

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan pesat teknologi digital telah mentransformasi pola komunikasi masyarakat di Indonesia. Media sosial seperti X (Twitter), Facebook, Tiktok dan Instagram tidak lagi sekadar menjadi sarana interaksi, tetapi telah menjadi ruang terbuka bagi masyarakat untuk menyebarkan opini, informasi, dan ekspresi personal secara cepat dan luas (Panjaitan & Patria, 2024; Fitriani, 2024). Namun, kebebasan dan ekspansi ruang digital ini juga diiringi dengan meningkatnya penyalahgunaan bahasa melalui ujaran kebencian dan pencemaran nama baik (defamasi), yang banyak dilakukan secara anonim (Ndatyapo et al., 2024; Aini, 2024; Khotimah et al., 2025).

Fenomena tersebut bukan sekadar asumsi teoretis, melainkan telah teridentifikasi secara empiris dalam sejumlah laporan pemantauan media digital di Indonesia yang menunjukkan maraknya penyalahgunaan media sosial untuk menyebarkan tindak defamasi, khususnya dalam konteks politik. Salah satunya adalah riset yang dilakukan oleh Aliansi Jurnalis Independen (AJI) bekerja sama dengan Monash Data & Democracy Research Hub menemukan bahwa dari sekitar 1,5 juta teks yang dipantau di platform Facebook, X (Twitter), dan Instagram selama periode September 2023 hingga Maret 2024, terdapat 200.000 unggahan (sekitar 13,82%) yang mengandung ujaran kebencian, dengan bentuk ujaran paling dominan (sekitar 65,5%) berupa serangan terhadap identitas (131.000 teks) dan sekitar 5,34% di antaranya berkaitan langsung dengan isu Pemilu 2024 (Idris et al., 2024).

Tingginya intensitas ujaran kebencian dalam konteks politik tersebut menunjukkan bahwa sebagian besar pengguna media sosial belum memiliki kemampuan literasi digital dan kesadaran linguistik yang memadai untuk menggunakan bahasa secara etis dan reflektif. Studi “Model Literasi Digital untuk Melawan Ujaran Kebencian di Media Sosial” menemukan bahwa rendahnya literasi digital di Indonesia menjadi faktor penyumbang utama munculnya ujaran kebencian, dan bahwa literasi digital efektif harus mencakup nilai moral seperti

perilaku etis dan motivasi sosial (Muannas & Mansyur, 2020).

Kondisi ini sejalan dengan pandangan Boeriswati yang menekankan pentingnya kesantunan berbahasa (Wildan et al., 2022; Ningsih et al., 2020, Ningsih et al., 2019; Anwar et al., 2021), kecerdasan berbahasa (Kusmana et al., 2018), pengembangan literasi kritis (Boeriswati, 2012), dan kesadaran linguistik (Darwin et al., 2023) dalam memahami serta memproduksi teks digital. Pemahaman terhadap struktur dan makna bahasa di ranah digital dengan demikian menjadi krusial bagi masyarakat dalam menghadapi informasi yang manipulatif, provokatif, maupun bersifat ofensif (Antoro et al., 2021).

Selain faktor rendahnya literasi digital, permasalahan lain yang turut meningkatkan frekuensi tindak defamasi di media sosial adalah anonimitas pengguna yang memungkinkan individu bertindak tanpa akuntabilitas. Anonimitas memungkinkan individu mengekspresikan pendapat ekstrem, ofensif, atau bahkan mencemarkan nama baik tanpa takut menanggung risiko sosial maupun hukum (Kim et al., 2023; Wang et al., 2024). Dalam konteks ini, identitas yang disamarkan atau sepenuhnya fiktif menciptakan apa yang disebut *online disinhibition effect*, yaitu kecenderungan pengguna untuk melanggar norma kesopanan dan etika komunikasi ketika merasa tidak dapat dilacak (Perera et al., 2023; Kim et al., 2023).

