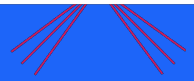
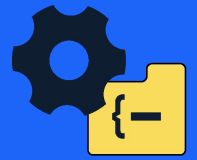
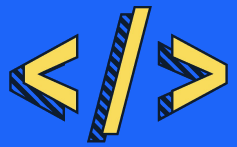




{codemotion} CONFERENCE MADRID 2025

LLM-Driven Attack Flow Generation from Security Publications



CONFERENCE
MADRID 2025

Juan Manuel Cristóbal

Senior Engineer

Senior Software Engineer con más de una década de experiencia diseñando y desplegando soluciones backend escalables para entornos de misión crítica.

Me apasiona la arquitectura en la nube, la automatización de procesos mediante prácticas DevOps y la adopción de metodologías ágiles para asegurar la calidad y el rendimiento en cada proyecto. A lo largo de mi trayectoria, he liderado implementaciones de principio a fin, garantizando que las soluciones cumplan altos estándares de seguridad y disponibilidad.

Disfruto compartiendo conocimientos sobre TDD, Python y la integración de tecnologías emergentes, con el objetivo de impulsar la innovación y el crecimiento continuo en cada equipo con el que colaboro.

Fran Gomez

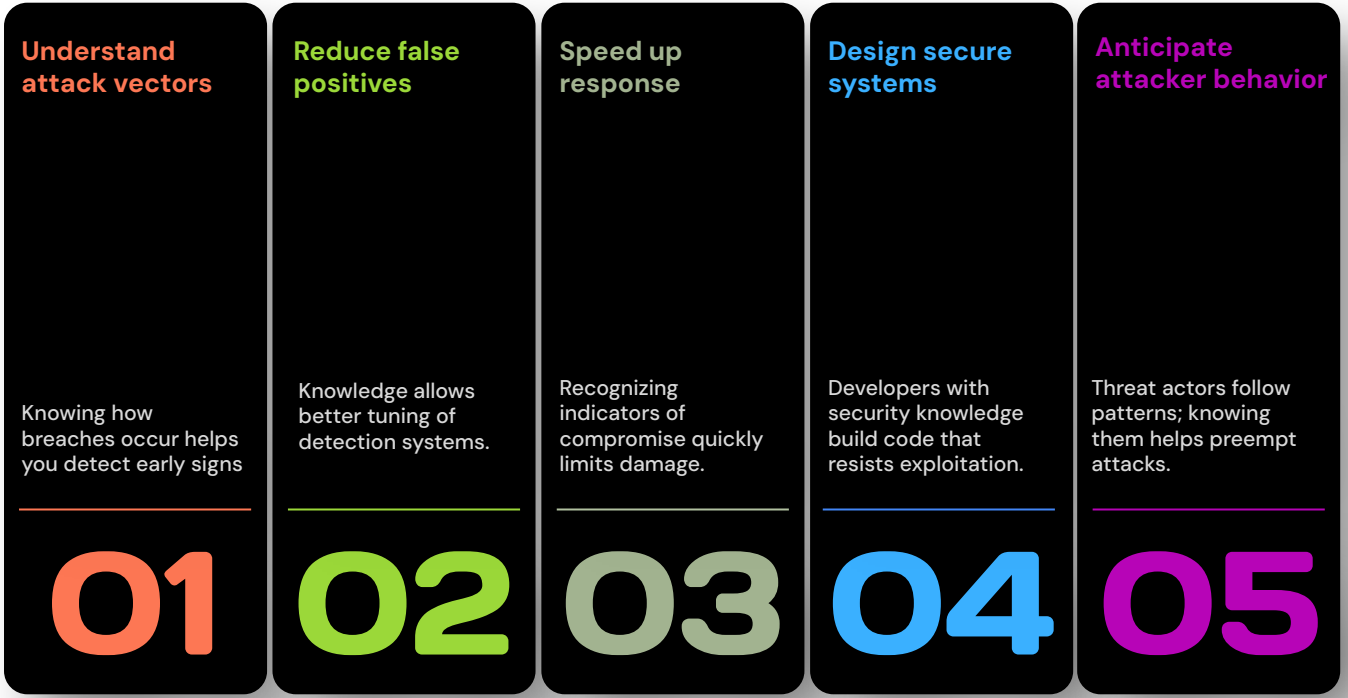
Security Research

Fran Gómez (@ffranz) is a cybersecurity researcher, builder, and speaker passionate about Threat Intelligence, OSINT/OSTI, IPv6, and Hacking. Founder of @MrLooquer, the creator of @e_Sinfonier, and currently leads Security Research at @devo_inc.

Previously involved with @OwaspMadrid and @TelefonicaTech, I’ve been sharing security insights since 2008 and has spoken at major conferences such as RootedCON, Black Hat, Spark Summit, CCNCERT, Codemotion Amsterdam, and NIAS (@nciagency).

How Do You Defend Against What You Don't Understand?

Cybersecurity Knowledge



Transforms raw data into Actionable Insight

Security Concepts & Tools

- Threat Modeling
- Security Controls
- Vulnerability Management
- Penetration Testing

Threat Intelligence

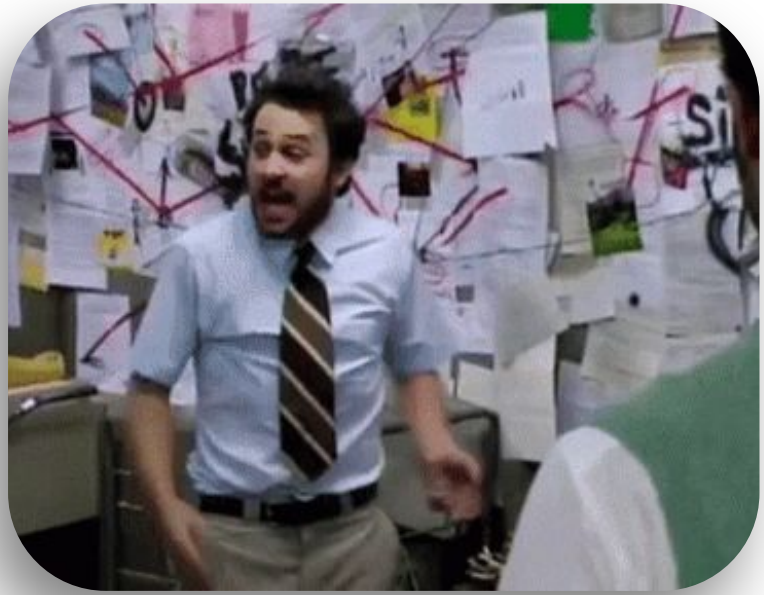
- TTPs
  - (Tactics, Techniques, Procedures)
- Indicators of Compromise
  - (IOCs)
- Threat Actor Profiling

