an masih jarang ditemukan.[9] Oleh karena itu, artikel ini bertujuan untuk menelaah bagaimana paradigma big data dapat mengubah metode analisis linguistik Qur'an sekaligus menyoroti keterbatasannya. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan metode analisis linguistik Qur'an yang lebih efektif dan efisien.

Penelitian ini berfokus pada penerapan pendekatan big data dalam analisis linguistik Qur'an, yang diharapkan dapat memberikan wawasan baru tentang struktur dan makna teks Qur'an. Dengan menggunakan metode analisis korpus, penelitian ini dapat

[8] Arfan Ahmed et al., "Freely Available Arabic Corpora: A Scoping Review," *Computer Methods and Programs in Biomedicine Update* 2 (2022).

[9] Ghuzayyil Mohammed Al-otaibi, "Semantic Prosody of Deictic Verbs in the Holy Qur ' ā N : A Corpus Study" (2025): 1–16.

mengidentifikasi pola-pola linguistik yang unik dalam Qur'an dan memahami bagaimana pola-pola tersebut berkontribusi pada makna keseluruhan teks. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi pada pengembangan teori semiotika Qur'an yang lebih komprehensif dan sistematis.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitataif dengan metode analisis linguistik korpus dan semiotik. Dalam penelitian ini, peneliti akan menganalisis makna kata "jaahada" baik dalam bentuk dasar maupun derivasinya sehingga maknanya kurang lebih tetap, seperti yang terdapat dalam kamus. Dalam memproses data ini, penelitian ini akan menganalisis makna dari kata "nashara" dalam perspektif semantik leksikal maupun penggunaannya dalam Al-Qur'an. kemudian menganalisis kolokasi dan konteks sintagmatik. Yaitu menelusuri kata yang sering muncul bersamaan denan jaahada. Adapun untuk mendapatkan ayat-ayat Al-Qur'an yang relevan dengan tema penelitian, maka penulis menggunakan korpus Al-Qur'an digital berbasis website, yaitu Quran Arabic Corpus (https://corpus.quran.com/) dan tanzil project (https://tanzil.net) dalam mengumpulkan data. Quran Arabic Corpus adalah salah satu website korpus Al-Qur'an yang mempunyai tampilan dan isi konten tentang teks Al Qur'an, sedangkan tanzil project menyediakan teks al-quran digital terstandar. Pendekatan ini bersifat deskriptif analitis, dimana data dari kopus digital digunakan untuk memperkuat penafsiran. Analitik pendekatan berfokus pada semantik leksikal.

HASIL PENELITIAN DAN PEMBAHASAN

Big Data Al-Qur'an

Big data merujuk pada sekumpulan data yang sangat besar, cepat berubah, dan beragam, yang tidak bisa dikelola dengan teknik pengolahan data tradisional.[10] Digitalisasi Al-Qur'an dalam dua dekade terakhir telah menciptakan ekosistem data linguistik yang sangat kaya, terstruktur, dan terbuka untuk dianalisis menggunakan teknik komputasional modern, yang dikenal sebagai Qur'anic Big Data Environment. Dua sumber data utama dalam ekosistem ini adalah Tanzil Project dan Quranic Arabic Corpus

[10] Uzair Nazri Siregar, "Penggunaan AI Dan Big Data Dalam Pembelajaran Nilai-Nilai Al-Qur'an Dan Hadis Di Madrasah," *Arba: Jurnal Studi Keislaman* 1, no. 3 (2025). 165

(QAC). Tanzil menyediakan teks Al-Qur'an dalam format digital yang telah melalui proses verifikasi ketat oleh para qārī dan pakar mushaf, menjadikannya rujukan standar global untuk penelitian komputer terhadap Qur'an (Kazim & Ghani, dalam Tanzil Documentation). Quranic Arabic Corpus memberikan struktur anotasi linguistik berlapis yang jauh lebih kaya, dengan setiap kata diberi label akar, lema, kelas kata, fitur morfologis, dan hubungan dependensi sintaktis. Dengan demikian, QAC merupakan salah satu korpus Arab klasik paling sistematis yang tersedia secara terbuka, dengan 77.430 kata yang dianotasi secara terstandar.[11]

