UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

itee PhD information technology electrical engineering

DIETI UNINA

Pellegrino Casoria

Advanced security methodologies for proactive simulation, detection and response to targeted attacks

Tutor: Simon Pietro Romano

Cycle: XXXIX

Year: Second

itee PhD information technology electrical engineering

accenture

My background

- **MSC DEGREE:** Computer Engineering
- **RESEARCH GROUP:** Dipartimento di Ingegneria elettrica e delle Tecnologie dell'Informazione (DIETI)
- **PHD START DATE:** 02-Nov-2023
- **YEAR:** Second
- **PARTNER COMPANY:** accenture
- **TUTOR:** Simon Pietro Romano
- **SCHOLARSHIP TYPE:** N/A

Research field of interest

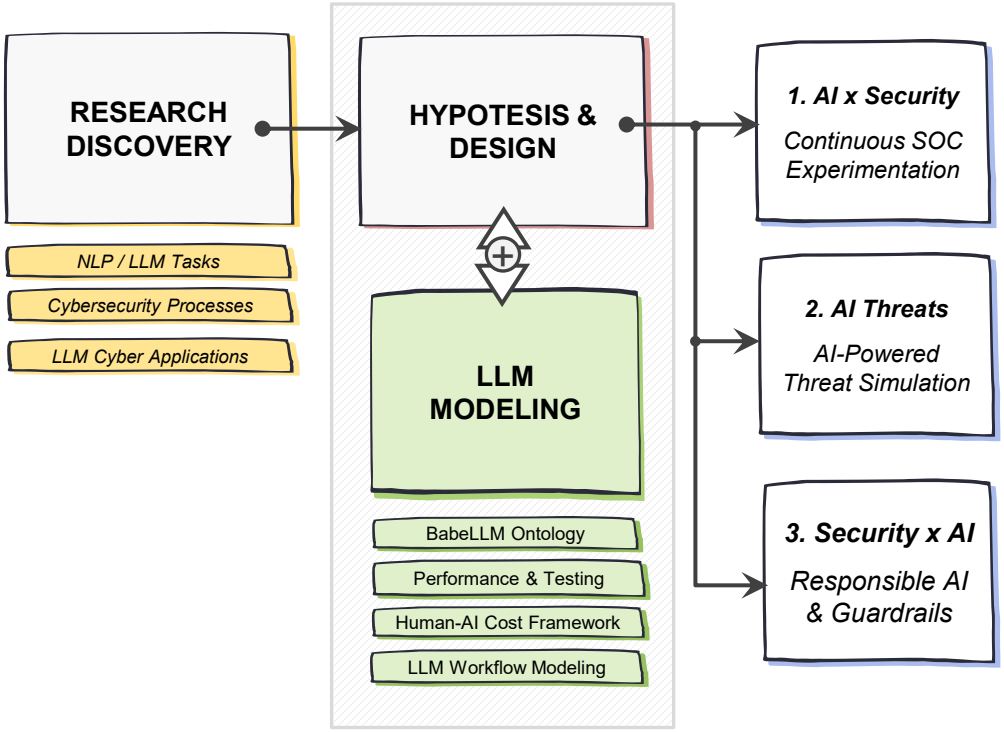
LOST IN TRANSLATION: WHEN LLM RESEARCH MISSES REAL-WORLD IMPACT

My research focuses on innovative application strategies for LARGE LANGUAGE MODELS to address emerging SECURITY CHALLENGES, while MANAGING EXPERIMENTAL NOISE and ensuring reliability and reproducibility of results

Research activity: Overview

CUTTING THROUGH THE NOISE: TOWARDS RELIABLE LLM APPLICATIONS FOR CYBERSECURITY

Y2 research has focused on **modeling the complexity** of LLM-based cybersecurity applications and **experimentally validating** new use cases to bridge prototyping into controlled adoption.