Regulations & Ethics

- Compliance Standards
- Security Policies

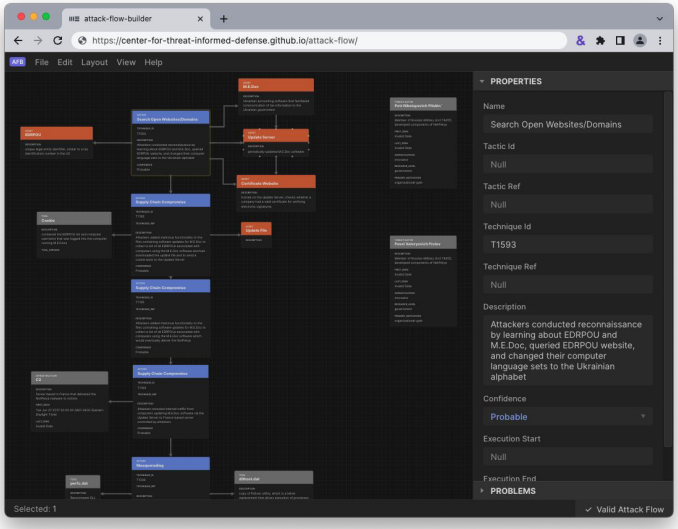
Crucial Knowledge Areas

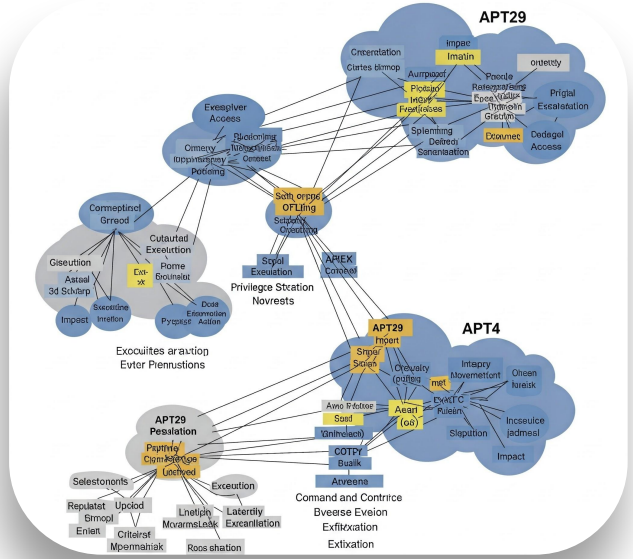
Attack Flow



- Powerful concept for a unified cyber framework
- Not just a threat, but a narrative: the how, why, and what.
- They make knowledge usable

Attack Graph vs. MITRE ATT&CK Flow





MITRE | ATT&CK®


NOW WHAT?

State of The Art

[1] P. Ammann, D. Wijesekera and S. Kaushik, "Scalable, Graph-Based Network Vulnerability Analysis," *Proceedings of the 9th ACM Conference on Computer and Communications Security*, 2002, pp. 217–224. [Online]. Available: https://www.researchgate.net/publication/3948684_Automated_Generation_and_Analysis_of_Attack_Graphs.

[2] A. Husari, C. N. Gutierrez, N. Harang and S. Chellappa, "TTPDrill: Automatic and Accurate Extraction of Threat Actions from Unstructured Text of CTI Sources," *Proceedings of the 2017 ACM Workshop on Artificial Intelligence and Security (AISec '17)*, 2017, pp. 59–69. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3134600.3134646.

[3] N. Soman, "Death to the IOC: What's Next in Threat Intelligence," *Black Hat USA 2019*, [Video and Slides]. [Online]. Available: https://www.youtube.com/watch?v=MJHSUsS9K7s and https://i.blackhat.com/USA-19/Thursday/us-19-Soman-Death-To-The-IOC-Whats-Next-In-Threat-Intelligence.pdf.

[4] S. Mittal, A. Joshi and A. Finin, "TIMiner: TTP-based Cyber Threat Intelligence Summarization from Unstructured Text," *Computers & Security*, vol. 96, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0167404820301395.

[5] V. Legoy, *rcATT: Mapping Unstructured Cyber Threat Intelligence Reports to the MITRE ATT&CK Framework Using Machine Learning*, Master's Thesis, University of Amsterdam, 2020. [Online]. Available: https://github.com/vlegoy/rcATT/blob/master/MScThesis_rcATT_VLegoy.pdf.

[6] MITRE Corporation, "Threat Report ATT&CK Mapper (TRAM)," [Online]. Available: https://ctid.mitre.org/projects/threat-report-attck-mapper-tram.

[7] S. Joshi, A. Mittal and A. Joshi, "EXTRACTOR: Extracting Attack Behavior from Threat Reports," *2021 IEEE European Symposium on Security and Privacy (EuroS&P)*, Vienna, Austria, 2021, pp. 207–221. [Online]. Available: https://ieeexplore.ieee.org/document/9581182/.

[8] B. Zhang et al., "Looking Beyond IoCs: Automatically Extracting Attack Patterns from External CTI," *arXiv preprint arXiv:2211.01753*, 2022. [Online]. Available: https://arxiv.org/pdf/2211.01753.

[9] A. Wahab et al., "CTINEXUS: A Framework for Scalable CTI Knowledge Graph Generation with LLM-Augmented Mapping to ATT&CK," *arXiv preprint arXiv:2410.21060*, 2024. [Online]. Available: https://arxiv.org/pdf/2410.21060.

[10] Y. Liu, "Threats To MITRE(CWD, ATT&CK) AI-LLM Mapper," GitHub Repository, 2024. [Online]. Available: https://github.com/LiuYuancheng/Threats_2_MITRE_AI_Mapper.

[11] B. Zhang et al., "AttackSeqBench: Benchmarking Sequential Threat Behavior Generation from Large Language Models," *arXiv preprint arXiv:2503.03170*, 2025. [Online]. Available: https://arxiv.org/pdf/2503.03170v1.