Qur'anic Big Data Environment menawarkan peluang baru untuk menganalisis Al-Qur'an secara komputasional, memungkinkan peneliti untuk mengidentifikasi pola-pola linguistik yang unik dan memahami makna teks Al-Qur'an secara lebih komprehensif. Integrasi antara Tanzil Project dan Quranic Arabic Corpus memungkinkan analisis linguistik yang lebih akurat dan sistematis, membuka jalan bagi penelitian Al-Qur'an yang lebih mendalam dan inovatif. Dengan demikian, peneliti dapat mengembangkan metode analisis yang lebih efektif dan efisien, serta meningkatkan pemahaman tentang struktur dan makna teks Al-Qur'an. Selain itu, Qur'anic Big Data Environment juga memungkinkan peneliti untuk melakukan analisis lintas disiplin, seperti linguistik, ilmu komputer, dan studi keislaman, sehingga dapat memberikan kontribusi yang lebih luas pada pemahaman tentang Al-Qur'an dan kebudayaan Islam. Oleh karena itu, pengembangan Qur'anic Big Data Environment merupakan langkah penting dalam memajukan penelitian Al-Qur'an dan meningkatkan pemahaman tentang agama Islam.

Metode korpus memberi peneliti cara yang lebih empiris untuk menelaah teks Al-Qur'an, mengalihkan fokus dari makna filosofis atau teologis ke pola numerik dan distribusi yang dapat diverifikasi. Sumber data yang ada memenuhi tiga ciri big data: (1) volume besar – seluruh teks telah terdigitalisasi, lengkap dengan tokenisasi dan anotasi per-kata; (2) variasi berlapis – meliputi morfologi, sintaksis, semantik, bahkan ontologi; (3) veracity – banyak korpus menyertakan verifikasi serta anotasi manual oleh ahli bahasa Arab dan dibangun di atas teks terkurasi seperti proyek Tanzil. Dengan demikian, "big data Qur'an" tidak hanya berarti kuantitas data, melainkan cara data tersebut disusun, dianotasi, dan disajikan sehingga memungkinkan analisis pola linguistik secara sistematis, objektif, dan terukur.

[11] Dukes and Habash, "Morphological Annotation of Quranic Arabic." 2535

Dengan memanfaatkan korpus, peneliti dapat mengamati Al-Qur'an dari perspektif yang lebih empiris, menyoroti kecenderungan numerik dan pola distribusi yang dapat diverifikasi, bukan sekadar makna filosofis atau teologis. Data yang tersedia memenuhi tiga kriteria big data: (1) volume besar – seluruh teks telah terdigitalisasi dan tersedia sebagai korpus lengkap beserta tokenisasi dan anotasi per-kata; (2) variasi berlapis – meliputi morfologi, sintaksis, semantik, bahkan ontologi; (3) veracity – banyak sumber korpus menyertakan proses verifikasi dan anotasi manual oleh ahli bahasa Arab serta dibangun di atas teks terkurasi seperti proyek Tanzil.[12] Jadi, big data Qur'an menekankan cara data disusun, dianotasi, dan disajikan untuk memungkinkan analisis pola linguistik secara sistematis, objektif, dan terukur.

