

Analisis Bibliometrik Publikasi tentang ChatGPT

Muhammad Irfan¹, Yogik Septiadi², Muhammad Alkautsar Rivandra³, Nur Aini Rakhmawati⁴

^{1,2,3,4}Departemen Sistem Informasi, Institut Teknologi Sepuluh Nopember
Keputih, Sukolilo, Surabaya, Jawa Timur, Indonesia, 60117

¹11irfan.205026@mhs.its.ac.id

²yogikseptiadi.205026@mhs.its.ac.id

³5026201048@mhs.its.ac.id

⁴nur.aini@is.its.ac.id

Abstrak— Analisis bibliometrik telah dilakukan terhadap 200 publikasi tentang “ChatGPT”, setelah dilakukan pengecekan ulang menggunakan software mendeley hanya 127 yang dapat digunakan, publikasi ini diambil melalui software Publish-or-Perish dengan menggunakan basis data Google Scholar. Penelitian ini bertujuan untuk menganalisis pembahasan ruang lingkup atau kata kunci “ChatGPT” dalam publikasi internasional menggunakan pendekatan bibliometrik. Bibliometrik merupakan metode statistik untuk menganalisis metadata publikasi yang tersedia dalam basis data publikasi dalam hal ini menggunakan aplikasi Publish-of-Perish kemudian mengekstrak hasil pencarian menggunakan VOSviewer. Melalui VOSviewer didapatkan sebelas cluster yang ditandai dengan kode warna merah, hijau, biru tua, kuning, ungu, biru muda, jingga, coklat, merah muda, peach, dan hijau muda. Berdasarkan hasil analisis dari VOSviewer, didapatkan topik yang paling sering dibahas adalah ChatGPT, artificial intelligence, dan natural language processing. Sementara itu, untuk tiga topik yang paling jarang dibahas adalah qualitative research, cybersecurity, dan library services.

Kata Kunci— ChatGPT, Bibliometrik, VOSviewer, Publish-or-Perish.

I. PENDAHULUAN

teknologi ChatGPT, termasuk meningkatkan kemampuan untuk memecahkan masalah dan memberikan solusi kreatif untuk situasi yang kompleks. Riset-riset yang telah dipublikasikan sangat beragam, berasal dari banyak disiplin ilmu. Beragam publikasi ditemukan dari berbagai tahun hingga yang terbaru.

Salah satu cara untuk menarik garis besar dari begitu banyak publikasi yang ada adalah dengan melakukan studi bibliometrik. Dalam studi bibliometrik, digunakan metode statistik untuk menganalisis metadata publikasi yang tersedia dalam basis data publikasi yang dikembangkan oleh banyak lembaga di dunia ini, analisis bibliometrik juga dapat digunakan untuk mengukur dampak dari publikasi tentang ChatGPT dan melacak perkembangan topik ini selama bertahun-tahun. Hal ini dapat memungkinkan mengidentifikasi tren, membandingkan publikasi, dan mencari hubungan antar publikasi. Basis data ini beberapa bisa diakses gratis dan beberapa harus berlangganan.

Meski dasar dari studi bibliometrik memerlukan pemahaman statistik yang cukup tinggi, banyak aplikasi yang tersedia untuk memudahkan prosesnya, salah satu di antaranya adalah VosViewer [1]. VosViewer merupakan singkatan dari Visualisations of Similarity atau Visualisasi kemiripan jurnal (VOSviewer - Publications), aplikasi ini telah melalui banyak

penelitian sehingga metodenya cukup teruji. VOSviewer merupakan perangkat lunak untuk membangun dan memvisualisasikan jaringan bibliometrik. Jaringan ini misalnya dapat mencakup jurnal, peneliti, atau publikasi individu dan mereka dapat dibangun berdasarkan kutipan, penggabungan bibliografi, kutipan bersama, atau hubungan penulisan bersama [2].