- TTPs extraction from text
- Not designed for unstructured text
- Cannot directly parse natural language narratives to extract attack steps or TTPs
- Achieving automation for its input type.
- Limited Training Data and Pattern Recognition
- Lack of Contextual Understanding
- IOC focused
- Traditional ML's attack flow accuracy is limited by its finite labeled data, while LLMs leverage vast knowledge to generate more diverse, novel, and contextually coherent scenarios.
- Automatic and Accurate Extraction of Threat Actions from Unstructured Text of CTI Sources Relies on fixed grammatical patterns (SVOI)
- Limited in capturing semantic nuance, context beyond sentence level, and implicit relationships
- Less adaptable to new or unseen attack descriptions.

DAQ

Feeds

Docs Storage

Extract

Nearly 400 data feeds are currently being monitored.

NER

Labelling

Classifying

Accurate labeling is essential for effective document classification.

RAG+

Vector Search

Docs Related

n8n

LLM Model

n8n

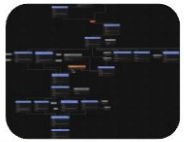
LLM Model

Storage

Each candidate document will retrieve *four* additional related documents from the document storage. This enables the system to process documents that were not initially selected as candidates.

OUTPUT

MITRE | ATT&CK



Attack Flow

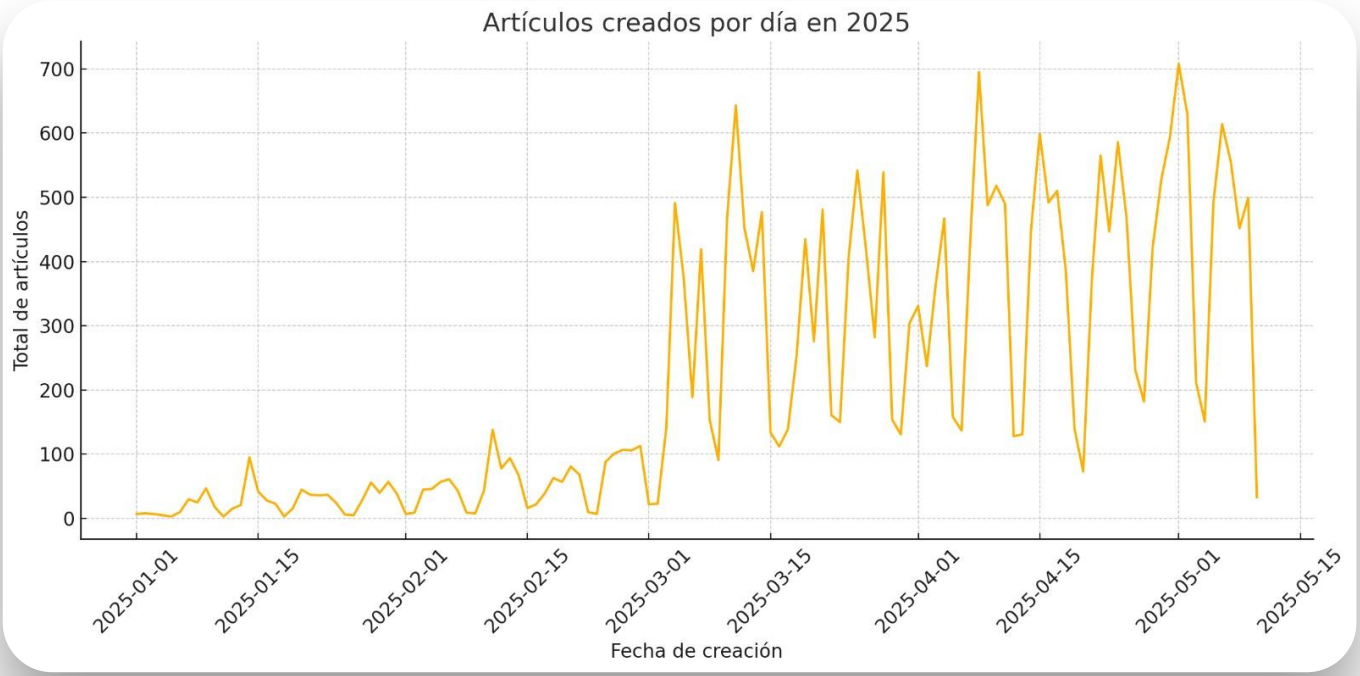
Knowledge Database

The system will generate an attack flow database compliant with the MITRE ATT&CK Flow standard

