10th LASIGE WORKSHOP

21 MAY 2025

POSTER SESSION

LASIGE driven by excellence

fct Fundação para a Ciência e a Tecnologia

Ciências ULisboa

ULISBOA

A Comparative Study of Reliable Predictive Models for Energy-Efficient MPC in Buildings

Alexandre Geraldo, António Casimiro, Alan Oliveira, José Cecílio, Pedro Ferreira

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

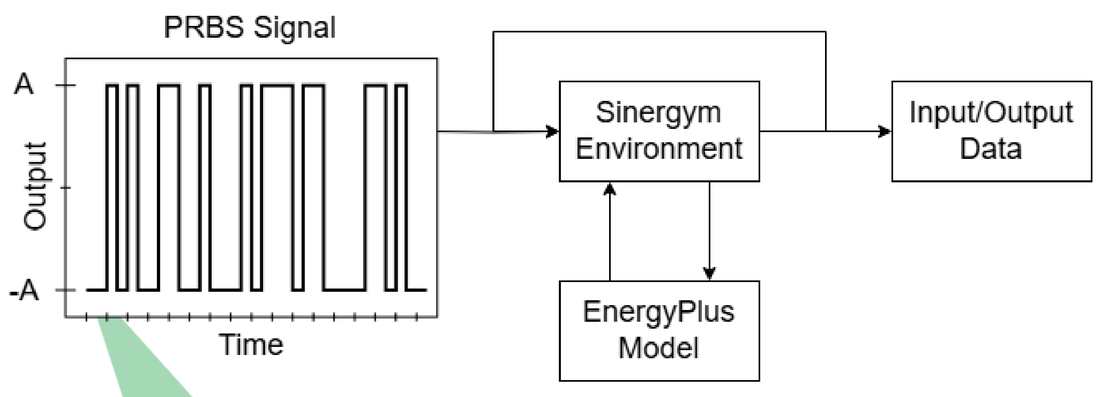
MOTIVATION & GOALS

Buildings consume ~40% of energy in the EU, yet most remain inefficient. Model Predictive Control (MPC) offers promising improvements for energy efficiency. However, its performance relies heavily on the accuracy of the predictive model used.

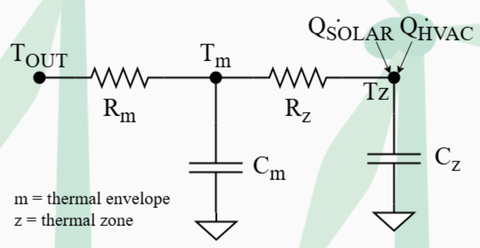
- Compare different models in terms of their forecasting reliability
- Determine how long each model remains reliable for MPC

DATA, MODELS, and METHODOLOGY

Input/ Output data generation:



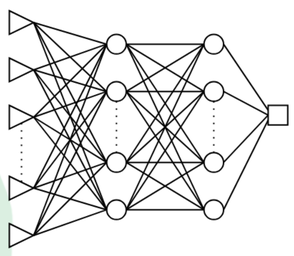
RC NETWORK



- **Simple 2R2C** physical model
- Parameters tuned via Particle Swarm Optimization (PSO)

Build RC network structure → Setup a PSO and define search space bounds → Multiple runs

NEURAL NETWORK

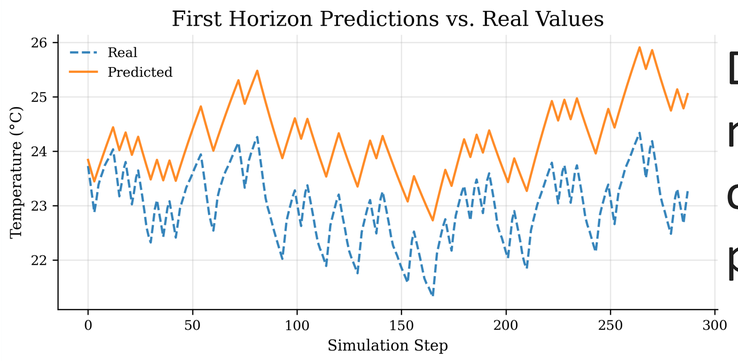


- **Feedforward ANN**
- Up to 24h lagged inputs: in/out temperatures, setpoints and solar radiation

Build NN structure → Try hidden layer and batch sizes → Grid search model design

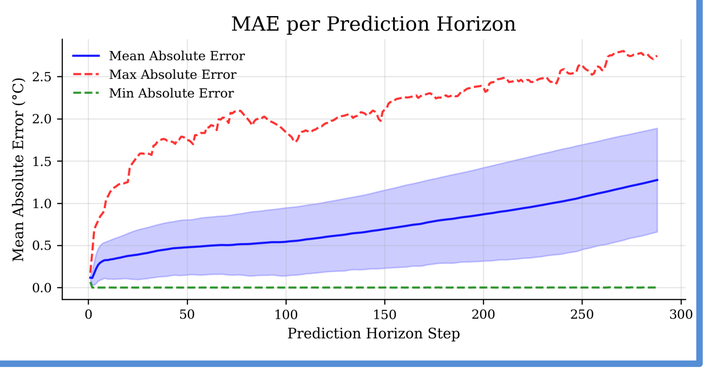
Evaluation of predictive reliability in a 24h prediction horizon

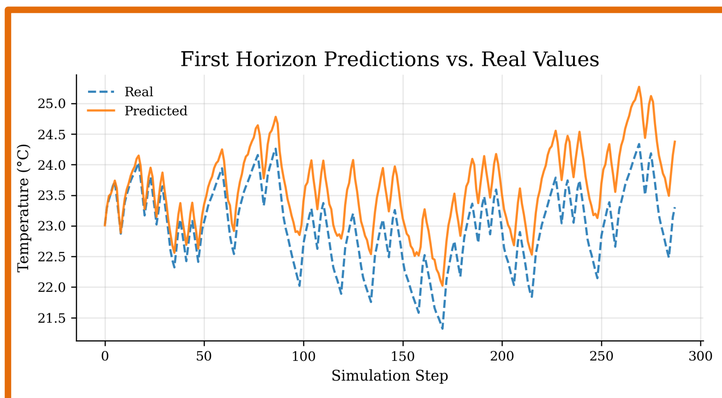
RESULTS



Deviates noticeably after only a small period of time

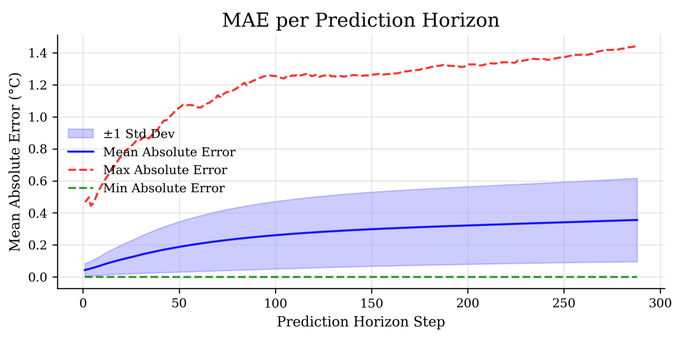
Steep MAE growth indicates short prediction reliability





Maintains accurate predictions over a greater period of time

Flatter MAE curve reflects higher long-term reliability



CONCLUSIONS

- The ANN model achieves significantly better long-term accuracy
  - MAE integral of 77
  - 207 for the RC model (indicates higher cumulative error)
- ANN remains reliable for the entire 24 hours
- RC surpasses the 1ºC threshold (for a typical ±0.5 °C sensor error) after ~9.5 hours

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Acknowledgements: This work was supported by the LASIGE Research Unit, ref. UID/00408/2025.

From 2D Nets to 3D Shapes: A GNN Approach

***N. Schmitz*[1]**, *N. A. M Araújo*[2,3] and Catia Pesquita[4]

[1]Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, P-1749-016 Lisboa, Portugal
[2]Centro de Física Teórica e Computacional, Faculdade de Ciências, Universidade de Lisboa, P-1749-016 Lisboa, Portugal
[3]Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, P-1749-016 Lisboa, Portugal
[4]LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, P-1749-016 Lisboa, Portugal

nmaraujo@fc.ul.pt

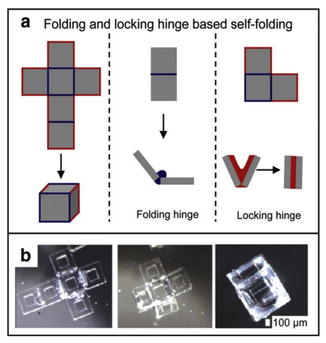
Ciências ULisboa

LASIGE driven by excellence

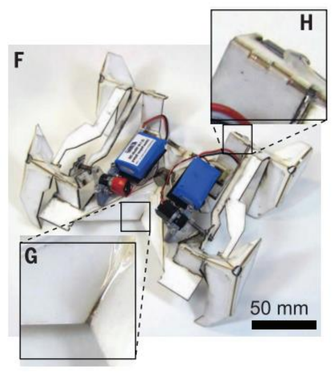
CFTC CENTRO DE FÍSICA TEÓRICA E COMPUTACIONAL

Why we care

The assembly of **3D polyhedral shells** is important for many areas, from robotics[1] to medicine[2].



Self-folding polymeric containers for encapsulation[2]



A self-folding crawler[1]

Current state: Most existing approaches use deterministic algorithms, which become **computationally infeasible as face count increases due to exponential time complexity.**

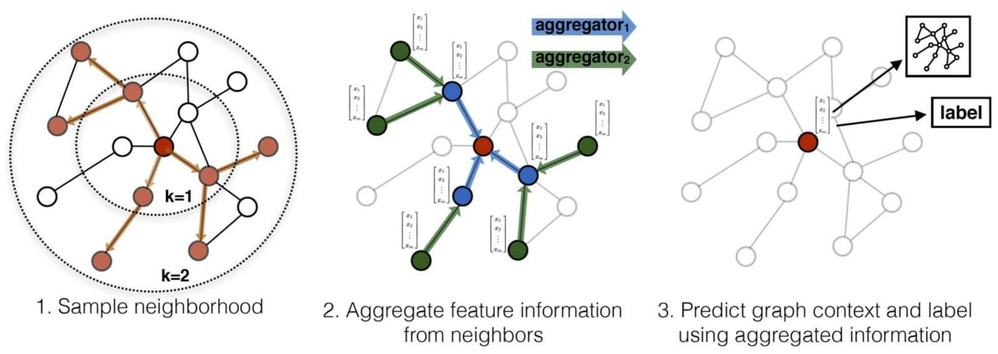
Motivation: ML-methods, especially **GNNs**, **replace costly rule-based algorithms** with **fast, scalable inference on graph-structured data.**

Objective: Develop a novel graph neural network (GNN) architecture for predicting the folding of polyhedra from their 2D unfolded representations into valid 3D structures.

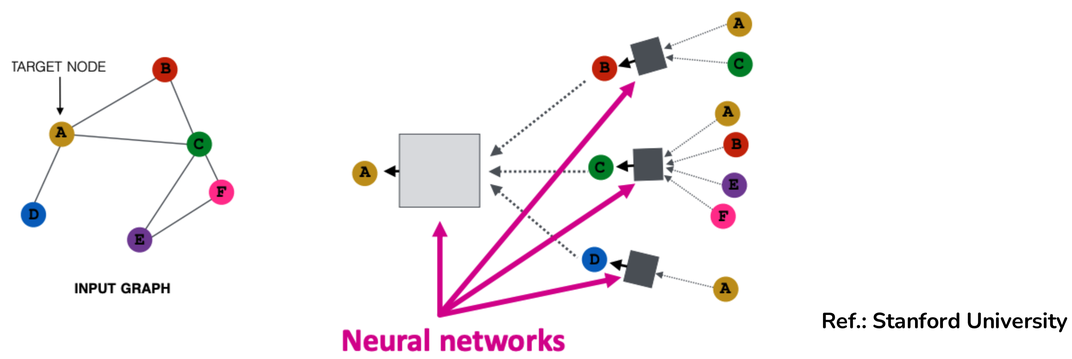
Predicting 3d polyhedral shells

To correctly predict possible 3d polyhedral shells, we employ **Graph Neural Networks (GNNs)** to predict **missing links in a graph**.

Idea of the GNN pipeline:

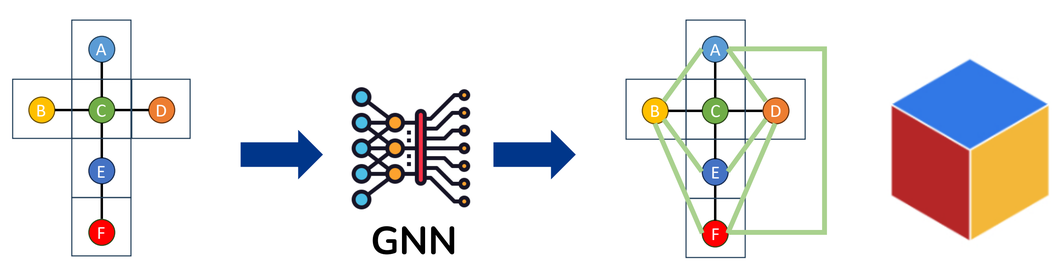


1. Sample neighborhood
2. Aggregate feature information from neighbors
3. Predict graph context and label using aggregated information



As one **GNN** consists of multiple neural networks, we expect, that the complexity of the physical world can be captured and understood by the model.

Example applied on the faces of a cube:

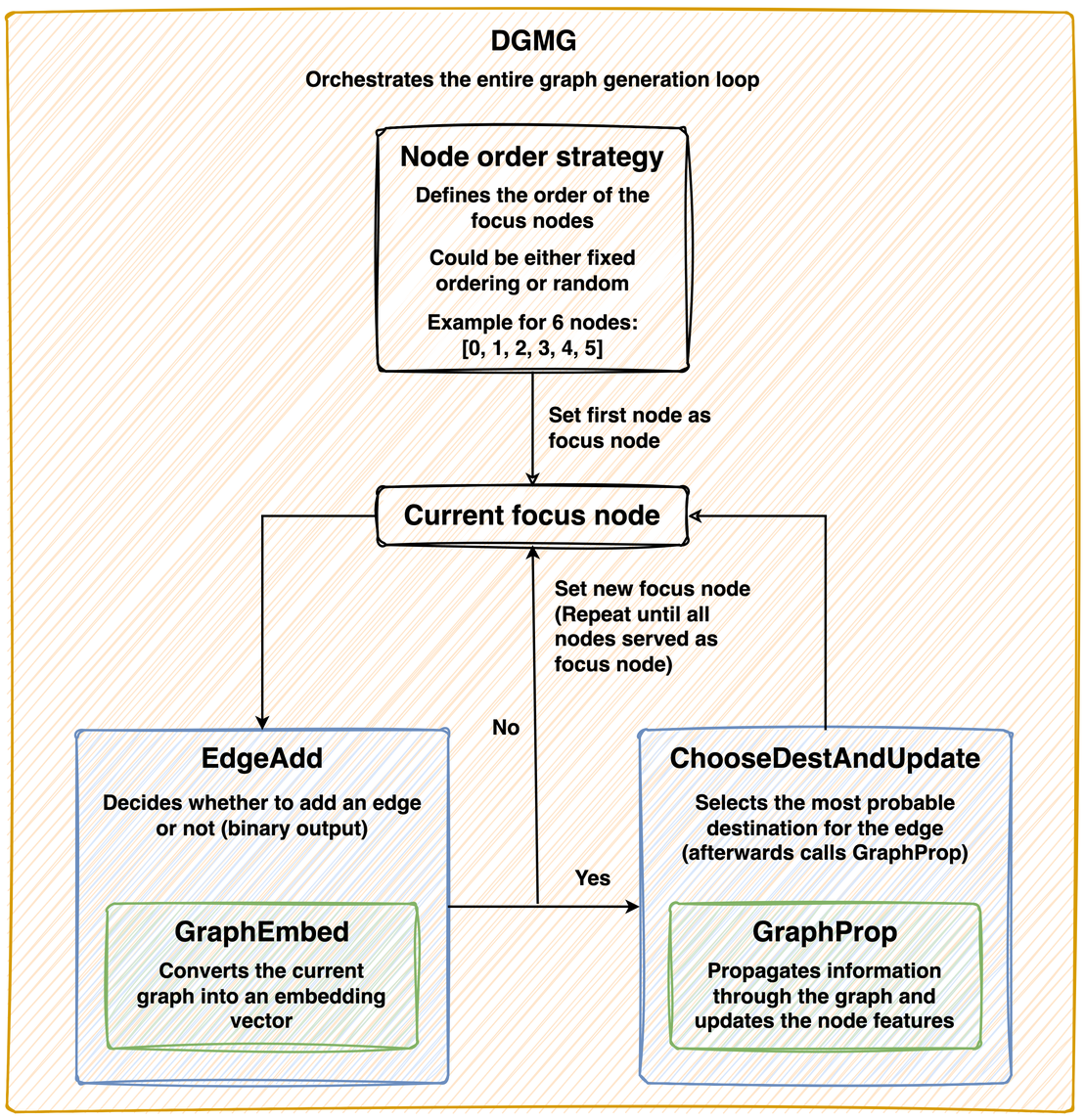


Methodology

Make use of **Deep Generative Model of Graphs (DGMG)**[3] model as inspiration and use it to **sequentially predict and update the graph with additional edges.**

Doing **stepwise predictions** and updating the graph after each prediction to take earlier made predictions into account **aligns with the underlying nature of the folding process.**

Implementation:



Model evaluation

To validate if the proposed GNN architecture performs better on the underlying task in comparison to existing methods, we define that a good model should have a **low number of False-Positives** (wrong predictions), therefore a **high Precision value**.

To achieve this, we employ the following metrics:

- **Precision**
- Recall
- **F1-Score**
- TP, **FP**, FN (measure avg. % to make comparable over different graph sizes)
- Precision-Recall AUC: tells best trade-off between Precision and Recall

References

[1] Felton, S., Tolley, M., Demaine, E., Rus, D., & Wood, R. (2014). A method for building self-folding machines. Science, 345(6197), 644-646

[2] Fernandes, R., & Gracias, D. H. (2012). Self-folding polymeric containers for encapsulation and delivery of drugs. Advanced drug delivery reviews, 64(14), 1579-1589

[3] Li Y., et al. (2018). Learning Deep Generative Models of Graphs.

Fundação para a Ciência e a Tecnologia

The authors acknowledge all the financial support from Fundação para a Ciência e a Tecnologia (FCT) under Contracts no. UIDP/00618/2020 (DOI: https://doi.org/10.54499/UIDP/00618/2020) and UIDB/00618/2020 (DOI: https://doi.org/10.54499/UIDB/00618/2020) and LASIGE Research Unit, ref. UID/00408/2025

cftc.ciencias.ulisboa.pt

Systematization of Knowledge: Self-Generated Nudes over Private Chats: How Can Technology Contribute to a Safer Sexting?

Joel Samper, Bernardo Ferreira

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE dependable and secure decentralized systems

Question

What technology-assisted permissive approaches (techniques, features...) provide: security, privacy or accountability benefits
to: users who send or receive nudes
in: 1-to-1 chats of mobile apps
?

Methodology

PRIOR LITERATURE
Adolescents, Breakups, Sex workers, Dating apps, Adults
user surveys
interviews
online posts
app analyses

→ user-centric harms →

actor	threat	attacks or incidents
Recipient	Interfere	Put pressure, Reproach
	Make copies	Screenshot, Screen-cam, Self-hack
	Disclosure	Post, Forward, Exploit, Extort
Sender	Self-leak	Metadata, Deanonymize, Slip-ups
	Deception	Synthetic, Others', Misrepresent
	Non-consensual	Cyberflashing, Unwanted
App	Covertly	Exploit, Scan gallery, Sponsors
Third parties	Grab device	Coerce, Stalkerware, Forensics, Lend
	Network	Eavesdrop, PiTM
	Servers	Vulnerability, Ex partner access
	Surrounding	Shoulder surfing, Bystanders, CCTV

Systematic literature review → 113 search strings in scholarly DB → 138 papers (Snowball)

Survey of systems and app features → 52 apps, UI options, Online guides, APIs & systems → 81 features & systems

Findings

Let's break them down into categories

POST EXPOSURE USER-TO-USER PRIVACY

Ephemeral messaging: *Time-expiring*, *View once*, *Unsending*, *Albums*
stopncii.org
takeitdown.ncmec.org
Inhibit screenshots

AUTHENTICITY

Camera firmware digitally signs
Recapture detection algorithms
Using AI to detect AI-generated images
Ownership verification systems

APP PLATFORM

E2EE messaging
Blocking ads with sandboxing or packet filtering
Middleware, third-party or attestation architectures
App permissions

OFFLINE TOOLS

Safety guidelines
Privacy and sexting education
Talk About Sex, iBuddy, FoxIT
Parental supervision with linked accounts
Peer support
Teen Talk app

PRE EXPOSURE USER-TO-USER PRIVACY

Confirmation prompts
Fake profile detection
Profile information and badges
Cropping, masking

ADVERSARIAL-RECIPIENT SECURITY

Anti-debugging
Middleware for image encryption
App integrity attestation
Certificate pinning

ACCOUNTABILITY

Reporting and blocking
Digital watermarking
Screenshot alerts
Visible watermarks

RECEPTION CONTROL

Blur, tap-to-open, warnings or policy-based filtering of nudes
First contact preferences and policies
NSFW preference

THIRD PARTIES

Hidden albums
Find & wipe my device
Visualisation privacy: screen protectors, gaze estimation,...
E2EE backups/messaging
2FA cloud access
Certificate pinning
Anonymity VPNs

Take-aways

Ephemeral messaging
- Popular: Snapchat, disappearing messages.
- 44 feature instances in our app survey.
- 2+ bn albums shared on Grindr in 2024.
- Post Exposure, Third Parties, App Platform.
- Most apps vulnerable to feature bypass.

Valuable even if not bullet-proof, but app integrity APIs increase security.

End-to-End Encryption
- ! Apps training nude detection AI models.
- ! Nation-states intercepting dating apps.
- ! Vulnerable servers, e.g. 2014 iCloud hack.
- ↑ amount of users impacted over ↓ time.
- App/Network threats: no or minimal trace.

E2EE as a fundamental S&P feature. On-device or encrypted-domain algorithms.

Threat: disclose or threaten to
- Most studied threat in our Related Work.
- Attacks escalation: fake profile, pressure, tamper with software, threaten, disclose.
- Protections layering: Offline Tools, Pre+Post Exposure, App Integrity, Accountability.

Central issue. Approaches go from sexting education to digital watermarking.

Usability difficulties in features
- Inconsistent UI with documentation naming.
- Non-uniform: Reception Controls, feature combinations, specifications, marketing.
- Parental supervision constanly evolving.
- Ephemeral feature roll-out median: ~2021.

Strategies to better assist users, e.g. safety settings preset configurations.

Privacy preservation
- No mutual information should be gained when third parties attest profile information.
- Digital watermarking, digital signatures, perceptual hashing and/or blockchain for ownership proof or leakage traceability.

Preserving privacy is a key challenge. Also: performance, usability, transparency.

- 19/138 papers mention private mesaging.
- Digital watermarking has focused more on copyright; image-encryption on social media posts; and blockchain on creative sector.
- Few sexting mentions in non-dating apps.

Use this paper to: apply sexting threat model, frame system proposals within a taxonomy, highlight value of features in sexting scenario.

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE | "la Caixa" Foundation

Acknowledgements: This work was supported by FCT through project APOSTLE (2023.12254.PEX), and the LASIGE Research Unit, ref. UID/00408/2025.

Sure! Here are a few title ideas for your poster on complex multi-ontology alignment

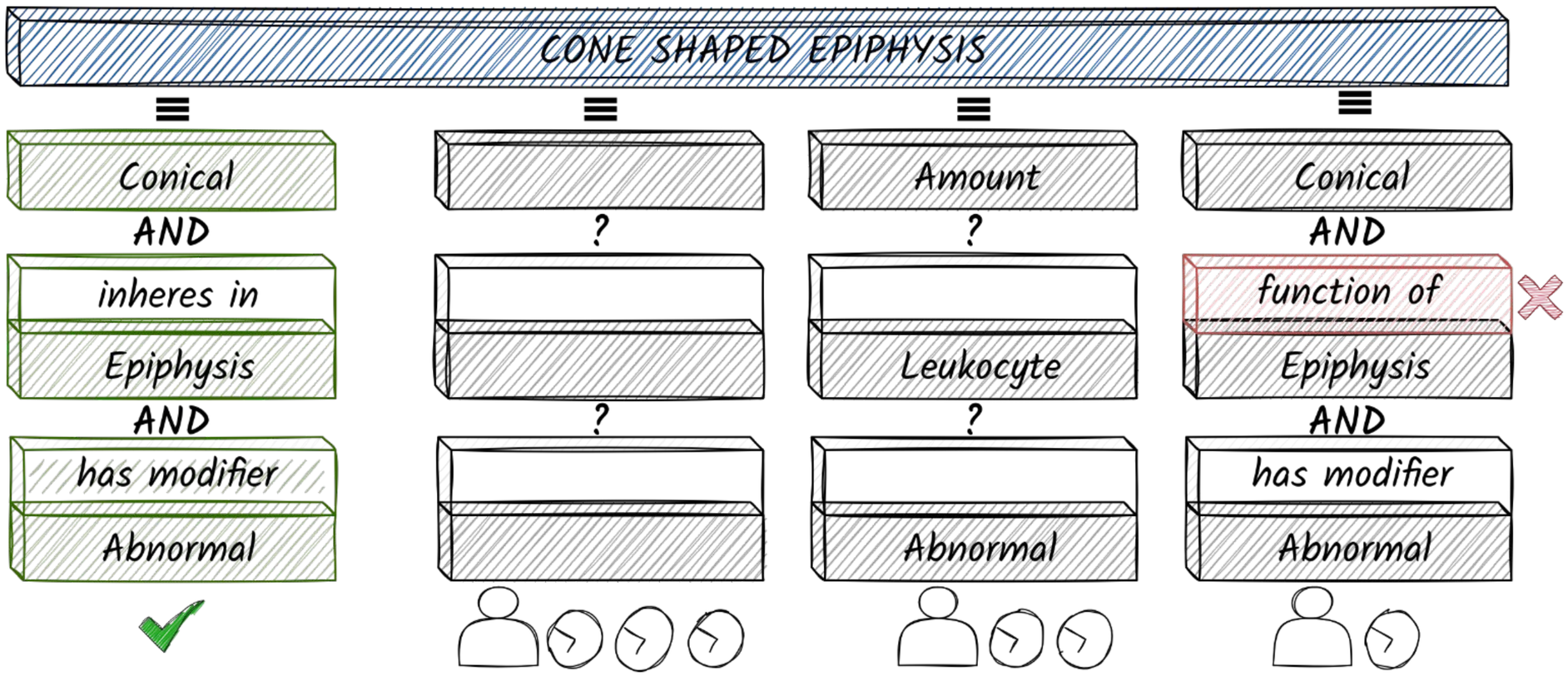
Marta C. Silva[1], Daniel Faria[2], Catia Pesquita[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] INESC-ID, Instituto Superior Técnico, Universidade de Lisboa, Portugal

LASIGE health and biomedical informatics

MULTI-DOMAIN INTEGRATION

Connecting diverse data requires concepts from multiple domains that are related but disjoint



CAN WE TEACH LLMs HOW TO SUCCESSFULY RECREATE THE PATTERNS WE WANT?