Terdapat berbagai penelitian yang menggunakan software VOSviewer yang mengkaji pada berbagai topik, seperti e-wallet, teknologi, dan masih banyak lagi. Sebagai contoh pada topik teknologi, salah satu penelitian bertujuan untuk mengetahui pemetaan riset penggunaan teknologi untuk deteksi berita palsu yang bisa dimanfaatkan untuk mengetahui perkembangannya sebagai referensi untuk mengembangkan penelitian lebih lanjut mengenai penggunaan teknologi untuk deteksi berita palsu. Selain itu, terdapat juga penelitian yang membahas topik kemampuan komunikasi matematis yang bertujuan untuk mengkaji dan menganalisis penelitian mengenai kemampuan komunikasi matematis siswa dalam pembelajaran matematika dan self-concept.

Penelitian ini bertujuan untuk (1) mengetahui peta perkembangan publikasi internasional berdasarkan kata kunci di platform Google Scholar menggunakan pendekatan bibliometrik. (2) Mengetahui jumlah citation metrics pada kata kunci “ChatGPT” dari tahun 2022-2023 di platform Google Scholar (3) Mengetahui berbagai kluster dalam pembahasan publikasi ChatGPT menggunakan VOSviewer (4) Mengetahui topik yang banyak dan paling jarang dibahas dalam paper dengan kata kunci “ChatGPT” pada tahun 2022-2023.

II. TINJAUAN PUSTAKA

A. ChatGPT OpenAI

OpenAI adalah perusahaan riset dan penyebaran AI, yang bertujuan memastikan bahwa kecerdasan umum buatan bermanfaat bagi seluruh umat manusia. ChatGPT OpenAI merupakan teknologi mesin berbasis kecerdasan buatan yang dilatih untuk bisa menirukan percakapan manusia menggunakan teknologi NLP (Natural Language Processing) [3]. Format dialog memungkinkan ChatGPT untuk menjawab pertanyaan tindak lanjut, mengakui kesalahan, menantang premis yang salah, dan menolak permintaan yang tidak pantas.

OpenAI adalah perusahaan riset dan penyebaran AI. Misi OpenAI adalah untuk memastikan bahwa kecerdasan umum buatan bermanfaat bagi seluruh umat manusia. OpenAI sedang membangun AGI yang aman dan bermanfaat, tetapi

juga akan menganggap misi kami terpenuhi jika pekerjaan kami membantu orang lain untuk mencapai hasil ini[4][5][6].

B. Bibliometrika

Bibliometrika berasal dari kata biblio atau bibliography dan metrics, biblio berarti buku dan metrics berarti mengukur. Jadi, bibliometrika dapat diartikan sebagai mengukur atau menganalisis buku atau literatur dengan menggunakan pendekatan matematika dan statistika [7]. Dengan pemanfaatan bibliometrika, kita dapat mengetahui bagaimana tingkat produktivitas dan sebaran publikasi ilmiah tersebut dalam bidang keilmuannya. Selain itu, hasil analisis bibliometrika dapat memberikan pemahaman yang lebih luas terhadap keseluruhan dari berbagai disiplin ilmu, mengungkapkan berbagai nama pengarang yang paling produktif dalam suatu disiplin ilmu [8].

C. VOSviewer

VOSviewer merupakan sebuah perangkat lunak yang digunakan untuk membuat dan memvisualisasikan peta jaringan bibliometrik [9]. Perangkat lunak ini digunakan untuk menganalisis bibliometrik dengan mencari berbagai jaringan bibliometrik yang mencakup, seperti jurnal ilmiah, peneliti, organisasi penelitian, negara, kata kunci, dll [10]. Pada software VOSviewer, terdapat tiga tampilan visualisasi berbeda dalam menganalisis bibliometrik, antara lain Network Visualisation, Overlay Visualization, dan Density Visualization. Network Visualisation berfungsi untuk memperlihatkan memperlihatkan jaringan antaristilah yang divisualisasikan. Apabila hubungan antaristilah satu dengan lainnya bercetak tebal dengan bulatan besar, maka hal tersebut menunjukkan adanya hubungan yang cukup banyak dan kuat antaristilah tersebut. Sebaliknya, apabila hubungan antaristilah satu dengan lainnya bercetak tipis dengan bulatan kecil, maka hal tersebut menunjukkan hubungan yang lemah antaristilah tersebut. Overlay Visualization berfungsi untuk menampilkan jejak historis penelitian. Apabila hasil visualisasi semakin gelap, maka hal tersebut menunjukkan penelitian yang telah dilakukan semakin lama dilakukan. Sebaliknya, apabila visualisasinya menunjukkan warna yang semakin terang, maka penelitian tersebut dilakukan pada waktu dekat. Visualisasi terakhir adalah Density Visualization yang berfungsi untuk menampilkan tingkat pembahasan pada kelompok penelitian. Apabila hasil visualisasi semakin terang, maka hal tersebut menunjukkan topik penelitian tersebut semakin sering dilakukan dan apabila pembahasan topik tersebut semakin sedikit, maka hasil visualisasi akan semakin gelap[9].

