

ABSTRAK

Kelangkaan data terstruktur pada ranah investasi Indonesia menghadirkan tantangan signifikan bagi proses analisis tren dan pengambilan keputusan. Meskipun *Temporal Knowledge Graph* (TKG) mampu menangkap dinamika pasar, dataset acuan yang ada saat ini sangat bias terhadap sumber berbahasa Inggris dan pasar negara maju, dan karenanya mengesampingkan konteks unik ekonomi berkembang seperti Indonesia. Penelitian ini menjawab kesenjangan tersebut dengan membangun IndovestDKG, dataset TKG pertama yang berfokus pada domain investasi Indonesia, yang bersumber dari portal berita Kompas. Dengan menggunakan *Large Language Model* (LLM), sebanyak 7.761 artikel berita diekstraksi menjadi 68.207 fakta temporal (*quadruplet*). Selanjutnya, lima metode *Temporal Knowledge Graph Completion* (TKGC) mutakhir dievaluasi pada IndovestDKG terhadap dua set data acuan lainnya (FinDKG dan ICEWS14). Hasil penelitian mengungkap performa yang secara signifikan lebih rendah dari semua model pada IndovestDKG, yang menyoroti ketidaksesuaian arsitektur model saat ini untuk pasar berkembang. Pasar ini memiliki karakteristik unik seperti tingkat *sparsity* (ketersebaran) yang ekstrem, dengan 70,8% entitas hanya muncul sekali (*singleton entities*), dan dinamika yang didorong oleh kebijakan. Model Re-Temp mencapai kinerja terbaik pada set data finansial, dengan MRR sebesar 8,39% pada IndovestDKG dan 28,12% pada FinDKG. Kontribusi utama dari penelitian ini adalah penyediaan set data acuan (*benchmark*) baru berskala kecil (IndovestDKG) dan analisis komparatif yang mengungkap keterbatasan metode TKGC yang ada dalam konteks pasar negara berkembang.

Kata Kunci: temporal knowledge graph completion, berita investasi indonesia, large language model