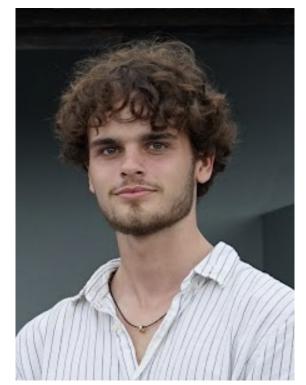
DIALOGIX

HackaTUM Winner 2023





Cross-Institutional Academic Partnership

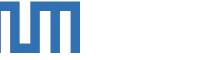








Jonas Matthaei



Philip Jünemann



Laurens Houtekamer



Noud Bonke







CONCISE COMPLETENESS



INTERACTIVITY

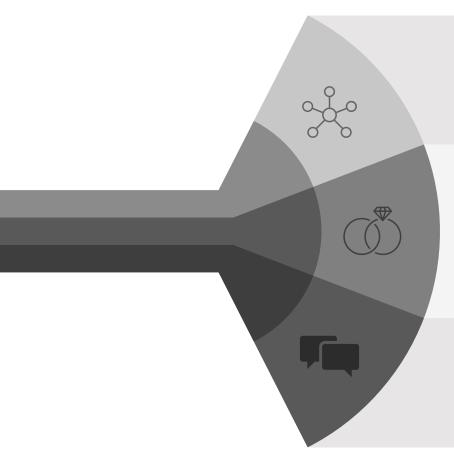


MEASURABLE SUCCESS



The approach

Our Product: A Bird's Eye View



Database Population with Clustered Logs

Grouping logs into clusters, storing cluster identifiers in a datastore, and saving fully tagged logs in a database.

Embedding-Based Question-Cluster Matching

Extract the embeddings for the cluster titles and the user's question, then match the question to the top 5 relevant cluster titles.

All-Encompassing Log Inclusion for Prompt Perfection

Use the cluster ID to access the complete logs and input them into an LLM prompt to give the LLM complete information

Log Parsing With Bidirectional Parallel Trees

S. Yu, P. He, N. Chen and Y. Wu, "Brain: Log Parsing With Bidirectional Parallel Tree"

- (1) Cutting edge method published in 2023.
- Paster and more efficient than neural networks, able to match the pace of log generation.
- Enhanced parsing accuracy over alternatives, proven by test on 16 benchmark datasets.
- Able to process approximately **one million** lines of logs in just **46 seconds**.

Optimizing through PERFORMANCE MEASUREMENT

Rating outcomes using semantic answer similarity and MLflow

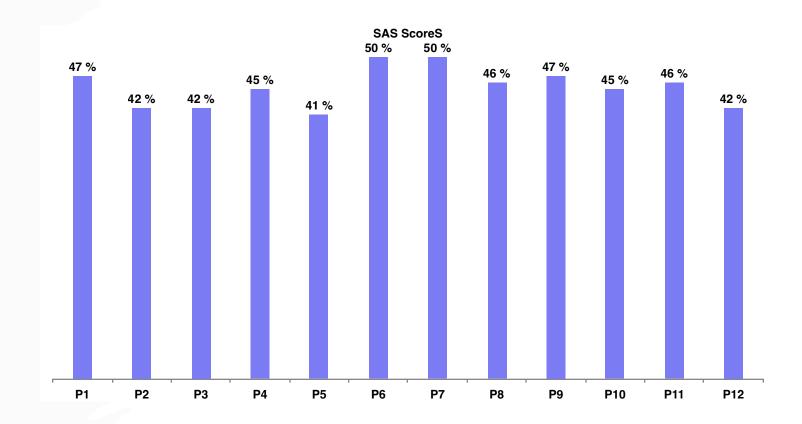
- Developed golden standard dataset based on wide array of potentially interesting logs.
- SAS to measure semantics instead of lexical overlap.
- SAS shown to correlate highly with human judgement on **three** different datasets.
- Created script to experiment with parameters in an attempt to achieve the optimal pipeline.