III. METODOLOGI PENELITIAN

Pencarian paper terkait dilakukan pada aplikasi Publish or Perish dengan menggunakan kata kunci "ChatGPT" dan sumber datanya didapatkan dari Google Scholar dalam jangka waktu pencarian mulai dari tahun 2022 hingga tahun 2023 serta batas jumlah papernya sebanyak 200 paper.

Hasil dari pencarian ini kemudian dilakukan evaluasi kembali terkait detail informasi pada Mendeley. Detail informasi yang dievaluasi kembali meliputi kesesuaian paper dengan topik yang dibahas serta penambahan manual untuk author keywords pada setiap paper. Setelah itu data yang telah dievaluasi, kemudian dilakukan analisis menggunakan VosViewer. Analisis yang dilakukan, yaitu analisis bibliometrik dengan melakukan analisis hubungan bibliometrik berdasarkan kata kunci atau biasa disebut co-occurrence.

Pada software VosViewer terdapat tiga kategori visualisasi berbeda, yaitu network visualization, overlay visualization, dan density visualization. Network visualization bertujuan untuk memperlihatkan gambaran dari hubungan antar topik berdasarkan judul artikel dan kata kunci dari jurnal. Overlay visualization bertujuan untuk melihat hubungan antar topik dengan keterangan tahun. Density visualization berfungsi untuk melihat topik pembahasan mana yang paling banyak atau paling sedikit dibahas dan ditampilkan dengan heatmap. Semakin banyak pembahasan topik tersebut, maka warnanya akan menjadi merah dan apabila pembahasan topik tersebut sedikit, maka akan semakin biru.

IV. HASIL DAN PEMBAHASAN

A. Hasil Pencarian Publish or Perish

Berdasarkan hasil penelusuran dengan kata kunci "ChatGPT" pada platform Google Scholar dengan membatasi 200 paper pada tahun 2022-2023, menggunakan Publish or Perish didapatkan citation metrics seperti pada Gbr 1.

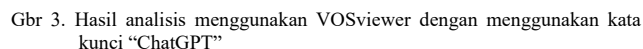
Citation metrics		Help
Publication years:	2022-2023	
Citation years:	1 (2022-2023)	
Papers:	200	
Citations:	656	
Cites/year:	656.00	
Cites/paper:	3.28	
Authors/paper:	2.89	
h-index:	13	
g-index:	19	
hI,norm:	8	
hI,annual:	8.00	
hA-index:	13	
Papers with ACC >= 1,2,5,10,20:	109,72,38,20,8	

Gbr 1. Hasil citation metrics dengan jumlah paper 200

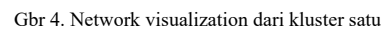
Setelah dilakukan pengecekan ulang menggunakan software Mendeley didapatkan 127 paper yang bisa digunakan, hal itu dilakukan dengan cara melihat jenis paper dan keyword yang ada. Dengan menggunakan Publish or Perish didapatkan citation metrics seperti pada Gbr 2.

