UNIVERSITAS
GADJAH MADA
**Pengembangan dan Evaluasi Model Generative AI Berbasis GPT Untuk Menghasilkan Rekomendasi Intervensi Ergonomis**
Rheina Khaisa Rhehani Putri, Ir. Ardiyanto, S.T., M.Sc., Ph.D., AEP., IPM.; Ir. Achmad Pratama Rifai, S.T., M.Eng., P
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

# INTISARI

Gangguan otot rangka (Gotrak) merupakan salah satu masalah kesehatan kerja yang umum terjadi pada sektor Usaha Mikro, kecil, dan Menengah (UMKM) akibat aktivitas kerja manual yang melibatkan postur tidak ergonomis, gerakan berulang, dan durasi kerja yang panjang. Kondisi ini dapat menurunkan produktivitas serta kualitas hidup pekerja. Meskipun intervensi ergonomis terbukti efektif dalam mengurangi risiko Gotrak, penerapannya masih terkendala oleh keterbatasan akses terhadap ahli ergonomi, biaya yag tinggi, dan keterbatasan sumber daya. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan solusi berbasis *Generative Artificial Intelligence* (GAI) menggunakan *Large Language Model* (LLM) berbasis *Generative Pre-trained Transformer* (GPT) untuk menghasilkan rekomendasi intervensi ergonomis secara otomatis dan kontekstual.

Sistem ini dirancang untuk menghasilkan rekomendasi intervensi ergonomis secara otomatis dan kontekstual berdasarkan deskripsi teks dan gambar mengenai situasi kerja. Pengembangan dilakukan menggunakan GPT *Builder* dengan *knowledge base* ergonomis yang kemudian diuji menggunakan studi kasus UMKM. Evaluasi dilakukan oleh tujuh evaluator terhadap rekomendasi yang dihasilkan oleh model *custom* GPT, *Basic* GPT, dan Pakar Ergonomika berdasarkan tujuh aspek, yaitu *accuracy*, *completeness*, *inclusion of rationale*, *provision of citation*, *human-likeness*, *readability*, dan *satisfaction*. Data kuantitatif dianalisis menggunakan uji *Kruskal-wallis* dan uji *Dunn's post hoc comparisons*, sedangkan data kualitatif dianalisis menggunakan analisis tematik dengan *inductive coding*. Selain itu, tiga studi kasus tambahan digunakan untuk mengevaluasi akurasi antar model GPT dengan menggunakan metode *BERTScore*, serta validitas referensi dianalisis secara manual.

Hasil menunjukkan adanya perbedaan signifikan pada variabel *inclusion of rationale* dan *readability*, sedangkan lima variabel lainnya tidak menunjukkan perbedaan signifikan antar model. Uji *Dunn's post hoc comparisons* menunjukkan bahwa *custom* GPT berbeda signifikan dengan Pakar Ergonomika pada variabel *inclusion of rationale*, serta berbeda signifikan dengan Pakar Ergonomika pada variabel *readability* bersama *basic* GPT. Hasil *BERTScore* menunjukkan *custom* GPT sedikit lebih akurat dibanding *basic* GPT (F1 = 0,6004), dengan keunggulan pada *precision*, sedangkan *basic* GPT unggul pada *recall*. Validasi referensi menunjukkan bahwa *custom* GPT memiliki proporsi referensi benar tertinggi (91,30%) dibanding Pakar Ergonomika (85,00%) dan *basic* GPT (57,69%), serta tidak menghasilkan referensi halusinasi. Analisis kualitatif menghasilkan lima tema utama, yaitu kualitas rekomendasi ergonomis, penjelasan dan justifikasi, bahasa dan penyajian, referensi ilmiah, serta kesesuaian dengan konteks UMKM. Temuan ini menunjukkan bahwa *generative* AI berpotensi menjadi alat bantu yang efektif untuk mendukung penyusunan rekomendasi ergonomis pada UMKM, meskipun evaluasi lebih lanjut dengan pendekatan yang lebih komprehensif masih diperlukan.

**Kata Kunci:** *Generative* AI, GPT, Rekomendasi Intervensi, dan Ergonomis.

UNIVERSITAS
GADJAH MADA
**Pengembangan dan Evaluasi Model Generative AI Berbasis GPT Untuk Menghasilkan Rekomendasi Intervensi Ergonomis**
Rheina Khaisa Rhehani Putri, Ir. Ardiyanto, S.T., M.Sc., Ph.D., AEP., IPM.; Ir. Achmad Pratama Rifai, S.T., M.Eng., P
**Universitas Gadjah Mada, 2026 | Diunduh dari http://etd.repository.ugm.ac.id/**

# *ABSTRACT*

Musculoskeletal disorders (MSDs) are among the most common occupational health problems in the Micro, Small, and Medium Enterprise (MSME) sector, resulting from manual work activities involving non-ergonomic postures, repetitive movements, and prolonged working durations. These conditions can reduce workers' productivity and quality of life. Although ergonomic interventions have been proven effective in reducing the risk of MSDs, their implementation remains constrained by limited access to ergonomics experts, high costs, and resource limitations. Therefore, this study aims to develop a Generative Artificial Intelligence (GAI)-based solution using a Large Language Model (LLM) based on Generative Pre-trained Transformer (GPT) to generate ergonomic intervention recommendations automatically and contextually.

The system was designed to generate ergonomic intervention recommendations automatically and contextually based on textual and visual descriptions of work situations. The development process was conducted using GPT Builder with an ergonomics knowledge base and subsequently evaluated using MSME-related case studies. The evaluation involved seven evaluators who assessed recommendations generated by the custom GPT, Basic GPT, and Ergonomics Expert models based on seven aspects: accuracy, completeness, inclusion of rationale, provision of citation, human-likeness, readability, and satisfaction. Quantitative data were analyzed using the Kruskal–Wallis test and Dunn's post hoc comparisons, while qualitative data were analyzed using thematic analysis with inductive coding. In addition, three supplementary case studies were used to evaluate accuracy between GPT models using the BERTScore method, and reference validity was analyzed manually.

The results showed significant differences in the inclusion of rationale and readability variables, while the other five variables showed no significant differences among the models. Dunn's post hoc comparisons indicated that the custom GPT differed significantly from the Ergonomics Expert in the inclusion of rationale variable and differed significantly from the Ergonomics Expert in the readability variable together with the Basic GPT. The BERTScore results showed that the custom GPT was slightly more accurate than the Basic GPT (F1 = 0.6004), with superior precision, whereas the Basic GPT demonstrated superior recall. Reference validation showed that the custom GPT had the highest proportion of correct references (91.30%) compared with the Ergonomics Expert (85.00%) and the Basic GPT (57.69%), and it generated no hallucinated references. Qualitative analysis identified five main themes: quality of ergonomic recommendations, quality of explanations and justifications, quality of language and presentation, quality of scientific references, and suitability for the MSME context. These findings indicate that generative AI has the potential to serve as an effective support tool for developing ergonomic recommendations in MSMEs, although further evaluation using more comprehensive approaches is still required.

**Keywords:** Generative AI, Intervention Recommendations, *and* Ergonomics.