IF YOU JUST ASK:

Create a complex mapping for class "Cone Shaped Epiphysis".

has part
Material Entity
AND
located in
Human Body

IF YOU PROVIDE THE CONCEPTS:

Create a complex mapping for "Cone Shaped Epiphysis" using the classes: "Conical", "Epiphysis", "Abnormal".

has morphology
Conical
AND
part of
Epiphysis

IF YOU PROVIDE EXAMPLES:

This is a list of possible examples of complex mappings. Create a complex mapping for "Cone Shaped Epiphysis".

Abnormal Shape
AND
inheres in
Bone Element
AND
has modifier
Abnormal

present in all of them!

OUR STRATEGY

IF YOU PROVIDE THE CONCEPTS AND THE EXAMPLES:

Conical
AND
inheres in
Epiphysis
AND
has modifier
Abnormal

:)

EVALUATION

		Precision	Recall	F1-score
HP	Just ask	0.210	0.203	0.207
	Give the pieces	0.275	0.255	0.265
	Give examples	0.617	0.617	0.617
	Give both!	0.634	0.632	0.633
MP	Just ask	0.203	0.193	0.198
	Give the pieces	0.287	0.270	0.278
	Give examples	0.645	0.642	0.643
	Give both!	0.666	0.662	0.664
WBP	Just ask	0.206	0.197	0.201
	Give the pieces	0.260	0.242	0.250
	Give examples	0.731	0.730	0.731
	Give both!	0.687	0.687	0.687

WHAT IF WE KNOW THE RIGHT CONCEPTS?

If we use the classes from the reference, we achieve a minimum of 0.87 in F1-score

"ANY SUFFICIENTLY ADVANCED TECHNOLOGY IS INDISTINGUISHABLE FROM MAGIC"
- Arthur C. Clarke

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE | KATY Clinical Knowledge for Humanity

Acknowledgements: This work was supported by FCT through the fellowship 2022.11895.BD, and the LASIGE Research Unit, ref. UID/000408/2025. It was also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453, and by project 41, HfPT: Health from Portugal, funded by the Portuguese Plano de Recuperação e Resiliência.

EXPLORING CONTINUITY IN GAMES:
Analog, Digital and Across Mediums

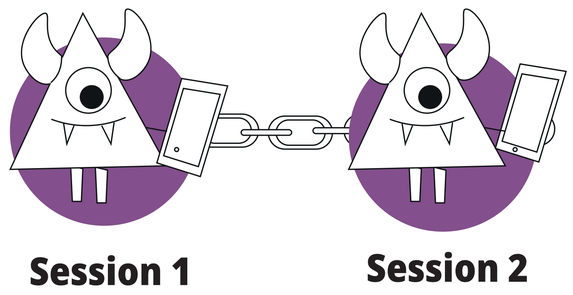
Inês Gil, David Gonçalves, João Guerreiro, André Rodrigues

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE inclusive human--computer interaction

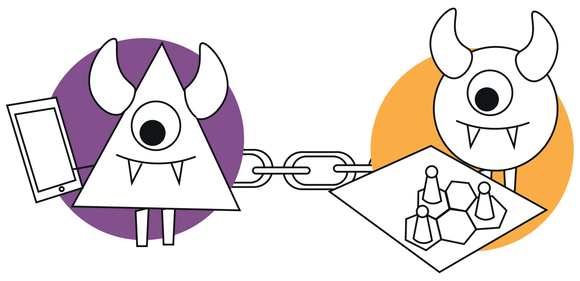
Game experiences can span multiple play sessions, which may evoke a sense of continuity

CONTINUOUS PLAY



Where game elements, such as **chapters, cards** or **characters carry over** to the next play session

CONTINUOUS TRANSMEDIAL PLAY



This feeling can appear when players play games of **different mediums** (e.g., connected by theme)

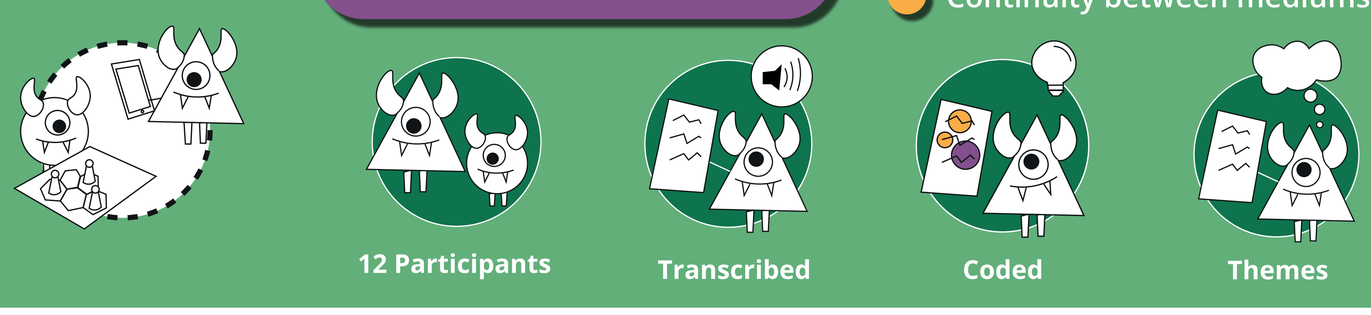
HOWEVER... IT IS NOT CLEAR

- What defines continuity
- Contributing elements
- Effect on the player experience
- What about Transmedial play?

Semi-structured interviews with **12 gamers** who play both **analog** and **digital games**

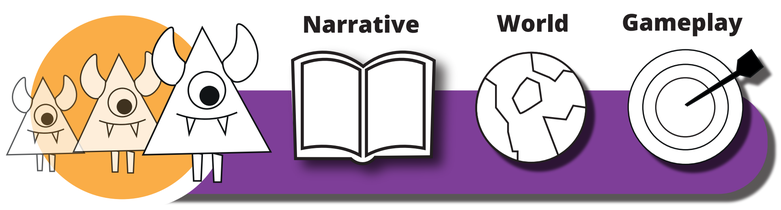
TOPICS

- Continuity in digital and analog games
- Continuity in shared play experiences
- Interruptions in continuity
- Continuity between mediums

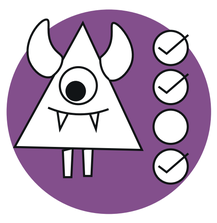


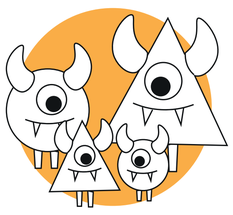
THE ELEMENTS OF CONTINUITY

Participants associated continuity with the **sense of progression**



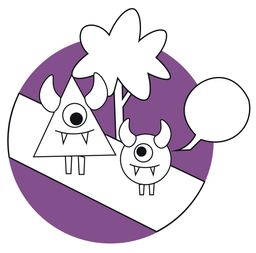
But also with gaining **mastery and knowledge** about a certain game.





And **with players** who share the same journey.

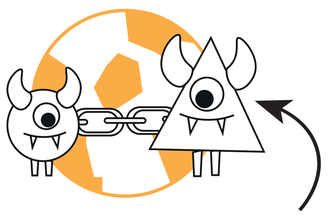
Even **outside of the game!** By talking about their experience with others.



TRADEOFFS OF CONTINUITY

Continuity can create a **stronger connection** with the game and also other players

Also allow players to experience a **unique journey!**



Especially when players **shape the world together!**

HOWEVER...

Finding time to play together can **be burdensome** as players have conflicting schedules.

IF TOO MUCH TIME PASSES

Players feel like they have **"wasted"** their time in a game they did not complete.

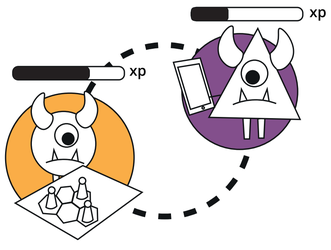
TRANSMEDIAL PLAY

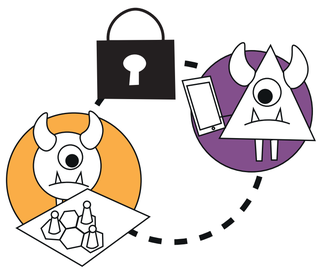
Continuous transmedial play as a way to address some of the challenges of continuity

- **Increased flexibility**
- **Complement each other**

BUT THERE ARE CHALLENGES!

Discrepancies in progress as players have different amounts of time to play





Medium dependency, if progression in one medium is mandatory for engaging with the other.

CHECK THE PAPER OUT!

Acknowledgements: This work was supported by FCT through project "Plug n' Play: Exploring Asymmetry and Modularity for Inclusive Game Design" ref. 2022.08895.PTDC (http://doi.org/10.54499/2022.08895.PTDC), the Institutional CEEC (https://doi.org/10.54499/CEECINST/00032/2018/CP1523/CT0003), scholarship ref. UI/BD/151178/2021 (https://doi.org/10.54499/UI/BD/151178/2021), and the LASIGE Research Unit, ref. UID/00408/2025.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Energy Aware Programming - A static Analysis Tool for Java

Afonso Carreira, Wellington Oliveira, Alcides Fonseca

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE reliable software systems

Motivation

Ever wondered how much energy your code uses?

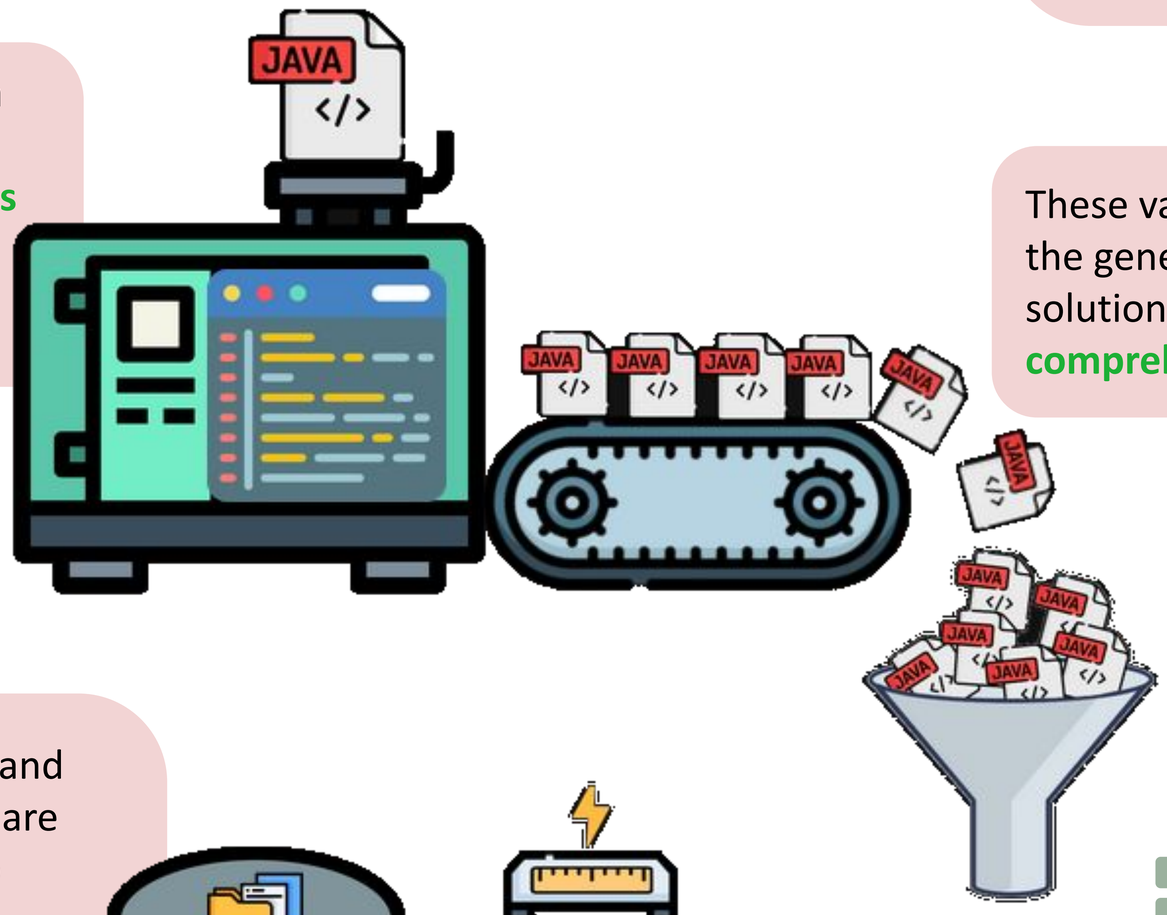
While developers often optimize for speed and memory, energy consumption is **overlooked**, despite its environmental impact, especially in data centers and AI. Developers lack tools and guidance to measure and reduce their code's energy use.

Objectives

Develop a tool that provides energy consumption estimates for Java code using static analysis

- Easy to use
- Fast estimations
- Accurate

From a single Java class, it generates **multiple programs** by varying input types and values.

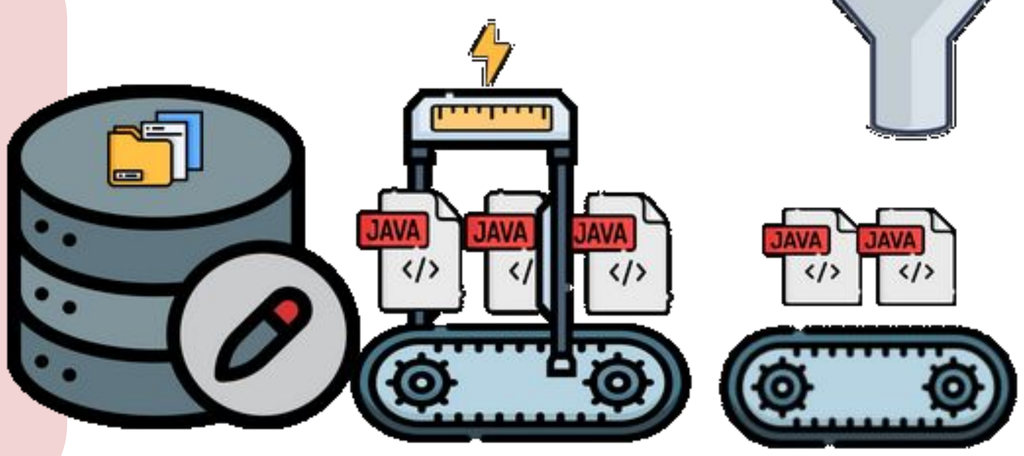


These variations enhance the generalization of our solution, aiming to build a **comprehensive dataset.**

The tool integrates with an IDE to use the best-performing model to make **real-time predictions for the energy cost**

The energy values and extracted features are stored and used to train **ML models**.

Multiple models are evaluated to identify the most accurate one for energy estimation.



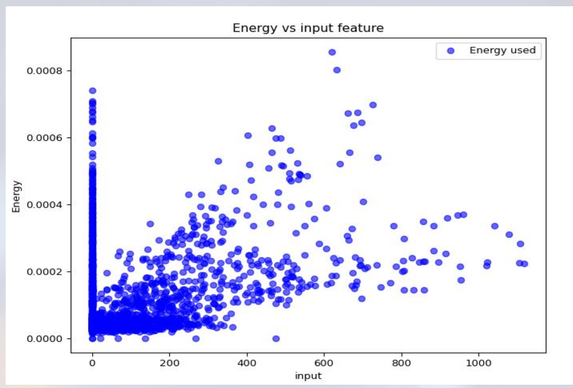

```
public class Main {
    Run | Debug
    public static void main(String[] args) {
cost: 0.5J  ArrayList<Integer> list = new ArrayList<>();
cost: 0.1J  int max = 1000;
            for (int i = 0; i < max; i++) {
cost: 0.1J      list.add(i);
            }
    }
}                                  total energy cost: 100.6J
```

For each program, **energy consumption** is measured, and features such as the number of conditions, loops, variables, and input size are extracted using a parser.

The developer can use a slider to specify expected input values for each function, which helps improve the **accuracy of the prediction**.

Results

Program energy consumption is influenced by multiple factors, with input data playing a key role. Each feature impacts energy use differently, for instance, as shown in the figure, energy consumption varies with list size. However, this relationship is not linear, as other factors like list implementation, data types, and garbage collection also contribute.



Performance metrics for energy estimation models on the *addAll* method of the list collection

Model	MSE	R^2
Decision Tree Regression	3.31×10^{-9}	0.825
Random Forest	2.09×10^{-9}	0.878
PySymbolic Regression	4.58×10^{-9}	0.8669

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Acknowledgements: This work was supported by the LASIGE Research Unit, ref. UID/000408/2025.

Mixture of Experts Technique for Image Classification in Autonomous Driving Systems

Bakary Badjie, José Cecílio, António Casimiro

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

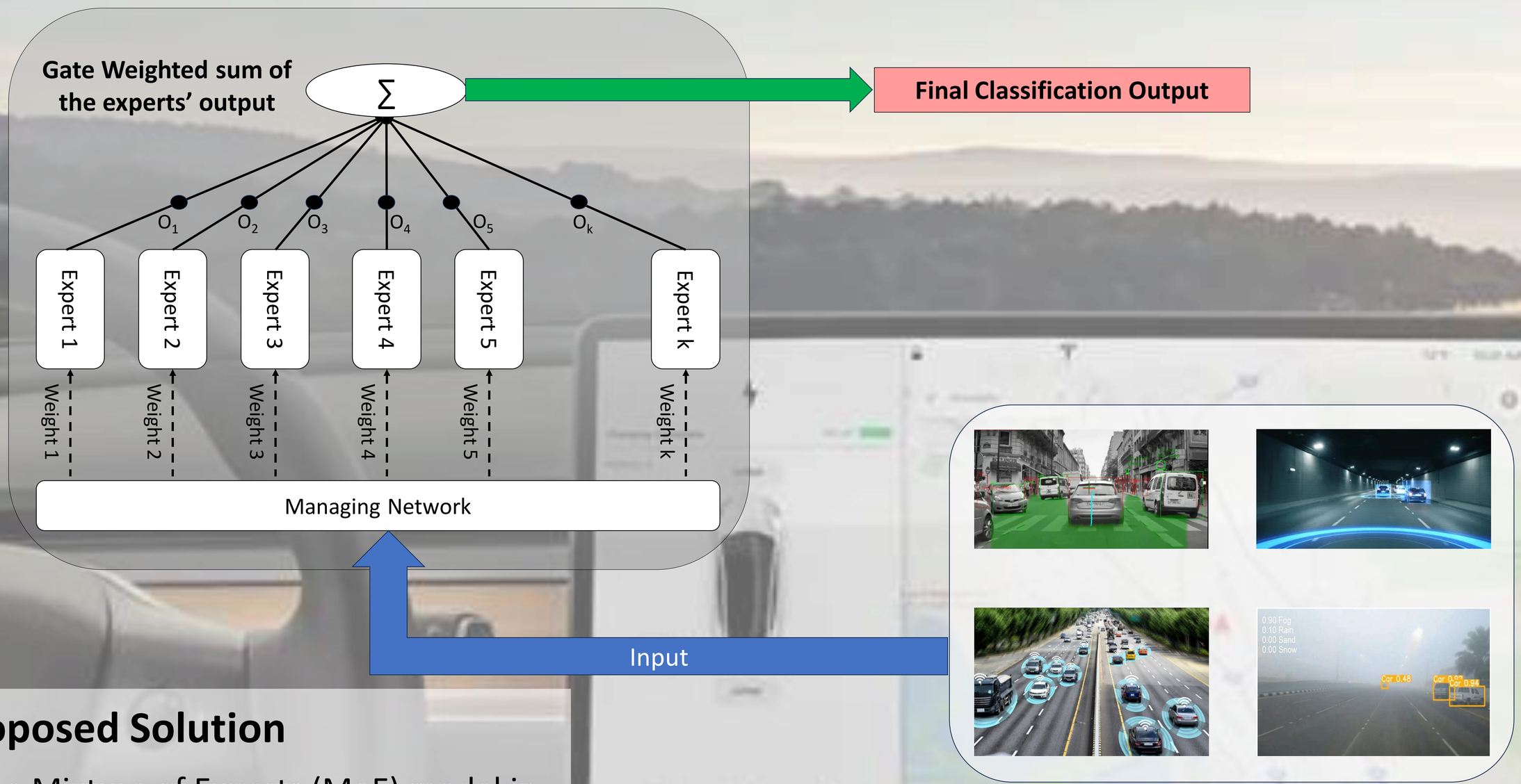
LASIGE cyber-physical systems

Motivation

Autonomous driving systems must accurately classify multiple types of objects in diverse and dynamic driving scenarios, and there is a need for adaptive, context-aware solutions for robust perception

Problem Statement

General-purpose image classification models (dense models) suffer performance loss in complex and dynamic driving environments due to lack of specialization in different scenarios (e.g., day/night, sunny/rainy weather, urban/highway driving)



Proposed Solution

Use a Mixture of Experts (MoE) model in which specialized expert networks handle different driving scenarios

A gate network dynamically assigns weights based on input features to optimize classification accuracy

Results

- Up to 27% accuracy improvement on GTSRB traffic sign classification compared to baseline dense models
- Demonstrated better adaptability in dynamic traffic sign inputs
- 41% efficiency improvement during inference

Conclusions

- Classification task is partitioned across multiple expert subnetworks
- Each expert is trained to specialize in specific traffic sign categories
- A dynamic gating mechanism analyzes input features in real-time
- The gate selectively activates the most relevant experts per input

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

Acknowledgements: This work was supported by Fundação para a Ciência e Tecnologia (FCT) under Ph.D. grant agreement No. 2023.02832.BD, through the LASIGE Research Unit, ref. UID/00408/2025.

Biclustering Data Analysis in a Nutshell

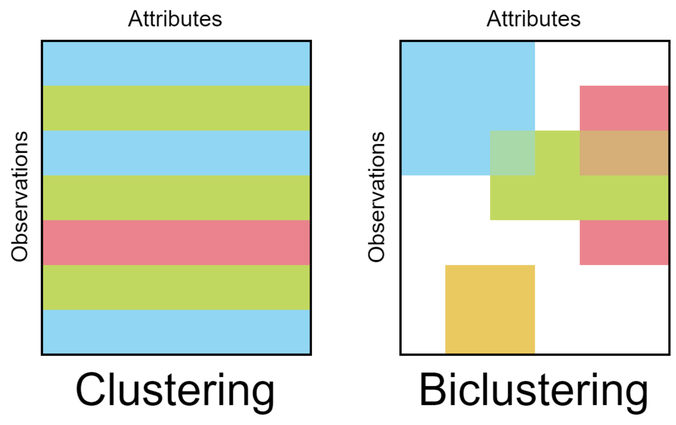
Eduardo N. Castanho[1], Helena Aidos[1], Sara C. Madeira[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE data and systems intelligence

What is Biclustering?

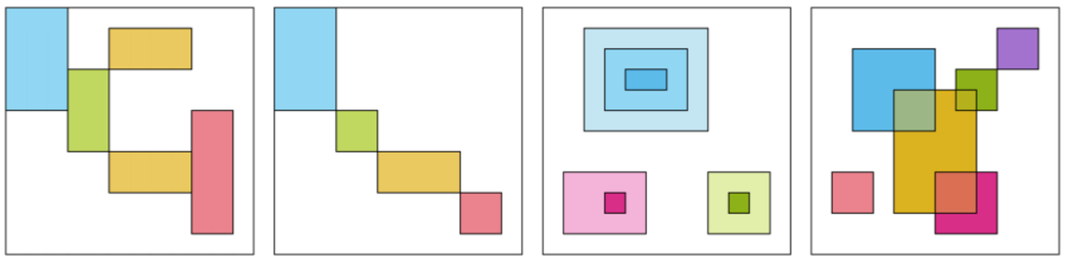
Biclustering is an unsupervised machine learning task that **generalizes clustering** by **simultaneously grouping** observations and features, discovering sub-matrices in a data matrix. It is **characterized** by **its flexibility**.



Biclustering **searches for local patterns** and discovers subtle interactions that traditional algorithms miss. **Consider biclustering** when **clustering underfits** and you suspect that **observations** behave similarly only under **specific attributes**.

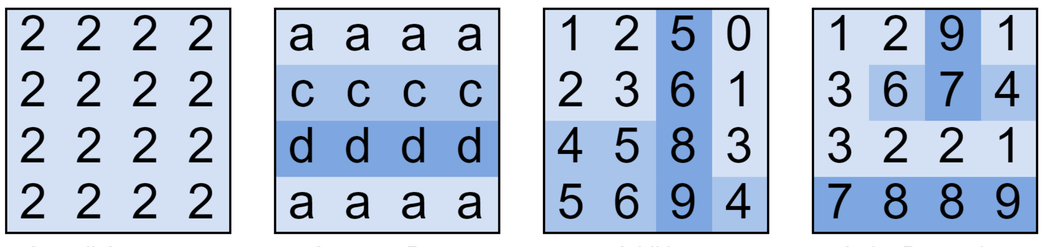
Why is Biclustering flexible?

