

1. Pendahuluan

1.1 Latar Belakang

Media sosial telah menjadi bagian penting dalam kehidupan sehari-hari, dan penggunaannya semakin meningkat. Saat ini terdapat 4,76 miliar pengguna media sosial di seluruh dunia, yang hampir setara dengan 60 persen dari total populasi global [1]. Media sosial berfungsi sebagai platform bagi pengguna untuk membuat, berbagi, dan mengonsumsi konten, berkontribusi pada berbagai aspek kehidupan sehari-hari, termasuk komunikasi, hiburan, dan berbagi informasi [2]. Namun, media sosial juga menghadirkan tantangan dalam forensik digital, di mana media sosial dapat menjadi sumber bukti yang berharga dan sarana untuk memfasilitasi kegiatan kriminal [2]. Data dari media sosial dapat memberikan informasi penting tentang identitas, aktivitas, hubungan, dan motif pengguna, yang membantu penyelidikan dan penegakan hukum [3].

Terlepas dari potensinya, data media sosial menghadirkan tantangan dalam pengumpulan, pemrosesan, dan analisis karena sifatnya yang sangat luas, beragam, dan dinamis [2]. Data dari media sosial dapat mencakup berbagai jenis konten, seperti teks, gambar, video, tautan, dan lain-lain [4]. Penelitian sebelumnya, seperti yang dilakukan oleh Powell dan Haynes [2] dan Nikolaidou et al. [3], telah mengeksplorasi isu-isu hukum dan kerangka kerja untuk menganalisis data media sosial, namun tidak membahas pengelompokan dan kategorisasi yang sistematis.

Salah satu metode untuk mengelola dan menganalisis data media sosial adalah dengan mengelompokkan dan mengkategorikan, yang mengurangi kompleksitas, redundansi, dan ketidakkonsistenan, sehingga data lebih terstruktur dan lebih mudah untuk dianalisis [5]. Namun, masih ada kesenjangan dalam metodologi yang dirancang untuk mengelompokkan dan mengkategorikan data secara sistematis dari Twitter dan Instagram pada platform web. Penelitian ini bertujuan untuk mengisi kesenjangan ini dengan mengembangkan metodologi untuk mengelompokkan dan mengkategorikan data media sosial secara efisien, menyederhanakan kompleksitasnya, dan memberikan informasi penting untuk analisis forensik.

Kebaruan dari penelitian ini terletak pada pendekatannya yang komprehensif dalam menangani data media sosial yang sangat banyak, beragam, dan dinamis. Dengan menggunakan alat seperti Apify untuk pengumpulan data dan PostgreSQL untuk manajemen data, sistem yang diusulkan menawarkan cara yang efektif untuk memproses dan menganalisis data. Kontribusi dari penelitian ini sangat penting karena mengembangkan metodologi terperinci untuk mengumpulkan dan melakukan pra-pemrosesan data media sosial, mengimplementasikan desain sistem yang memfasilitasi pengelompokan dan kategorisasi data yang efisien, dan menunjukkan keefektifan sistem dalam skenario forensik dunia nyata, yang menyoroti kegunaan praktisnya bagi penegak hukum dan analis forensik.

1.2 Topik dan Batasannya

Penelitian ini melibatkan pengumpulan data dari Twitter dan Instagram pada platform web, serta pengelompokan dan pengkategorian data untuk analisis forensik digital. Input dari penelitian ini adalah data dari media sosial tersebut, dan outputnya adalah kategori dan kelompok data yang mendukung analisis forensik. Data yang diambil mencakup tweet, postingan, profil pengguna, dan metadata lainnya. Pengumpulan data hanya dari platform web, tanpa melibatkan aplikasi mobile, serta fokus pada data publik yang tersedia secara umum dan data pribadi peneliti yang diambil dari platform tersebut.

1.3 Tujuan

Penelitian ini bertujuan untuk mengidentifikasi data yang dapat dikumpulkan dari Twitter dan Instagram pada platform web. Selain itu, penelitian ini juga mengembangkan skema pengelompokan dan pengkategorian data dalam proses analisis forensik guna menjawab pertanyaan penyidikan. Skema yang telah dikembangkan akan diuji berdasarkan parameter akurasi dan kelengkapan data.

1.4 Organisasi Tulisan

Laporan penelitian ini terdiri dari lima bagian. Bagian pertama membahas latar belakang, menguraikan tantangan yang dihadapi dalam mengelola data Twitter dan Instagram yang besar dan beragam, serta kebutuhan akan pengelompokan dan pengkategorian data. Bagian kedua menyediakan penjelasan tentang istilah-istilah terkait dan ulasan literatur yang terkait. Bagian ketiga menjelaskan deskripsi sistem, perancangan, dan pengujian sistem yang telah dibangun penelitian ini. Bagian keempat memaparkan hasil dari pengelompokan dan pengkategorian data Twitter dan Instagram pada sistem. Bagian kelima menyajikan kesimpulan serta harapan untuk penelitian di masa depan.