Dampak dari anonimitas tersebut terlihat nyata pada meningkatnya kasus ujaran kebencian dan pencemaran nama baik yang sulit ditangani aparat penegak hukum karena pelaku tidak dapat diidentifikasi dengan cepat. Studi terbaru menunjukkan bahwa investigasi kasus kejahatan daring yang melibatkan akun anonim umumnya memerlukan waktu lebih lama dan sering berakhir tanpa putusan hukum yang tegas (Rocha et al., 2017). Ketidakjelasan identitas pelaku juga menyebabkan lemahnya efek jera karena publik melihat bahwa tindakan serupa dapat dilakukan tanpa konsekuensi langsung (Kim et al., 2023; Mukred et al., 2024; Jaidka et al., 2022; Muannas & Mansyur, 2020). Selain itu, berbagai platform media sosial cenderung memiliki kebijakan privasi yang melindungi pengguna anonim, yang meskipun penting untuk kebebasan berekspresi, justru menyulitkan pelacakan pelaku defamasi secara forensik (Huang et al., 2025).

Fenomena ini menunjukkan bahwa anonimitas bukan sekadar persoalan teknis atau legal, tetapi juga merupakan masalah sosiolinguistik yang kompleks

karena hal tersebut memengaruhi cara individu membentuk identitas, menyesuaikan gaya bahasa, dan mengatur relasi kuasa dalam komunikasi daring (Kumar & Shah, 2018; D'Anna et al., 2023). Dalam konteks Indonesia, kasus akun anonim seperti “Fufufafa” di media sosial Kaskus menjadi contoh konkret bagaimana anonimitas mempersulit identifikasi pelaku ujaran defamasi meskipun terdapat jejak linguistik yang khas. Kasus ini menegaskan bahwa dalam menghadapi persoalan anonimitas digital, diperlukan pendekatan linguistik forensik yang mampu menelusuri ciri-ciri idiolektal penulis anonim melalui analisis stilometrik dalam bentuk atribusi kepengarangan (*authorship attribution*) (Juola, 2021).

Dalam konteks linguistik forensik, atribusi kepengarangan dipandang sebagai sarana ilmiah untuk mengidentifikasi individu penulis melalui analisis terhadap ciri-ciri kebahasaan yang bersifat konsisten dan khas, yang disebut sebagai idiolek, yang berangkat dari prinsip dasar bahwa setiap individu memiliki gaya linguistik unik yang secara tidak disadari tercermin dalam pilihan leksikal, sintaksis, dan struktur teks yang mereka hasilkan (McMenamin et al., 2002; Coulthard, 2004; Juola, 2006; Stamatatos, 2009; Grant, 2010; Fobbe, 2020). Dalam konteks digital forensik modern, atribusi kepengarangan tidak hanya berfungsi sebagai alat bantu linguistik, tetapi juga sebagai instrumen pendukung investigasi hukum untuk membuktikan keterhubungan antara teks anonim dan individu tertentu secara objektif dan dapat diuji ulang (Rocha et al., 2017; He et al., 2024).

Perkembangan atribusi kepengarangan menunjukkan adanya pergeseran paradigma dari pendekatan linguistik deskriptif menuju pendekatan linguistik komputasional yang semakin terukur dan dapat direplikasi. Pada tahap awal perkembangannya, identifikasi penulis dilakukan secara manual melalui analisis kualitatif terhadap karakteristik kebahasaan yang muncul dalam suatu teks. Seorang ahli bahasa membandingkan pilihan kosakata, pola sintaksis, struktur wacana, kebiasaan ejaan, penggunaan tanda baca, hingga ekspresi pragmatik yang dianggap mencerminkan idiolek penulis. Pendekatan tersebut didasarkan pada asumsi bahwa setiap individu memiliki kebiasaan berbahasa yang relatif konsisten sehingga dapat digunakan sebagai dasar untuk membedakan satu penulis dengan penulis lainnya (Coulthard, 2004; Grant, 2010; McMenamin et al., 2002). Dalam praktik linguistik

forensik, analisis semacam ini telah lama dimanfaatkan sebagai salah satu bentuk pendapat ahli untuk membantu penyelidikan maupun pembuktian perkara yang melibatkan dokumen anonim, surat ancaman, pemerasan, plagiarisme, hingga tindak pencemaran nama baik.