USE CASES

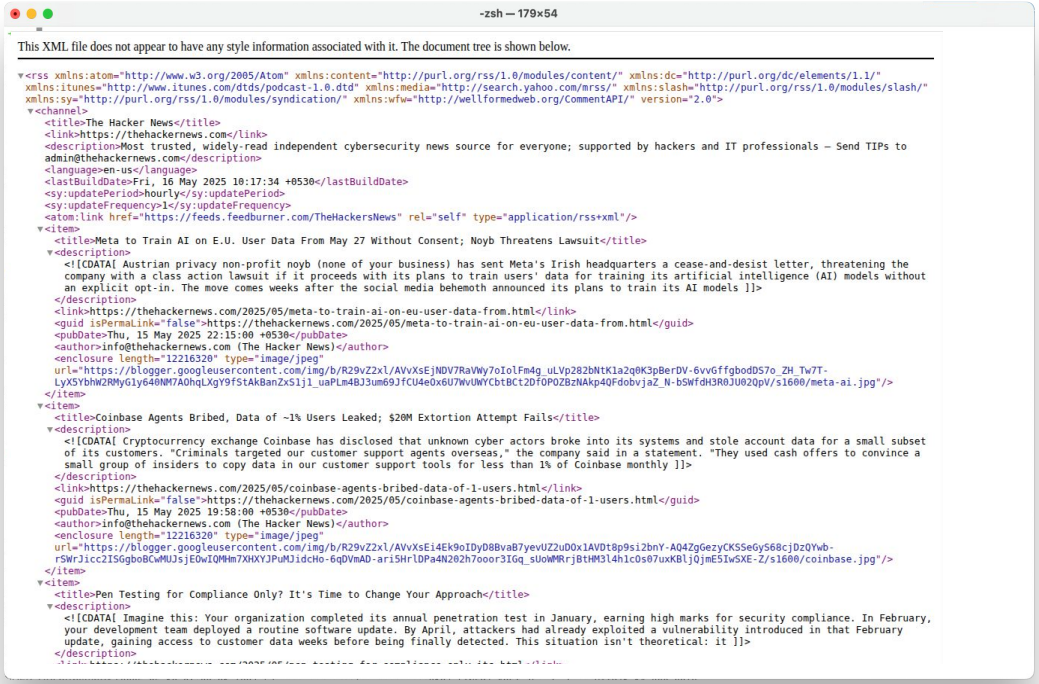
SIEM

Data Acquisition (DAQ): Article Volume



400 new documents per day from almost 400 feeds

Data Acquisition (DAQ): RSS Feed

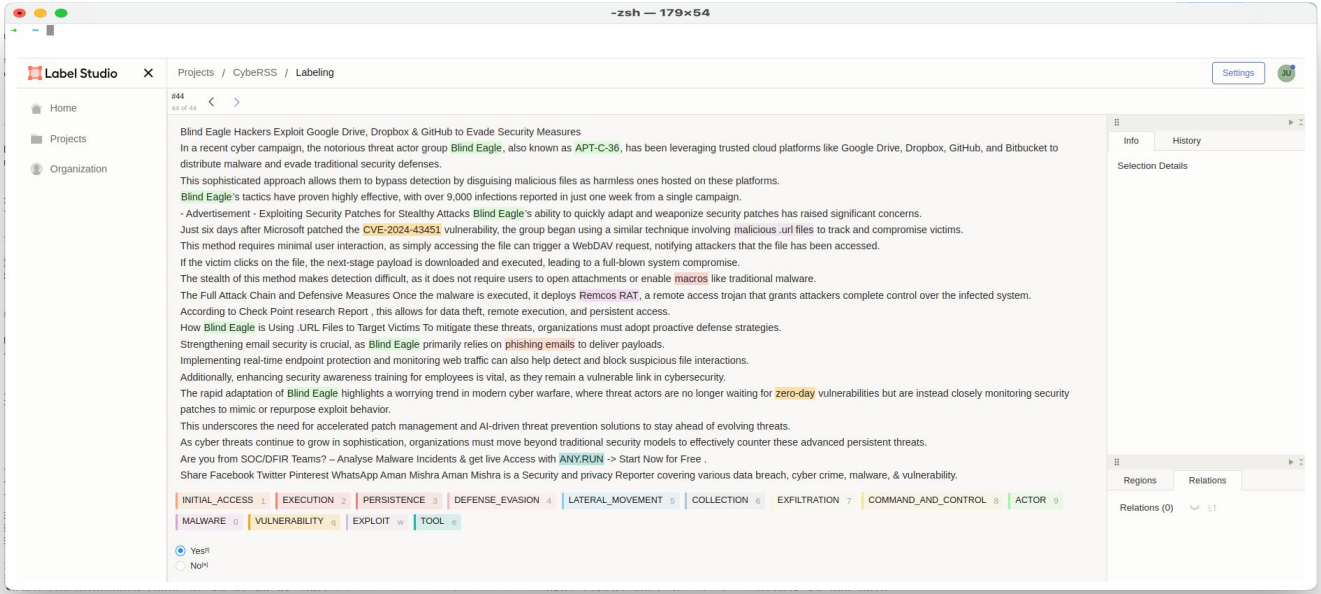


We pull in RSS feeds—glitches and occasional issues like embedded code—to keep our cybersecurity info fresh.

Turning Raw Content into Clean Input

- Extract main content (remove menus, ads, etc.)
- spaCy + NLP models
- Output: clean, structured text ready for analysis

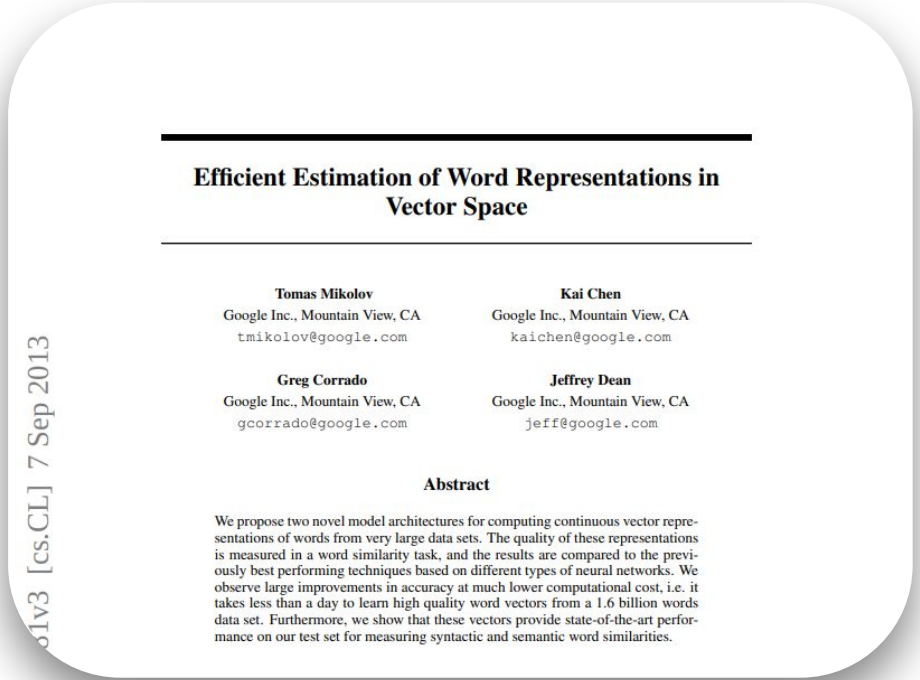
Automatic Structured Annotation



- Using **Label Studio** for automatic entity annotation in CTI texts
- Integrated with models to suggest tags automatically
- Custom NER model detects: **actors, vulnerabilities, tactics, tools**, etc.
- We only need to review the entities suggested by the model

From Text Input to Ranked Results

This post might ring a bell—it changed the game by showing how to represent words as vectors and even do meaningful math with them.