Melalui analisis korpus, struktur dan makna Al-Qur'an dapat dipahami secara lebih komprehensif. Metode ini memungkinkan identifikasi pola linguistik unik, seperti frekuensi kata, kolokasi, dan jaringan semantik, yang pada gilirannya memperkaya pemahaman tentang makna teks dan relevansinya dengan kehidupan modern. Selain itu, penggunaan korpus membuka peluang kolaborasi lintas disiplin antara linguistik, ilmu komputer, dan studi keislaman. Dengan menggabungkan keahlian tersebut, peneliti dapat mengembangkan teknik analisis yang lebih efektif dan efisien, serta memperdalam pemahaman tentang struktur dan makna Al-Qur'an. Pemanfaatan big data Qur'an juga meningkatkan wawasan tentang budaya Islam dan kontemporer, sehingga penelitian Al-Qur'an memberikan kontribusi yang lebih luas pada pemahaman agama dan kebudayaan Islam.

Temuan Linguistik (Kata Kunci, Pola Semantik)

Melalui website korpus Al-Qur'an, kata "nashara" dalam Al-Qur'an memiliki bentuk kata yang beragam, di antaranya adalah kata yang berbentuk fi'il (verba), yakni verba perfektif, verba imperfektif, dan verba imperatif. Selain itu, ada juga dari bentuk isim (nomina), yaitu nomina, partisipal aktif, dan partisipal pasif. Kata "nashara" dengan berbagai derivasinya disebutkan dalam Al-Qur'an sebanyak 158 kali yang tersebar dalam 46 Surah.[13]

[12] Mohamed Ahmed Sherif and Axel-cyrille Ngonga Ngomo, "Semantic Quran" 1 (2009): 1–5. 2

[13] Luthfia Khoiriyatunnisa, Dedy Ari Asfar, and Agus Syahrani, "Analisis Semantik Makna Kata /Nashara/ نصر Dan Derivasinya Dalam Al-Qur'an / Semantic Analysis of the Meaning of the Word 'نصر' and Its Derivation in the Qur'an," *Diwan : Jurnal Bahasa dan Sastra Arab* 8, no. 2 (2022).212

Analisis korpus Al-Qur'an memungkinkan peneliti untuk memahami struktur dan makna kata "nashara" dalam konteks yang lebih luas. Dengan menggunakan metode analisis korpus, peneliti dapat mengidentifikasi pola-pola linguistik yang unik dalam penggunaan kata "nashara" dalam Al-Qur'an, seperti frekuensi kata, kolokasi, dan jaringan semantik. Hal ini dapat membantu meningkatkan pemahaman tentang makna kata "nashara" dan relevansinya dengan kehidupan modern. Selain itu, analisis korpus juga memungkinkan peneliti untuk memahami bagaimana kata "nashara" digunakan dalam konteks yang berbeda-beda dalam Al-Qur'an, seperti dalam konteks sejarah, sosial, dan teologis. Dengan demikian, peneliti dapat mengembangkan pemahaman yang lebih komprehensif tentang makna kata "nashara" dan relevansinya dengan kehidupan modern.

Temuan berdasarkan kolokasi kata menunjukkan bahwa "nashara" dalam Al-Qur'an memiliki makna yang beragam dan terkait dengan konteks yang berbeda-beda. Kolokasi mengacu pada kecenderungan suatu kata bersanding dengan kata lain, sehingga membantu memahami makna kata tersebut dalam kerangka yang lebih luas.[14] Analisis kolokasi mengungkapkan bahwa "nashara" muncul bersamaan dengan istilah-istilah yang mencerminkan bantuan, kemenangan, atau dukungan, baik dalam konteks historis, moral, maupun spiritual. Misalnya, kata ini sering berpasangan dengan istilah yang mengindikasikan pertolongan Allah, dukungan manusia, atau kemenangan dalam perjuangan. Variasi kolokasi ini memperkaya pemahaman tentang dimensi semantik "nashara" dan menegaskan bahwa makna kata tersebut tidak bersifat statis, melainkan dipengaruhi oleh jaringan hubungan leksikal di sekitarnya.