PHD YR2 START

PHD COURSES & SEMINARS ATTENDANCE

11.4 CFUs

SYSTEMATIC LITERATURE REVIEW

1170+ Publications

ENGINEERING LLM x CYBERSECURITY

5 Products

CONFERENCES & TUTORSHIP

- ***ITASEC x SERICS 2025***
- ***Red Hot Cyber 2025***
- ***Rome Technopole***

CO-INNOVATION WITH CLIENTS

powered by accenture

PHD YR2 END

almost there... ^_^''

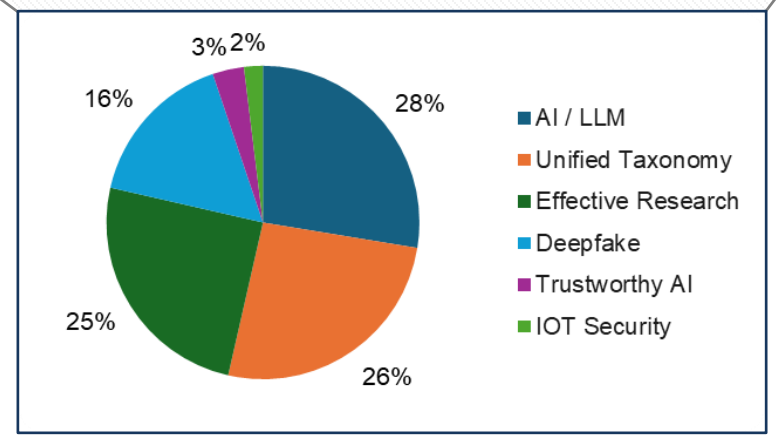
YEAR 2 RESEARCH TIMELINE

Summary of study activities

FROM LEARNING TO ACTUAL RESEARCH ON LARGE LANGUAGE MODELS FOR CYBERSECURITY

ACTIVITY CREDITS OVERVIEW

REPORT	COURSES	SEMINARS	RESEARCH	TUTORSHIP	TOTAL
YEAR 1	**24**	**5.1**	**33.2**	**0**	**62.3**
Bimonth 1	0	0	10	0	10
Bimonth 2	8	0.2	6	0	14.2
Bimonth 3	0	0.2	8	0	8.2
Bimonth 4	0	0	8	1.6	9.6
Bimonth 5	3	0	6	1.6	10.6
Bimonth 6	0	0	9	0	9
YEAR 2	**11**	**0.4**	**47**	**3.2**	**61.6**
TOTAL	**35**	**5.5**	**80.2**	**3.2**	**123.9**
Expected	**30 - 70**	**10 - 30**	**80 - 140**	**0 – 4.8**	



ATTENDED CONFERENCES & TUTORSHIP

- [ATTENDEE] ITASEC25 & SERICS, JOINT NATIONAL CONFERENCE ON CYBERSECURITY (Bologna)
- [SPEAKER] RED HOT CYBER CONFERENCE 2025 (Rome)
  >> Doxing with Langflow: Are We Building the End of Privacy?
  >> Social Engineering 2.0: Uncovering Emerging Deepfake Threats
- [MENTOR] ACCENTURE CYBER HACKADEMY, 4th ed. (Naples)
  >> GenAI: The revolution in corporate language
- [TUTORSHIP] ROME TECHNOPOLE, MICROCREDENZIALI DIGITAL (Rome)
  >> Cybersecurity Fundamentals Module (20h)
  >> Cybersecurity Advanced Module (20h)

in collaboration with:

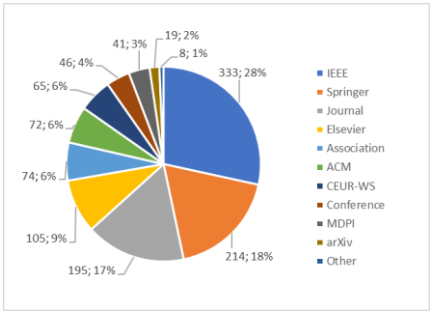
BabeLLM: Systematic Literature Review

AN UMBRELLA SYSTEMATIC REVIEW FOR LLM-BASED CYBERSECURITY TASKS

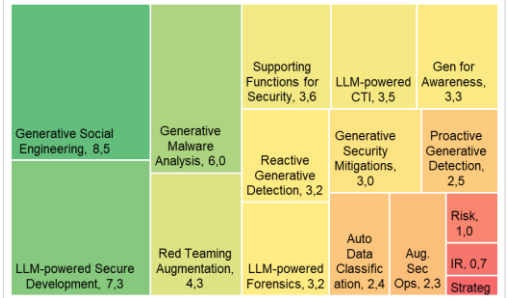
Umbrella Systematic Literature Review (SLR) of 1400+ unique publications over the last 10 years on **NLP Tasks** (RQ.1), **Cybersecurity Processes** (RQ.2) and **LLM Cyber Applications** (RQ.3).

ID	AREA	TIME	# IDENTIFY	# SCREEN	# SELECT
RQ.1	NLP Tasks	2015-2025	587	44	27
RQ.2	Cybersec Processes	2015-2025	658	62	6
RQ.3	LLM Cyber Apps	2015-2025	247	36	13
-	TOTAL	-	**1492**	**142**	**46**

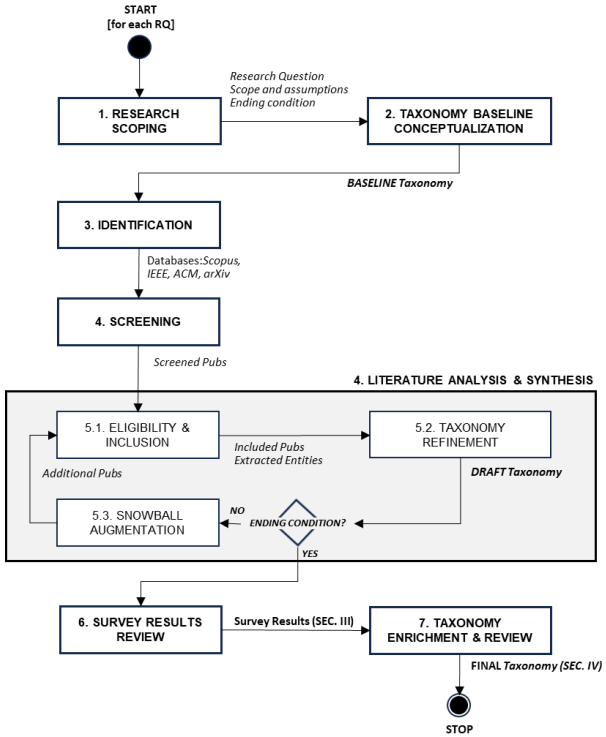
SOA DATASET BY PUBLISHER (OVERALL)



LLM-APPLICATIONS LV2 CATEGORIES FREQ. RATE



PENDING PUBLICATION: *P. Casoria, S.P. Romano, "BabeLLM: Umbrella Systematic Review for a Standardized Taxonomy of LLM-based Cybersecurity Tasks", Computer Science Review, Elsevier, pending review as of October 2025.*



PROPOSED METHODOLOGY *

* ***Sources:*** *Webster & Watson; Brocke et al; PRISMA; Nickerson et al.*

ID	AUTHORS	YR	FOCUS	TYPE	PUB	AUX		DET		OFF	RES			STR	RSP	#ENTITY	#MATCH
						S	C	P	R		X	D	W				
[149]	Chen et al.	2024	Task	journal	Elsevier	●	●	↔	●	●	↔	↔	○	○	↔	4 (Sub: 10)	35 (38%)
[150]	Xu et al.	2025	Task, Model, Dataset	preprint	arXiv	↔	↔	●	●	●	↔	○	●	●	↔	5 (Sub: 29)	33 (35%)
[151]	Hasanov et al.	2024	Task, Threat	journal	IEEE	↔	↔	↔	↔	↔	●	↔	○	↔	●	4 (Sub: 6)	33 (35%)
[152]	Motlagh et al.	2024	Task, Model	preprint	arXiv	↔	●	○	↔	●	●	●	↔	○	●	2 (Sub: 18)	32 (34%)