Efficient Estimation of Word Representations in Vector Space

Tomas Mikolov
Google Inc., Mountain View, CA
tmikolov@google.com

Kai Chen
Google Inc., Mountain View, CA
kaichen@google.com

Greg Corrado
Google Inc., Mountain View, CA
gcorrado@google.com

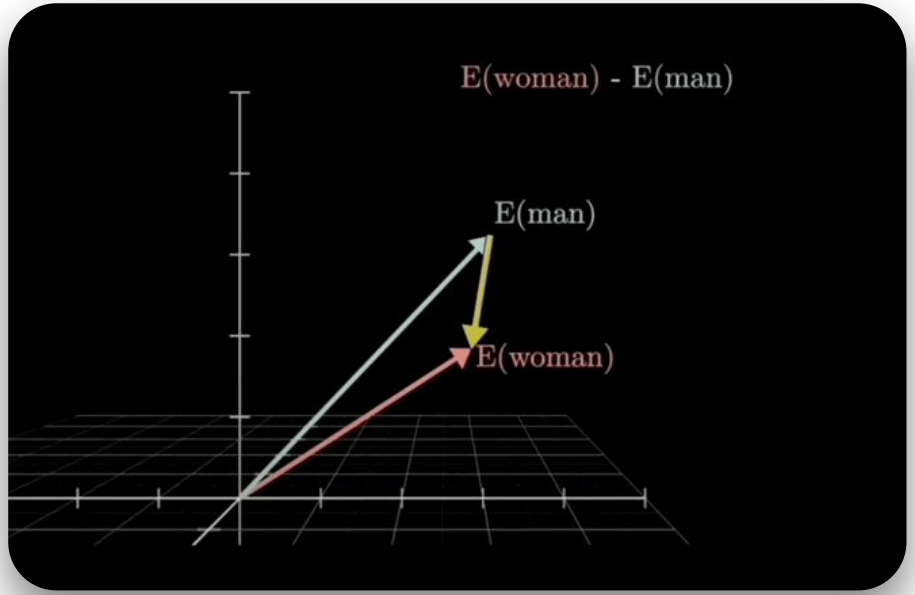
Jeffrey Dean
Google Inc., Mountain View, CA
jeff@google.com

Abstract

We propose two novel model architectures for computing continuous vector representations of words from very large data sets. The quality of these representations is measured in a word similarity task, and the results are compared to the previously best performing techniques based on different types of neural networks. We observe large improvements in accuracy at much lower computational cost, i.e. it takes less than a day to learn high quality word vectors from a 1.6 billion words data set. Furthermore, we show that these vectors provide state-of-the-art performance on our test set for measuring syntactic and semantic word similarities.

https://arxiv.org/pdf/1301.3781

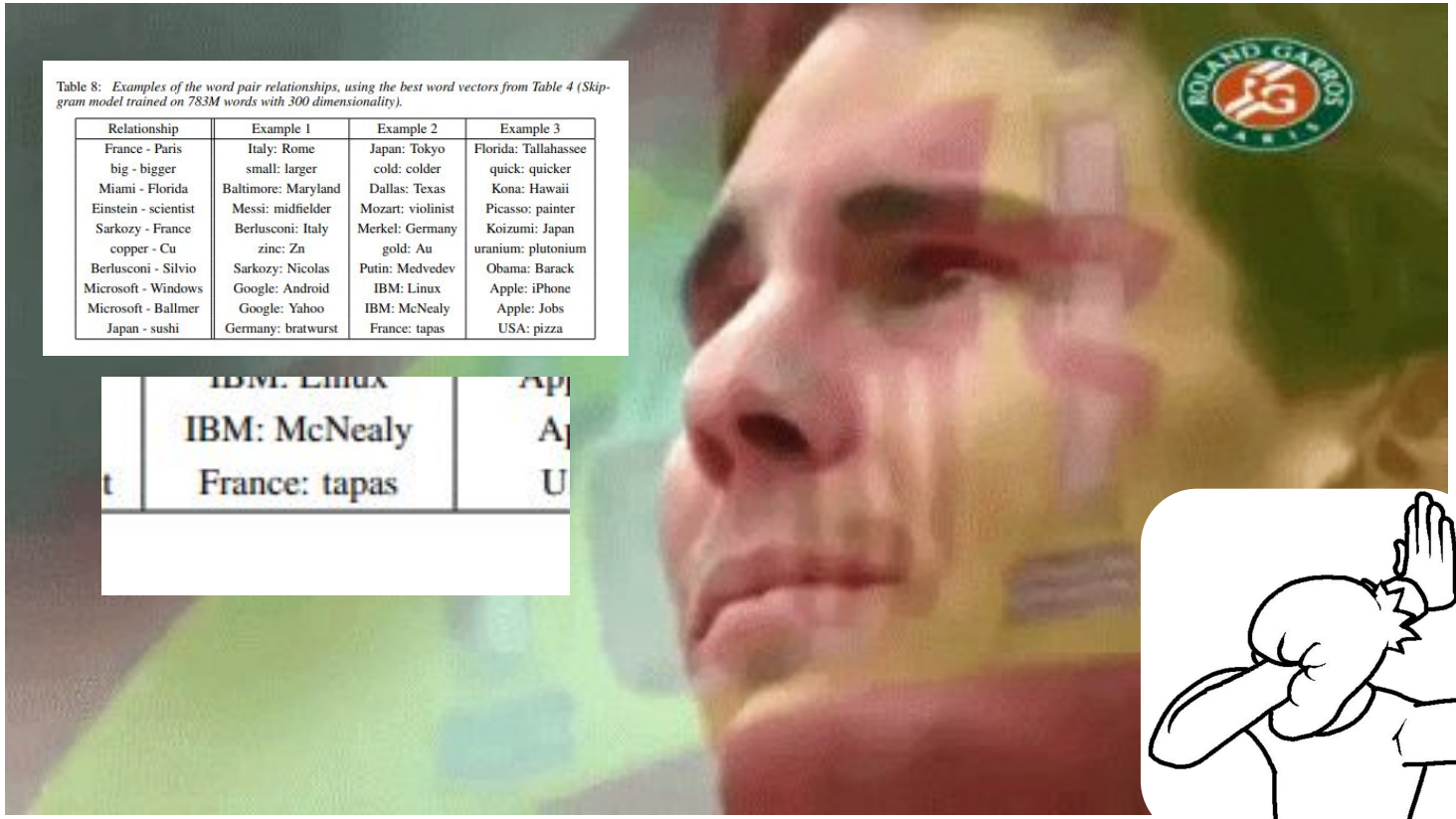
From Text Input to Ranked Results



Word2Vec models capture word meanings with math—like how king - man + woman gets you close to queen—revealing not just meanings, but also cultural, geographic, even foodie associations.