Different Structures



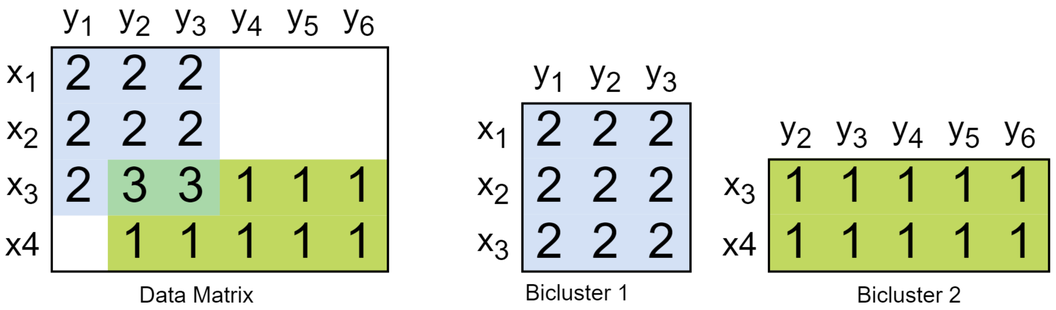
Exclusive Columns | Exclusive Rows and Columns | Hierarchical Structure | Arbitrarily Overlapping

Flexible measures of similarity



Overall Constant | Constant Rows | Additive | Order Preserving

Admits overlapping



How many lines of code to use it?

```
from sklearn.cluster import SpectralCoclustering
import pandas as pd
X = pd.read_csv("./dataset.csv")
biclustering = SpectralCoclustering(n_clusters = 2).fit(X)
print(biclustering.row_labels_)
print(biclustering.column_labels_)
```

What is the data to apply biclustering?

Biclustering **can be** applied to **analyze** multiple **data types**

Data **must be** processed as a **data matrix**.

What are popular applications?

- Biological and Biomedical
  - Functionally annotate unclassified genes
  - Discover submodules in biological networks
  - Identify profiles of clinical progression
- Other
  - Identify spatio-temporal patterns
  - Recommend items based on user preferences
  - Discover groups of words that appear together

Read More

Acknowledgements: This work was supported by Fundação para a Ciência e a Tecnologia (Grant Nos. 2021.07810.BD, DSAIPA/AI/0033/2019 and PTDC/CCI-CIF/4613/2020), H2020-RIA (Grant No. 818290), and the LASIGE Research Unit, ref. UIDB/00408/2020 (https://doi.org/10.54499/UIDB/00408/2020) and ref. UIDP/00408/2020 (https://doi.org/10.54499/UIDP/00408/2020).

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

Probabilistic Chained Blockchain Consensus

André Santos, Hasan Heydari. Alysson Bessani

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE dependable and secure decentralized systems

What Is Consensus?

Consensus refers to the problem in which several participants in a distributed system attempt to agree on a single common value.

Consensus

Agreement
All correct nodes agree on the same value

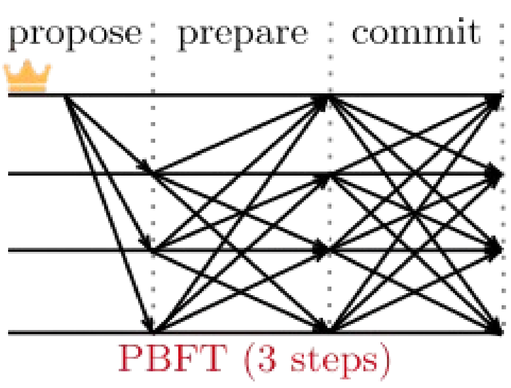
Termination
Every correct node eventually decides

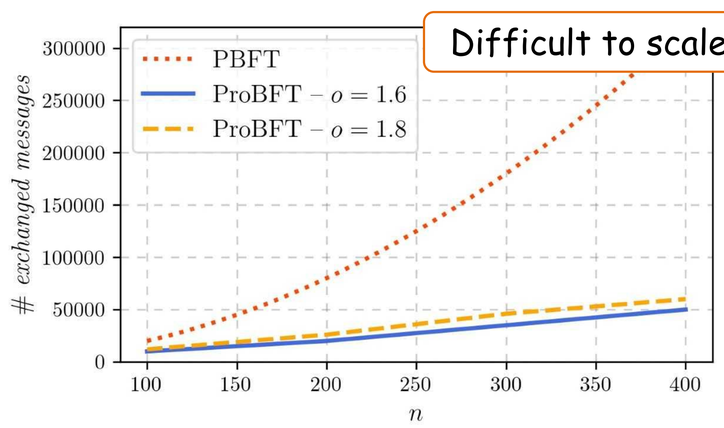
Validity
The agreed value was proposed and is valid

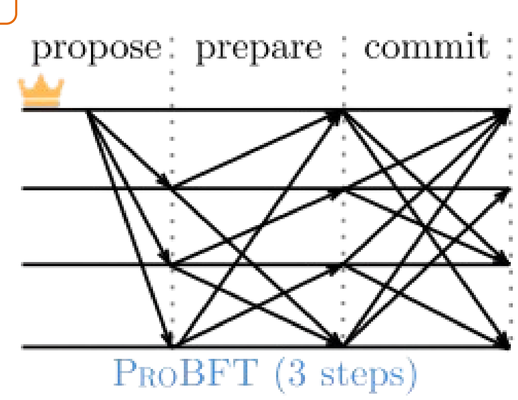
Probabilistic Consensus

Probabilistic Agreement
All correct nodes agree on the same value with high probability

Probabilistic Termination
Every correct node eventually decides with probability 1







ProBFT is a recently developed Probabilistic Consensus protocol at LASIGE.

Achieves sub-quadratic message communication by selecting samples of receptors randomly.

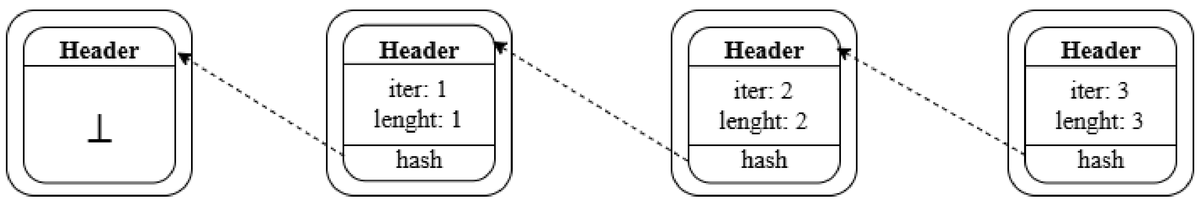
Problem!

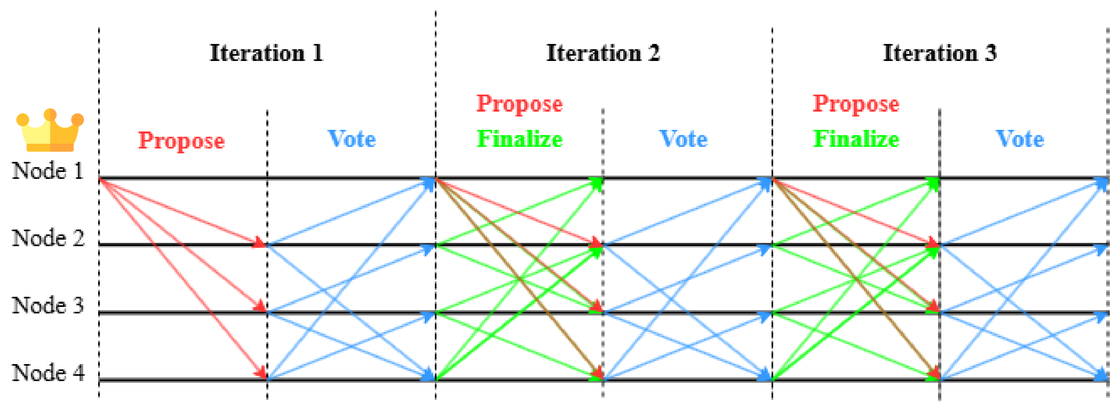
ProBFT has never been implemented or evaluated experimentally and only solves the single-shot consensus problem.

Approach

1. Develop a SMR and Blockchain Probabilistic Consensus protocol capable of ordering blocks of transactions.
2. Integrate ProBFT's core mechanisms to a simpler and recent Byzantine Consensus protocol, Simplex.
3. Evaluate and compare this new protocol with Streamlet, Simplex and BFT-SMaRt.

ProSimplex





A practical implementation of Simplex was also developed beforehand using Rust!
Available at: https://github.com/AndreSantos0/Probabilistic-Chained-Blockchain-Consensus

Work Plan

Development
Integrate ProBFT's mechanisms into a practical implementation of Simplex

Evaluation
Evaluate and compare the newly implemented protocol with others

Document
Write protocol proofs

Patching and Improvement
Fix latency issues and bugs found with the current implementation of the protocol

Acknowledgements: This work was supported by the LASIGE Research Unit, ref. UID/00408/2025.
This work was supported by FCT through project SMaRtChain, ref. 2022.08431.PTDC, DOI 10.54499/2022.08431.PTDC (https://dx.doi.org/10.54499/2022.08431.PTDC) and the LASIGE Research Unit, ref. UID/000408/2025.

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

Discord, Conflict and Contradiction in AI: A benchmark for learning with contradictory facts

Beatriz Moreira, Catia Pesquita

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE health and biomedical informatics

Motivation

Machine Learning models enhanced with **Knowledge Graphs'** semantic background, typically use a majority rule approach to resolve **contradictions**, which can **perpetuate misconceptions** and **hide different viewpoints**.

The creation of **benchmark datasets based on contradictory information** to train these models is then crucial to ensure they can **foster different perspectives** and **enhance their accuracy and inclusivity**.

The **intuition** behind this work is that **language models** can be used to create **plausible contradictions** based on triples extracted from the base KGs to create **contradiction-based benchmark datasets**.

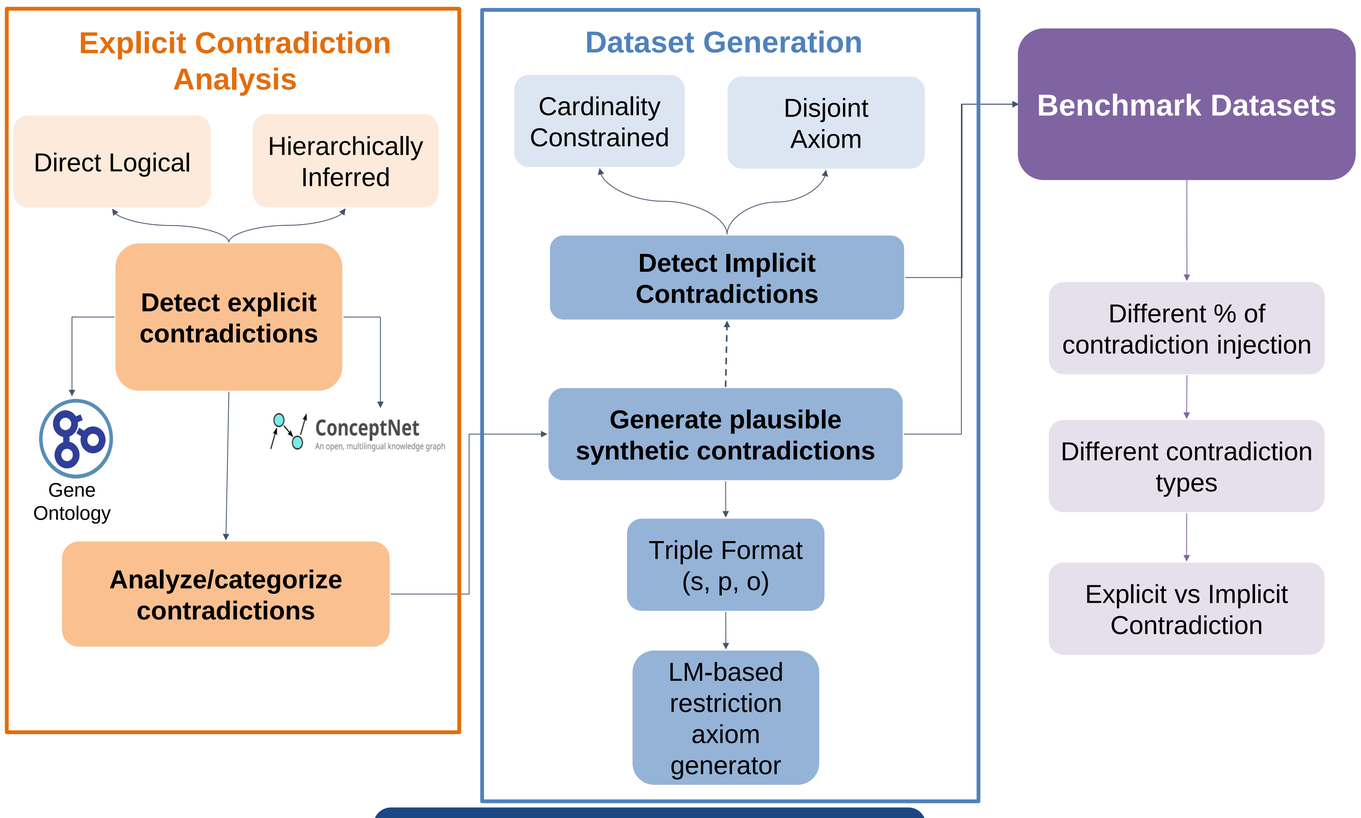
Objectives

The goal of this work is **to create a benchmark to evaluate the impact of contradictory facts** in Artificial Intelligence (AI) applications based on Knowledge Graphs (KGs).

Existing contradictions will guide the **generation of novel contradictions using language models** based on facts extracted from KGs.

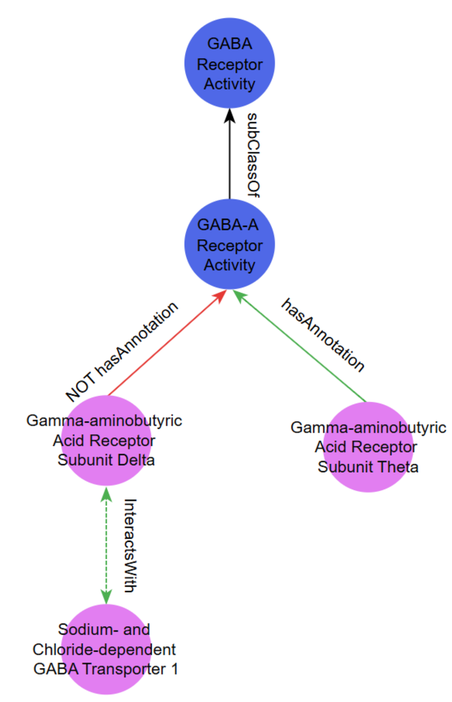
The methodology will support the creation of **benchmark datasets** with **different percentages and types of contradictions**.

Methodology

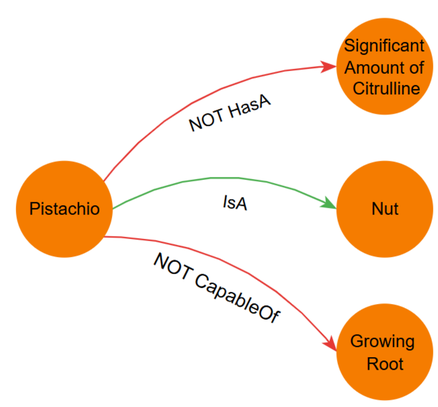


Data

Gene Ontology (GO) KG + PPIs



ConceptNet + Uncommonsense



GO terms · Proteins · GO Edges · Positive Annotations · Negative Annotations · PPIs · Nodes · Positive Edges · Negative Edges

Contradiction Types

Explicit contradictions: Contradictions that can be detected by inspecting the KG or through reasoners.

Implicit Contradictions: Contradictions that cannot be detected by reasoners but are considered as such in real-world situations (found with pattern mining).

Contradiction Types	Description Logics Formulae	Examples
Direct Logical	$(a \subset B) \sqcap (a \not\subset B)$ or $(B(a)) \sqcap (\neg B(a))$	Pizza subClassOf Food; Pizza NOT subClassOf Food
Hierarchically Inferred	$(B(a)) \sqcap (\neg C\ (a)) \sqcap (B \subseteq C)$	Margherita hasBase PizzaBase; Margherita NOT hasIngredient Base; hasBase subPropertyOf hasIngredient
Disjoint Axiom	$(B(a)) \sqcap (C\ (a)) \sqcap (B \perp C)$	VegeterianPizza hasTopping only Vegetables; VegetarianPizza hasTopping Meat; Meat disjointWith Vegetables
Cardinality Constrained	$a \sqsubseteq (\leqslant nB(a, c)) \sqcap (\geqslant mB(a, c))$, $n \in$ N, $m \in$N	Margherita hasBase ThinPizzaBase; Margherita hasBase DeepPizzaBase; hasBase maxCardinality 1

Preliminary Results

Positive Annotations	Negative Annotations	Direct Logical Contradictions	Hierarchically Inferred Contradictions
74902	69211	6574	1831

Table 1: Statistics for the **GO KG + PPIs** dataset.

Positive Triples (ConceptNet)	Negative Triples (Uncommonsense)	Direct Logical Contradictions	Hierarchically Inferred Contradictions
3423004	13643924	2555	8981

Table 2: Statistics for the **ConceptNet+Uncommonsense** dataset.

Both base datasets contain **explicit contradictions**, which will be used to guide the creation of **novel synthetic contradictions**.

Acknowledgements: This work was supported by FCT through LASIGE Research Unit, ref. UID/00408/2025. It was also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Typing Patterns as Digital Footprints: Enhancing Health Monitoring

Diana Pimentel, Diogo Branco, Filipa Ferreira-Brito, André Santos, Tiago Guerreiro

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE inclusive human-computer interaction

Data created while interacting with digital devices, known as **digital footprints**, can reveal patterns and behaviours that can be leveraged for health monitoring

Typing is one of the most frequent task when using digital devices

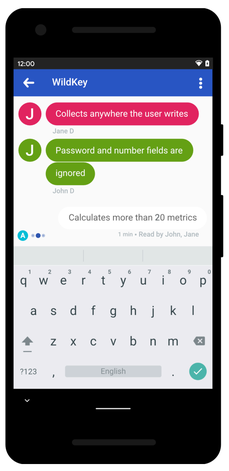
- Typing speed (words per minute)
- Errors
- Touch dynamics (flight and hold time)

Typing involves coordinated **motor skills and cognitive functions**

Text-Entry Patterns

Early Disease Detection and Health State Assessment

Most studies are based on controlled environments and a shift to **in-the-wild** assessment is desirable



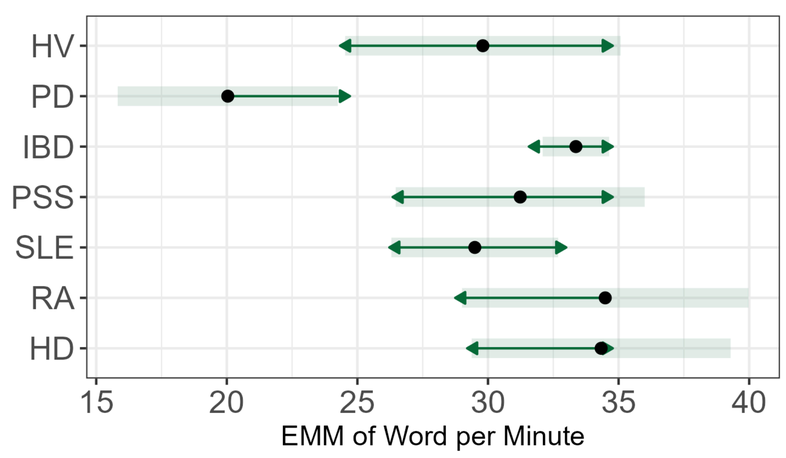
WildKey toolkit

- Open-source, privacy-aware keyboard toolkit designed to collect text-entry data in the wild
- Captures data in natural typing environments, enabling real-world monitoring
- Used to study typing behaviour in neurodegenerative and inflammatory diseases
- Promising digital biomarker for continuous, non-invasive daily health assessment

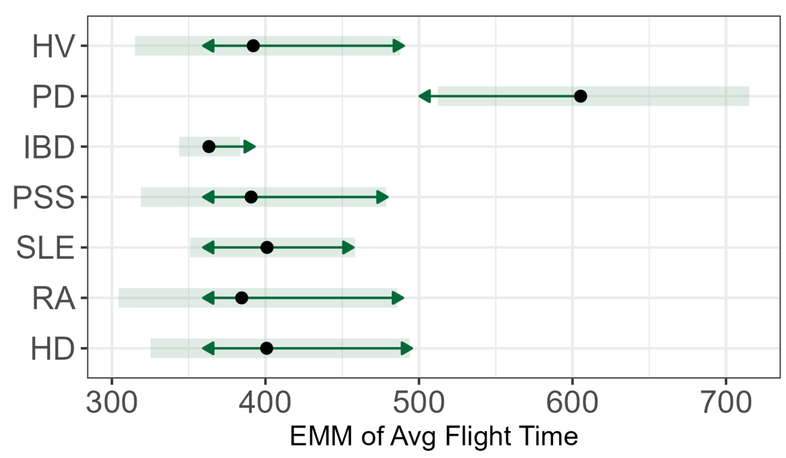
WildKey

Participants with Parkinson's disease are slower writers and take more time changing keys.

Estimated Marginal Means (EMMs). Arrows represent the pairwise comparisons; overlapping arrows indicate no significant differences (Tukey's adjustment, $\alpha = 0.05$). HV - Healthy Volunteers, PD - Parkinson's disease, IBD - Inflammatory Bowel Disease, PSS - Primary Sjogren's Syndrome, SLE - Systemic Lupus Erythematosus, RA - Rheumatoid arthritis, HD - Huntington's disease.



Age = 40.8



Age = 40.8

Opportunities and **Challenges** of Digital Footprints

- Enable continuous, real-world data collection with minimal effort and financial investment
- Additional source of information, acting as a detailed health backlog over time
- Enable remote monitoring, allowing early intervention measures and supporting medical decision

- In-the-wild data collection is often incomplete and noisy
- Users' perception of the symptoms might be influenced by the data
- Concerns regarding privacy and data security

Reach us here

Check full paper

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · Health Portugal Agenda HfPT · IDEA FAST · PRR Plano de Recuperação e Resiliência · República Portuguesa · Financiado pela União Europeia NextGenerationEU

Acknowledgements: This work was supported by Project 41, HfPT: Health from Portugal, funded by the Portuguese Plano de Recuperação e Resiliência; Fundação para a Ciência e a Tecnologia, LASIGE Research Unit, ref. UID/00408/2025; and IDEA-FAST project, which has received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No. 853981. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA and associated partners.

Usability Barriers for Liquid Types

Catarina Gamboa[1,2], Abigail Reese[2], Alcides Fonseca[1], Jonathan Aldrich[2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] School of Computer Science, Carnegie Mellon University, USA

LASIGE reliable software systems

Liquid Types have been introduced in 2008, have been implemented across many languages (Haskell, Java, C, JavaScript, Rust) to statically detect bugs ranging from simple errors to security issues and protocol violations.
They add logical predicates to the type system, e.g.:

```
{-@ abs :: Int -> {v: Int | v >= 0} @-}
abs :: Int -> Int
abs n = if n > 0 then n else (-n)
```

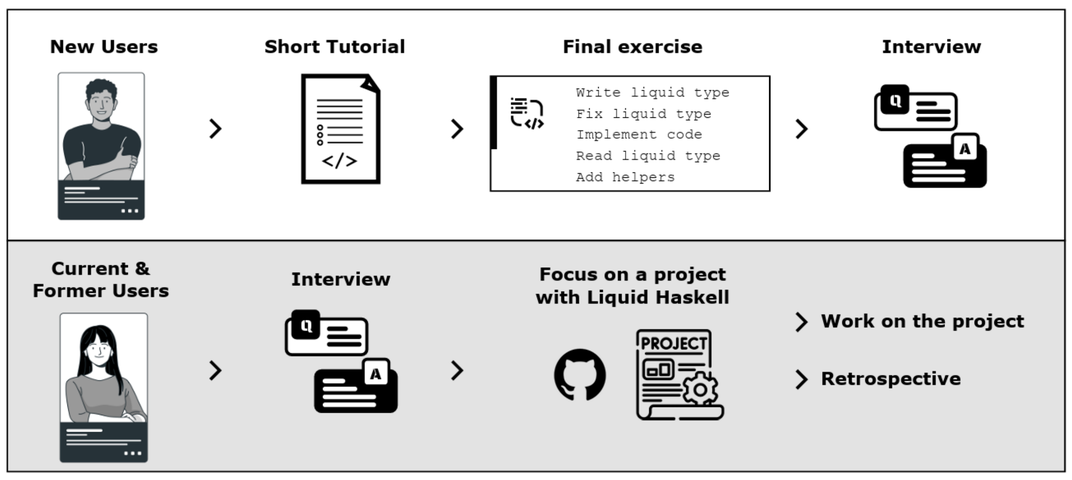
Liquid Type
Implementation

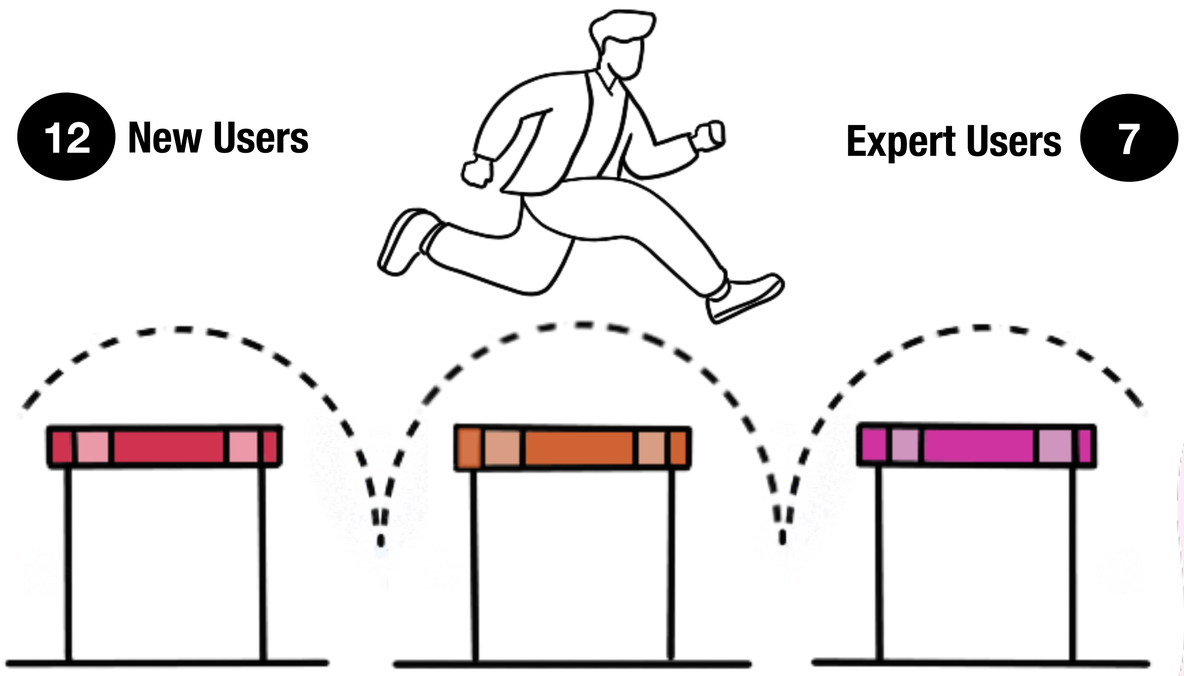
These guarantees can be checked before your program runs, and they have been used to find bugs in MVC frameworks, object protocols, and database queries.
However, they have yet to gain widespread adoption in mainstream programming.