COMPARING PARAMETER COMBINATIONS

- (1) Retriever model
- (2) Reader model

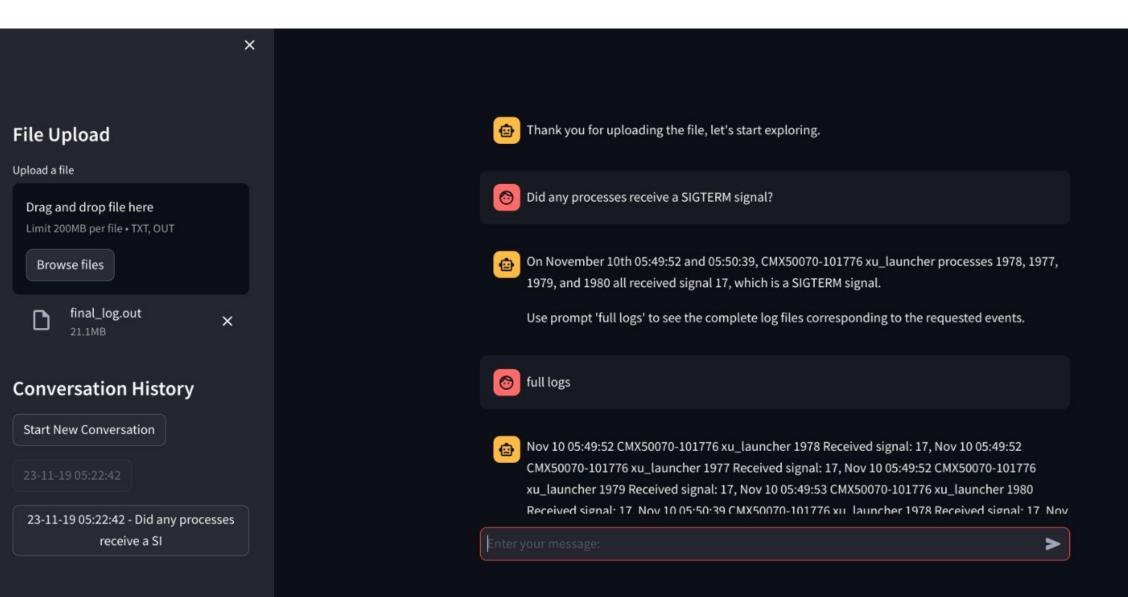
(3) Similarity threshold

(4) Prompt engineering

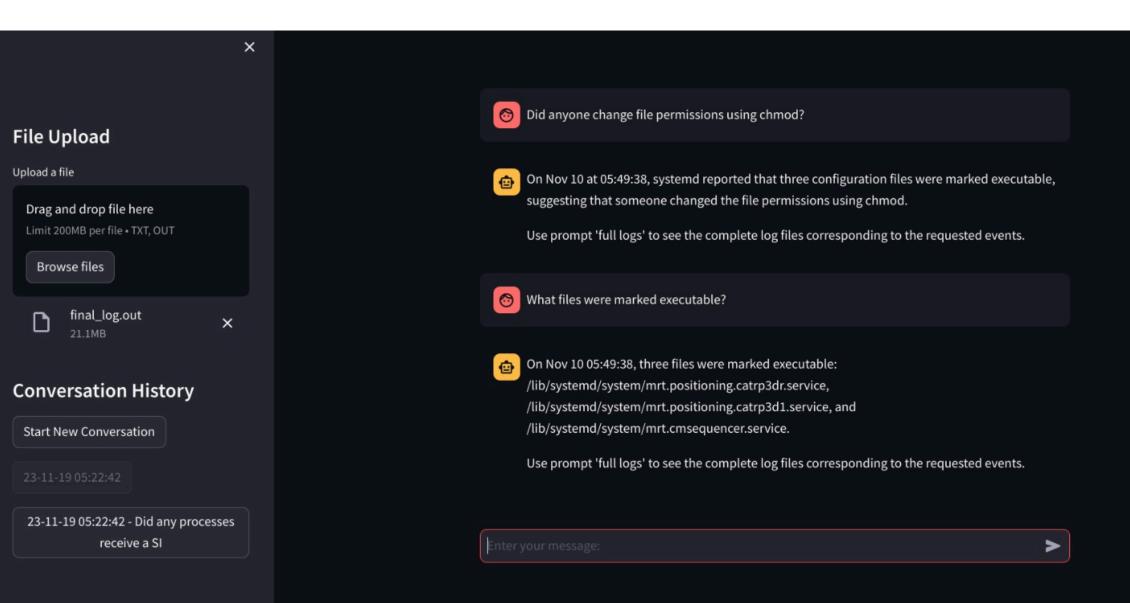




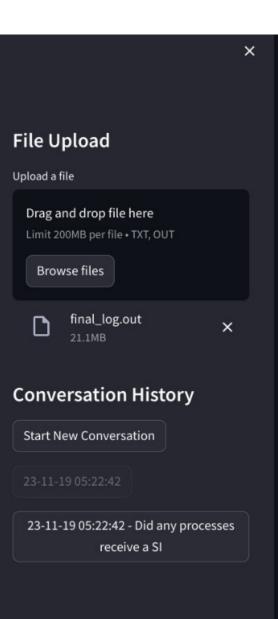
The results



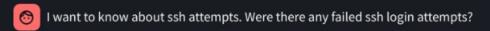








CMX50070-101776 xu_launcher 1978 Waiting for any signal to shut down..., Nov 10 05:49:52 CMX50070-101776 xu_launcher 1977 Waiting for any signal to shut down..., Nov 10 05:49:53 CMX50070-101776 xu_launcher 1979 Waiting for any signal to shut down..., Nov 10 05:50:41 CMX50070-101776 xu_launcher 1980 Waiting for any signal to shut down..., Nov 10 05:50:41 CMX50070-101776 xu_launcher 1981 Waiting for any signal to shut down... Nov 10 05:49:37 localhost kernel pcieport 0000:00:01.1: PME: Signaling with IRQ 25, Nov 10 05:49:37 localhost kernel pcieport 0000:00:01.3: PME: Signaling with IRQ 26, Nov 10 05:49:37 localhost kernel pcieport 0000:00:01.5: PME: Signaling with IRQ 27, Nov 10 05:49:37 localhost kernel pcieport 0000:00:08.1: PME: Signaling with IRQ 28, Nov 10 05:49:37 localhost kernel pcieport 0000:00:08.2: PME: Signaling with IRQ 29 Nov 10 05:49:54 CMX50070-101776 sh 3390 cmseq | CMsequencer starter processes IDs:



From Nov 10 06:09:41 to Nov 13 17:14:04, there were 14 failed authentication attempts from rhost 10.0.2.217. Additionally, there were 8 successful publickey authentication attempts from various rhosts for root.

Use prompt 'full logs' to see the complete log files corresponding to the requested events.

Enter your message:





The VISION

OUR FUTURE VISION

- 1 Intelligent agents selecting optimal tools.
- Creating a SQuAD format dataset specialized on logs for further pipeline optimization.
- The optimal user interface, communicating inside the logs.
- 4 Transforming the logs into a story.