Gbr 2. Hasil citation metrics setelah pengecekan

Berdasarkan dari hasil pencarian yang menggunakan kata kunci “ChatGPT” pada VOSviewer yang terindeks pada database Google Scholar tahun 2022-2023 menghasilkan sebelas kluster berbeda, sebelas kluster itu ditandai dengan kode warna merah, hijau, biru tua, kuning, ungu, biru muda, jingga, coklat, merah muda, peach, dan hijau muda. Setiap warna mempunyai item yang menjelaskan tentang keyword yang digunakan didalam paper yang menjadi sumber penelitian. Hal ini memudahkan pembaca untuk mencari informasi yang spesifik dan relevan dengan topik yang sedang mereka teliti, seperti pada Gbr 3

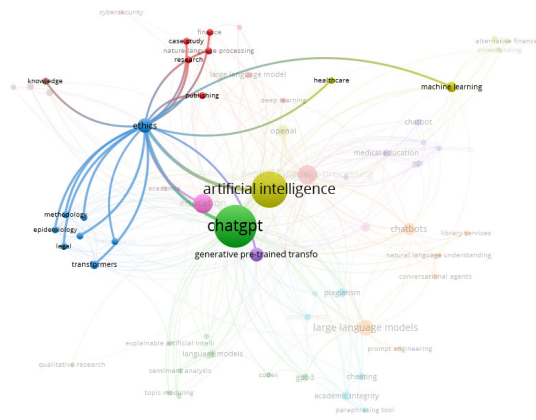


natural language processing, nature language processing, publishing, dan research, seperti pada Gbr 4.



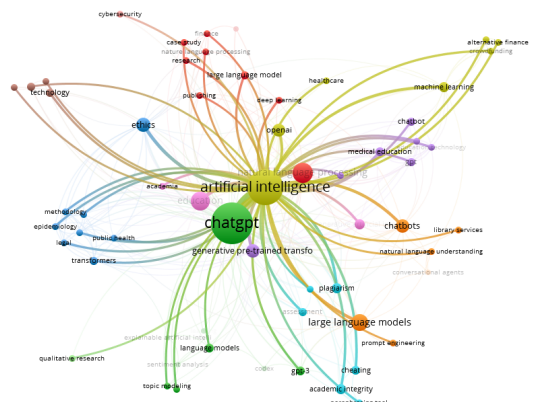
Gbr 5. Network visualization yang didapat dari kluster dua

2) Kluster tiga ditandai dengan warna biru tua dan terdiri dari delapan item, yaitu epidemiology, ethics, legal, methodology, public health, scientific research, strobe dan transformers, seperti pada Gbr 6.



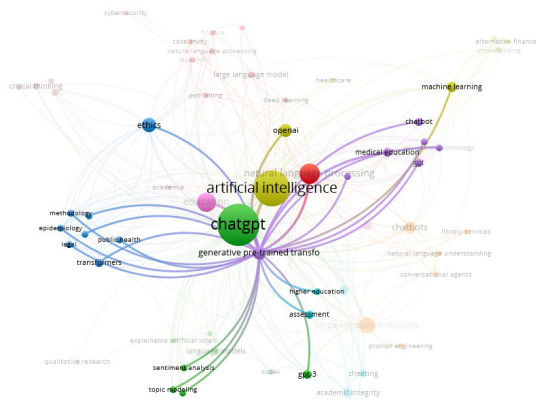
Gbr 6. Network visualization yang didapat dari cluster tiga

3) Kluster empat ditandai dengan warna kuning dan terdiri dari tujuh item, yaitu alternative finance, artificial intelligence, community finance, crowdfunding, healthcare, machine learning, dan openai, seperti pada Gbr 7.



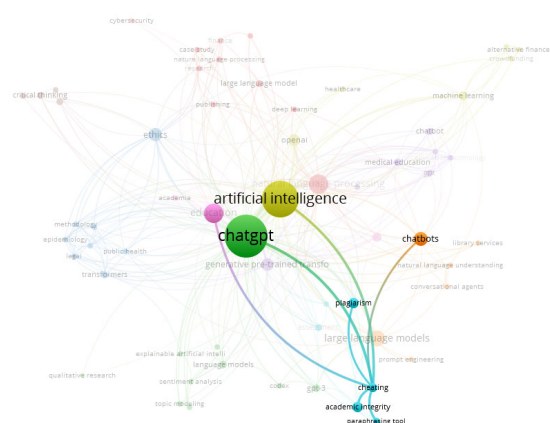
Gbr 7. Network visualization yang didapat dari cluster empat

4) Kluster lima ditandai dengan warna ungu dan terdiri dari tujuh item, yaitu chatbot, education technology, generative pre-trained transformers, gpt, medical education, medqa, dan ophthalmology, seperti pada Gbr 8.