Analisis semantik prosodi berupaya mengungkap muatan afektif atau evaluatif yang secara konsisten terikat pada sebuah leksem melalui pola distribusinya dalam korpus. Secara metodologis, kajian ini memanfaatkan tiga instrumen utama: konkordansi untuk menelusuri semua kemunculan kata dalam konteks; ukuran kolokasi seperti mutual information atau t-score untuk mengukur kekuatan asosiasi antar-kata; serta representasi vektor (word embeddings) yang memungkinkan penilaian apakah suatu kata cenderung muncul dalam konteks bernada positif, negatif, atau netral. Sebagai ilustrasi, meskipun

[14] From Wikipedia, "Kolokasi," Kolokasi - Wikipedia bahasa Indonesia, ensiklopedia bebas . diakses pada 05 November 2025 pada pukul 08.15

secara denotatif waqaʿa berarti “jatuh” atau “terjadi”,[15] analisis korpus menunjukkan bahwa leksem ini sering muncul bersamaan dengan istilah yang menyiratkan kerugian, bencana, atau hukuman, sehingga prosodi semantiknya berwarna negatif. Temuan ini tidak bersifat sekadar statistik; mereka mengungkap pola pragmatis-retoris yang mengatur bagaimana teks menimbulkan reaksi emosional atau penilaian terhadap peristiwa.

Selain itu, teknik komputasional lain memperkaya analisis: text-mining mengekstrak pola frekuensi kata; topic modeling (misalnya Latent Dirichlet Allocation) memetakan koherensi tematik lintas ayat atau surat; analisis jaringan kolokasi menyoroti pasangan kata yang membentuk nuansa makna bersama; dan embedding mengidentifikasi dimensi makna yang muncul dalam konteks serupa.[16] Kombinasi metode ini menyediakan bukti empiris mengenai kecenderungan makna dan fungsi retoris, namun hasilnya harus ditafsirkan dalam kerangka historis-kontekstual dan hermeneutik untuk menghindari reduksi makna menjadi sekadar angka.

Dengan mengintegrasikan pendekatan kuantitatif dan kualitatif, analisis semantik prosodi tidak hanya memperluas pemahaman tentang struktur bahasa Al-Qur’an, tetapi juga menawarkan model interpretatif yang menyeimbangkan keunggulan teknologi modern dengan kekayaan tradisi keilmuan Islam klasik. Pendekatan ini berpotensi menghasilkan wawasan baru tentang cara teks suci mengorganisir nilai evaluatif dan emosional, sekaligus menjaga sensitivitas terhadap konteks historis dan budaya yang melasarinya.

Implikasi Tafsir

Relasi metodologis antara pendekatan digital dan pendekatan tradisional perlu ditinjau ulang untuk memahami bagaimana kedua pendekatan ini dapat saling melengkapi dan meningkatkan pemahaman kita tentang Al-Qur'an. Penelitian linguistik digital terhadap Al-Qur'an telah menunjukkan potensi besar dalam memahami struktur dan makna Al-Qur'an. Dengan menggunakan metode yang lebih objektif dan sistematis, penelitian ini dapat membantu memvalidasi dan memperluas pemahaman tentang Al-

[15] Eman Abdulaziz Alhusaiyan, “Semantic Prosody in Holy Qur’an: A Corpus Study of the Collocations and Meanings of the Verb Waqa’a‘Fell’” 7, no. 4 (2023): 19–33. 25

[16] Mhd Ammar Alsalka Menwa Alshammeri, Eric Atwell, “Quranic Topic Modelling Using Paragraph Vectors,” *Intelligent Systems Conference (IntelliSys)* (2020). 4

Qur'an. Penelitian korpus dan linguistik komputasional telah menunjukkan bahwa pola semantik yang terdeteksi melalui analisis frekuensi, kolokasi, embedding, dan ontologi cenderung menyelaraskan atau bahkan menguatkan temuan-temuan mufasir klasik, meskipun dalam kerangka metodologi yang berbeda.[17] Hal ini menunjukkan bahwa penelitian linguistik digital dapat menjadi salah satu alat yang berguna dalam pengembangan tafsir dan ulum al-Qur'an.