Meskipun memberikan kontribusi penting terhadap perkembangan linguistik forensik, pendekatan manual memiliki sejumlah keterbatasan. Proses identifikasi sangat bergantung pada pengalaman dan intuisi analis sehingga hasil analisis berpotensi berbeda ketika dilakukan oleh ahli yang berbeda. Selain itu, pendekatan tersebut memerlukan waktu yang relatif lama, sulit diterapkan pada korpus digital yang berukuran besar, serta menghadapi tantangan dalam menjaga konsistensi dan keterulangan hasil analisis. Dalam konteks pembuktian hukum modern, kondisi tersebut menjadi perhatian karena bukti ilmiah dituntut tidak hanya akurat, tetapi juga transparan, dapat diuji ulang, dan menghasilkan kesimpulan yang konsisten apabila dianalisis menggunakan prosedur yang sama (Chaski, 2001; Grant, 2022; Juola, 2006).

Perkembangan teknologi komputasi kemudian mendorong lahirnya pendekatan stilometri yang memanfaatkan teknik statistik untuk mengukur kemiripan gaya bahasa secara lebih objektif. Berbeda dengan pendekatan manual yang bertumpu pada interpretasi ahli, stilometri menghitung distribusi fitur-fitur linguistik, seperti function words, karakter n-gram, kata n-gram, panjang kalimat, kekayaan kosakata, serta berbagai indikator linguistik lainnya yang terbukti relatif stabil sebagai representasi idiolek penulis (Stamatatos, 2009; Eder et al., 2016; Juola, 2006). Integrasi linguistik dengan metode komputasional tersebut memungkinkan analisis dilakukan terhadap ribuan dokumen secara cepat, konsisten, dan memiliki tingkat replikasi yang lebih tinggi dibandingkan analisis manual.

Transformasi tersebut menunjukkan bahwa fokus penelitian atribusi kepengarangan tidak lagi hanya diarahkan pada kemampuan mengidentifikasi penulis secara akurat, tetapi juga pada bagaimana memastikan bahwa hasil identifikasi tersebut memiliki kualitas ilmiah yang dapat dipertanggungjawabkan. Dalam konteks linguistik forensik, keluaran analisis stilometri tidak dapat diterima semata-mata karena menghasilkan tingkat akurasi yang tinggi, melainkan juga

harus memenuhi prinsip validitas, reliabilitas, transparansi metodologis, dan interpretabilitas agar dapat digunakan sebagai alat bantu pembuktian di pengadilan (Chaski, 2013; Grant, 2022; Ishihara et al., 2024). Dengan demikian, perkembangan atribusi kepengarangan pada era komputasional tidak hanya menimbulkan kebutuhan akan algoritme yang semakin canggih, tetapi juga memunculkan kebutuhan baru berupa kerangka penilaian yang mampu mengevaluasi kualitas hasil analisis secara sistematis dan memenuhi standar ilmiah maupun forensik. Kebutuhan inilah yang menjadi landasan utama bagi penelitian ini untuk mengembangkan desain penilaian atribusi kepengarangan berbasis Stylo R Package, sehingga hasil analisis tidak hanya akurat secara komputasional, tetapi juga memiliki kredibilitas sebagai bukti linguistik dalam konteks forensik.

Keberhasilan atribusi kepengarangan pada dasarnya tidak hanya ditentukan oleh kecanggihan algoritma komputasional, tetapi juga oleh ketepatan dalam memilih dan menginterpretasikan ciri-ciri linguistik yang merepresentasikan identitas penulis. Dalam linguistik forensik, ciri linguistik (*linguistic features*) dipahami sebagai unsur-unsur kebahasaan yang muncul secara relatif konsisten dalam produksi bahasa seseorang sehingga dapat berfungsi sebagai representasi idiolek. Oleh karena itu, analisis atribusi kepengarangan tidak bertujuan mencari kesamaan isi atau topik pembicaraan, melainkan mengidentifikasi pola penggunaan bahasa yang cenderung muncul secara tidak disadari dan relatif stabil pada diri seorang penulis (Grant, 2010; Juola, 2006; Stamatatos, 2009)