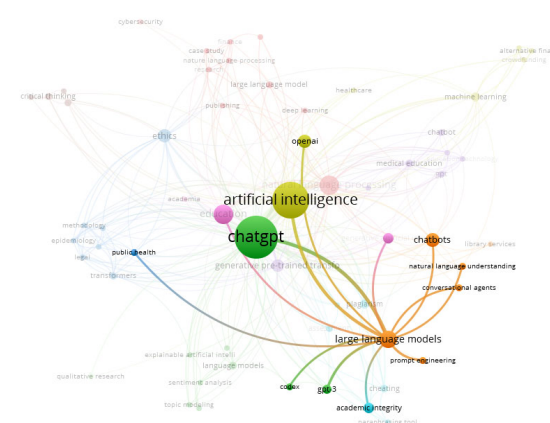
Gbr 8. Network visualization yang didapat dari cluster lima

5) Kluster enam ditandai dengan warna biru muda dan terdiri dari enam item, yaitu academic integrity, assessment, cheating, higher education, paraphrasing tool, dan plagiarism, seperti pada Gbr 9.



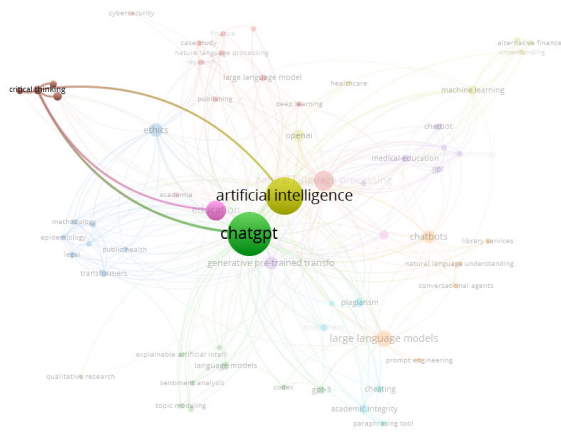
Gbr 9. Network visualization yang didapat dari cluster enam

6) Kluster tujuh ditandai dengan warna jingga dan terdiri dari enam item, yaitu chatbots, conversational agents, large language models, library services, crowdfunding, natural language understanding, dan prompt engine, seperti pada Gbr 10.



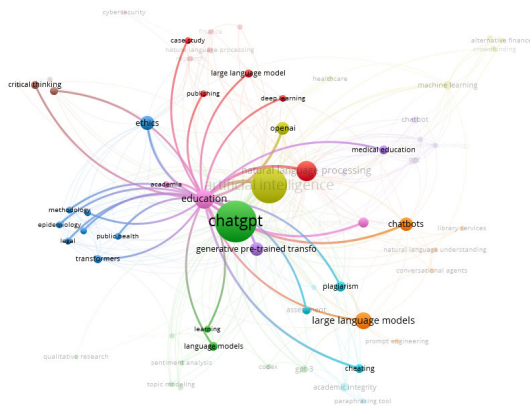
Gbr 10. Network visualization yang didapat dari cluster tujuh

7) Kluster delapan ditandai dengan warna coklat dan terdiri dari empat item, yaitu covid-19, critical thinking, knowledge, dan technology, seperti pada Gbr 11.



