

BAB I PENDAHULUAN

1.1. Latar Belakang

Perkembangan ilmu pengetahuan dan alat-alat modern saat ini telah mendorong lahirnya banyak penelitian lintas bidang, terutama dalam memahami berbagai fenomena keagamaan dalam Islam. Kajian-kajian ini mencakup bidang seperti hukum Islam, pemanfaatan teknologi berbasis syariah, otoritas keagamaan, hubungan sosial umat, dan berbagai aspek lain yang terus berkembang dalam masyarakat global. Agama tetap memainkan peran signifikan dalam membentuk nilai, perilaku, dan struktur sosial di berbagai belahan dunia. Agama Islam, sebagai salah satu agama terbesar dengan lebih dari 2 miliar pengikut [1], menunjukkan perkembangan dinamis dalam kehidupan modern. Tren pemikiran dalam publikasi ilmiah terkait Islam telah meningkat, dengan fokus pada isu-isu yang relevan seperti sosial, ekonomi, politik, pendidikan, keuangan, hukum Islam, psikologi, dan teknologi, terutama selama pandemi COVID-19. Penelitian-penelitian ini memberikan temuan penting yang dapat membantu memahami dampak agama dalam konteks kehidupan masyarakat.

Namun, seiring dengan meningkatnya jumlah publikasi artikel dan penelitian terkait Islam setiap tahun, muncul permasalahan baru yang dikenal sebagai *Information Overload*. Fenomena ini terjadi karena banyaknya publikasi yang bertumpuk, sehingga pembaca atau peneliti kesulitan dalam memfilter informasi yang relevan. Akibatnya, para akademisi dan peneliti sering kali kesulitan memperoleh data terkini serta mengidentifikasi topik-topik inovatif yang dapat mendukung penelitian mereka. Meskipun kemajuan signifikan telah dicapai, masih banyak area yang belum terjangkau dan memerlukan penelitian lebih lanjut. Di tengah meningkatnya publikasi, platform-platform terpercaya seperti SINTA dan GARUDA memfasilitasi penyebaran penelitian di bidang Islam. **Namun**, belum ada penelitian yang mengangkat topik analisis tren topik penelitian

Islam dari tahun ke tahun, khususnya di platform-platform Indonesia yang menjadi rujukan utama bagi peneliti di tanah air.

Seiring dengan perkembangan publikasi yang semakin pesat, data yang diambil dari situs SINTA yang terhubung dengan GARUDA menunjukkan tren yang jelas terkait publikasi artikel dan penelitian di bidang Islam. Jumlah publikasi dalam bidang ini menunjukkan tren kenaikan yang cukup menonjol dalam beberapa tahun terakhir. Pada 2019, terdapat 2.808 publikasi; angka ini naik menjadi 3.060 pada 2020 (peningkatan 8,9%), kemudian mencapai 3.600 pada 2021 (kenaikan 17,6%). Tahun 2022 mencatat 4.103 publikasi (bertambah 13,9%). Namun, pada 2023, terjadi penurunan cukup tajam dengan hanya 3.289 publikasi (penurunan 19,8%). Tren ini menunjukkan bahwa publikasi di bidang Islam cenderung meningkat setiap tahun [2], [3]. Adapun data tahun 2024 belum dapat dianalisis dalam penelitian ini karena saat pengumpulan data dilakukan (tahun 2024), data publikasi tahun tersebut belum tersedia secara lengkap dan belum terverifikasi sepenuhnya di platform SINTA maupun GARUDA.

Tren pesat dalam penelitian terkait topik Islam dalam beberapa tahun terakhir telah memicu penerapan berbagai metode pemodelan topik, terutama yang berbasis NLP (*Natural Language Processing*). Model klasik seperti TF-IDF, BoW, dan LSA/LSI telah banyak digunakan, namun memiliki keterbatasan dalam menangani kompleksitas data teks. Menurut Church dan Kordoni (2022), pendekatan *State-of-the-Art* (SOTA) menawarkan solusi yang lebih canggih dalam pemodelan topik, meskipun mereka juga mengingatkan akan potensi risiko dari mengejar SOTA secara berlebihan, seperti kurangnya wawasan mendalam dan tantangan dalam replikasi hasil penelitian [4]. Dalam konteks penelitian Islam, pendekatan SOTA seperti LDA dan BERTopic penting untuk mengungkap tren dan pola tersembunyi yang tidak terdeteksi oleh metode tradisional.

Namun demikian, hasil tinjauan tersebut juga mengungkapkan beberapa keterbatasan: (i) Belum adanya analisis komprehensif yang

memetakan tren topik penelitian Islam dari tahun ke tahun, khususnya menganalisis tren topik penelitian Islam di Indonesia, (ii) Mayoritas penelitian berfokus pada analisis teks Arab atau Al-Qur'an, namun belum banyak yang mengkaji abstrak publikasi ilmiah terkait Islam secara luas, (iii) Penelitian sebelumnya umumnya hanya menggunakan satu metode pemodelan topik, baik LDA atau BERTopic, namun belum ada yang membandingkan kinerja kedua metode tersebut dalam konteks analisis tren pemikiran Islam.

Menyadari hal ini, dan dengan mempertimbangkan pentingnya pendekatan State-of-the-Art (SOTA) dalam memberikan solusi yang lebih efektif untuk pemetaan tren, penelitian ini bertujuan untuk mengisi kekosongan kajian yang ada dengan mengkaji perkembangan pemikiran Islam melalui pemetaan isi abstrak dari publikasi ilmiah yang terbit antara tahun 2019 hingga 2023. Data yang digunakan diperoleh melalui teknik scraping dari situs SINTA yang terhubung dengan portal GARUDA. Dengan demikian, penelitian ini diharapkan mampu memberikan sumbangsih dalam mengevaluasi kinerja dua pendekatan pemodelan topik, yakni LDA dan BERTopic, khususnya dalam aspek visualisasi tren topik, skor *coherence*, serta keterkaitan hasil interpretasi dengan konteks studi keislaman. Temuan ini diharapkan dapat memperkaya pemahaman mengenai dinamika dan isu-isu sentral dalam penelitian Islam.

1.2. Rumusan Masalah

- Bagaimana pemetaan topik penelitian agama Islam dengan menggunakan algoritma BERTopic dan LDA selama periode 2019-2023?
- Apakah ada perbedaan yang signifikan antara nilai *coherence* dari LDA dan BERTopic dalam menemukan tren topik penelitian pada publikasi ilmiah tentang Islam?

1.3. Tujuan Penelitian

- Menganalisis tren topik dengan menggunakan algoritma BERTopic dan LDA dalam penelitian di bidang agama Islam selama periode 2019-2023.
- Menganalisis dan membandingkan nilai *coherence* antara metode LDA dan BERTopic dalam menemukan tren topik penelitian pada publikasi ilmiah tentang Islam.

1.4. Batasan Penelitian

- Menggunakan data abstrak publikasi ilmiah dari tahun 2019 hingga 2023, khusus pada bidang agama Islam.
- Data diambil dari platform publikasi ilmiah Indonesia yaitu SINTA dan GARUDA dengan cara teknik scraping.
- Menggunakan dua algoritma pemodelan: BERTopic dan LDA.
- Dataset terdiri dari 16.860 abstrak dari 179 jurnal terakreditasi SINTA.

