Text Analytics: Topic Modeling, Vocabulary Structure, and Document Embeddings

This section examines the latent semantic structure of the Active Inference corpus through complementary text-analytic methods: non-negative matrix factorization for topic discovery, TF-IDF vocabulary analysis, document embedding projections, and term co-occurrence patterns. Together, these analyses reveal thematic structure that cuts across the keyword-based domain taxonomy presented in Section 3.

Topic Modeling: Latent Structure

Non-negative matrix factorization (NMF) applied to the TF-IDF matrix identifies five latent topics:

Topic	Top Terms	Interpretation
0	learning, agent, model, agents, active, environments, aif,	Agent-environment modeling and robotic applications