Visualizing transformers and attention | Talk for TNG Big Tech Day '24. Grant Sanderson. https://www.youtube.com/watch?v=KJtZARuO3JY

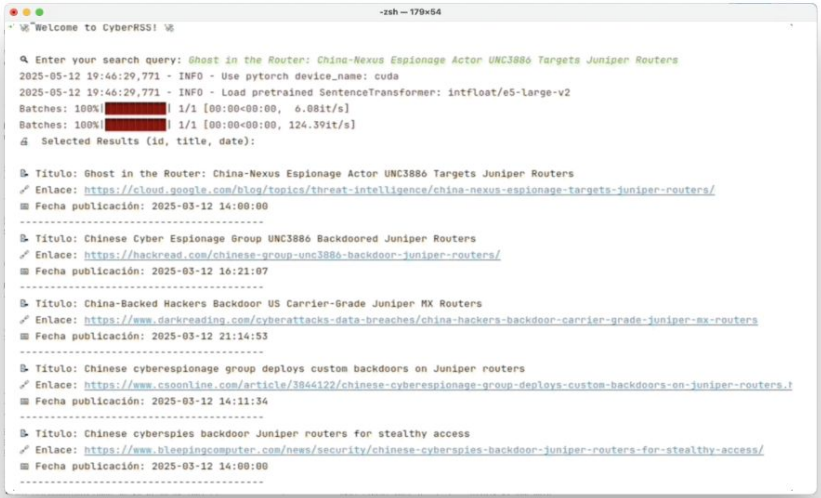
From Text Input to Ranked Results



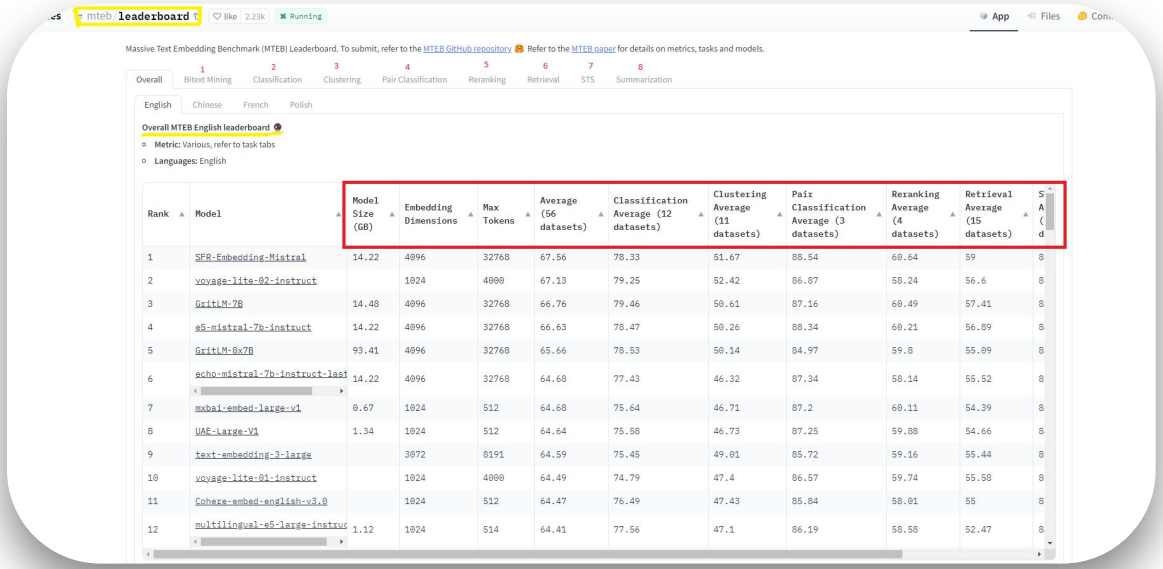
From Text Input to Ranked Results

It lets us compare texts by meaning, not just matching words—helping us rank articles, spot similarities.

Here, the system turns a headline into a vector and finds semantically similar articles, even if they use different words or angles.



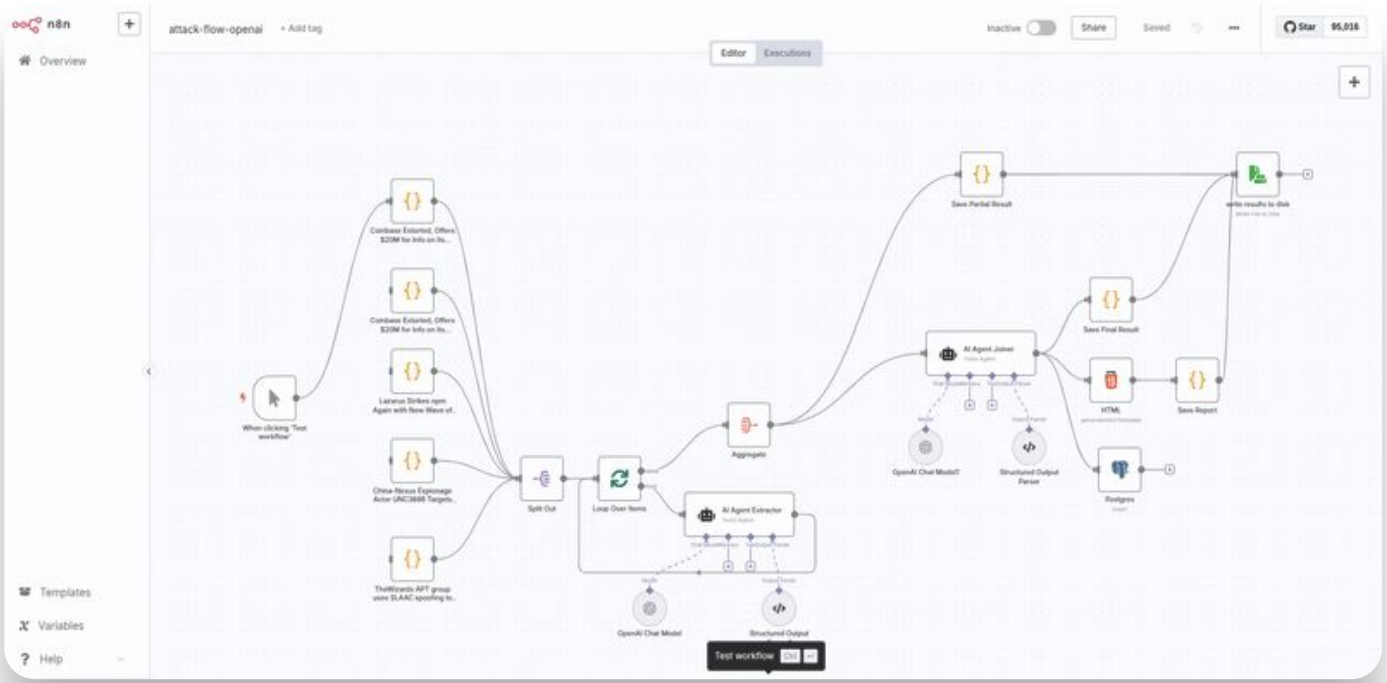
Comparing Embedding Models (MTEB Leaderboard)



https://huggingface.co/spaces/mteb/leaderboard

Benchmarks show models like e5-large and bge excel at semantic search—just what we need to organize and enrich articles.