BabeLLM: The OWL Ontology

AN UMBRELLA SYSTEMATIC REVIEW FOR LLM-BASED CYBERSECURITY TASKS

To demonstrate practical applicability, we developed **a machine-readable OWL implementation named BabeLLM**, and validated it through **GPT-based use cases** for automated inference.

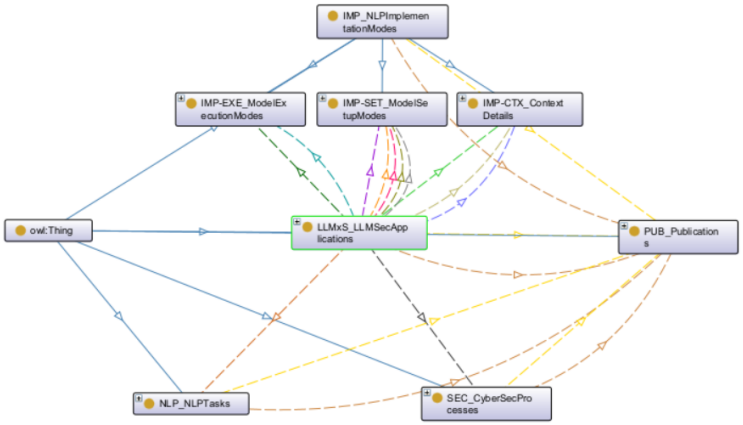
Classes	81	Individual count	303
Object properties	18	ObjectPropertyAssertion	2,957
Data properties	29	DataPropertyAssertion	2,856
Annotation Properties	10	AnnotationAssertion	3,681
Total Axioms			**10,478**

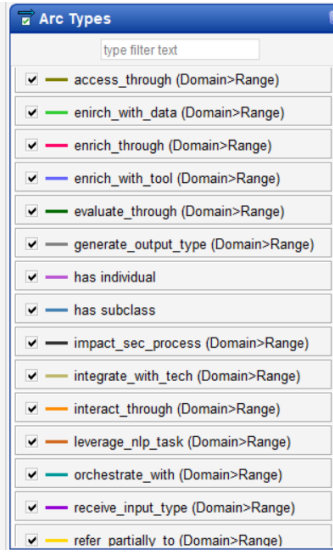
DATASETS: *OWL ontology, scripts and datasets are publicly available upon publication on* ***GitHub*** *and* ***Mendeley Data****.*

IMPLEMENTED LLM-BASED USE CASES

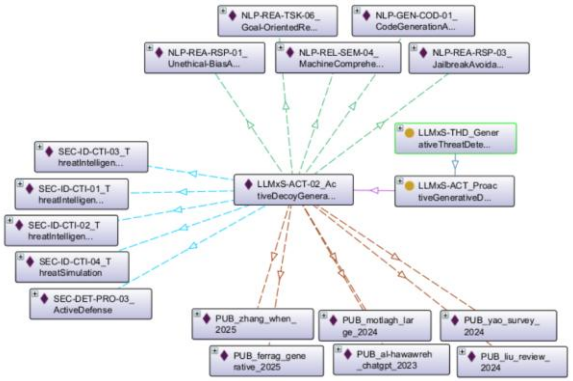
1. LLM-powered Data Properties Enrichment
2. LLM-powered Object Properties Inference

BABELLM ONTOLOGY GRAPH (PROTÉGÉ)