RQ: What are the current barriers developers face in adopting and using liquid types?

Developed a study focused on:

LiquidHaskell





Verification Challenges

- Unclear divide between Haskell and LiquidHaskell — 11 New, 5 Exp
- Confusing verification features — 11 New, 4 Exp
- Unfamiliarity with Proof Engineering — 7 New, 5 Exp

> *I think that liquid types have, like, its own way to program. (CR1)*

Example from the tutorial for New Users

```
{-@ rot :: f:SList a
     -> b:SListN a {1 + size f}
     -> acc:SList a
     -> SListN a {size f + size b + size acc} @-}
rot front back acc
 | size front == 0 = hd back `cons` acc
 | otherwise =
    hd front `cons`
       rot (tl front) (tl back) (hd back `cons` acc)
```

Think about

- Where are `f` and `b` in the code?
- Why don't we need to check for an empty `back`?
- Is this an idiomatic way to write code?
- How do you create these requirements for your code?
- What can you write in the specifications?

> *I think that if I had to come up with the type of rotate by myself, trying to figure out what these invariants should be might be quite difficult. (NC6)*

Scalability Challenges

- Limitations of Automation and Manual Proof Flexibility — 7 Exp
- Scalability and Solver limitations — 6 Exp

5+ hours to verify the full project (FM4)

When automation fails, what do you need to add, and how do you get to a point where it works?

The solver itself has limitations:
- Function translation;
- Bugs…

Developer Experience Challenges

- Unhelpful Error Messages — 9 New, 4 Exp
- Limited IDE Support — 11 New, 4 Exp
- Insufficient learning and reference resources — 7 New, 5 Exp
- Complex Installation and Setup — 7 New, 5 Exp

> *I tried to install Liquid Haskell (...) before with this interview today, and I failed miserably. So that's one reason why I might not be too inclined to use it in the future. (NC17)*

Example from the tutorial for New Users

```
{-@ measure qsize @-}
qsize          :: Queue a -> Int
qsize (Q f b) = size f + size b

{-@ type QueueN a N = {v:Queue a | qsize v = N} @-}

-- Testing
✖ {-@ emp :: QueueN _ 0 @-}

{-@ example2Q :: QueueN _ 2 @-}
✖ example2Q = Q (1 `cons` (2 `cons` nil)) nil
```

Error indirection & code mismatch

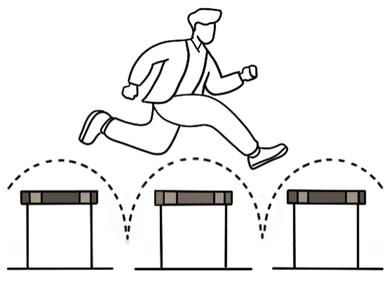
No syntax highlighting

Using comments feels "fake"

Error message

```
Error: Illegal type specification for
Tutorial_09_Case_Study_Lazy_Queues.emp
  Tutorial_09_Case_Study_Lazy_Queues.emp :: forall a .
  {VV : (Tutorial_09_Case_Study_Lazy_Queues.Queue a) | qsize == 0}

Sort Error in Refinement: {VV#0 :
(Tutorial_09_Case_Study_Lazy_Queues.
Queue a##a2cb) |
Tutorial_09_Case_Study_Lazy_Queues.qsize == 0}
    The sort func(0 , [(Tutorial_09_Case_Study_Lazy_Queues.Queue @(42));
int])
is not numeric
    because
Cannot unify func(0 , [(Tutorial_09_Case_Study_Lazy_Queues.Queue @(42));
int]) with int in expression: Tutorial_09_Case_Study_Lazy_Queues.qsize
== 0
    because
Invalid Relation Tutorial_09_Case_Study_Lazy_Queues.qsize == 0 with
operand types func(0 , [(Tutorial_09_Case_Study_Lazy_Queues.Queue
@(42)); int]) and int
```



Are these barriers unique to LiquidHaskell?

In Liquid Types: LiquidJava, Flux

In other verification tools: JML, Dafny, Agda, KeY, Frama-C, Why3

Expert survey on formal methods in 2020

Together, we can cross these barriers and empower developers to build better quality software.

Read our Paper
Accepted at
PLDI '25
In Seoul, South Korea, in June 2025

LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · S3D · Carnegie Mellon Portugal

Acknowledgements: This work is supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the Carnegie Mellon Portugal Program under the fellowship PRT/BD/154254/2021, the LASIGE Research Unit, ref. UID/00408/2025, and the RAP project (https://doi.org/10.54499/EXPL/CCI-COM/1306/2021)

SafeRoads - Toward next generation of safe roads for people and biodiversity

Lara Ângelo[1], César Capinha[3], Rui Rebelo[2], Ana Paula Afonso[1], Carlos Duarte[1], Fernando Ascensão[2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] cE3c, Faculdade de Ciências, Universidade de Lisboa, Portugal
[3] IGOT, Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

Problem

Traditional methods of roadkill mitigation usually fail to account for dynamic factors like time of day, weather, and animal behavior

Amphibians are the most endangered vertebrates globally!

Why?

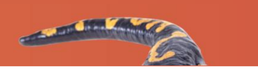
Amphibians face high mortality rates when migrating across roads, threatening local populations and contributing to the biodiversity crisis

Objective

Develop a mobile application that helps reduce roadkill incidents, particularly amphibians, by providing real-time, dynamic, data-driven solutions

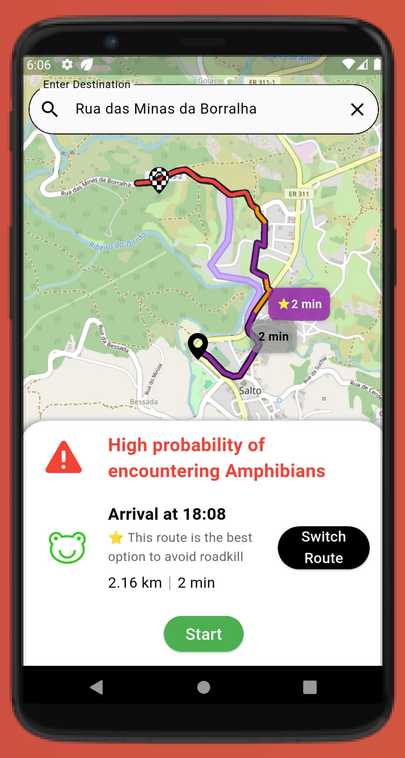
How?

- Integrating a machine learning-based predictive risk model that updates daily and informs users about high-risk zones for wildlife collisions
- Make drivers aware of potential roadkill risks and guide them to safer routes



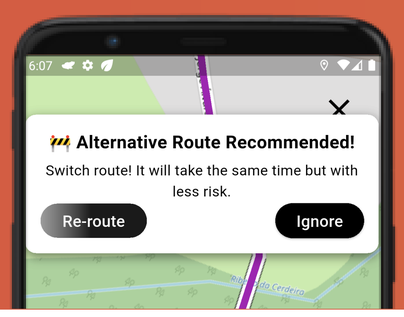
Daily updates on roadkill risks using factors like weather, and time of day

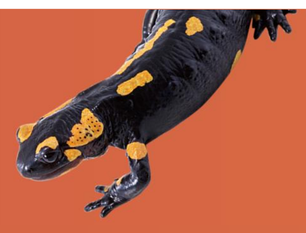
Roadkill Risk Assessment

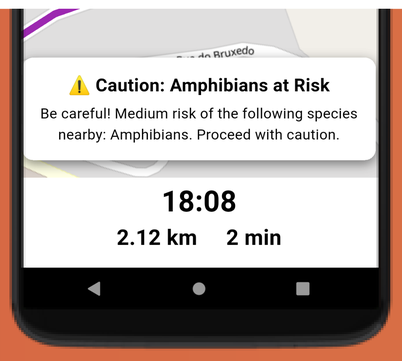


Re-routing options to avoid higher-risk areas

Alternative Route Suggestions







Real-Time Driver Alerts

Visual, auditory and haptic signals warn users when approaching high-risk zones

Expansion & Evaluation

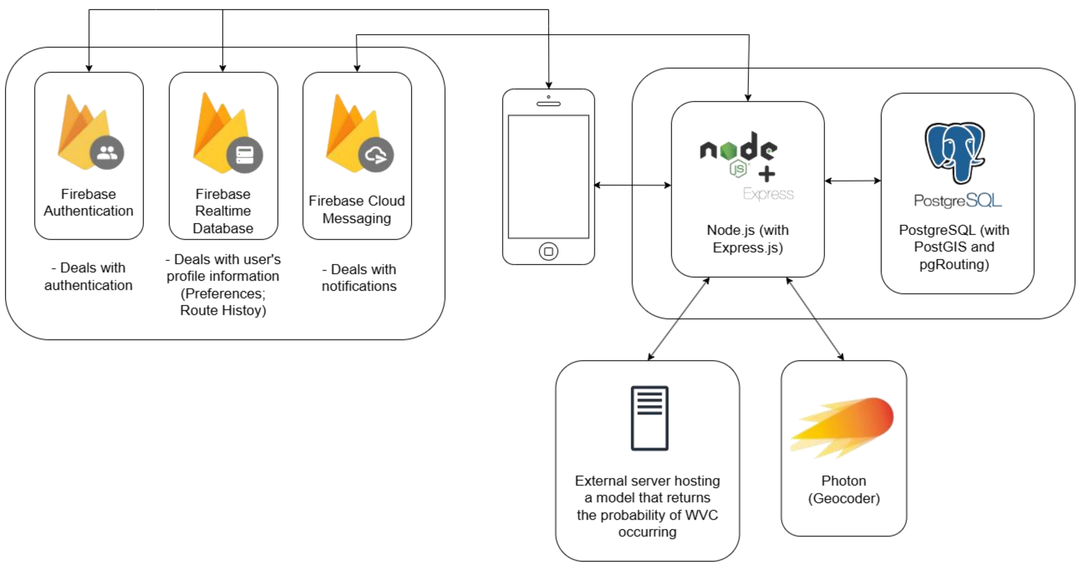
SafeRoads can extend by integrating more species, such as snakes and hedgehogs

The app also has the potential to expand beyond Portugal

Help us save these animals! Sign up to test our app!

SafeRoads

Architecture



Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UID/00408/2025 and the CE3C Research Unit, ref. 10.54499/UIDB/00329/2020.

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE | ce3c centre for ecology, evolution and environmental changes

Exploring Explainable AI in Emotion Recognition

Rute Patuleia[1], Cláudia Ribeiro[2], Soraia M. Alarcão[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] IADE, Faculdade de Design Tecnologia e Comunicação, Portugal

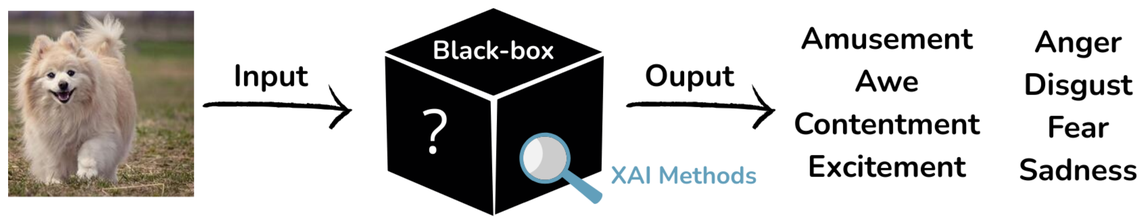
LASIGE data and systems intelligence

Motivation

- AI-based emotion recognition models have a **"black-box" nature**, which limits trust in their predictions
- It remains unclear why a specific image is associated with a given emotion in emotion recognition models. This gap between low-level image features and high-level emotional response is known as the **affective gap**
- **Explainable AI (XAI)** methods attempt to bridge this gap, but their effectiveness in emotion recognition tasks remains underexplored

Aim

- **Explore benchmark XAI methods** for emotion recognition tasks
- **Evaluate the quality of explanations** through quantitative, semi-quantitative, and qualitative analysis



Methodology

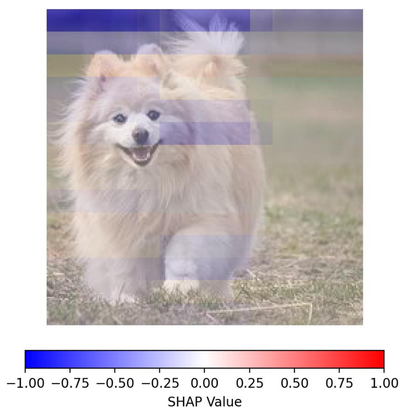
Model Training: **VGG16** on **EmoSet dataset**

XAI Methods: Each method was applied to test images to generate **saliency maps** that highlight the regions influencing the model's predictions

Grad-CAM

LIME

SHAP



Evaluation:

1. Quantitative Analysis

The quality of the explanations will be assessed using four **XAI methods' properties** to measure how informative, focused, necessary, and sufficient an explanation is

Property	Implementation Strategy
Fidelity	Remove the highlighted region and observe if the model prediction changes
Sparsity	Calculate the number of non-zero pixels of the highlighted region in the saliency map
Necessity	Remove only the highlighted region and check if the model still predicts the same class
Sufficiency	Keep only the highlighted region and check if the model still predicts the same class

2. Semi-Quantitative Analysis

This analysis focuses on using **EmoSet's attribute annotations** to assess the XAI explanations. The goal is to analyse how often highlighted regions overlap with emotion-relevant attributes

3. Qualitative Analysis

Conduct a user study with adult participants to complement the quantitative evaluation by assessing explanation quality from a human-centered perspective. Users will:

- Evaluate different XAI methods through misclassified and misexplained images
- Rate the clarity and usefulness of each explanation
- Evaluate how they feel when viewing the visual explanation in the saliency map

Contributions

- A **systematic benchmarking** of commonly used XAI methods in the context of image emotion recognition
- A novel use of **emotion-related attributes** to interpret and evaluate explanations as well as understand and mitigate the affective gap
- A **user-centric assessment** of explanation quality when applied to emotion classification

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · IADE

Acknowledgements: This work was supported by FCT through LASIGE Research Unit, ref. UID/00408/2025.

Processing Web Applications using NLP for Vulnerability Identification

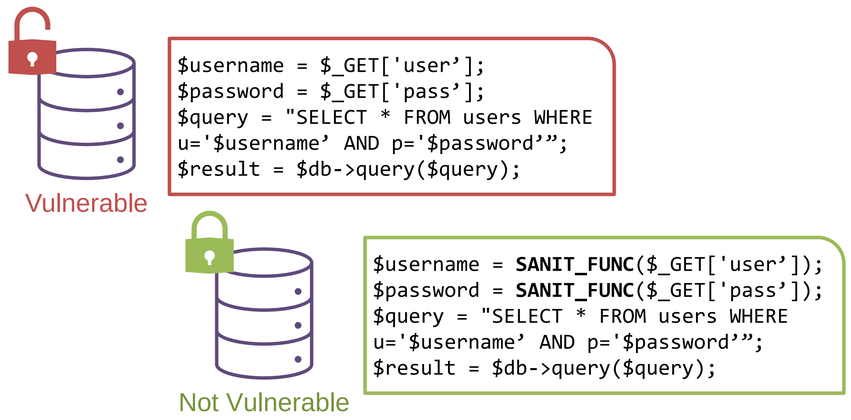
Jorge Guerreiro, Ibéria Medeiros

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE dependable and secure decentralized systems

Web Application Vulnerabilities

- Web applications are **compromised** when the source code has **vulnerabilities**;
- **SQL Injection** is a **common** and **critical** vulnerability that when exploited allows attackers to **steal data** from the database, **compromising** the web application.

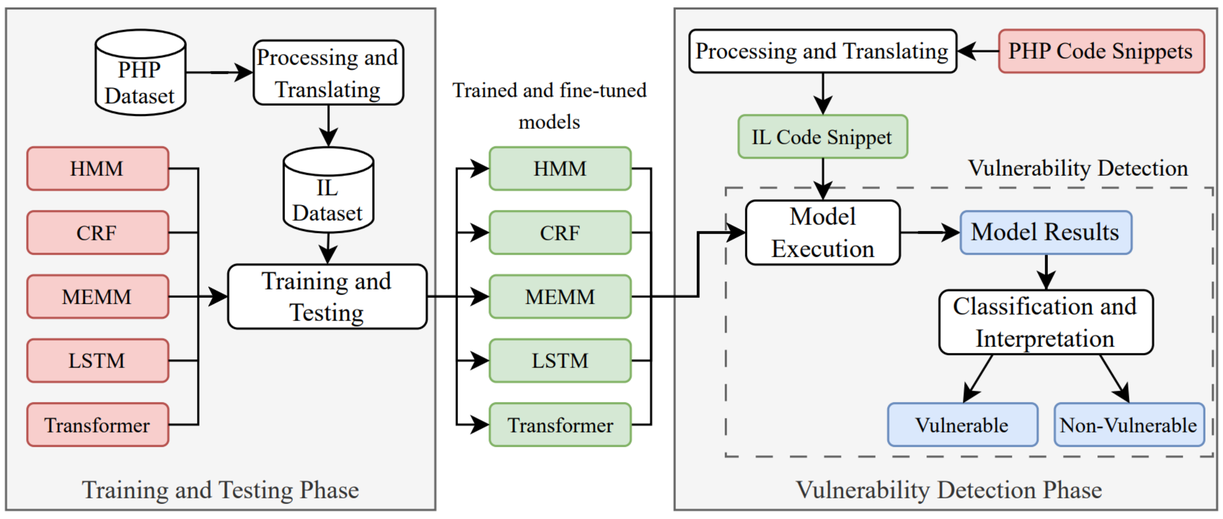


Approach

Objectives

- Represent PHP code similarly to a natural language to identify and explain vulnerabilities with NLP models – **Neural Networks** and **Sequence models**.
- Sequence models will explain/justify the vulnerabilities found by neural networks.

Architecture



A tool uses the models to find vulnerabilities in Intermediate Language code snippets. The models results undergo into a final step of classifying and interpreting to declare the snippets as "Vulnerable" or "Non-Vulnerable" and explain the classification.

Natural Language Processing (NLP)

- NLP **represents the meaning** of natural language text, allowing translation, classification, PoS tagging, and more.
- Transformers use high-dimensional vectors, where context, surrounding words, position, and other features are encoded into a vector representation.

Translation

PHP vulnerable to SQL Injection

```
$username = $_GET['user'];
$password = $_GET['pass'];
$query = "SELECT * FROM users WHERE u = '$username'
          AND p = '$password'";
$result = $db->query($query);
```

Intermediate Language

```
var1 GET
var2 GET
var3 SQL_QUERY var1 var2
var4 var5->query var3
```

Both snippet for classification and the dataset files to be used in the training cycle to find the best models, go through the translator. The models are executed in IL.

Preliminary Results

Dataset

1,362 PHP code snippets:
- 860 SQLi
- 502 safe

Transformers

Base Model	Accuracy	Precision	Recall	F1-Score
CodeBERT	0.93431	0.90426	**1.00000**	0.94972
GraphCodeBERT	**0.97080**	**1.00000**	0.95181	**0.97531**
RoBERTa	0.94161	**1.00000**	0.90909	0.95238
GPT-2	0.93431	0.90426	**1.00000**	0.94972

LSTM

Optimizer	Accuracy	Precision	Recall	F1-Score
Adam	**0.90511**	**0.94048**	0.90805	**0.92398**
AdaDelta	0.63504	0.63504	**1.00000**	0.77679
RMSProp	0.66423	0.65649	0.98851	0.78899

Conclusions & Future Work

- Applying NLP to find and explain vulnerabilities is a promising approach, and the results showcase the potential of this methodology;
- A dataset with more instances and classes of vulnerabilities can improve the results and assess the efficacy of this approach.

Acknowledgements: This work was partially supported by P2030 through project I2DT, ref. COMPETE2030-FEDER-00389100, an ITEA4 European project (ref. 22025), and by FCT through the LASIGE Research Unit, ref.UIDB/00408/2025.

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE EUREKA ITEA4 I2DT

Neuro-Symbolic AI for modelling Bio-Knowledge Graphs with contradictions

Laura Balbi, Catia Pesquita

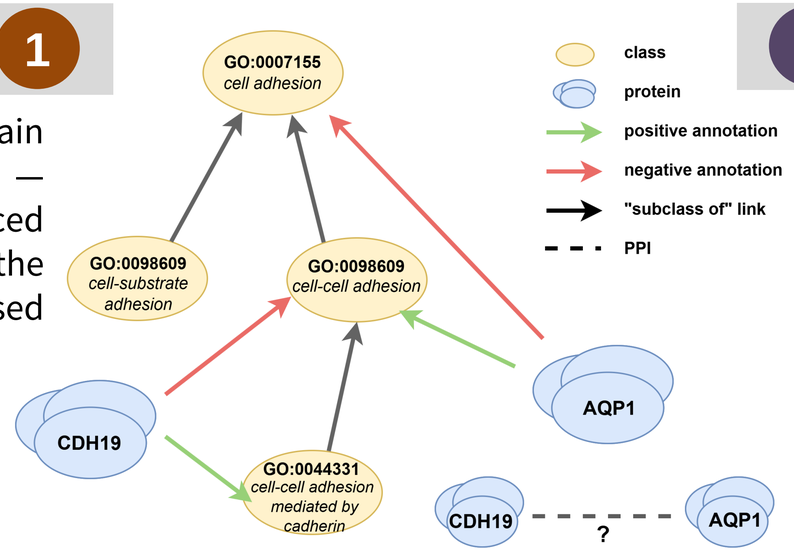
LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

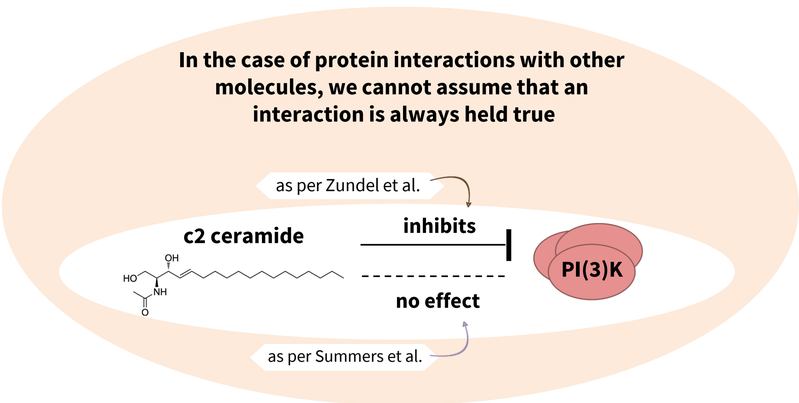
LASIGE health and biomedical informatics

1 Motivation

Knowledge Graphs (KGs) often contain contradictions that are not necessarily errors — they can represent polyvocal truths or nuanced realities that, if properly modeled, can enhance the completeness and expressiveness of KG-based machine learning.

In biomedical KGs this may be due to the universe of experimental conditions in which biological entities perform specific roles, functions and interactions.





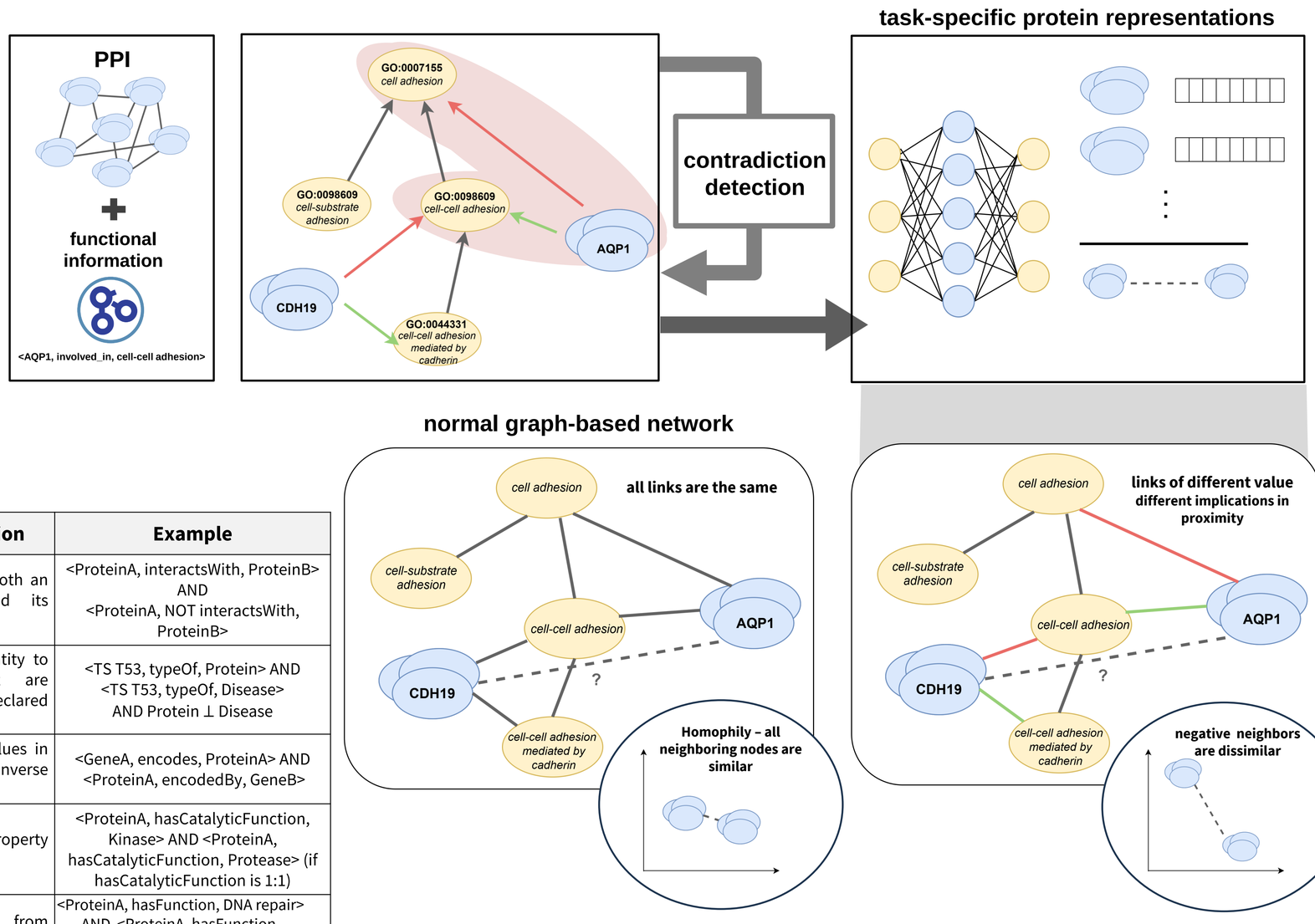
Current KG Representation Learning is not well-equipped to handle contradictory facts within KGs, especially those that are implicit and context-dependent. There is a lack of mechanisms to explicitly model or leverage contradictions during learning.