From article content to structured attack flow



Constructing the Attack Flow

UNC3886 Chinese APT Group Installs Custom Backdoors on Juniper Routers for Persistent Espionage

Published: 12/03/2025

Category: APT

Summary: Chinese APT UNC3886 compromised Juniper routers since 2024, using custom backdoors to maintain persistent espionage access to defense, telecom, and ISP networks.

Actor

Actor: UNC3886

Aliases: China-nexus threat group | APT41 (tooling overlap) | UNC3886

Origin: China

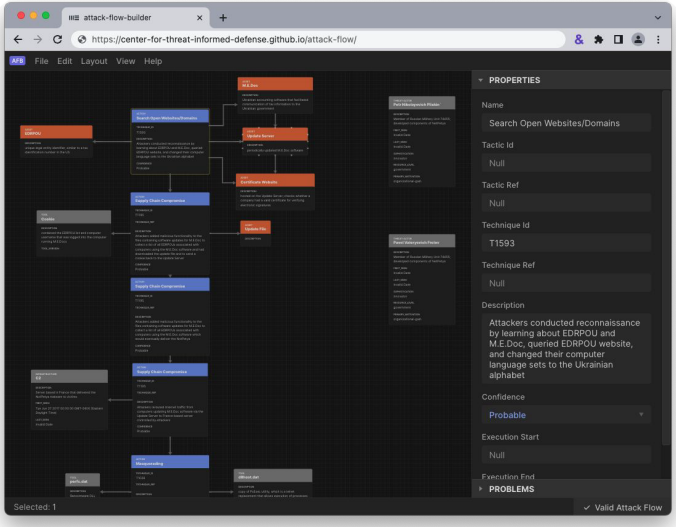
Motivation: Cyber espionage targeting defense, technology, telecommunications, and ISPs via persistent access to network devices.

Discovered by: Mandiant

Target Profile

Victim Organizations:

- Juniper Networks



Attack Flow https://ctid.mitre.org/projects/attack-flow

Use Cases

What are the potential uses for the knowledge produced?

- Threat Detection & Alert Tuning
- Threat Hunting
- Incident Response & Root Cause Analysis
- Security Content Validation & Adversary Emulation
- SOC Playbook Automation & Enrichment
- Threat Intelligence Enrichment
- AI/ML Model Training for Detection

Use Case	Challenge Addressed	How an Attack Flow Database Helps	Real Life
Threat Detection & Alert Tuning	High alert volume, false positives	Matches alerts with known malicious sequences to prioritize true positives	Triage Re-Rank, Attack Pattern.
Threat Hunting	Lacks context for proactive searches	Provides hypothesis-driven paths based on known TTP sequences. Present and Past.	Hunting for post-exploitation actions. Rollback through know attack flows
Incident Response & Root Cause Analysis	Hard to reconstruct multi-stage attacks	Maps observed evidence to known flows, accelerating investigation	Mapping events to Hafnium attack flow to identify initial access method
Security Content Validation & Adversary Emulation	In some cases detections could fail to detect multi-stage attacks and Simulations are too generic	Simulates real attacks to validate security control effectiveness. Enables red teams to emulate real-world attack chains; helps blue teams test detection/response	Combine with Attack Simulation Solutions
SOC Playbook Automation & Enrichment	Playbooks lack context	Enhances playbooks with dynamic branches based on attack stage and context	Improve Automatic Case Creation and allow for Threat Hunting
Threat Intel & Enrichment	IOCs lack context	Places indicators and TTPs within a larger narrative for better relevance	Context for OSINT/OSTI IOCs
Model Training for Detection	Models need labeled, high-fidelity attack data	Supplies curated attack sequences to train and validate machine learning detection models	Using known flows to build a labeled dataset for behavioral anomaly detection models

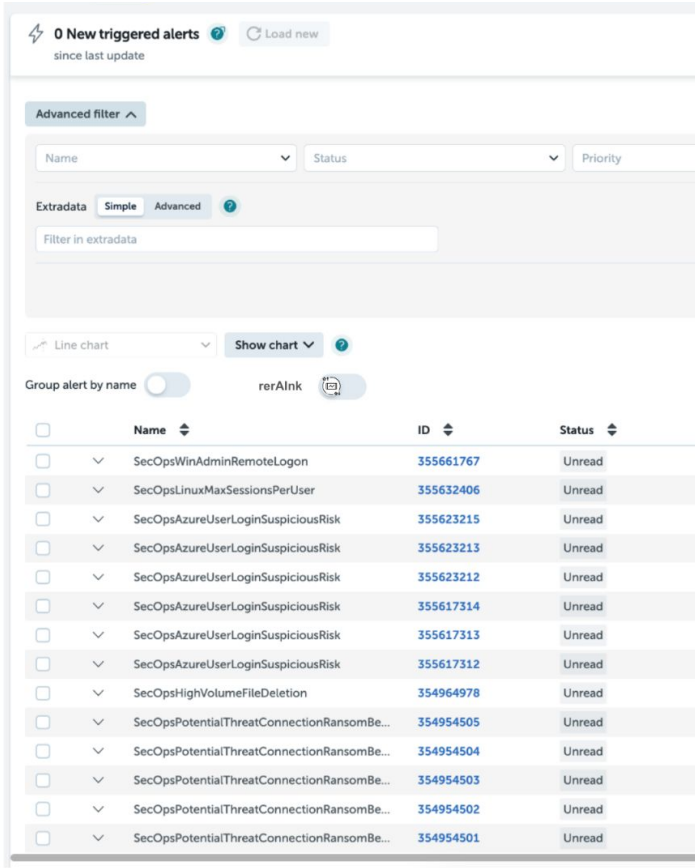
Alert Triage Re-Rank

Threat Detection & Alert Tuning

High alert volume, false positives

Matches alerts with known malicious sequences to prioritize true positives

Triage Re-Rank, Attack Pattern.