Secara umum, ciri linguistik yang digunakan dalam atribusi kepengarangan dapat dikelompokkan ke dalam beberapa tingkatan analisis bahasa. Pada tingkat leksikal, identifikasi dilakukan terhadap pilihan kosakata, frekuensi penggunaan kata fungsi (*function words*), kekayaan kosakata (*type-token ratio*), serta kecenderungan penggunaan kata-kata tertentu. Pada tingkat ortografis, analisis mencakup kebiasaan penggunaan huruf kapital, tanda baca, singkatan, emotikon, maupun variasi ejaan yang sering kali muncul secara konsisten pada penulis tertentu. Pada tingkat morfologis dan sintaktis, perhatian diarahkan pada pola pembentukan kata, struktur frasa, panjang kalimat, kompleksitas sintaksis, serta kecenderungan penggunaan konstruksi gramatikal tertentu. Sementara itu, pada tingkat stilistika dan pragmatik, analisis dapat mencakup gaya bertutur, strategi

penekanan, pilihan ekspresi evaluatif, pola kesantunan, penggunaan metafora, maupun kecenderungan retorika yang membentuk karakteristik individual penulis (Coulthard & Johnson, 2010; McMenamin et al., 2002; Grant, 2022)

Dalam praktik analisis komputasional, berbagai karakteristik linguistik tersebut kemudian ditransformasikan menjadi fitur-fitur kuantitatif yang dapat dihitung secara statistik. Sebagai contoh, Stylo R Package mengekstraksi representasi linguistik melalui fitur seperti function words, character n-grams, word n-grams, distribusi frekuensi kata, ukuran kekayaan kosakata, dan parameter stilometrik lainnya yang terbukti efektif dalam membedakan gaya bahasa antarpengarang (Eder et al., 2016). Dengan demikian, proses komputasi pada dasarnya bukan menggantikan analisis linguistik, melainkan mengoperasionalkan karakteristik kebahasaan ke dalam bentuk data numerik sehingga pola idiolektal dapat dibandingkan secara objektif dan dapat direplikasi.

Meskipun demikian, hasil ekstraksi fitur linguistik tersebut tidak dapat langsung ditafsirkan sebagai bukti kepengarangan. Setiap indikator linguistik perlu dievaluasi berdasarkan kualitas data, konsistensi kemunculan antar-teks, stabilitas hasil analisis, signifikansi linguistik, reliabilitas replikasi, serta relevansinya terhadap konteks forensik. Dengan kata lain, indikator linguistik tidak hanya berfungsi sebagai masukan (input) bagi proses komputasional, tetapi juga menjadi dasar interpretasi ilmiah dalam menarik kesimpulan mengenai kemungkinan kepengarangan suatu teks. Oleh sebab itu, penelitian ini memandang bahwa pengembangan desain penilaian atribusi kepengarangan perlu dibangun di atas indikator-indikator linguistik yang jelas, terukur, dan memiliki landasan teoretis yang kuat sehingga hasil analisis tidak hanya akurat secara statistik, tetapi juga dapat dipertanggungjawabkan sebagai bukti linguistik dalam praktik forensik.

Dalam konteks digital, penelitian tentang atribusi kepengarangan telah mengalami perkembangan signifikan dalam dua dekade terakhir, terutama melalui integrasi antara linguistik forensik dan komputasi teks (Rayhan et al., 2025a). Kajian awal berfokus pada fitur-fitur linguistik seperti leksikon, morfologi, dan sintaksis yang mencerminkan idiolekt penulis (Juola, 2006; Stamatatos, 2009; Grant, 2010), sedangkan pendekatan modern menggabungkan metode statistik dan machine learning untuk meningkatkan objektivitas dan akurasi identifikasi (Eder et