Dengan mempertimbangkan kedua pendekatan ini secara bersama-sama, kita dapat memperoleh pemahaman yang lebih komprehensif dan mendalam tentang Al-Qur'an. Penelitian linguistik digital dapat membantu mengidentifikasi area-area yang memerlukan penelitian lebih lanjut dan memperluas cakupan penelitian Al-Qur'an, sehingga kita dapat meningkatkan kualitas penelitian dan pemahaman kita tentang wahyu. Penelitian ini juga dapat membantu memperkuat hubungan antara linguistik dan ilmu-ilmu keislaman, serta meningkatkan pemahaman tentang Al-Qur'an dan ajaran Islam. Dengan demikian, penelitian linguistik digital dapat menjadi salah satu kunci untuk memahami Al-Qur'an secara lebih mendalam dan komprehensif.

Kata "qalb" (hati) dalam Al-Qur'an memiliki makna yang lebih luas daripada sekadar organ biologis. Dalam banyak kasus, kata "qalb" digunakan secara metaforis untuk merujuk pada pemahaman, keimanan, emosi, dan niat. Hal ini menunjukkan bahwa kata "qalb" memiliki makna yang kompleks dan terkait dengan aspek spiritual dan emosional manusia.

Dalam Al-Qur'an, kata "qalb" digunakan dalam berbagai konteks, seperti dalam Surah Al-Baqarah ayat 7:

خَتَمَ اللّٰهُ عَلٰى قُلُوْبِهِمْ وَعَلٰى سَمْعِهِمْ ۗوَعَلٰٓى اَبْصَارِهِمْ غِشَاوَةٌ وَّلَهُمْ عَذَابٌ عَظِيْمٌ

(Allah telah mengunci hati mereka, dan pendengaran mereka, dan penglihatan mereka, dan bagi mereka azab yang besar). Dalam ayat ini, kata "qalb" digunakan untuk merujuk pada pemahaman dan keimanan. Dalam Surah Al-Shu'ara ayat 89:

اِلَّا مَنْ اَتَى اللّٰهَ بِقَلْبٍ سَلِيْمٍ ۗ

[17] Ensaf Hussein Mohamed and Eyad Mohamed Shokry, "QSST: A Quranic Semantic Search Tool Based on Word Embedding," *Journal of King Saud University - Computer and Information Sciences* 34, no. 3 (2022). 937

(kecuali orang yang datang kepada Allah dengan hati yang bersih). Dalam ayat ini, kata "qalb" digunakan untuk merujuk pada keimanan dan niat. Dalam Surah Al-Mu'minun ayat 60:

وَالَّذِيْنَ يُؤْتُوْنَ مَآ اٰتَوْا وَّقُلُوْبُهُمْ وَجِلَةٌ اَنَّهُمْ اِلٰى رَبِّهِمْ رٰجِعُوْنَ ۙ

(dan orang-orang yang memberikan apa yang mereka berikan, sedang hati mereka penuh dengan rasa takut, karena mereka akan kembali kepada Tuhan mereka). Dalam ayat ini, kata "qalb" digunakan untuk merujuk pada emosi dan rasa takut.

Dengan demikian, kata "qalb" dalam Al-Qur'an memiliki makna yang kompleks dan terkait dengan aspek spiritual dan emosional manusia. Hal ini menunjukkan bahwa kata "qalb" tidak hanya merujuk pada organ biologis, tetapi juga pada aspek yang lebih dalam dan spiritual dari manusia.Dalam sebuah artikel (lihat: https://afosa.org/the-meaning-of-heart-qalb-in-quran/), Kata "qalb" dalam Al-Qur'an dan literatur Islam memiliki makna yang kompleks dan terkait dengan aspek spiritual dan emosional manusia. Selain sebagai organ fisik, kata "qalb" juga digunakan untuk merujuk pada emosi, pemahaman, niat, dan keimanan.