2 Hypothesis

An approach that integrates symbolic and sub-symbolic representations can bridge KGs and LLMs to answer:

- **RQ1:** How do contradictions in a KG, implicit and explicit, impact the SOTA KGRL performance for PPIs?
- **RQ2:** Does modelling contradictions onto protein representations improve ML?
- **RQ3:** Can external sources of knowledge such as LLMs be explored to detect implicit contradictions?

3 Methodology



Taxonomy of Contradictions

Explicit

Direct, logic-based conflicts that can be detected via rule violations or ontology constraints.

Implicit

Semantically conflicting or context-sensitive facts that do not violate formal logic alone, but appear contradictory under additional assumptions or external knowledge.

	Subtype	Description	Example
Direct	Negated Assertion	Presence of both an assertion and its negation.	<ProteinA, interactsWith, ProteinB> AND <ProteinA, NOT interactsWith, ProteinB>
Direct	Disjoint Classes	Assigns an entity to classes that are explicitly declared disjoint.	<TS T53, typeOf, Protein> AND <TS T53, typeOf, Disease> AND Protein ⊥ Disease
Inverse Property Conflict		Conflicting values in mutually inverse properties.	<GeneA, encodes, ProteinA> AND <ProteinA, encodedBy, GeneB>
Indirect	Cardinality/ Domain/Range Violation	Violation of property constraints.	<ProteinA, hasCatalyticFunction, Kinase> AND <ProteinA, hasCatalyticFunction, Protease> (if hasCatalyticFunction is 1:1)
Indirect	Hierarchically Inferred	Conflict from inherited logical constraints.	<ProteinA, hasFunction, DNA repair> AND <ProteinA, hasFunction, Negative regulation of DNA repair> AND DNA repair ⊥ Negative regulation of DNA repair

4 Baselines

ML	KG	# Edge Types	Precision	Recall	ROC-AUC
HGCN	Full	1	0.9440	0.3193	**0.9023**
		2	**0.9574**	0.2316	0.9002
		5	0.9402	**0.3536**	0.8916
	no LC	1	0.9444	0.3153	**0.8998**
		2	**0.9546**	0.2675	0.8976
		5	0.9441	**0.3844**	0.8888
	Negative Annots	1	0.9346	0.3369	**0.8954**
		1	0.9391	0.2869	0.8909
		4	**0.9398**	**0.3640**	0.8904
	Positive Annots	1	**0.9477**	0.3045	0.9007
		2	0.9475	0.3259	**0.9010**
		4	0.9429	**0.3687**	0.8912

ML	KG	# Edge Types	Precision	Recall	ROC-AUC
HGAT	Full	1	0.9375	**0.1546**	0.6913
		2	**0.9452**	0.1438	0.6642
		5	0.9386	0.1414	**0.7030**
	no LC	1	0.9310	**0.1491**	0.6778
		2	0.8853	0.1140	0.6255
		5	**0.9391**	0.1415	**0.6985**
	Negative Annots	1	0.9269	**0.2175**	0.6777
		1	0.9105	0.1945	0.6990
		4	**0.9393**	0.1639	**0.7021**
	Positive Annots	1	0.9487	0.1419	**0.6888**
		2	0.9470	**0.1504**	0.6636
		4	**0.9599**	0.1057	0.6775

Preliminary discussion

value of including negative knowledge
only negative annotations outperform

advantage in actively modelling opposing relations and contradictions in separate
consensual increase in performance from the GCN to the HGCN

GAT weighs edges according to their importance across edge types
GAT overperformed Bi- and HGAT

Acknowledgements: This work was supported by FCT through the fellowship 2024.01208.BD, and the LASIGE Research Unit, ref. UID/00408/2025. It was also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453.

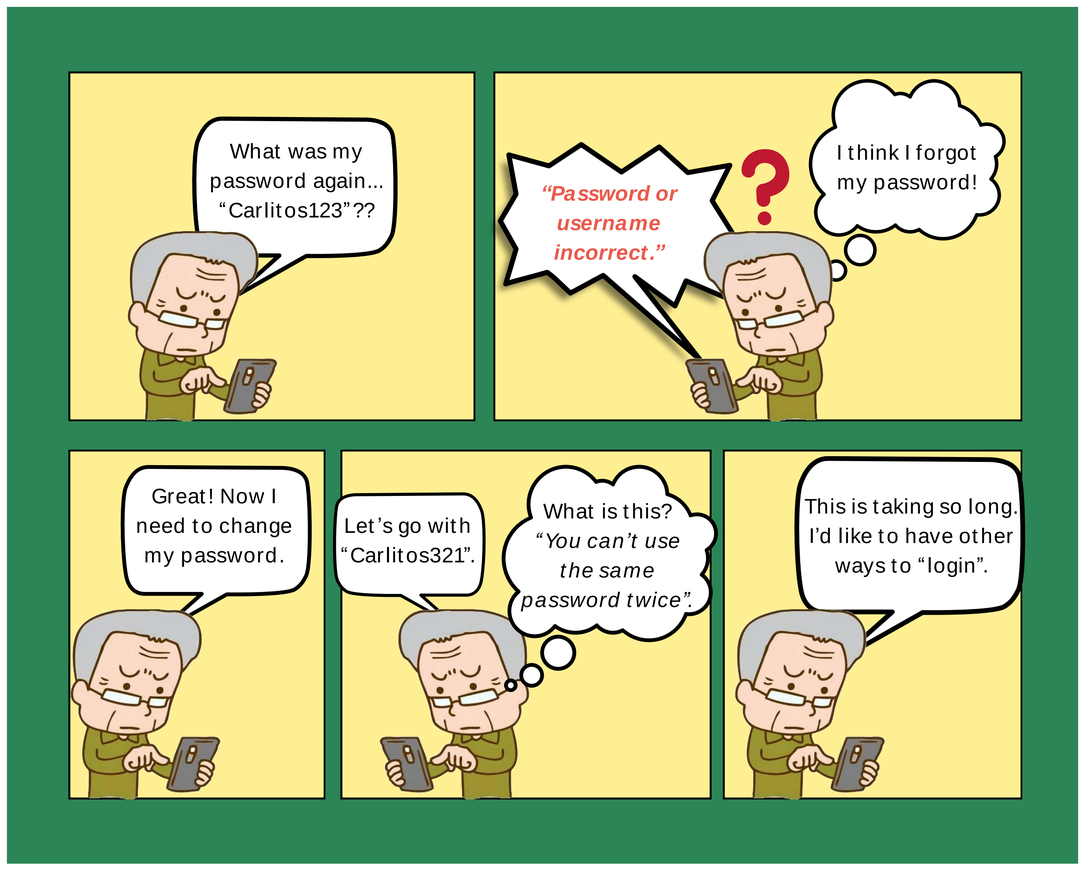
U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Studying the Accessibility of Mobile Authentication for Older Adults

Joaquim Simões[1], Vânia Mendonça[1], Soraia M. Alarcão[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE inclusive human-computer interaction



Goal

To analyse different authentication methods with older adults:

- **RQ1:** How do different methods compare in terms of usability?
- **RQ2:** How do different methods compare in terms of users' perception of security?
- **RQ3:** What are the main challenges experienced by older adults when faced with the methods studied?

Proposal

Method selection:

- *what you know* (e.g., password)
- *what you have* (e.g., SMS token)
- *who you are* (e.g., fingerprint)

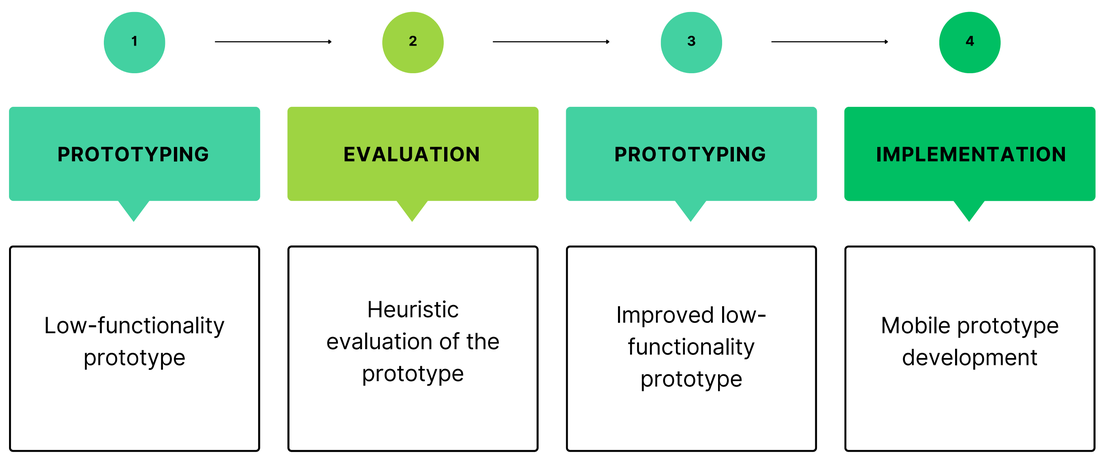
Two sessions:

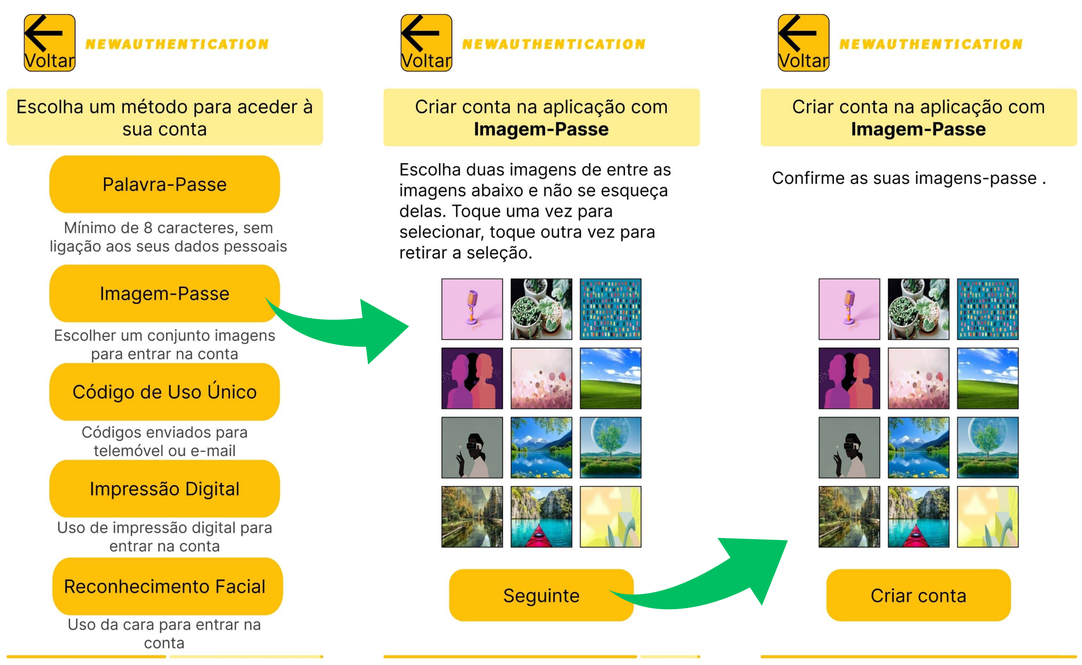
- #1: creating accounts and logging into them
- #2: one week later, seeing if the participants remember their credentials

Metrics and instruments:

- Time taken to create accounts and login, as well the errors made
- STAM, ASQ, MDPQ questionnaires

Work in Progress





Creating an account with pass-image

Next Steps

- Final prototype implementation
- Pilot testing
- Participant recruiting
- Evaluation with older adults

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UID/00408/2025

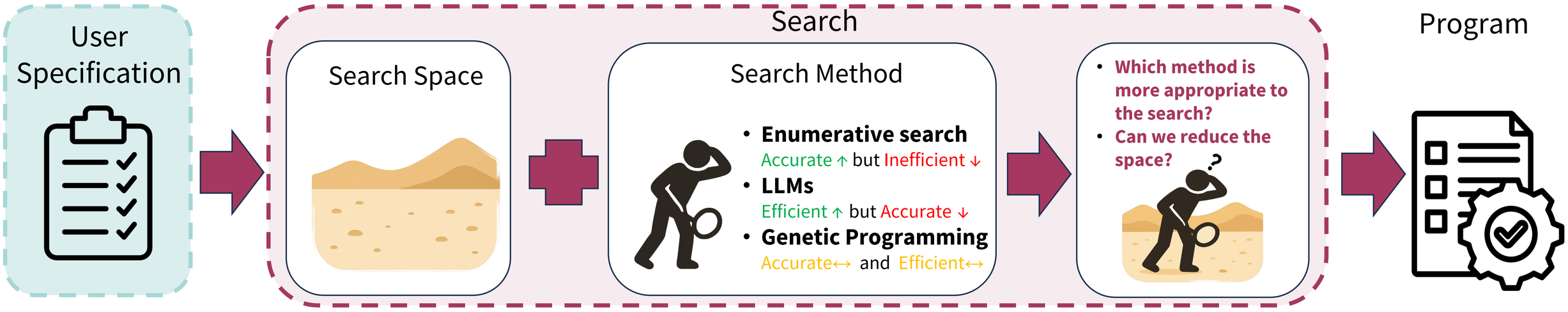
The Impact of Dependent Types on Program Synthesis Efficiency

Lishun Su, Diogo Sousa, Alcides Fonseca

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE reliable software systems

Program Synthesis



Aeon Programming Language

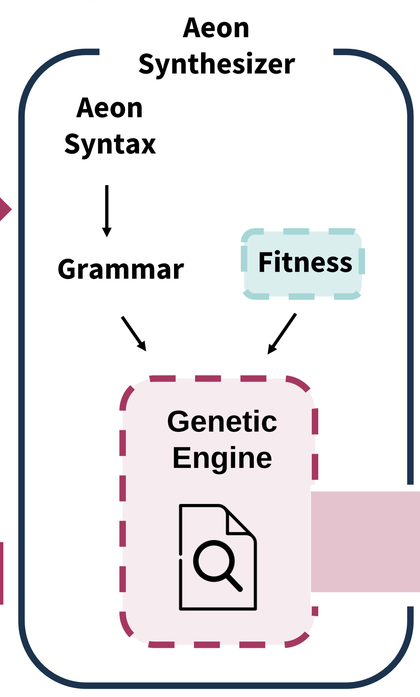
- Built for program synthesis with native support for dependent types.

```
@maximize(len subset)
def subsetSum
    (ps : List[int])
    (target : Int | target > 0):
    (subset : List[int]
        | target == sum subset)
{
    (?hole: List[int])
}
```

LSP

```
?hole   : List[int]
?values = [1,1,1,1]
          [1,1,2]
          [1,2,1]
          [2,1,1]
          ...
```

Possible Solutions



Hybrid Approach

LLM Part

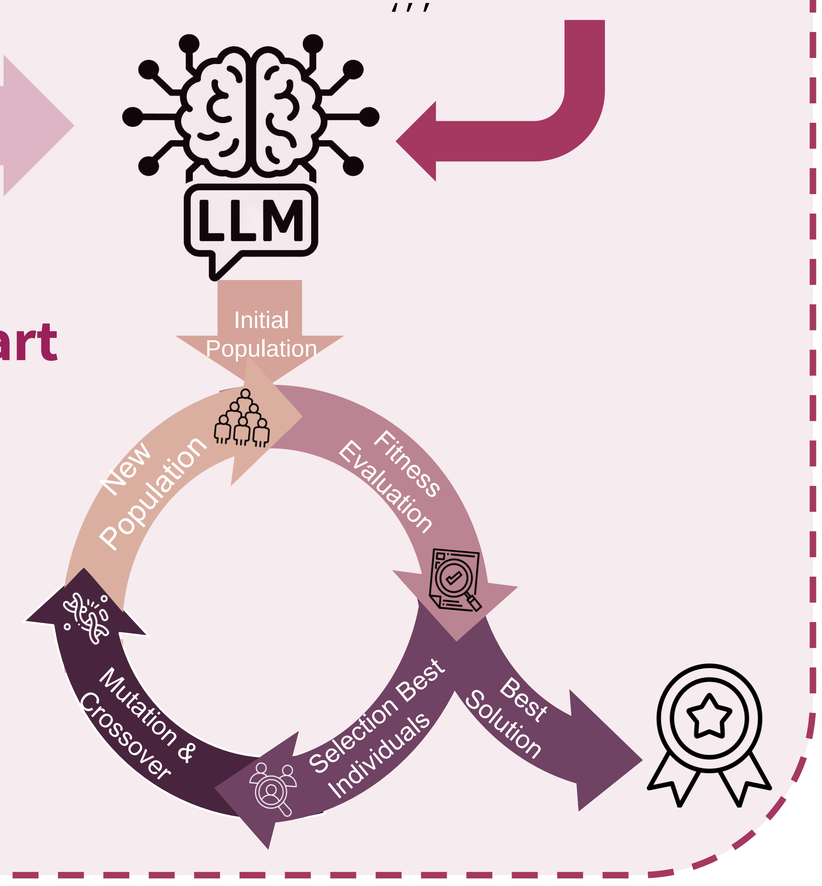
Data Class

```
@dataclass
class Queen:
    posX: Annotated[int, IntRange[0,8]]
    posY: Annotated[int, IntRange[0,8]]

@dataclass
class NQueens:
    queens: Annotated[list, ListSizeBetween(8,8)]
```

Prompt

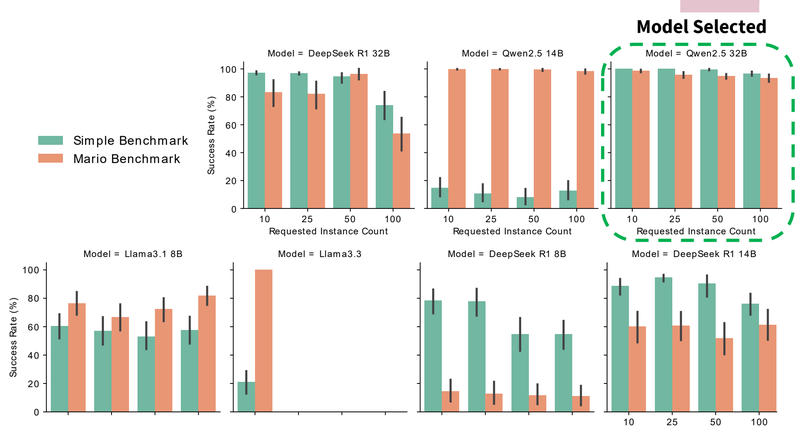
```
Given the following
Python code, that uses
dataclasses and custom
annotations, please
create exactly 100
instantiations of class
Nqueens, show all of
them, one per line,
don't show keyword
argument, with diversity
and don't generate
python code.
'''
[code of DataClass]
'''
```



LLM Selection

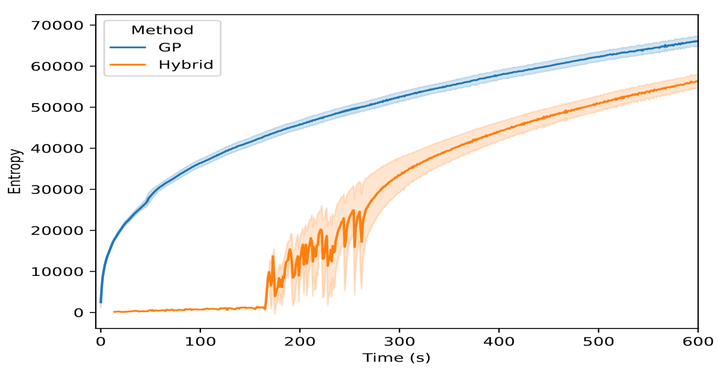
Model Evaluated

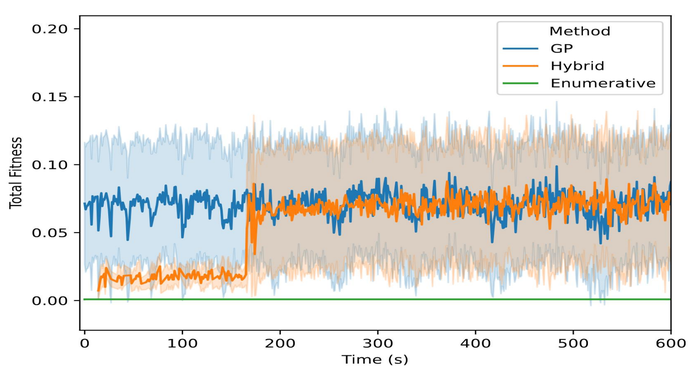
deepseek
LLaMA
Qwen



- **Simple Benchmark** – a data class with four variables and three are integers and a list of integer
- **Mario Benchmark** – multiple data classes and hierarchy between the classes

Evaluation of Hybrid Approach(↑ Better)





LLMs introduced **less diversity** and **added ~180s of overhead**, offering **limited improvement to the initial population**.

Based on this result, we chose GP as our search method.

Acknowledgements: This work was supported by Fundação para a Ciência e Tecnologia (FCT) in the LASIGE Research Unit under the ref. (UID/00408/2025)

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

Vehicle-to-Vehicle Flooding Datasets using MK5 On-board Unit Devices

Breno Sousa[1], Naercio Magaia[2], Sara Silva[1], Nguyen Thanh Hieu[3], Yong Liang Guan[3]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] School of Engineering and Informatics, University of Sussex, United Kingdom
[3] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

Full Paper

LASIGE cyber-physical systems

Introduction

Motivation

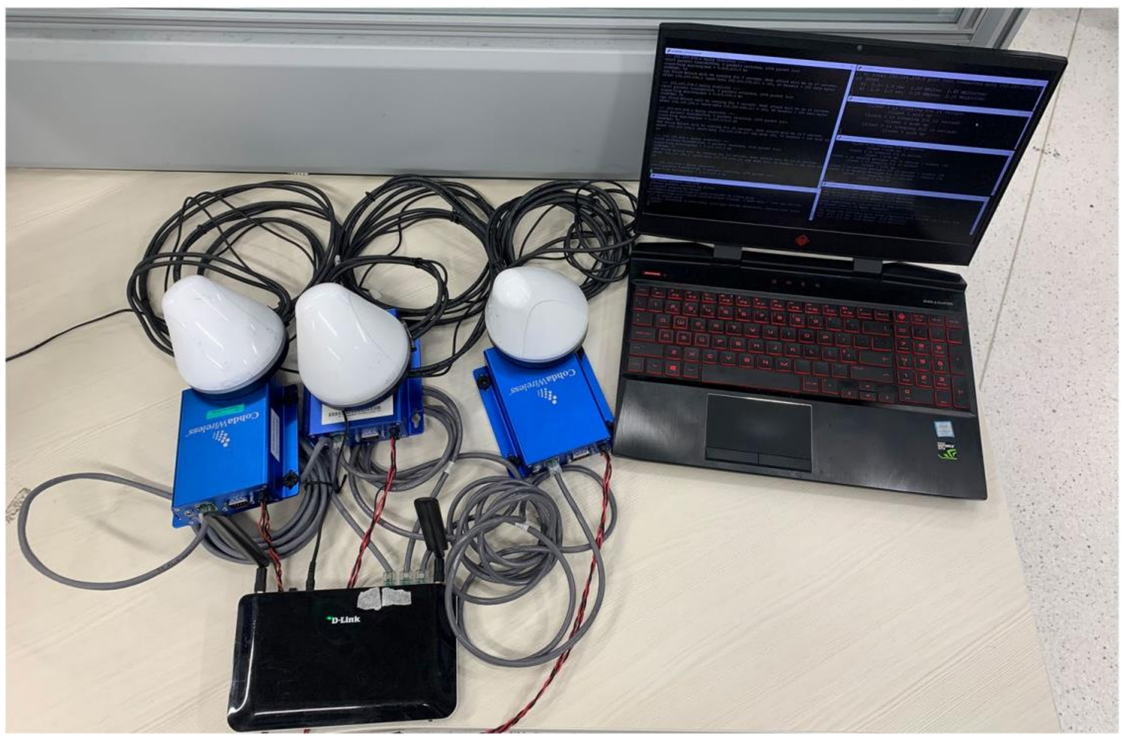
Ensuring information availability is critical in vehicular networks, where disruptions can compromise safety. Flooding attacks have been widely studied, but most datasets rely on simulations and lack real-world behavior.