al., 2016; Savoy, 2020). Berbagai studi mutakhir menegaskan efektivitas algoritma seperti Burrows's Delta, nearest shrunken centroid (NSC), support vector machine (SVM), dan k-nearest neighbor (k-NN) dalam mengenali gaya linguistik individu, bahkan pada teks pendek dan informal seperti unggahan media sosial (Rocha et al., 2017; Sarwar et al., 2024; Misini et al., 2024). Tren terbaru juga menunjukkan pemanfaatan deep learning dan transformer-based language models untuk mendeteksi gaya penulisan penulis manusia maupun teks yang dihasilkan AI (Huang et al., 2025; Huertas-Tato et al., 2023; Berriche & Larabi-Marie-Sainte, 2024). Selain aspek teknis, beberapa penelitian menyoroti potensi penerapan atribusi kepengarangan dalam ranah forensik digital untuk membantu identifikasi pelaku ujaran kebencian, disinformasi, dan kejahatan berbasis teks anonim di platform daring (Grant, 2010; Cafiero & Camps, 2023; Juola, 2021).

Meskipun berbagai penelitian menunjukkan kemajuan signifikan dalam pengembangan metode atribusi kepengarangan, sebagian besar studi tersebut masih berfokus pada korpus berbahasa Inggris dan konteks akademik atau eksperimen terkontrol (Rocha et al., 2017; Kumarage & Liu, 2023; Savoy, 2020). Belum banyak penelitian yang secara khusus menyoroti penerapan teknik ini dalam konteks hukum digital di Indonesia, khususnya terhadap teks anonim singkat berbahasa Indonesia yang beredar di media sosial dan kerap mengandung unsur defamasi.

Selain itu, desain penilaian atribusi kepengarangan pada teks anonim di media sosial sangat dibutuhkan untuk memastikan bahwa hasil analisis stilometri yang dihasilkan bersifat objektif, reliabel, dan dapat dipertanggungjawabkan secara hukum. Dalam konteks forensik digital, penilaian semacam ini berfungsi sebagai jembatan antara pendekatan linguistik dan metode komputasional, sehingga bukti kebahasaan dapat memenuhi standar ilmiah dan forensik seperti kriteria Daubert yang menuntut validitas, replikasi, serta tingkat kesalahan yang diketahui (Chaski, 1998; Ishihara et al., 2024). Dalam konteks tersebut, Stylo R Package menawarkan platform analisis yang transparan dan dapat direplikasi, sehingga mendukung prinsip reliabilitas metodologis yang menjadi tuntutan utama dalam linguistik forensik modern (Eder et al., 2016).

Dengan demikian, desain penilaian atribusi kepengarangan ini tidak hanya menghasilkan laporan komputasional, tetapi juga memberikan kerangka

interpretatif yang memungkinkan ahli bahasa forensik, peneliti, atau aparat hukum menilai derajat keabsahan hasil secara sistematis. Pendekatan ini diharapkan dapat menjembatani kesenjangan antara analisis stilometrik berbasis komputasi dan praktik pembuktian hukum yang menuntut transparansi metodologis, objektivitas, dan akuntabilitas ilmiah.

1.2 Fokus Penelitian

Berdasarkan latar belakang tersebut, maka penelitian ini berfokus pada “Bagaimana desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package?”. Adapun subfokus penelitian ini adalah sebagai berikut:

- (a) Analisis kebutuhan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (b) Rancangan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (c) Pengembangan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (d) Tingkat kelayakan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.

1.3 Pertanyaan Penelitian

Berdasarkan latar belakang dan fokus penelitian tersebut, maka pertanyaan penelitian ini dirumuskan sebagai berikut:

- (a) Bagaimana kebutuhan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (b) Bagaimana Rancangan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (c) Bagaimana Pengembangan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (d) Bagaimana Tingkat kelayakan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.

1.4 Tujuan Penelitian

Berdasarkan pada rumusan masalah yang telah ditetapkan, maka tujuan dari penelitian ini adalah sebagai berikut:

- (a) Mendeskripsikan kebutuhan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (b) Merancang desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (c) Mengembangkan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.
- (d) Mendeskripsikan tingkat kelayakan desain penilaian atribusi kepengarangan pada teks defamasi anonim di media sosial berbasis Stylo R Package.