Analisis korpus dapat membuka ruang interpretasi baru mengenai kata “qalb” dalam Al-Qur’an. Dengan menelusuri penggunaan kata tersebut dalam berbagai ayat, peneliti dapat mengungkap nuansa makna serta signifikansinya yang lebih mendalam. Pendekatan ini juga memungkinkan identifikasi kolokasi, embedding, dan pola linguistik lain yang memperkaya pemahaman tentang dimensi spiritual dan emosional yang diwakili oleh “qalb”. Hasilnya memberikan gambaran komprehensif tentang peran kata tersebut dalam teks suci serta memperluas wawasan tentang aspek manusia yang dibahas dalam Al-Qur’an.

Keterbatasan Big Data

Penelitian ini menegaskan bahwa big data bukanlah solusi tunggal dalam memahami Al-Qur'an. Meskipun QAC (Quranic Arabic Corpus) menyediakan anotasi morfologi, sintaksis, dan ontologi semantik, namun tidak tersedia kolom resmi yang merekam konteks historis-budaya, sebab-turunnya ayat, atau latar sosial-keagamaan sebagai bagian dari data korpus. Fokus QAC masih sangat tekstual, sehingga data digital ini sering kali terbatas karena hanya berfokus pada teks dan bukan konteks. Konteks historis-budaya, seperti perang Badar, perselisihan Yahudi Madinah, atau situasi sosial Quraisy, sangat penting dalam memahami makna Al-Qur'an. Namun, QAC tidak menyediakan informasi ini secara langsung. Oleh karena itu, penelitian ini menekankan

bahwa big data harus digunakan sebagai alat bantu, bukan sebagai satu-satunya sumber kebenaran. Big data dapat membantu memahami struktur dan pola bahasa Al-Qur'an, namun tidak dapat menangkap makna yang lebih dalam dan spiritual dari wahyu.

Selain itu, QAC disusun melalui anotasi manual yang juga rentan terhadap perbedaan interpretasi. Verifikasi sintaksis masih terus direvisi hingga kini, sehingga masih ada kemungkinan kesalahan. Hal ini menunjukkan bahwa big data bukanlah solusi yang sempurna dan harus digunakan dengan hati-hati. Penelitian ini juga menunjukkan bahwa fokus hanya pada angka/kuantifikasi dapat menyebabkan reduksi makna Al-Qur'an. Mengabaikan aspek kontekstual, spiritual, historis, dan nilai normatif wahyu dapat membuat kita kehilangan makna yang lebih dalam dari Al-Qur'an. Penelitian ini juga menekankan bahwa big data harus digunakan dalam kerangka yang lebih luas, yaitu dengan mempertimbangkan konteks historis-budaya, spiritual, dan nilai normatif wahyu. Dengan demikian, kita dapat memperoleh pemahaman yang lebih komprehensif dan mendalam tentang Al-Qur'an. Penelitian ini juga menunjukkan bahwa big data dapat menjadi salah satu alat yang berguna dalam pengembangan tafsir dan ulum al-Qur'an, namun harus digunakan dengan hati-hati dan dalam kerangka yang lebih luas.

Studi ini menegaskan bahwa big data bukanlah jawaban tunggal dalam memahami Al-Qur'an, melainkan salah satu alat yang dapat mendukung pengembangan tafsir dan ilmu al-Qur'an. Oleh karena itu, data besar sebaiknya dipakai sebagai pendukung, bukan sebagai satu-satunya sumber kebenaran, dan harus diintegrasikan dengan pertimbangan historis-budaya, spiritual, serta nilai normatif yang terkandung dalam wahyu. Hanya dengan kerangka yang lebih luas inilah analisis digital dapat memberikan kontribusi yang berarti tanpa mengabaikan dimensi kontekstual yang esensial.