Contribution

We generated flooding attack datasets using realistic vehicular devices (MK5 OBUs). These are the first public V2V flooding datasets based on real devices. Repository for download: https://doi.org/10.5281/zenodo.13220637

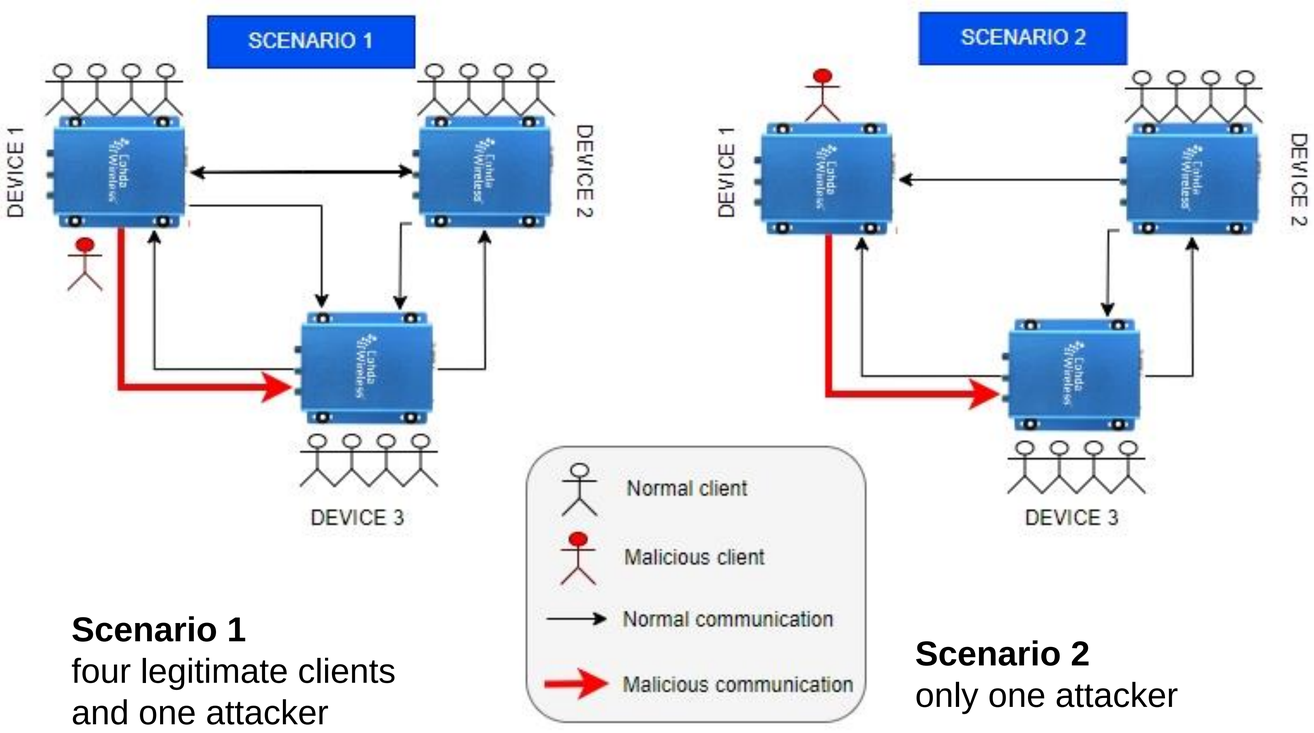
Our Testbed – Cohda MK5 OBU

Test bed consisting of three MK5 On-Board Unit (OBU) devices. They support Dedicated Short-Range Communications (DSRC), hence are equipped with an IEEE 802.11p radio.
Here, we emulated Vehicle-to-Vehicle (V2V) communication.



Attacking Scenarios

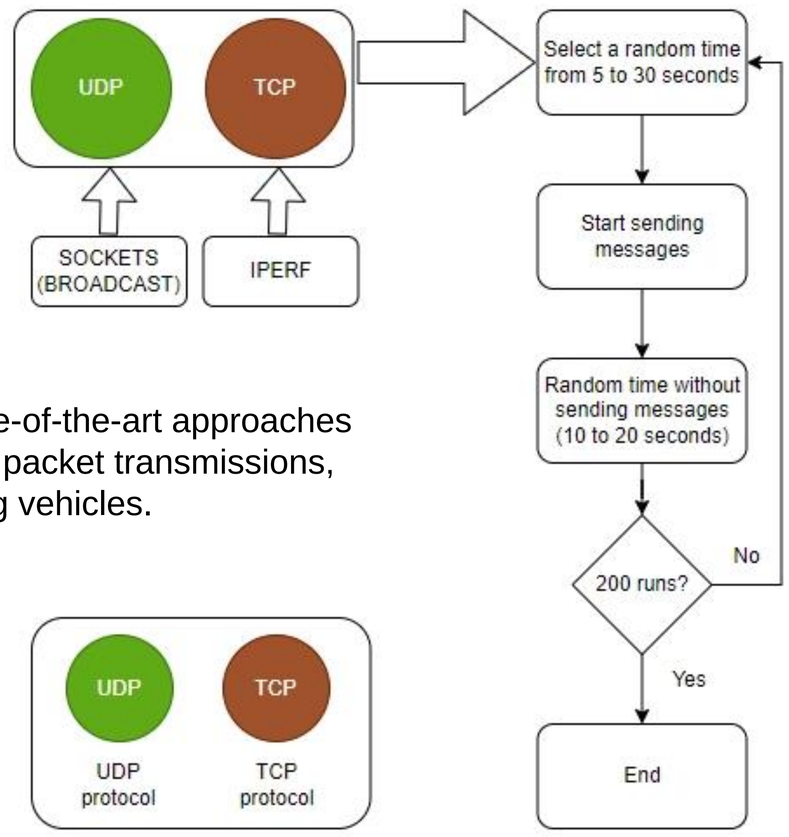
We generated three V2V flooding datasets, each with 91 features.
Two datasets from Scenario 1 and one dataset from Scenario 2.



Scenario 1
four legitimate clients and one attacker

Scenario 2
only one attacker

Communication Strategy



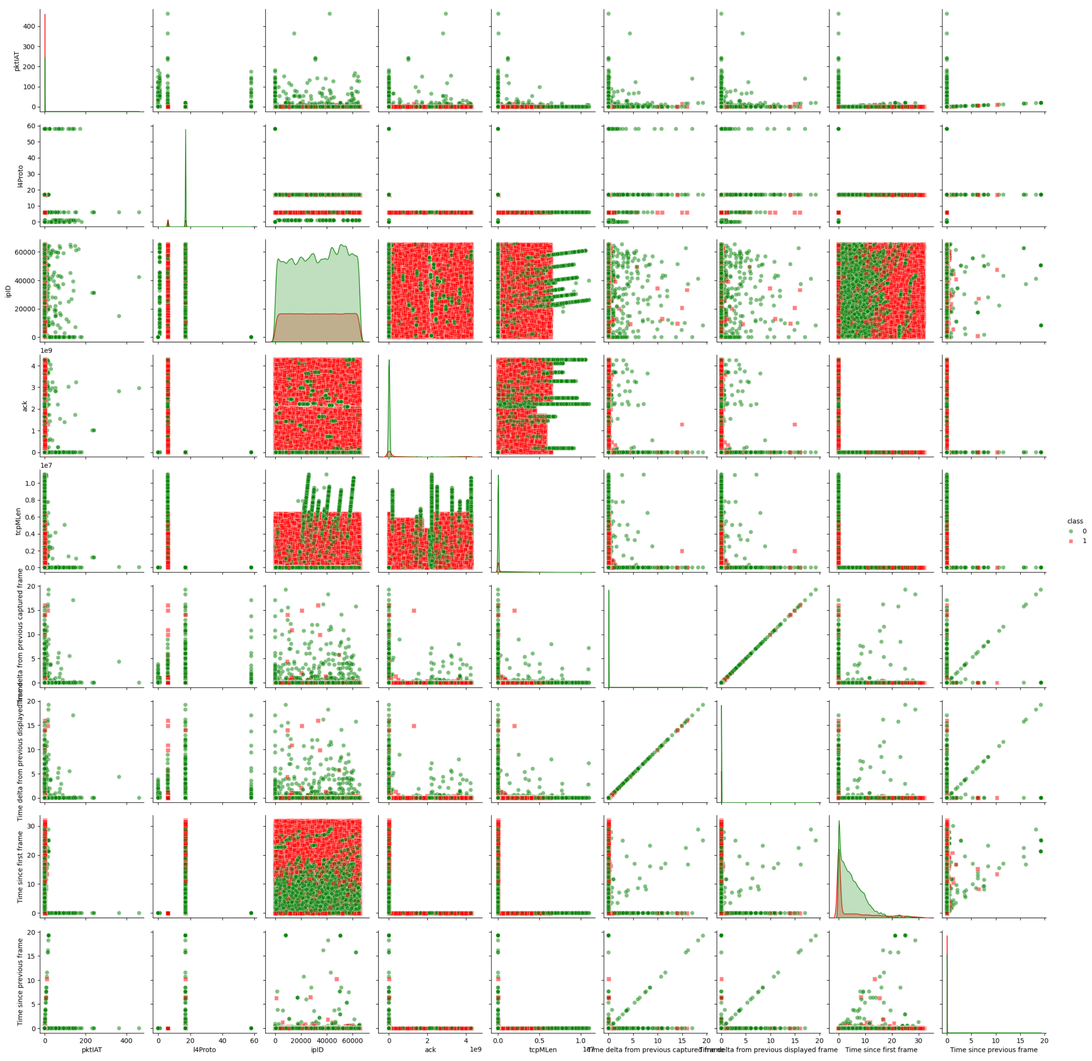
We followed state-of-the-art approaches to randomly time packet transmissions, emulating moving vehicles.

Datasets Overview

Scenario	Number of Rows Class 0	Number of Rows Class 1	Total Number of Rows
1a	2,151,768 (75%)	708,106 (25%)	2,859,874
1b	2,795,001 (80%)	695,784 (20%)	3,490,785
2	2,185,811 (58%)	1,572,289 (42%)	3,758,100

Data Visualization

Pairwise scatterplots and class distributions of 23 selected features.



This work was supported by:
H2020-MSCA-RISE under grant No. 101006411;
Fundação para a Ciência e a Tecnologia (FCT), Portugal, through doctoral grant SFRH/BD/151413/2021 (https://doi.org/10.54499/SFRH/BD/151413/2021) under the MIT Portugal Program, and through funding of the LASIGE Research Unit, ref. UID/00408/2025;
Temasek Laboratories@NTU seed research project No. TLSP24-05.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Can Language Models Align Biomedical Ontologies?: Evaluating Retrieval-Augmented Prompt Strategies in Bio-ML

Lucas Ferraz, Pedro Giesteira Cotovio and Catia Pesquita

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE data and systems intelligence

1. The problem and hypothesis

What is **Ontology Alignment**?

Ontology alignment is the process of identifying semantic correspondences (mappings) between entities in different ontologies.

It enables **interoperability between systems** that use distinct vocabularies to represent similar or overlapping domains, a critical need in biomedical data integration.

Why is it **difficult**?

Most ontologies are developed independently, using varying labels, structures, and modeling assumptions.

Many **mappings require deep domain knowledge and context** that is not captured by simple lexical or structural features alone.

Could a direct **LMs + RAG** approach help?

Language Models (LMs) can understand complex semantic relationships between concepts beyond surface text.

Retrieval-Augmented Generation (RAG) enhances LMs by injecting contextual ontology data into the prompt, helping models account for meaning, hierarchy, and equivalence.

2. Our proposed approach

To explore how semantic prompting affects biomedical ontology alignment, we developed a **modular prompt design framework**.

This framework builds prompts by **combining reusable components**, such as task instructions, source/target labels, and contextual information, such as hierarchical relations.

Our approach operates in two stages:

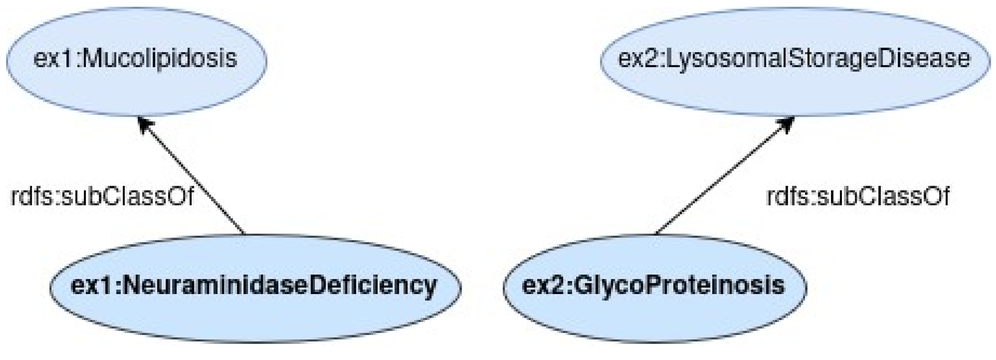
- **Static Stage:** We create generic template structures using symbolic placeholders (e.g., $S, $T, $CTX_S, $CONF).
- **Dynamic Stage**: We inject ontology-specific labels and hierarchy data into these templates to form complete, context-aware prompts.

Each prompt is used to query a language model about the equivalence between two ontology classes. The model's textual response is then parsed to extract binary decisions and confidence scores.

Element Tag	Description
$TC (Task Context)	A description of the task the model should perform.
$I (Instruction)	A description of the nature of the question that will be asked and the expected answer format.
$CONF (Confidence)	The type of confidence that the model should output (if any).
$S (Source)	The main label(s) of the source ontology entity.
$CTX_S (Source Context)	Labels of potentially meaningful entities to the source entity.
$T (Target)	The main label(s) of the target ontology entity.
$CTX_T (Target Context)	Labels of potentially meaningful entities to the target entity.
$TYPE (Equivalence Type)	The type of equivalence to be assessed.

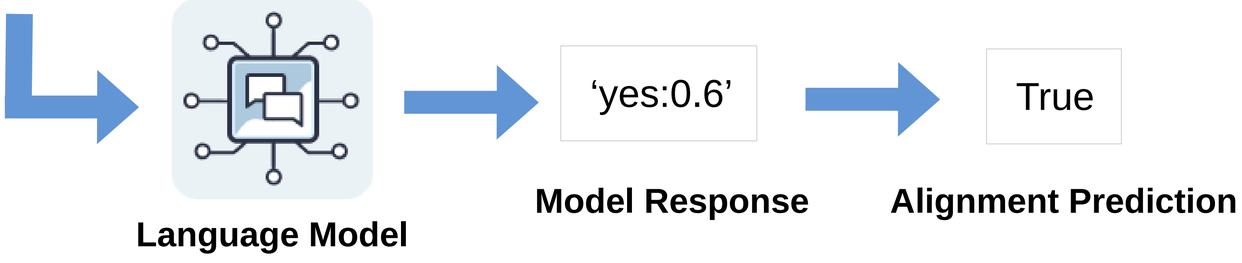
3. Individual alignment prediction example

Extraction of labels and hierarchy data from the ontologies



Prompt construction from labels + hierarchical info

You are doing an ontology alignment task. I am going to ask you a question and you should answer 'yes' or 'no', followed by your confidence in your answer as a score from 0 to 1, like this: 'yes:0.8'. Question: Are 'Neuraminidase Deficiency' (a kind of Mucolipidosis) and 'glycoproteinosis' (a kind of lysosomal storage disease) equivalent?



4. Results and conclusions

MODEL		w/o HC		w/ HC (kind of)		w/ HC (subclass of)	
		Pred: 1	Pred: 0	Pred: 1	Pred: 0	Pred: 1	Pred: 0
flan-t5-base	Actual: 1	1	5	3	3	1	5
	Actual: 0	0	6	1	5	0	6
Claude 3.7 Sonnet	Actual: 1	3	3	5	1	4	2
	Actual: 0	0	6	1	5	0	6
OpenAIo1	Actual: 1	4	2	4	2	3	3
	Actual: 0	0	6	0	6	0	6
GPT4o	Actual: 1	5	1	4	2	3	3
	Actual: 0	1	5	0	6	0	6
GPT4	Actual: 1	3	3	5	1	3	3
	Actual: 0	1	5	0	6	0	6

The value of adding semantic context is heavily influenced by the verbalisation used when designing prompts.

Smaller models like OpenAIo1 and GPT4o outperformed larger models like GPT4 when no hierarchical context was included in the prompt.

Future work: In-context learning + extra ontology features

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

This work was supported by FCT through the fellowships 2022.10557.BD (Pedro Cotovio), and the LASIGE Research Unit, ref. UID/000408/2025. It was also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453. This work was also supported partially by project 41, HfPT: Health from Portugal, funded by the Portuguese Plano de Recuperação e Resiliência.

Improving the Efficiency of Multi-Cloud Storage

Guilherme Santos, Alysson Bessani

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE dependable and secure decentralized systems

Multi-Cloud Advantages

No vendor lock-in

Combines **multiple cloud providers** into **one unified storage system**

aws, Azure, BACKBLAZE, ORACLE CLOUD, Google Cloud Platform, wasabi

Cost-Effective

Reduce storage costs by **using erasure codes** instead of replication

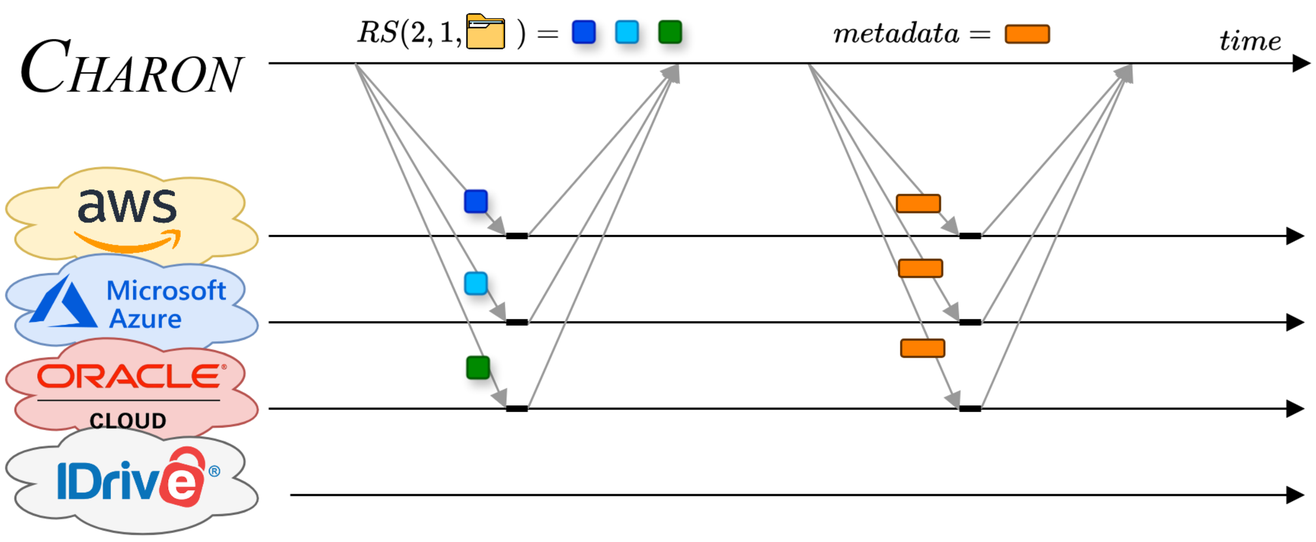
Resilience

Data remains **accessible and secured** even during failures

CHARON

- Cloud-backed storage system capable of storing data using multiple untrusted cloud providers
- Metadata and data are written in **separate quorums accesses**
- Uses **Byzantine quorum systems** and **erasure codes**

For some cloud provider combinations, it is possible **to reduce storage costs** by uploading the data across **all available clouds** instead of a quorum

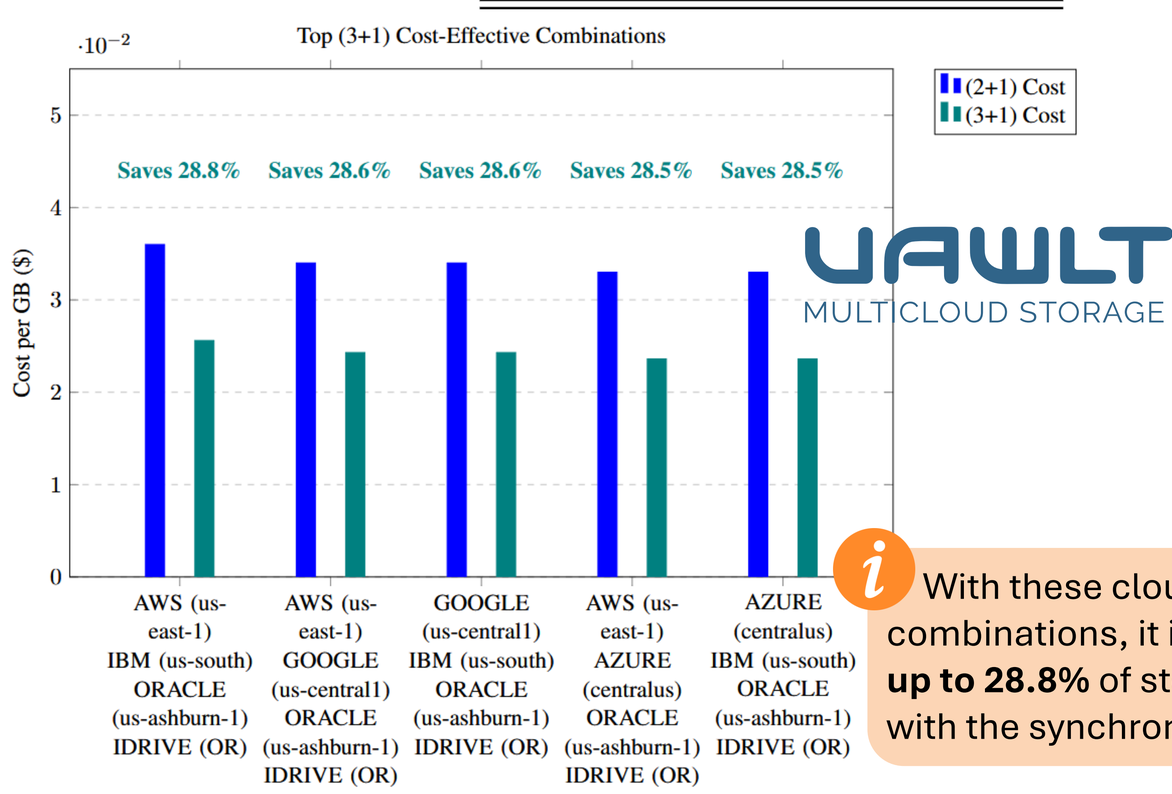


Analysis

Storage Overhead

The synchronous approach has a **lower** storage overhead

	Synchronous f			Asynchronous f		
n	0	1	2	q	1	2
4	0%	33.3%	100%	3	50%	-
5	0%	25%	66.7%	4	33.3%	-
6	0%	20%	50%	4	33.3%	-
7	0%	16.7%	40%	5	25%	66.7%
8	0%	14.3%	33.3%	6	25%	50%

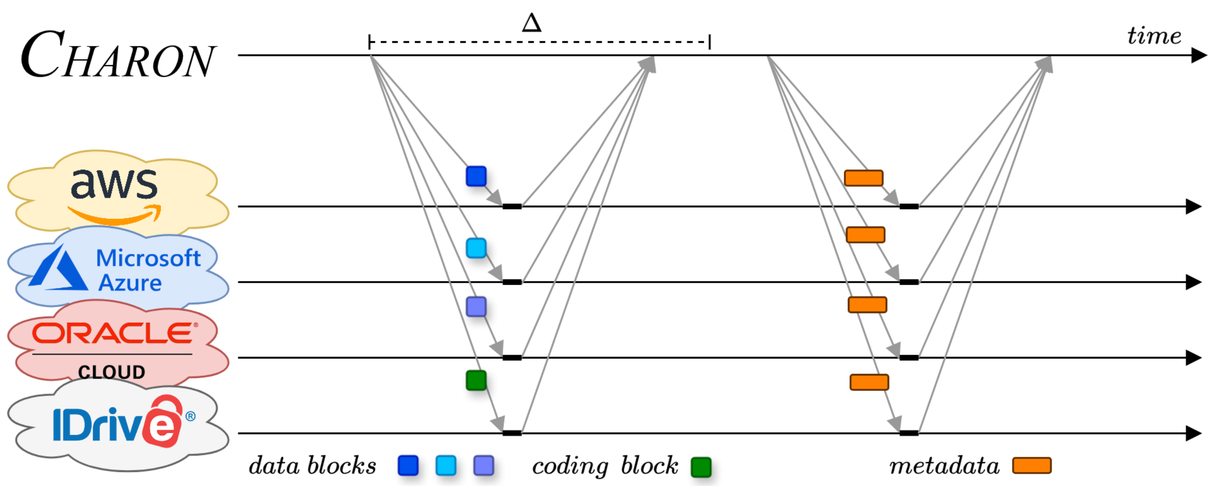


UAWLT MULTICLOUD STORAGE

With these cloud provider combinations, it is possible to **save up to 28.8%** of storage costs per GB with the synchronized approach

Approach

- CHARON computes the most cost-effective approach
- In case it is the synchronous approach, use RS(3,1) **(33% overhead)** instead of RS(2,1) **(50% overhead)**
- Upload data to **all *n* clouds** within Δ timeout
- A timeout reverts the write to the quorum approach



Acknowledgements: This work was supported by the Digital European Programme under the Grant Agreement 101083770 and from the Recovery and Resilience Plan (PRR) within the scope of the Recovery and Resilience Mechanism (MRR) of the European Union (UE), framed in the Next Generation EU, for the period 2021-2026, within project ATTRACT, with reference 774, and by the FCT through the LASIGE Research Unit, ref. UID/000408/2025

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · ATTRACT

Negative Statements in Protein-Protein Interaction Networks

Pedro Serrano, Cátia Pesquita

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

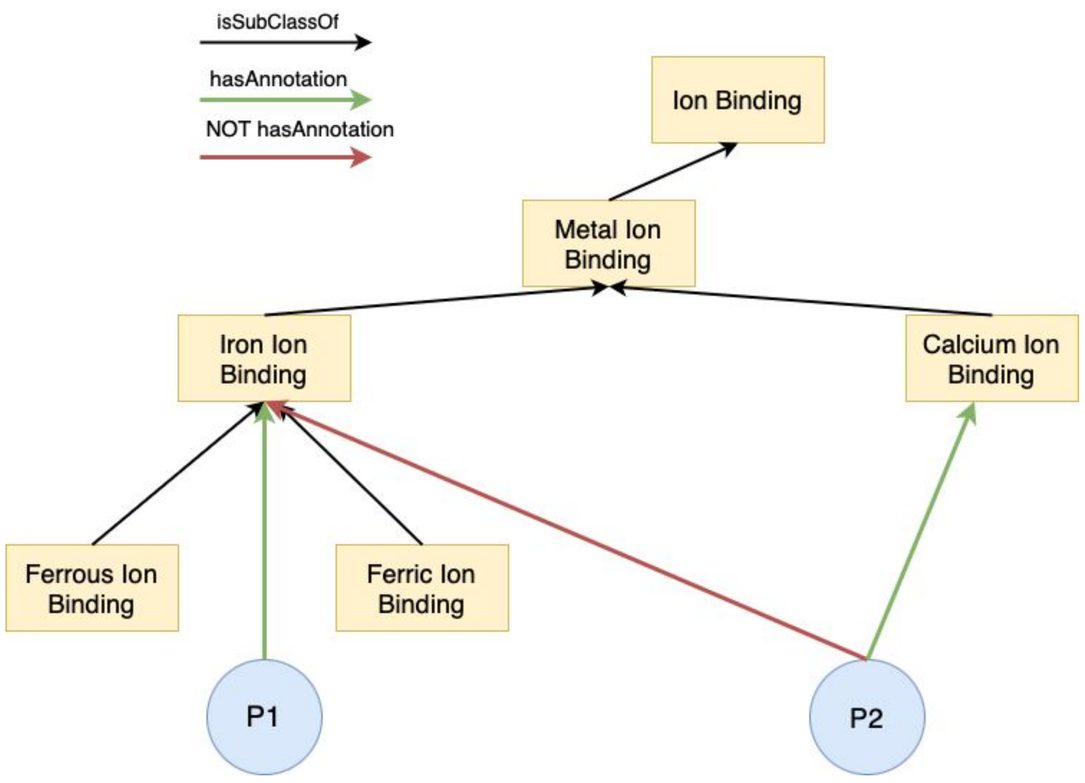
LASIGE health and biomedical informatics

Motivation

Knowledge Graphs have been used for a variety of different Machine Learning Tasks. Recently, the relevance and importance of negative statements has been increasingly recognized.