1.5 Kegunaan Penelitian

Penelitian ini diharapkan memberikan manfaat baik secara teoretis maupun praktis yang berkaitan dengan pengembangan kajian linguistik forensik dan implementasinya dalam penanganan kasus defamasi anonim pada media sosial di Indonesia.

1.5.1 Kegunaan Teoretis

Secara teoretis, penelitian ini diharapkan memberikan kontribusi terhadap pengembangan ilmu linguistik terapan, khususnya bidang linguistik forensik. Penelitian ini memperkaya kajian atribusi kepengarangan berbasis stilometri melalui integrasi konsep idiolek, pola kebahasaan individual, dan pendekatan analisis kuantitatif menggunakan Stylo R Package. Integrasi tersebut diharapkan dapat memperluas pemahaman mengenai bagaimana karakteristik kebahasaan individu direpresentasikan dalam teks digital serta dimanfaatkan untuk mengidentifikasi kepengarangan secara ilmiah.

Selain itu, penelitian ini menghasilkan suatu desain penilaian (assessment design) yang dapat dijadikan sebagai model evaluatif dalam menilai validitas, reliabilitas, dan kredibilitas hasil analisis stilometrik. Desain tersebut diharapkan mampu mengisi kekosongan kajian yang selama ini belum banyak membahas parameter evaluasi yang sistematis dan terstandar, khususnya dalam konteks

atribusi kepengarangan berbahasa Indonesia.

Penelitian ini juga menawarkan kerangka metodologis berbasis Design-Based Research (DBR) yang dapat dijadikan acuan dalam pengembangan perangkat maupun instrumen analisis linguistik forensik. Kerangka tersebut diharapkan dapat direplikasi dan dikembangkan lebih lanjut pada penelitian-penelitian berikutnya, baik dalam lingkungan akademik maupun institusi profesional yang bergerak di bidang analisis kebahasaan forensik.

Lebih lanjut, penelitian ini memperkuat landasan teoretis mengenai peran ciri-ciri idiolektal dalam teks digital, khususnya pada wacana media sosial yang umumnya berciri teks pendek, informal, dinamis, dan sangat kontekstual. Karakteristik tersebut masih relatif kurang dieksplorasi dalam literatur linguistik forensik, baik pada tingkat nasional maupun internasional, sehingga penelitian ini diharapkan dapat memperkaya khazanah keilmuan pada bidang tersebut.

1.5.1 Kegunaan Praktis

Secara praktis, penelitian ini diharapkan memberikan manfaat bagi berbagai pemangku kepentingan yang terlibat dalam analisis maupun penanganan kasus defamasi anonim di media sosial. Bagi aparat penegak hukum, seperti kepolisian, kejaksaan, dan lembaga peradilan, desain penilaian yang dihasilkan dapat digunakan sebagai acuan dalam mengevaluasi hasil analisis atribusi kepengarangan sehingga bukti linguistik menjadi lebih mudah dipahami, dinilai, dan dipertanggungjawabkan dalam proses pembuktian, termasuk dalam pengujian validitas ilmiah sesuai dengan prinsip-prinsip pembuktian forensik, seperti Daubert Standard.

Bagi ahli linguistik forensik, penelitian ini menyediakan desain penilaian yang dapat dimanfaatkan sebagai alat bantu profesional dalam mengevaluasi kualitas analisis stilometri. Penggunaan desain tersebut diharapkan mampu meningkatkan transparansi prosedur analisis, memperkuat konsistensi hasil, serta meminimalkan bias interpretatif dalam penyusunan pendapat ahli.

Bagi peneliti dan akademisi, penelitian ini menyediakan kerangka metodologis beserta parameter dan indikator evaluatif yang dapat dijadikan rujukan dalam penelitian lanjutan mengenai atribusi kepengarangan, analisis idiolek,

linguistik forensik, maupun bidang computational linguistics. Kerangka tersebut juga membuka peluang pengembangan instrumen evaluasi yang lebih komprehensif sesuai dengan karakteristik bahasa Indonesia.