Dengan demikian, penelitian ini dapat membantu meningkatkan pemahaman kita tentang Al-Qur'an dan ajaran Islam, serta membantu memperkuat hubungan antara linguistik dan ilmu-ilmu keislaman. Penelitian ini juga dapat membantu mengidentifikasi area-area yang memerlukan penelitian lebih lanjut dan memperluas cakupan penelitian Al-Qur'an, sehingga kita dapat memperoleh pemahaman yang lebih komprehensif dan mendalam tentang Al-Qur'an.

Sebagaimana dicatat oleh Ni'am (2025),[18] algoritma digital dapat mengganggu tradisi hermeneutik Islam dengan menyederhanakan makna teks. Demikian pula, Shahid

[18] Mohammad Fattahun Ni'am, "From Sanad to Syntax : The Disruption of Islamic Hermeneutics by Qur ' AnGPT," *Basmala Journal of Qur'an and Hadith* 1, no. 1 (2025): 46–67. 52

(2025)[19] memperingatkan bahwa penerjemahan dan pemrosesan komputasional tak selalu menangkap nuansa asli bahasa Arab. Oleh karena itu, penelitian ini menekankan bahwa big data harus digunakan sebagai alat bantu yang dapat membantu memahami Al-Qur'an, namun tidak dapat menggantikan penelitian tradisional dan kontekstual. Dengan menggunakan big data secara bijak, kita dapat memperoleh pemahaman yang lebih dalam tentang Al-Qur'an dan meningkatkan kualitas penelitian kita.

PENUTUP

Penelitian ini menunjukkan bahwa pendekatan big data, melalui korpus digital seperti Tanzil dan Quranic Arabic Corpus, mampu menghadirkan cara baru dalam menelaah struktur kebahasaan Al-Qur'an. Melalui analisis frekuensi, kolokasi, dan prosodi semantik, penelitian ini berhasil mengidentifikasi pola-pola linguistik kata *nashara* secara lebih sistematis, serta menunjukkan bagaimana pola tersebut membentuk jaringan makna yang konsisten di berbagai konteks ayat. Meskipun demikian, temuan penelitian juga menegaskan bahwa metode komputasional tidak dapat berdiri sendiri. Data digital tidak menyediakan konteks historis, sosiologis, dan teologis yang menjadi fondasi utama tafsir dan ulum al-Qur'an. Ketergantungan pada kuantifikasi linguistik semata berisiko mereduksi kompleksitas makna wahyu. Karena itu, diperlukan integrasi metodologis yang memadukan ketelitian analisis korpus dengan kedalaman pendekatan hermeneutik.

Dengan menggabungkan kedua pendekatan tersebut, penelitian ini menunjukkan bahwa big data tidak menggantikan metode tafsir tradisional, tetapi memperkaya dan melengkapinya. Hasil penelitian ini diharapkan memberi kontribusi pada pengembangan metodologi analisis linguistik Qur'an yang lebih komprehensif, kritis, dan relevan dengan kebutuhan penelitian kontemporer.

[19] William Sayers Shahid, Usama, Muhammad Zunnurain Hussain, "Computational Analysis of Quran Text Using Machine Learning and Large Language Models," *universitas of gloucestershire* (2025).

DAFTAR RUJUKAN

.

Ahmed, Arfan, Nashva Ali, Mahmood Alzubaidi, Wajdi Zaghouani, Alaa A. Abd-alrazaq, and Mowafa Househ. "Freely Available Arabic Corpora: A Scoping Review." *Computer Methods and Programs in Biomedicine Update* 2 (2022).

Al-otaibi, Ghuzayyil Mohammed. "Semantic Prosody of Deictic Verbs in the Holy Qur ' ā N : A Corpus Study" (2025): 1–16.

Al-turas, Buletin, and Kajian Semantik Arab. "Medan Makna Jihad Dalam Perspektif Bahasa Arab Dan Islam" XXIV, no. 1 (2018).