These negative statements have different semantic implications than positive statements. Some of these implications go against core concepts of some ML tasks.

GCNs have been used for predicting new protein-protein interactions. In this work, we will expose what concepts GCNs are misinterpreting, and propose solutions



Problems with negative statements

Inheritance

Inheritance:
(Ferritin, hasFunction, Iron Ion Binding)
(Iron Ion Binding, isSubClassOf, Metal Ion Binding)

(Ferritin, hasFunction, Metal Ion Binding)

Inheritance FAILS:
(Ferritin, NOT hasFunction, Iron Ion Binding)
(Iron Ion Binding, isSubClassOf, Metal Ion Binding)

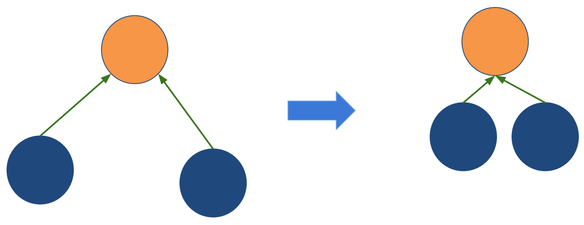
(Ferritin, NOT hasFunction, Metal Ion Binding)

Reverse Inheritance:
(Ferritin, NOT hasFunction, Iron Ion Binding)
(Ferrous Ion Binding, isSubClassOf, Iron Ion Binding)

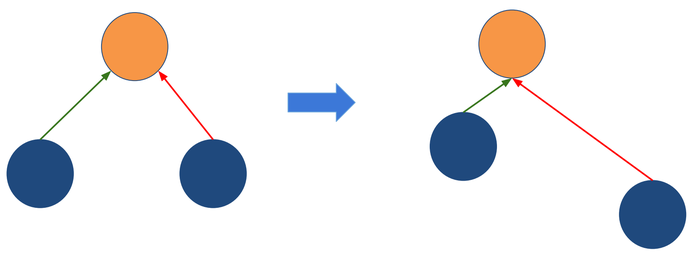
(Ferritin, NOT hasFunction, Metal Ion Binding)

Distance

Guilty By Association



Connected nodes aren't always similar



Objectives

Heterogeneous GCNs were built to handle different edge types. However, we do not expect them to handle negative edge types. We want to produce a baseline that can expose this flaw.

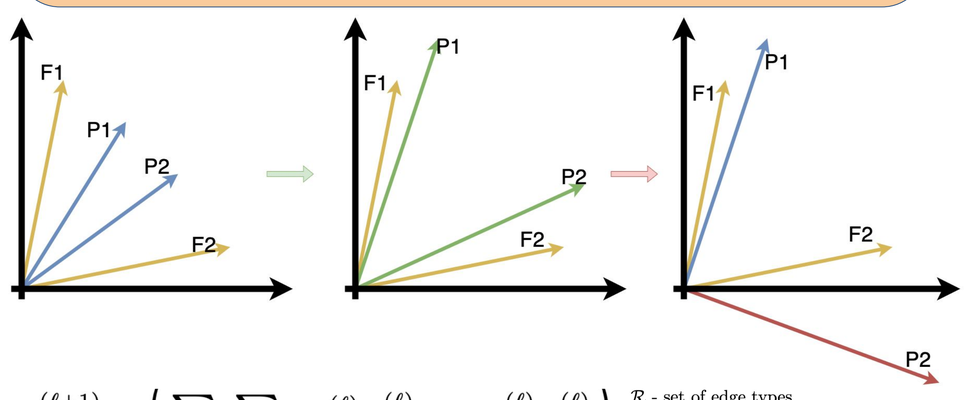
- Do H-GCNs benefit from the presence of negative statements
- How do H-GCNs compare against other PPI models that do handle negative edges

Can we improve H-GCNs performance with our proposal?

Distance Proposal

With Relational GCNs we can attribute weights to each edge type.

Can we bind positive weights to positive edges, and negative weights to negative edges?



$$h_i^{(\ell+1)} = \left(\sum_{r \in \mathcal{R}} \sum_{j \in \mathcal{N}_i^r} W_r^{(\ell)} h_j^{(\ell)} + W_0^{(\ell)} h_i^{(\ell)} \right)$$

$\mathcal{R}$ - set of edge types
$\mathcal{N}_i^r$ - neighbours of node i via relation r
W_r - weight assigned to relation r
W_0 - weight assigned to self relation

Inheritance Proposal

Each node will have a dual representation:
- A positive, for positive edges and inheritance
- A negative, for negative edges and reverse inheritance

BASELINE	Model	avg pr ± stdev	avg re ± stdev	avg f1 ± stdev
TrueWalks	TrueWalks	**0,952 ± 0,001**	**0,952 ± 0,001**	**0,966 ± 0,0009**
H-GCN	PPI+POS	0,633 ± 0,003	0,990 ± 0,0014	0,587 ± 0,0074
	PPI_POS	0,679 ± 0,0004	0,967 ± 0,0007	0,679 ± 0,0008
	PPI+POS+NEG	0,611 ± 0,001	0,996 ± 0,0003	0,526 ± 0,0040
	PPI_POS+NEG	0,678 ± 0,0009	0,967 ± 0,0004	0,677 ± 0,0015
	PPI+POS_NEG	0,677 ± 0,0002	0,9677 ± 0,0006	0,676 ± 0,0005
	PPI_POS_NEG	0,6788 ± 0,0002	0,967 ± 0,0006	0,677 ± 0,0005
	PPI+POS+GO	0,631 ± 0,0045	0,990 ± 0,0014	0,584 ± 0,0112
	PPI_POS_GO	0,683 ± 0,0002	0,963 ± 0,0006	0,685 ± 0,0005
	PPI+POS+NEG+GO	0,612 ± 0,0013	0,995 ± 0,0003	0,531 ± 0,004
	PPI_POS+NEG _GO	0,678 ± 0,002	0,967 ± 0,0004	0,679 ± 0,003
	PPI+POS+GO_NEG	0,634 ± 0,0024	0,99 ± 0,0007	0,586 ± 0,005
	PPI_POS_NEG_GO	0,633 ± 0,0004	0,99± 0,0004	0,586 ± 0,001

Acknowledgements: This work was supported by FCT through project <title>, ref.\ <project reference> (*<project DOI>*), and the LASIGE Research Unit, ref. UIDB/00408/2020 (*https://doi.org/10.54499/UIDB/00408/2020*) and ref. UIDP/00408/2020 (*https://doi.org/10.54499/UIDP/00408/2020*).

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

AutoUsability: Identifying the Relation Between Physiological Signals and the Usability of Interactive Systems

João Gonçalves, Manuel J. Fonseca

LASIGE
inclusive human-computer interaction

MOTIVATION

- Usability is usually assessed through user tests and standard questionnaires
- Filling in the questionnaires takes time and can be boring
- Questionnaires results may be inaccurate

OBJECTIVES

- Measuring usability without using questionnaires
- Investigate the relationship between the usability of the interface as perceived by users and their physiological signals as they interact with the application
- Building machine learning models to predict usability from physiological signals

DATA

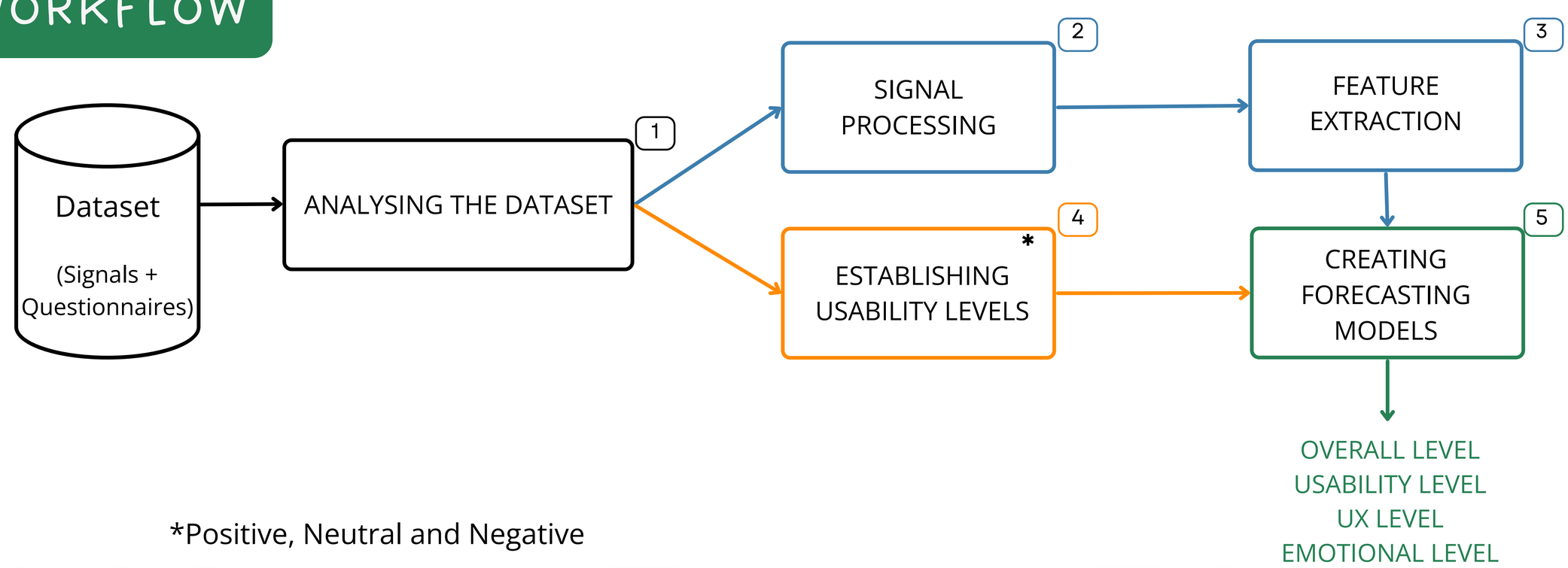
Physiological Signals

- Gyroscope (GYR)
- Accelerometer (ACC)
- Electroencephalography (EEG)
- Photoplethysmogram (PPG)

Questionnaires

- System Usability Scale (SUS)
- User Experience Questionnaire (UEQ)
- Single Ease Question (SEQ)
- NASA Task Load Index (TLX)
- Usability Metric for User Experience LITE (UMUX-LITE)
- Self-Assessment Manikin (SAM)

WORKFLOW



*Positive, Neutral and Negative

PRELIMINARY RESULTS

Offline

- Using dataset

Prediction	f1	acc	Frequency
Overall	61,1	62,1	Every 30 seconds In the end
Usability	66,5	68,3	In the end
UX	48,0	50,0	In the end
Emotional	71,7	71,6	Every 10 seconds In the end

FINAL EVALUATION

Online

- User study using real physiological signals
- Verification that the predictions align with the users' perception
- Expand the original dataset with new data

Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UIDB/00408/2020 (*https://doi.org/10.54499/UIDB/00408/2020*) and ref. UIDP/00408/2020 (*https://doi.org/10.54499/UIDP/00408/2020*).

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

DevSecOps with CI/CD: Integrating security in every step

Alexandre Figueiredo[1], Ricardo Ramalho[2], Ana Respício[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] MEO

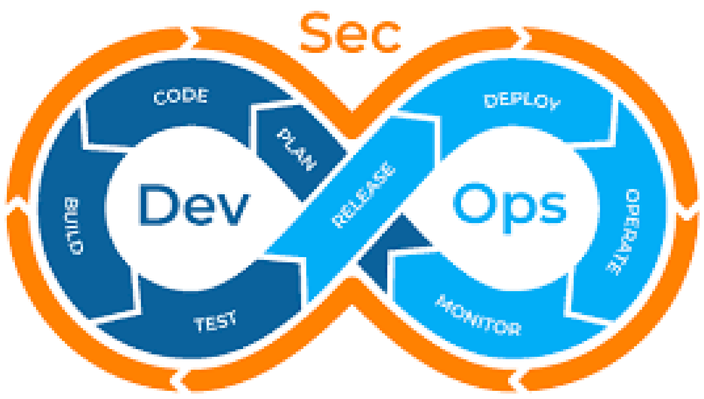
LASIGE reliable software systems

How to enhance security in application development and maintenance stages?

Security is often addressed in the final stages of application development. With DevSecOps security is taken into account in every stage of the development lifecycle.

Shitf-Left promotes early security testing, allowing teams to find and fix vulnerabilities earlier. This can be achieved with CI/CD security-focused pipelines.

Applying these security practices leads to fast vulnerability resolution, reduced risks and lower costs.



Methodology

Select Scanners

CI/CD Integration
Coverage
Maintenance
Report Quality
Configurations

CI/CD Components

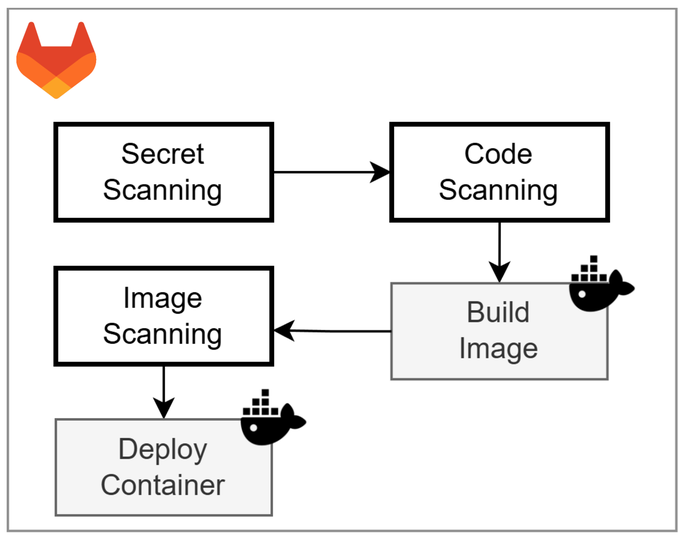
GitLab CI/CD Catalog:

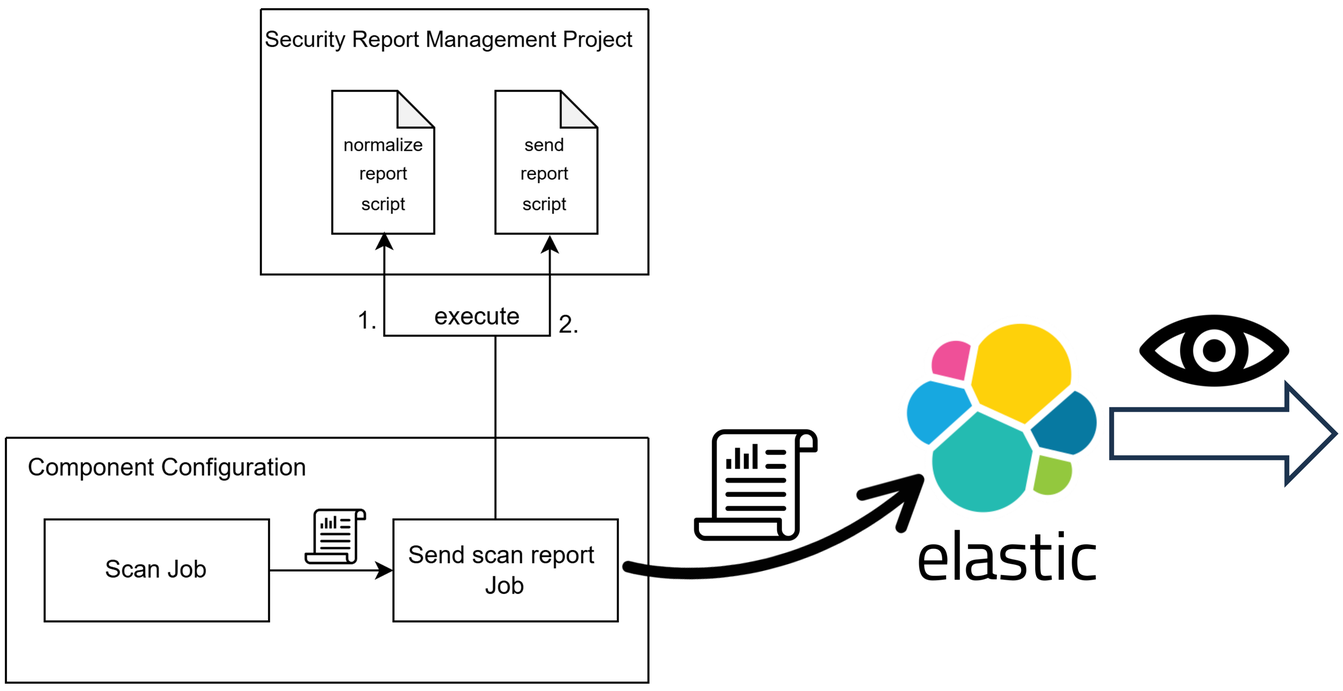
- Secret Scan
- Code Scan
- Image Scan

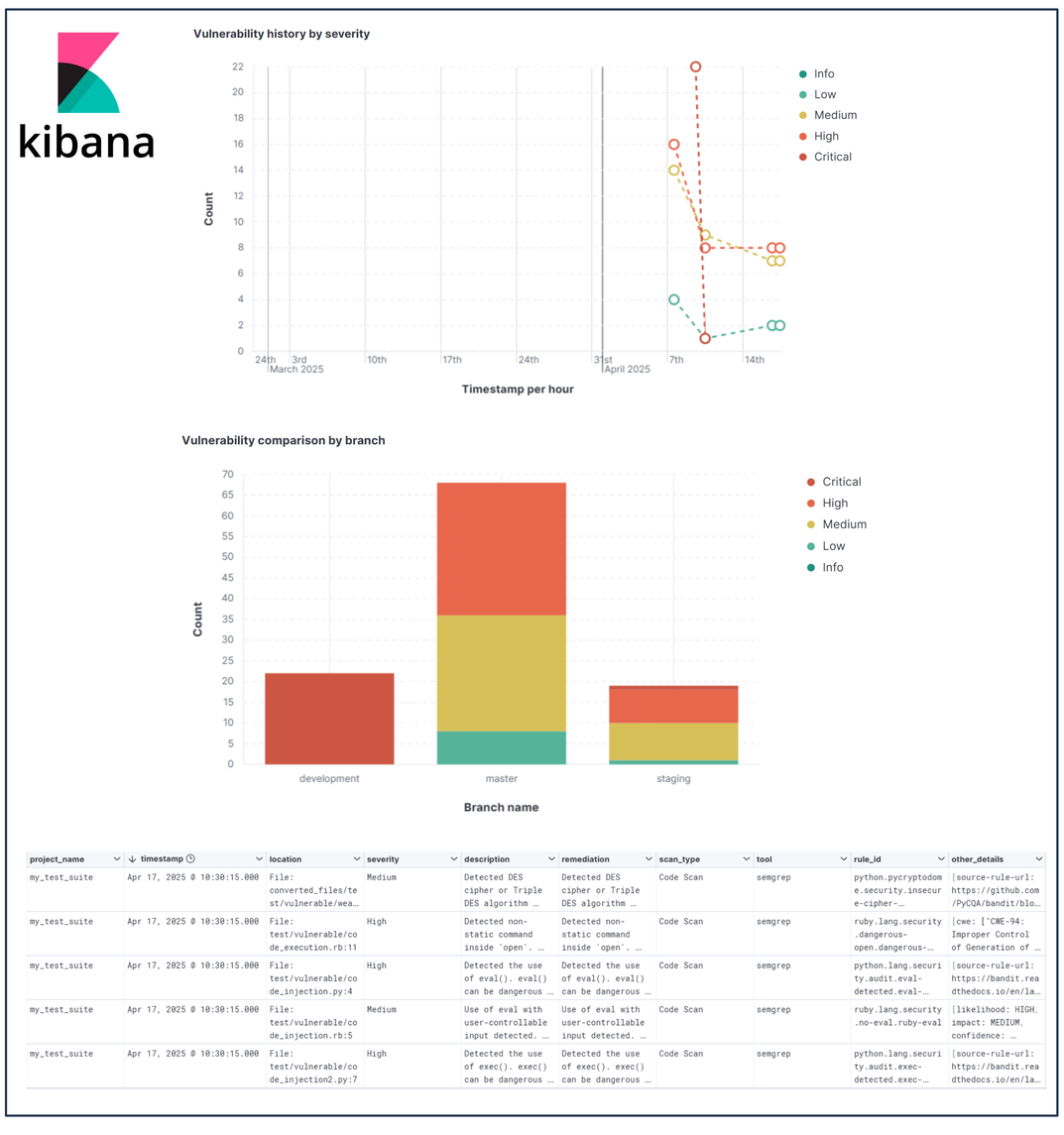
Continuous Monitoring

Leverage CI/CD mechanisms to guarantee continuous vulnerability scanning (scheduled pipelines).

Security-focused CI/CD Pipeline







project_name	timestamp	location	severity	description	remediation	scan_type	tool	rule_id	other_details
my_test_suite	Apr 17, 2025 @ 10:30:15.000	File: converted_files/test/vulnerable/wea…	Medium	Detected DES cipher or Triple DES algorithm …	Detected DES cipher or Triple DES algorithm …	Code Scan	semgrep	python.pycryptodome.security.insecure-cipher-…	[source-rule-url: https://github.com/PyCQA/bandit/blo…
my_test_suite	Apr 17, 2025 @ 10:30:15.000	File: test/vulnerable/code_execution.rb:11	High	Detected non-static command inside 'open'. …	Detected non-static command inside 'open'. …	Code Scan	semgrep	ruby.lang.security.dangerous-open.dangerous-…	[cwe: ["CWE-94: Improper Control of Generation of …
my_test_suite	Apr 17, 2025 @ 10:30:15.000	File: test/vulnerable/code_injection.py:4	High	Detected the use of eval(). eval() can be dangerous …	Detected the use of eval(). eval() can be dangerous …	Code Scan	semgrep	python.lang.security.audit.eval-detected.eval-…	[source-rule-url: https://bandit.readthedocs.io/en/la…
my_test_suite	Apr 17, 2025 @ 10:30:15.000	File: test/vulnerable/code_injection.py:5	Medium	Use of eval with user-controllable input detected. …	Use of eval with user-controllable input detected. …	Code Scan	semgrep	ruby.lang.security.no-eval.ruby-eval	[likelihood: HIGH, impact: MEDIUM, confidence: …
my_test_suite	Apr 17, 2025 @ 10:30:15.000	File: test/vulnerable/code_injection2.py:7	High	Detected the use of exec(). exec() can be dangerous …	Detected the use of exec(). exec() can be dangerous …	Code Scan	semgrep	python.lang.security.audit.exec-detected.exec-…	[source-rule-url: https://bandit.readthedocs.io/en/la…

GitLab CI/CD Components:

- Centralized catalog: + visibility and usability
- Easy security integration in CI/CD pipelines

Continuous Vulnerability Monitoring:

- Security even after development
- Newly published vulnerabilities can be detected

Acknowledgements: This work was supported by FCT through project Intelligi, ref. 2024.07506.IACDC, and the LASIGE Research Unit, ref. UID/00408/2025, and by Meo through project SecContainers, Lasige#1388.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · MEO

Automated Execution of Attack Chains to Train Data-driven IDS in Cyber-physical Systems

Diogo Pedro, José Cecílio, Pedro Ferreira, Alan Oliveira

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

Motivation

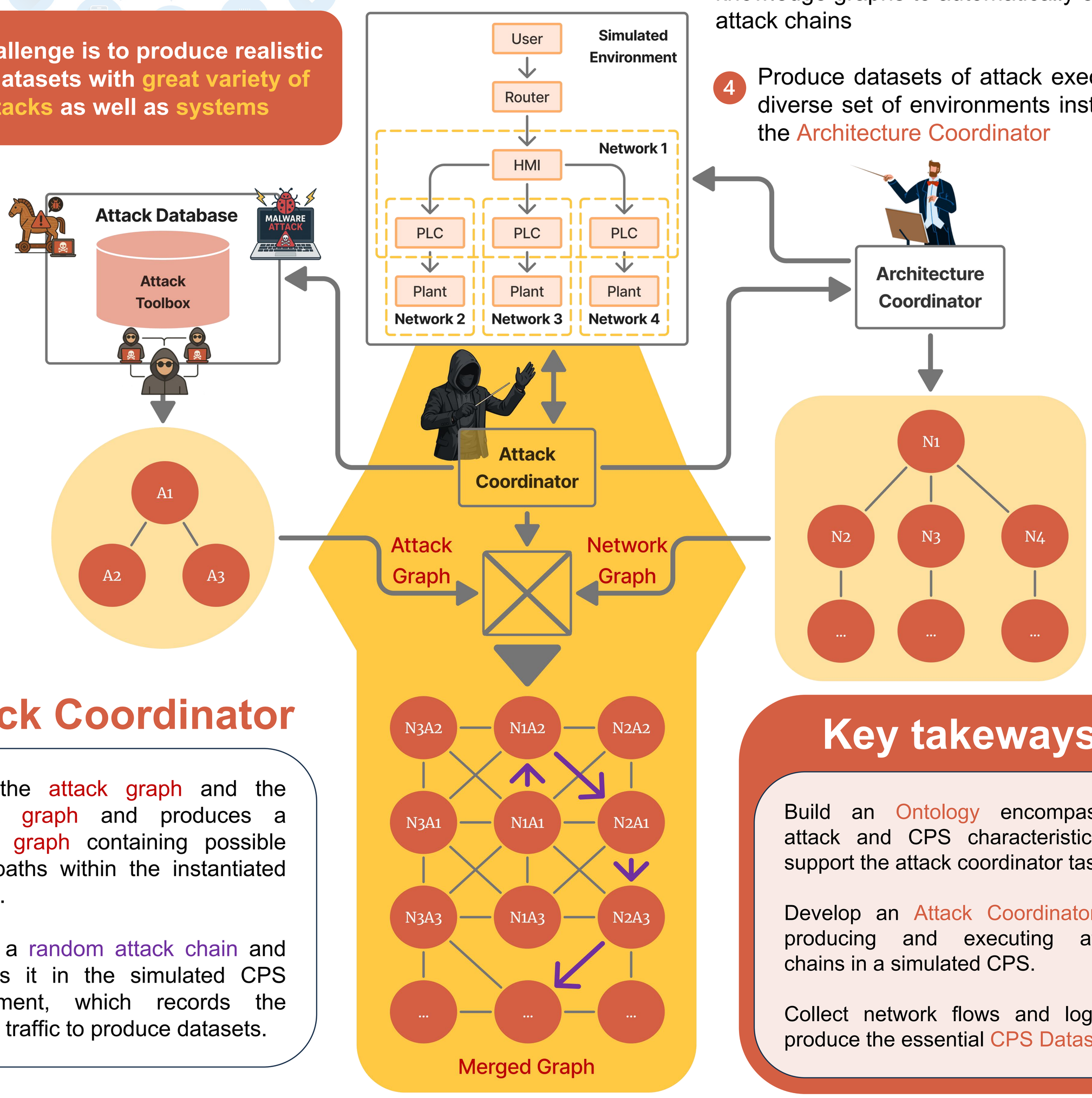
The number of attacks on cyber-physical systems (CPS) is growing, and modern security tools are struggling to keep up.