Selain itu, penelitian ini diharapkan memberikan manfaat bagi pengembang perangkat komputasional dengan menyediakan informasi mengenai kebutuhan fitur, parameter teknis, serta aspek evaluatif yang diperlukan dalam pengembangan aplikasi stilometri berbahasa Indonesia. Temuan penelitian ini dapat menjadi dasar awal untuk menghasilkan perangkat yang lebih ramah pengguna, terstandar, dan sesuai dengan kebutuhan investigasi forensik digital di Indonesia.

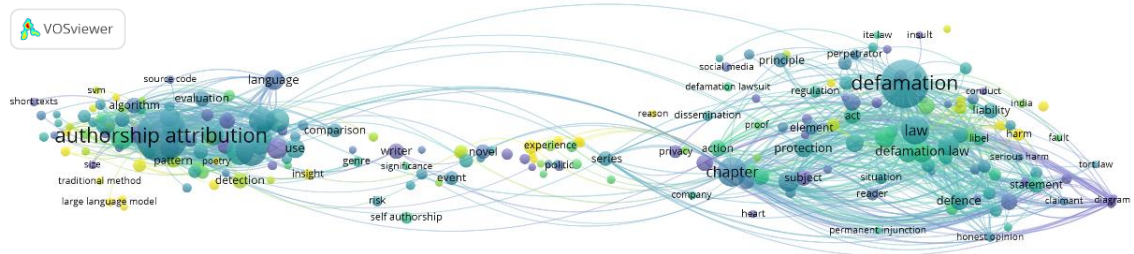
Bagi masyarakat secara umum, penelitian ini diharapkan dapat meningkatkan literasi digital dan kesadaran linguistik mengenai risiko penggunaan anonimitas di media sosial, potensi terjadinya tindak defamasi, serta pentingnya penggunaan bahasa yang bertanggung jawab dalam ruang digital. Dengan demikian, hasil penelitian ini tidak hanya berkontribusi terhadap pengembangan ilmu pengetahuan, tetapi juga mendukung terciptanya ekosistem komunikasi digital yang lebih etis dan bertanggung jawab.

1.6 Kebaruan Penelitian

Pemetaan bibliometrik menggunakan VOSviewer menunjukkan bahwa penelitian mengenai atribusi kepengarangan dan defamasi anonim memiliki lanskap yang terstruktur dalam beberapa klaster utama. Melalui *network visualization*, terlihat bahwa bidang ini terkonsentrasi pada dua garis besar kajian. Kelompok pertama berfokus pada aspek teknis dan komputasional, ditandai oleh kehadiran kata kunci seperti *authorship attribution*, *algorithm*, *feature*, *stylometry*, *model*, dan *dataset*. Pola ini menunjukkan bahwa sebagian besar penelitian bertujuan mengembangkan model klasifikasi, mengevaluasi fitur stilometrik (misalnya kata fungsi, n-gram, atau ciri sintaktis), serta membangun kerangka kerja analitis untuk meningkatkan akurasi identifikasi penulis. Kelompok kedua menggambarkan dimensi hukum dan sosial, di mana kata kunci seperti *defamation*, *court*, *liability*, dan *legal responsibility* menegaskan pentingnya atribusi kepengarangan dalam konteks pembuktian hukum dan pengambilan keputusan yudisial. Keterhubungan kedua klaster ini menandakan meningkatnya perhatian terhadap integrasi metode