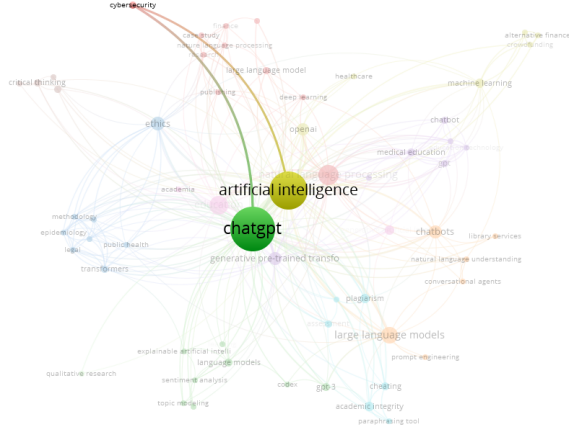
Gbr 11. Network visualization yang didapat dari cluster delapan

8) Kluster sembilan ditandai dengan warna merah muda dan terdiri dari tiga item, yaitu academia, education, knowledge, dan generative artificial intelligence, seperti pada Gbr 12.



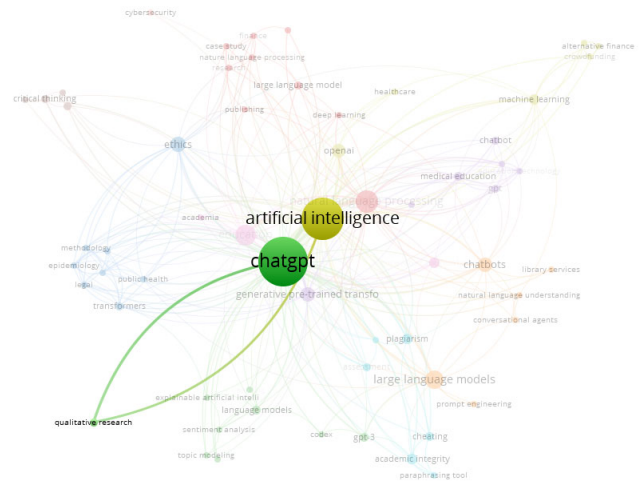
Gbr 12. Network visualization yang didapat dari cluster sembilan

9) Kluster sepuluh ditandai dengan warna peach dan terdiri dari satu item, yaitu cybersecurity, seperti pada Gbr 13.



Gbr 13. Network visualization yang didapat dari cluster sepuluh

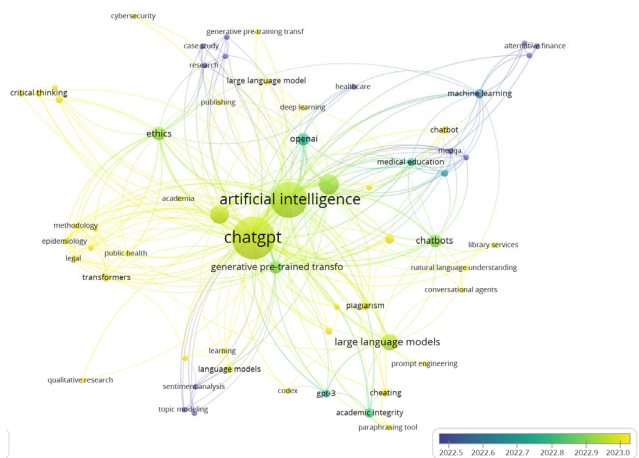
10) Kluster sebelas ditandai dengan warna hijau muda dan terdiri dari satu item, yaitu qualitative research, seperti pada Gbr 14.



Gbr 14. Network visualization yang didapat dari cluster sebelas

C. 4.3. Overlay Visualization

Pada tipe visualisasi ini, kita bisa melihat pada tahun dimana penelitian ini dilakukan dengan rentang yang digunakan antara pertengahan tahun 2022 sampai dengan tahun 2023. Dalam hal ini, dapat kita lihat, bahwa topik community finance, alternative finance, dan crowdfunding diteliti pada awal tahun 2022. Kemudian, untuk topik artificial intelligence, chatgpt, education, dan natural language processing diteliti pada akhir tahun 2022. Selain dari beberapa topik tersebut diteliti pada awal tahun 2023, seperti epidemiology, public health, plagiarism, critical thinking, dll, seperti pada Gbr 15.



