Document Validation Report for ReDKG

Correspondence: Yes
Percentages: 92.0%

Conclusion:

The codebase closely implements the system architecture, algorithms, and experimental pipeline detailed in the Russian documentation. The described components—such as genetic algorithms for topic model hyperparameter optimization, surrogate modeling for acceleration, and dynamic knowledge graph construction using attention matrices—are all reflected in the code. There are modules for graph and hypergraph construction, knowledge graph embedding training (e.g., TransE on OGBL-BioKG), genetic optimization, and evaluation scripts (MRR, Hits@K). Preprocessing, configuration, and support for custom visualization and physical layout models are implemented. Input formats (pd.DataFrame, torch Data objects, etc.), dataset handling, and core evaluation/metric scripts are present and consistent with the documentation. The main gap is the absence of explicit experimental result output or reporting modules; while all building blocks are present to reproduce results, there is no evidence of scripts to generate final performance numbers or integrate all steps in an automated benchmark. Overall, the code is highly aligned with the documented algorithms and procedures, with only minor missing glue for full end-to-end reproducibility.