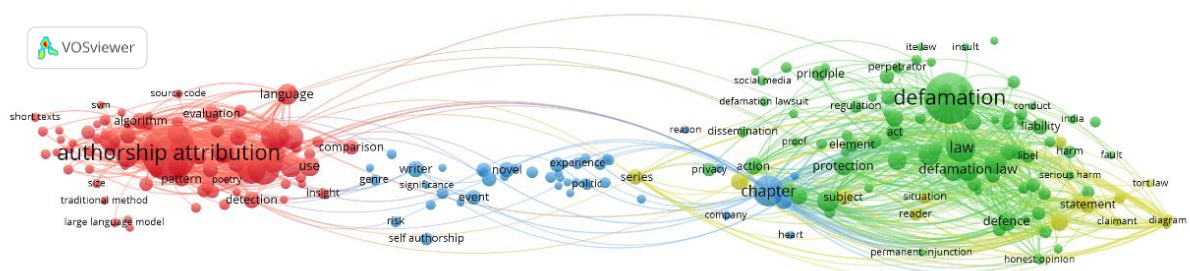
komputasional dengan kebutuhan praktik forensik.

Gambar 1.1. Hasil Analisis *Network Visualization*



Visualisasi kedua, yaitu *overlay visualization*, memperlihatkan perkembangan temporal penelitian. Istilah-istilah yang lebih awal seperti *stylometry*, *function words*, atau *Bayesian models* menunjukkan fokus awal pada pendekatan statistik klasik. Namun dalam perkembangan dua hingga tiga tahun terakhir, warna pada peta beralih ke istilah seperti *machine learning*, *deep learning*, *transformer*, *BERT*, dan *short text classification*. Pergeseran ini menunjukkan kecenderungan menuju model berbasis neural dan representasi vektor yang lebih kompleks. Meskipun demikian, istilah terkait bahasa tertentu, terutama untuk konteks non-Inggris atau teks pendek media sosial, tidak muncul sebagai pusat penelitian yang padat. Hal ini mengindikasikan adanya kebutuhan besar untuk kajian atribusi kepengarangan pada teks pendek, informal, dan multilingual sebagaimana karakteristik media sosial Indonesia.

Gambar 1.2. Hasil Analisis *Overlay Visualization*



Visualisasi ketiga, yaitu *density visualization*, menunjukkan intensitas penelitian pada area tertentu. Titik-titik dengan kepadatan tertinggi tampak pada kata kunci *authorship attribution*, *model*, *feature*, dan *defamation*, yang berarti bahwa keempat konsep tersebut merupakan pusat diskusi dan landasan metodologis

bagi penelitian saat ini. Namun, beberapa aspek seperti *assessment*, *evaluation framework*, atau *forensic validation* tidak muncul sebagai area yang padat, sehingga dapat disimpulkan bahwa penelitian terdahulu lebih banyak berfokus pada pengembangan alat dan algoritme daripada pengembangan kerangka penilaian yang menyatukan validitas teknis dan kebutuhan forensik. Celah ini menunjukkan bahwa meskipun metode atribusi kepengarangan semakin canggih, upaya untuk menstandarkan prosedur penilaian dan memastikan keterterimaan hasil dalam konteks hukum masih terbatas.

Gambar 1.3. Hasil Analisis *Density Visualization*



Secara keseluruhan, peta penelitian tersebut menunjukkan bahwa bidang atribusi kepengarangan modern telah berkembang pesat dari sisi teknik dan algoritme, tetapi belum sepenuhnya menjangkau kebutuhan praktik hukum, terutama pada kasus defamasi anonim dalam konteks lokal seperti Indonesia. Belum banyak penelitian yang mengembangkan desain penilaian yang dapat menjembatani analisis komputasional dengan standar forensik yang menuntut transparansi, reliabilitas, keterulangan, dan interpretabilitas hasil. Sejalan dengan itu, penelitian ini menempati posisi strategis untuk mengisi kekosongan tersebut melalui perancangan desain penilaian atribusi kepengarangan berbasis Stylo R Package, yang diarahkan untuk menangani karakteristik teks pendek media sosial serta memenuhi tuntutan metodologis dan kebutuhan pembuktian hukum.

1.7 Peta Jalan Penelitian

Peta Jalan penelitian pada ilustrasi di bawah ini menggambarkan proses atau perjalanan penelitian yang telah dan sedang dilakukan. Peta perjalanan berikut ini

berisi alur penelitian yang dimulai sejak awal tahun 2025.

Gambar 1.4. Peta Jalan Penelitian