SOC Automation & Enrichment

SOC Playbook Automation & Enrichment	Playbooks lack context	Enhances playbooks with dynamic branches based on attack stage and context	Improve Automatic Case Creation and allow for Threat Hunting

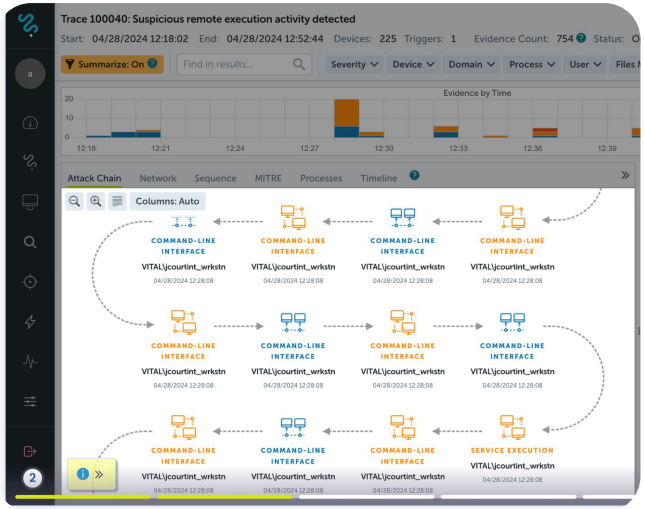
MITRE | ATT&CK®

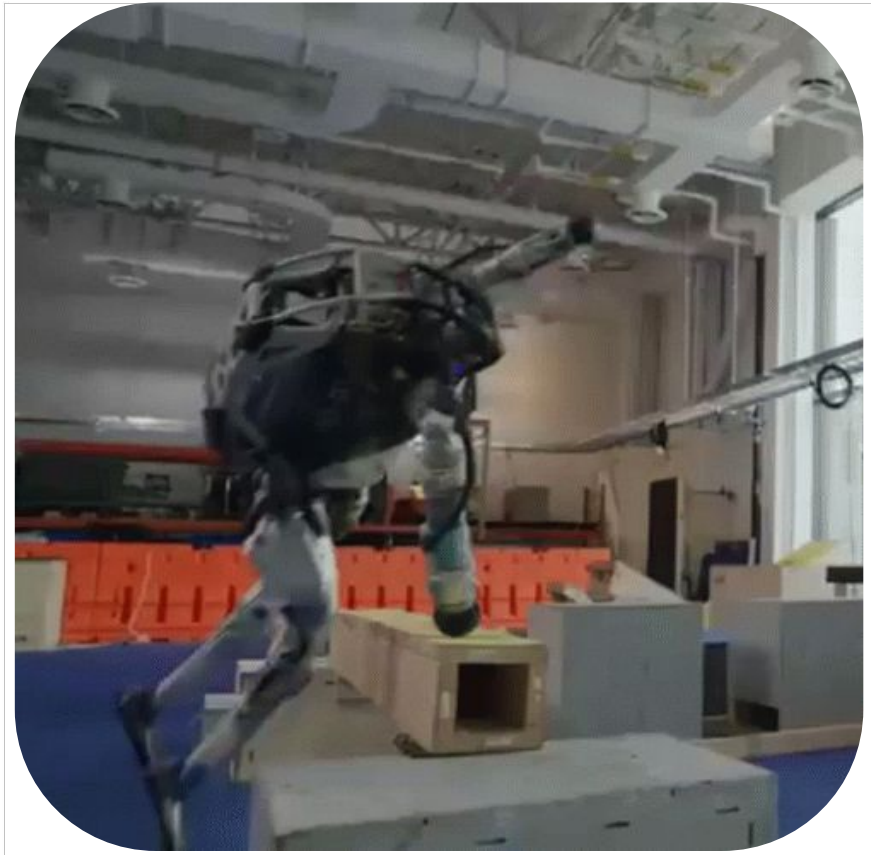
Technique Inference Engine

Given a list of observed techniques, infer the next most likely techniques.

Threat Hunting

Threat Hunting	**Lacks context for proactive searches**	**Provides hypothesis-driven paths based on known TTP sequences. Present and Past.**	**Hunting for post-exploitation actions. Rollback through know attack flows**





CONFERENCE
MADRID 2025

Fails & Future

All that glitters is not gold

- Ads
- Vulnerability advisory lack of details

Named Entity Recognition (NER)

Future Uses

- Automatically filter content based on detected entities
- Enrich the database with structured information
- Automatically map content to **MITRE ATT&CK** tactics and techniques
- Generate training datasets for new models
- Improve prompts

DAQ & SELF-IMPROVEMENT

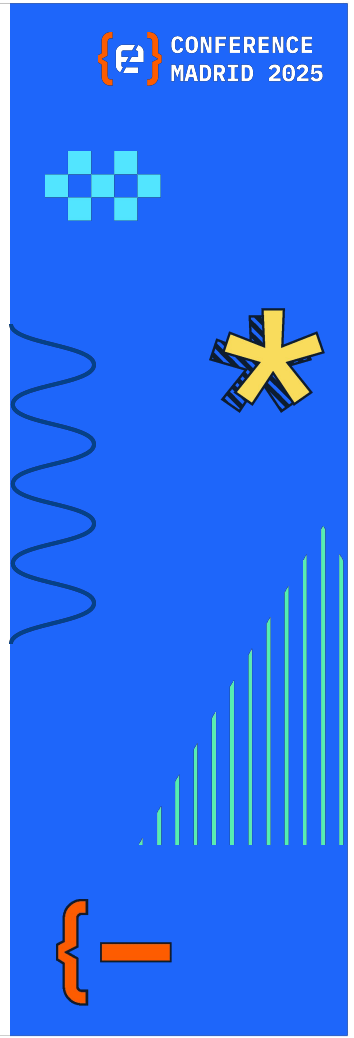
Future Improvements

- Result validation
- Auto-discovery of feeds
- Feed ranker and new acquisition
- Attack Pattern prediction

Production Ready

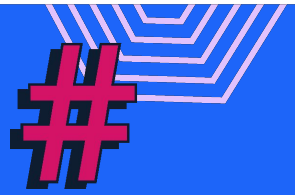
Future Improvements

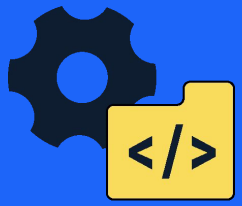
- UI + API
- Self deployment



{codemotion} CONFERENCE MADRID 2025

Don't forget to rate the talk!





{codemotion} CONFERENCE MADRID 2025

Thanks!