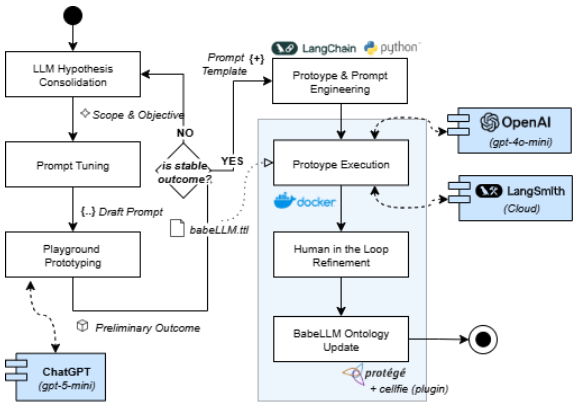


ontology integration with GPT model



SAMPLE OBJECT PROPERTY RESULTS: ACTIVE DECOY GENERATION

automatic OWL metadata inference



LLM-BASED EXPERIMENTATION APPROACH

Products

CONTINUOUS EXPERIMENTATION ON LARGE LANGUAGE MODELS

[P1]	**(Publication) SYSTEMATIC LITERATURE REVIEW:** P. Casoria, S.P. Romano, *"BabeLLM: Umbrella Systematic Review for a Standardized Taxonomy of LLM-based Cybersecurity Tasks"*, Computer Science Review, Elsevier, pending review as of October 2025. **MAIN TECHS:** Overleaf ELSEVIER draw.io **JOURNAL:** Computer Science Review (Elsevier)
[P2]	**(Ontology) BABELLM ONTOLOGY:** Machine-readable representation of LLM-based cybersecurity tasks, integrating concepts from natural language processing, cybersecurity processes, and model implementation modes. **MAIN TECHS:** protégé cellfie WebVOWL GitHub MENDELEY **REPOSITORY:** Github, Mendeley Data
[P3]	**(Prototype) LLM-POWERED METADATA INFERENCE:** Langchain integration with GPT models to automatically infer BabeLLM ontology metadata (i.e., data and object properties) and enrich NLP tasks, cybersecurity processes and LLM cybersecurity applications. **MAIN TECHS:** LangChain LangSmith OpenAI python docker **JOURNAL:** Computer Science Review (Elsevier)
[P4]	**(Prototype) CYBER-HUMINT RECONNAISSANCE:** Automated pipeline for scraping social network data to build user profiles through LLM-based interpretation of publicly exposed information. **MAIN TECHS:** Langflow Selenium OpenAI **CONFERENCE:** Red Hot Cyber 2025
[P5]	**(Prototype) PROMPT GUARDRAILING:** Integrated workflow for detecting and mitigating anomalous LLM prompts - such as prompt injections or jailbreak attempts - using contextual and semantic filtering techniques. **MAIN TECHS:** Streamlit docker Gemini PRESIDIO Guardrails AI LLaMA NVIDIA **CONFERENCE:** Accenture Cyber Hackademy (4th edition)

YEAR 1

YEAR 2

PHD YR1 START

PHD YR2 START

LLM INDUSTRY & MARKET ANALYSIS

PHD COURSES & SEMINARS ATTENDANCE

STATE OF THE ART STUDY FOR LLM/GENAI SECURITY

SYSTEMATIC LITERATURE REVIEW

PHD COURSES & SEMINARS ATTENDANCE

ENGINEERING LLM x CYBERSECURITY

SOLUTION PROTOTYPING

CONFERENCES & TUTORSHIP

CONFERENCES ATTENDANCE

CO-INNOVATION WITH CLIENTS

PHD YR1 END

PHD YR2 END

Next Year

ROAD TO YEAR THREE

- Keep publishing to conferences and journals
- Period abroad in international research institutions
- AI Security performance, cost and modeling framework
- GenAI experimentation on augmented SOC
- Trustworthy AI and Deepfake simulation
- IOT hacking and next gen communications
- Finish PhD! :)

Thank You!