Gbr 15. Overlay visualization yang didapat dari kata kunci "ChatGPT"

D. Density Visualization

Pada tipe visualisasi ini, kita bisa melihat seberapa jarang dan seberapa sering keyword “ChatGPT” dibahas. Tipe ini menampilkan tingkat seringnya pembahasan suatu topik dengan heatmap, semakin sering pembahasan topik tersebut, maka warnanya akan menjadi merah dan apabila pembahasan topik tersebut jarang, maka akan semakin biru. Dari hasil visualisasi tersebut, dapat kita lihat, bahwa tiga topik yang paling sering dibahas adalah ChatGPT, artificial intelligence, dan natural language processing. Sementara itu, untuk tiga topik yang paling jarang dibahas adalah qualitative research, cybersecurity, dan library services, seperti pada Gbr 16.



Gbr 16. Density visualization yang didapat dari kata kunci “ChatGPT”.

V. KESIMPULAN

Berdasarkan dari hasil dan pembahasan di atas dapat disimpulkan, bahwa pertumbuhan publikasi mengenai “ChatGPT” selama periode tahun 2022-2023 mengalami kenaikan. Selain itu, hasil yang didapatkan pada VOSviewer untuk Network Visualization menunjukkan, bahwa hubungan antartopik dikelompokkan menjadi tujuh kluster dan ditandai dengan berbagai warna, mulai dari warna merah, hijau, biru tua, kuning, ungu, biru muda, jingga, coklat, merah muda,

peach, dan hijau muda. Untuk Overlay Visualization menunjukkan, bahwa topik dengan keyword “ChatGPT” ini mulai diteliti pada tahun 2022 sampai tahun 2023, contohnya seperti topik finance yang mulai diteliti pada awal tahun 2022, topik chatgpt diteliti pada akhir tahun 2022, dan topik critical thinking diteliti pada awal tahun 2023. Adapun Density Visualization menunjukkan, bahwa tiga topik yang paling sering dibahas adalah ChatGPT, artificial intelligence, dan natural language processing. Sementara itu, untuk tiga topik yang paling jarang dibahas adalah qualitative research, cybersecurity, dan library services.

REFERENSI

- [1] van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>.
- [2] Effendy, F., Gaffar, V., Hurriyati, R. and Hendrayati, H., 2021. Analisis Bibliometrik Perkembangan Penelitian Penggunaan Pembayaran Seluler Dengan Vosviewer. *Jurnal Interkom: Jurnal Publikasi Ilmiah Bidang Teknologi Informasi Dan Komunikasi*, 16(1), pp.10-17.
- [3] Aydin, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT Generated Literature Review : Digital Twin in Healthcare. In *Emerging Computer Technologies 2* (Vol. 2). İzmir Akademi Derneği.
- [4] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8), 9.
- [5] Setiawan, A., & Luthfiyanti, U. (2023). Penggunaan ChatGPT Untuk Pendidikan di Era Education 4.0: Usulan Inovasi Meningkatkan Keterampilan Menulis. *JURNAL PETISI (Pendidikan Teknologi Informasi)*, 4(1), 49-58. <https://doi.org/10.36232/jurnalpetisi.v4i1.3680>
- [6] Kirmani, A. R. (2022). Artificial intelligence-enabled science poetry. *ACS Energy Letters*, 8, 574-576.
- [7] Diodato, V. (1994). *Dictionary of bibliometrics*. New York: The Haworth Press.
- [8] Nuryudi. (2016). Analisis Bibliometrika Islam: Studi Kasus Dokumentasi Publikasi Ilmiah di UIN Syarif Hidayatullah Jakarta. *Al-Maktabah: Jurnal Komunikasi dan Informasi Perpustakaan*, 15, 1, 41-55.
- [9] S. H. H. Shah, S. Lei, M. Ali, D. Doronin, and S. T. Hussain, “Presumption: bibliometric analysis using HistCite and VOSviewer,” *Kybernetes*, 2019, doi: 10.1108/K12-2018-0696.
- [10] Zakiyyah, F. N., Winoto, Y., & Rohanda, R. (2022). Pemetaan bibliometrik terhadap perkembangan penelitian arsitektur informasi pada Google Scholar menggunakan VOSviewer. *Information: Journal of Library and Information Science*, 2(1), 43-60.