Alhusaiyan, Eman Abdulaziz. "Semantic Prosody in Holy Qur'an: A Corpus Study of the Collocations and Meanings of the Verb Waqa'a'Fell'" 7, no. 4 (2023): 19–33.

Dukes, Kais, and Nizar Habash. "Morphological Annotation of Quranic Arabic." *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010* (2010): 2530–2536.

Hassani, Hossein, Christina Beneki, Stephan Unger, and Maedeh Taj Mazinani. "Text Mining in Big Data Analytics" (n.d.): 1–34.

Herpindo, Herpindo, Astuty Astuty, Mursia Ekawati, Gilang Fadhilia Arvianti, Miftahula Rizqin Nikmatullah, and Muhammad Nur Afiq. "Pembelajaran Dan Pengajaran Tata Bahasa Berdasarkan Korpus." *Risenologi* 8, no. 2 (2023): 25–37.

Khoiriyatunnisa, Luthfia, Dedy Ari Asfar, and Agus Syahrani. "Analisis Semantik Makna Kata /Nashara/ نصر Dan Derivasinya Dalam Al-Qur'an / Semantic Analysis of the Meaning of the Word 'نصر' and Its Derivation in the Qur'an." *Diwan : Jurnal Bahasa dan Sastra Arab* 8, no. 2 (2022).

Menwa Alshammeri, Eric Atwell, Mhd Ammar Alsalka. "Quranic Topic Modelling Using Paragraph Vectors." *Intelligent Systems Conference (IntelliSys)* (2020).

Millah, Ahmad Sihabul. *Semiotika Al-Qur'an Mohammad Arkoun*. Yogyakarta: lintang hayuning buwana, 2022.

Mohamed, Ensaf Hussein, and Eyad Mohamed Shokry. "QSST: A Quranic Semantic Search Tool Based on Word Embedding." *Journal of King Saud University - Computer and Information Sciences* 34, no. 3 (2022).

Munawar, Ali Mahfuz, Muhammad Diaz Supandi, Alfan Arif Rabbani, Abdillah Isnainurrahman, and Muhammad Muzaini Asyauqillah. "Epistemological Distinctions between the Living Qur ' an Methodology and Hermeneutics : A Theoretical Study in Contemporary Tafsir Studies" 1, no. 4 (2025): 461–476.

Ni'am, Mohammad Fattahun. "From Sanad to Syntax : The Disruption of Islamic Hermeneutics by Qur ' AnGPT." *Basmala Journal of Qur'an and Hadith* 1, no. 1 (2025): 46–67.

Nur Hizbullah, Fazlurrahman, Fuzi Fauziah. "Penyusunan Model Korpus Al-Qur'an Digital." *AL-AZHAR INDONESIA SERI HUMANIORA* 3, no. 3 (2016).

Rozi, A Fahrur, and Niswatur Rokhmah. “Tafsir Klasik: Analisis Terhadap Kitab Tafsir Era Klasik.” *Jurna Kaca* 9 (2019): 33–58.

Shahid, Usama, Muhammad Zunnurain Hussain, William Sayers. “Computational Analysis of Quran Text Using Machine Learning and Large Language Models.” *universitas of gloucestershire* (2025).

Sherif, Mohamed Ahmed, and Axel-cyrille Ngonga Ngomo. “Semantic Quran” 1 (2009): 1–5.

Sunan, U I N, Ampel Surabaya, and Indonesia Email. “THE ROLE OF DIGITAL TOOLS IN QUR ’ ANIC RESEARCH AND TAFSIR : Improving Methodology and Interpretation” 12, no. 1 (2025): 124–136.

Uzair Nazri Siregar. “Penggunaan AI Dan Big Data Dalam Pembelajaran Nilai-Nilai Al-Qur’an Dan Hadis Di Madrasah.” *Arba: Jurnal Studi Keislaman* 1, no. 3 (2025).

Wikipedia, From. “Kolokasi.” https://id.wikipedia.org/wiki/Kolokasi.