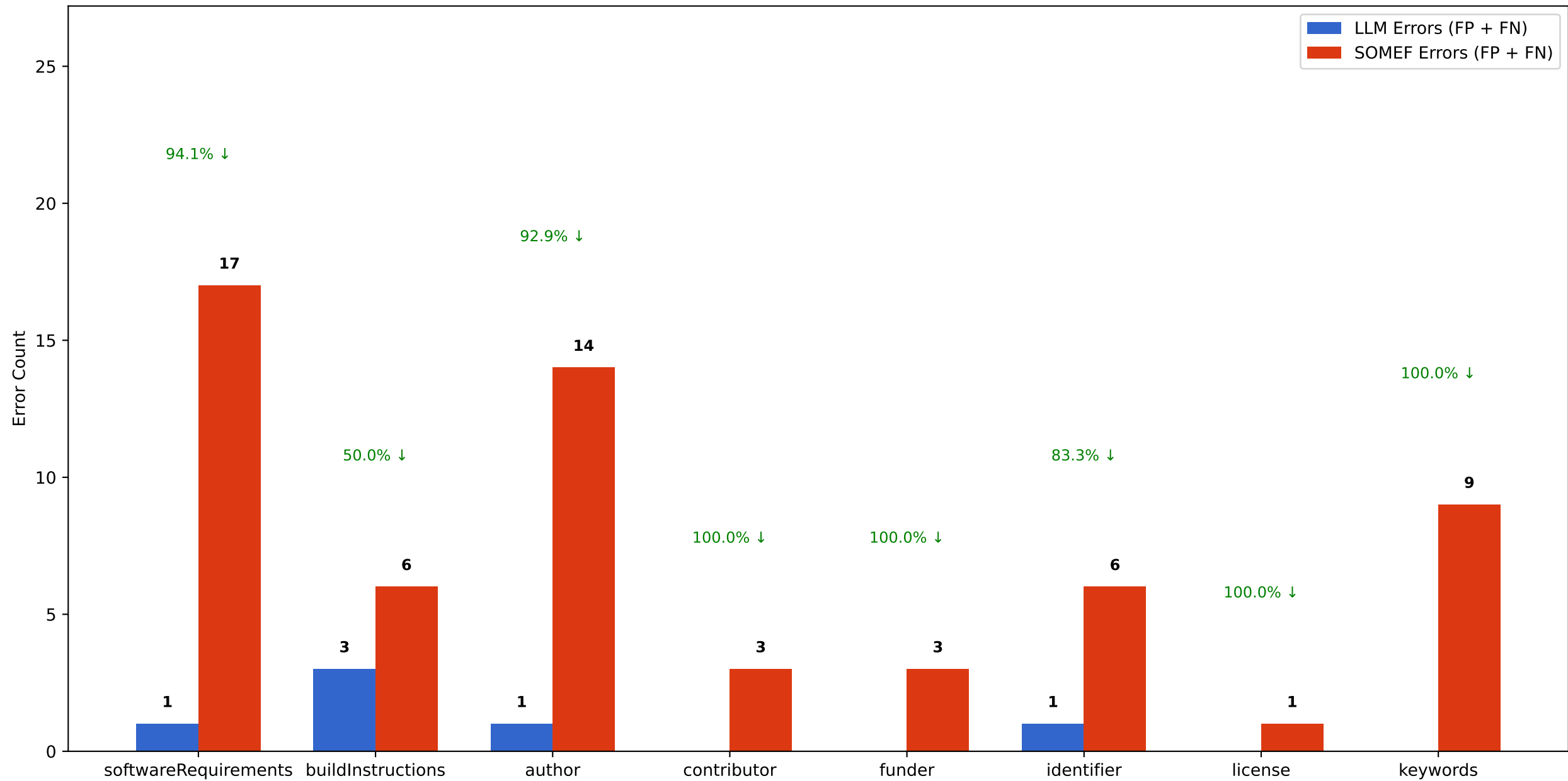


LLM vs SOMEF Total Errors (FP + FN)



↓ indicates percentage fewer errors compared to SOMEF