Many Intrusion Detection Systems (IDS) rely on Machine Learning techniques; However, the current datasets for CPS are limited in representativeness and extensiveness.

The challenge is to produce realistic CPS datasets with great variety of attacks as well as systems

Objectives

1. Select and test a variety of singular attacks and compile them into an Attack Database
2. Produce an Ontology to establish a common way to represent both CPS and attacks
3. Design an Attack Coordinator that interprets knowledge graphs to automatically execute attack chains
4. Produce datasets of attack executions in a diverse set of environments instantiated by the Architecture Coordinator



Attack Coordinator

Takes the attack graph and the network graph and produces a merged graph containing possible attack paths within the instantiated network.

Selects a random attack chain and executes it in the simulated CPS environment, which records the network traffic to produce datasets.

Key takeways

Build an Ontology encompassing attack and CPS characteristics to support the attack coordinator tasks.

Develop an Attack Coordinator for producing and executing attack chains in a simulated CPS.

Collect network flows and logs to produce the essential CPS Datasets.

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

Acknowledgements: This work was supported by the 'Companhia Paranaense de Energia - COPEL' research and technological development program, through the PD-02866-0533/2021 project, regulated by ANEEL, and by FCT through the LASIGE Research Unit, ref. UID/000408/2025.

Deep Generative Modeling for Longitudinal Clinical Data Generation

Ruben Branco[1], Marta Gromicho[2], Mamede de Carvalho[2], Piero Fariselli[3], Sara C. Madeira[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Faculdade de Medicina, Universidade de Lisboa, Portugal
[3] Department of Medical Sciences, University of Torino, Italy

LASIGE data and systems intelligence

Deep Generative Modeling

- **Precision Medicine** holds the potential to revolutionize healthcare through patient-tailored diagnosis, prognosis, and disease management.
- **Machine Learning** methods are driving innovation in many fields, but their application in healthcare is often constrained by scarcity of data, common in many disease contexts.
- **Proposed solution:** learning to generate **realistic** synthetic data using **Deep Generative Models**, thereby increasing data availability without compromising **patient privacy**.

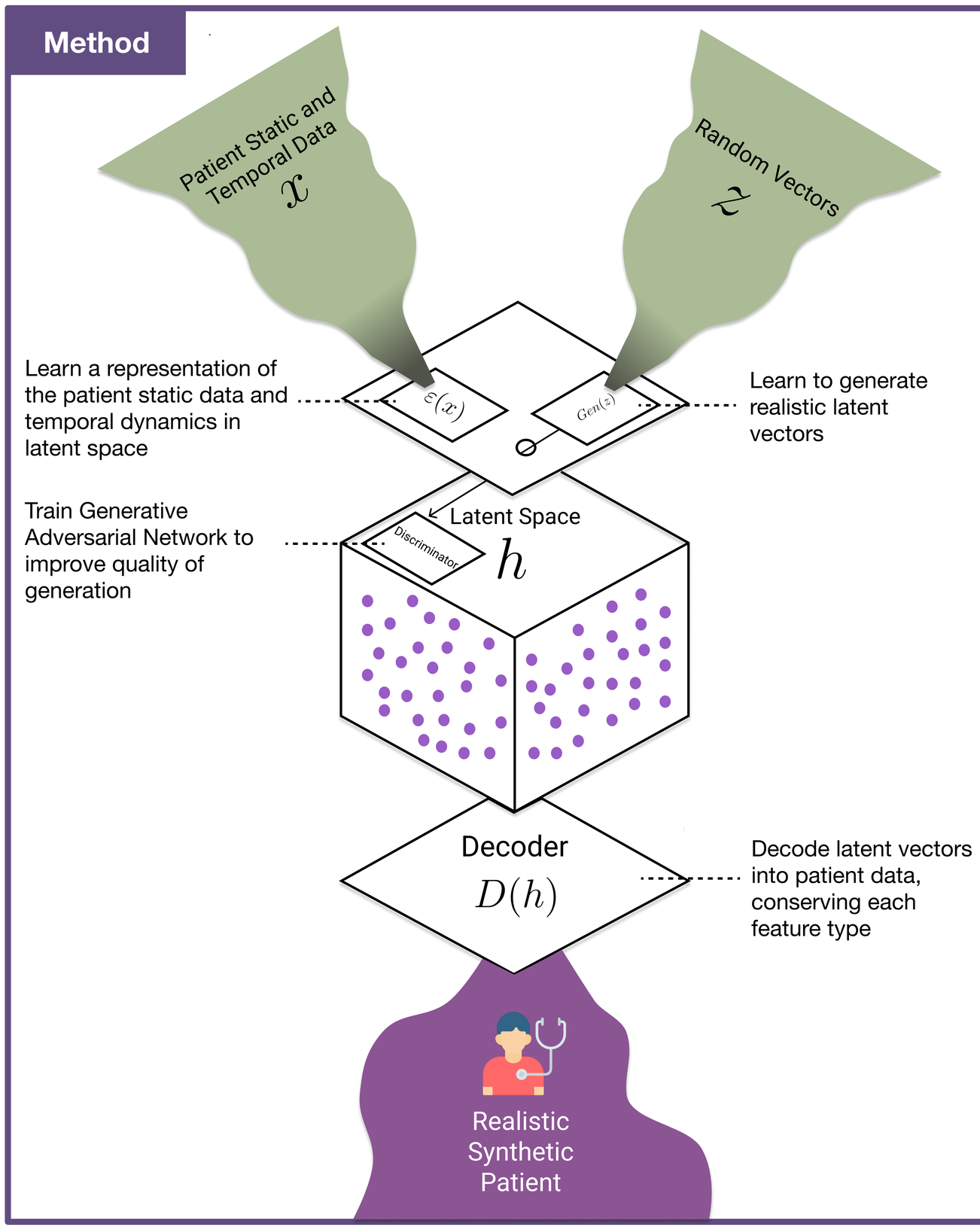
Case Study: Amyotrophic Lateral Sclerosis (ALS)

- **ALS** is a neurogenerative disease that affects motor neurons, causing a progressive muscular paralysis and leading to the patient's death due to respiratory failure.
- No known cure; treatments try to improve the quality of life and the life expectancy of patients.

Challenges

- Longitudinal data (static and temporal)
- Mixed-type Features
- Variable-sized Clinical History
- Low Resource (small cohort)

Method



Results

- Descriptive Analysis shows good distribution approximation
- High similarity on distribution similarity measures
- Low disclosure risk (preserves privacy)
- Able to deceive expert human clinician
- Temporal Dynamics can be unrealistic and chaotic
- Usability for tasks like prognosis shows potential but could benefit from more realistic data

Contributions

Innovative Method for Synthetic Data Generation: Developed a deep generative approach tailored specifically for mixed-type longitudinal datasets, capturing both static and temporal patient features

Latent Space Representation Learning: Our model not only generates realistic synthetic data but also learns meaningful latent representations of patient trajectories.

Rigorous Evaluation: Designed and implemented a comprehensive evaluation framework, integrating both established and novel metrics, to thoroughly assess data quality, privacy, and usability.

U LISBOA · Ciências ULisboa · LASIGE · Università di Torino · Faculdade de Medicina Lisboa · fct Fundação para a Ciência e a Tecnologia

Acknowledgements: This work was supported by FCT through project AIpALS ref. PTDC/CCI-CIF/4613/2020 (https://doi.org/10.54499/PTDC/CCI-CIF/4613/2020), the European Union Horizon 2020 through the BRAINTEASER project (GA no. 101017598), the LASIGE Research Unit, ref. UIDB/00408/2020 (https://doi.org/10.54499/UIDB/00408/2020) and ref. UIDP/00408/2020 (https://doi.org/10.54499/UIDP/00408/2020), and a PhD research scholarship (2022.10727.BD) to RB.

Secure, Resilient Streaming Communication and Processing in Challenging Environments

Luís Viana, Vinícius Cogo

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE dependable and secure decentralized systems

Motivation

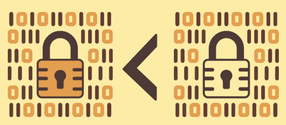
IoT data processing is increasingly offloaded to the cloud to support constrained devices and enable collaborative tasks.

Encrypted computing schemes address the privacy concerns regarding secure data processing and the trustworthiness of single-cloud solutions.

However, many of the existing approaches, such as fully homomorphic encryption and related techniques, remain computationally intensive, being unsuitable for latency-sensitive applications.

Background

Order revealing encryption(ORE) enables order comparison of ciphertexts using a publicly computable function without decrypting them.

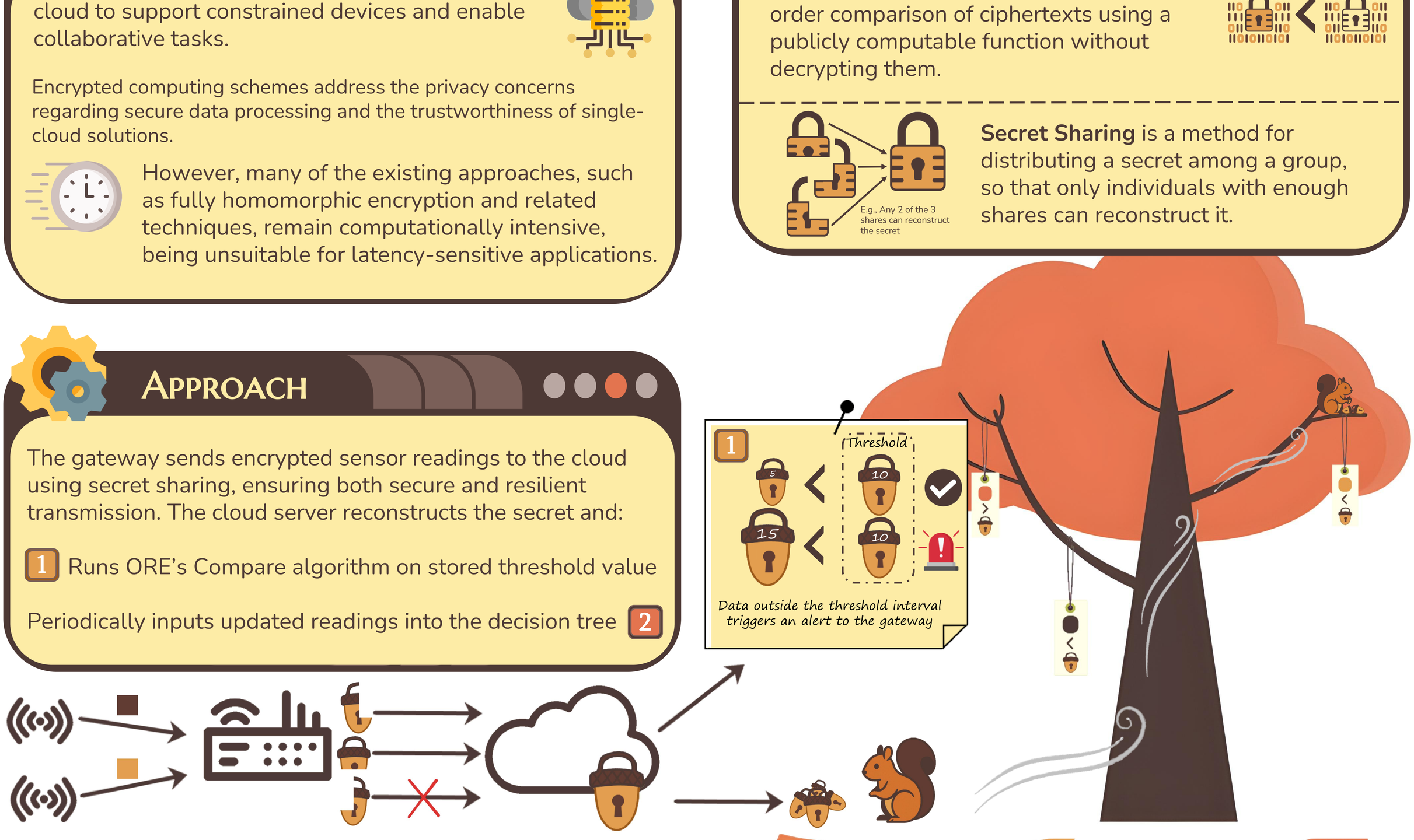


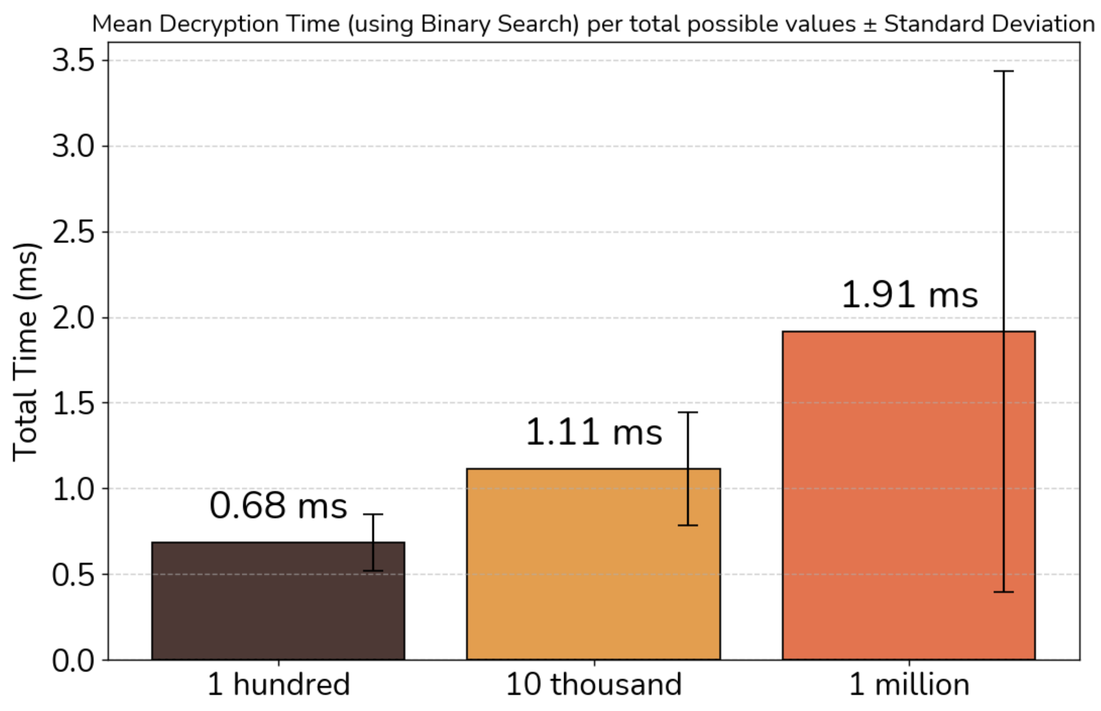
Secret Sharing is a method for distributing a secret among a group, so that only individuals with enough shares can reconstruct it.

Approach

The gateway sends encrypted sensor readings to the cloud using secret sharing, ensuring both secure and resilient transmission. The cloud server reconstructs the secret and:

1. Runs ORE's Compare algorithm on stored threshold value
2. Periodically inputs updated readings into the decision tree





* The performance test demonstrates that decryption time remains highly efficient even with large datasets of numerical values, making it a competitive alternative to maintaining a local cache at the gateway, regardless of the sensor's precision and range

2

- ✓ Encrypted decision tree learning
- ✓ Encrypted decision tree thresholds are updated to each party's local encryption key using Updatable ORE, enabling the same data to be independently compared against thresholds while remaining compatible with the collaboratively learned decision tree

Next Steps

- Integration of the encrypted decision tree
- Evaluation of the system's performance and scalability
- Development and integration of lightweight secret sharing followed by further performance testing

Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UID/00408/2025.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Comparative evaluation of protein classification performance using ProtBERT embedding variants and one hot encodings

Ana F. Rodrigues, J. Lucas Ferraz , Pedro Cotovio, Catia Pesquita

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

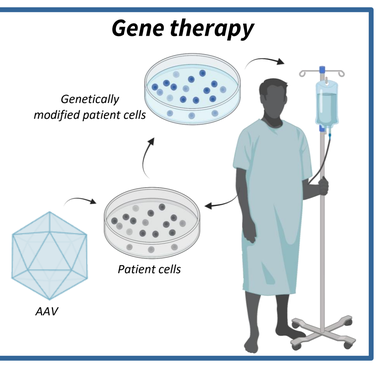
LASIGE health and biomedical informatics

Motivation

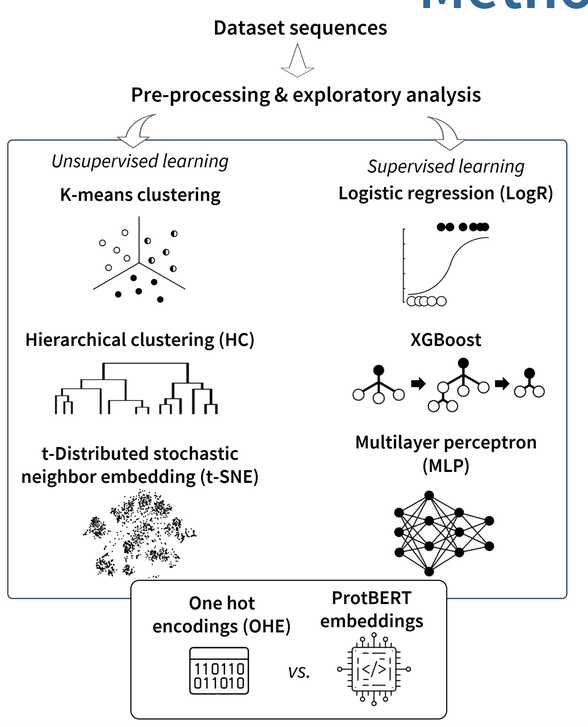
Effective representations of protein sequences are key to the success of machine learning (ML) in protein design. However, **comprehensive comparative studies remain scarce**, despite their crucial role in providing *a priori* insights into which representations are most likely to capture richer information, facilitate learning, and ultimately drive superior predictive performance. Here, we conduct a **systematic comparative analysis** of protein classification using **multiple embedding variants from ProtBERT**[1] and **benchmark them against one-hot encodings**, to assess the benefits of modern representation learning. Our results offer insights to **support informed decisions in representation selection** for protein classification pipelines in biotechnology and bioengineering applications.

Context & Dataset

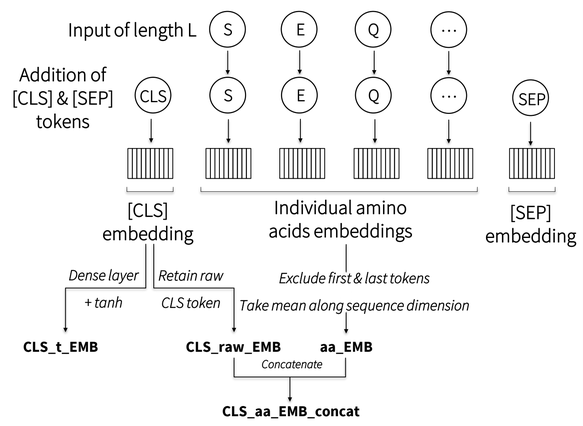
- Adeno-associated viral vectors (AAVs) are one of the most used vectors in human gene therapy with AAV capsid bioengineering being a major focus in current R&D.
- Machine-guided protein design, specifically, classification-based approaches helps identifying promising sequence variants in bioengineering approaches.
- The bioengineering target here is viability, i.e., the ability to form capsids that deliver the therapeutic gene. The dataset[2] includes ~270K AAV capsid sequence variants (52% are viable) which were designed using either ML- or non-ML-based strategies.



Methodology

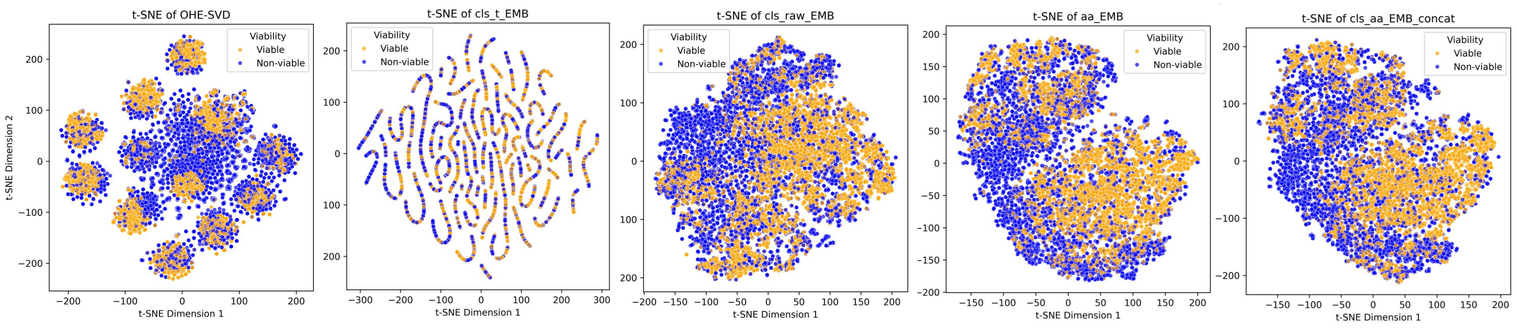


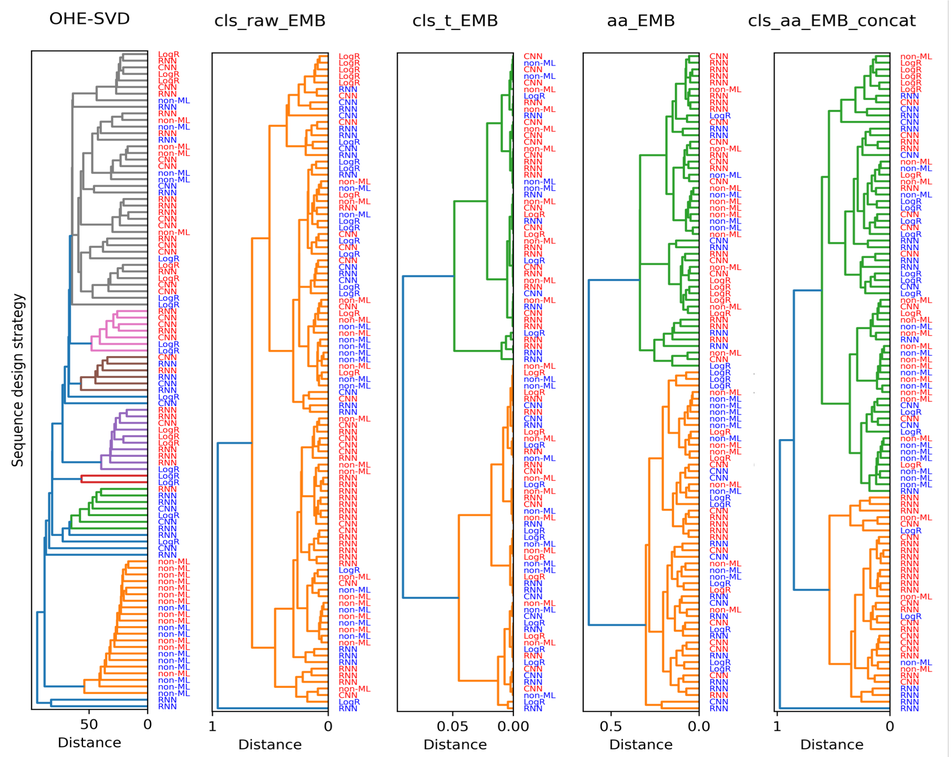
Variants of ProtBERT embeddings



Results

☞ Unsupervised learning





- ✓ No representation clusters sequences by viability (same for K-means, data not shown)
- ✓ All representations (except cls_t_EMB) show some ability to reflect the sequence design strategy
- ✓ OHE yields many fine-grained clusters with ProtBERT embeddings form fewer/tighter clusters
- ✓ cls_t_EMB flattens the data and reduces variance

☞ Supervised learning

Representation	XGBoost							
	Precision		Accuracy		F1 Score		ROC AUC	
	Train-val	Test	Train-val	Test	Train-val	Test	Train-val	Test
OHE-SVD	**0.9146**	0.8915	**0.9278**	0.8963	**0.9316**	0.9008	**0.9803**	0.9590
cls_raw_EMB	0.8923	0.8919	0.8944	0.8950	0.8986	0.8994	0.9627	0.9626
cls_t_EMB	0.5664	0.5659	0.5769	0.5760	0.6549	0.6544	0.5908	0.5893
aa_EMB	0.8952	0.8928	0.8976	0.8966	0.9017	0.9010	0.9648	0.9644
cls_aa_EMB_concat	0.8997	**0.8987**	0.9017	**0.9013**	0.9057	**0.9053**	0.9672	**0.9668**

	MLP							
	Precision		Accuracy		F1 Score		ROC AUC	
	Train-val	Test	Train-val	Test	Train-val	Test	Train-val	Test
OHE-SVD	**0.938**	0.827	**0.937**	0.771	**0.940**	0.761	**0.983**	0.872
cls_raw_EMB	0.878	0.897	0.885	0.892	0.891	0.895	0.962	0.961
cls_t_EMB	0.518	0.518	0.518	0.518	0.682	0.682	0.412	0.415
aa_EMB	0.910	**0.927**	0.908	0.908	0.913	0.909	0.972	0.972
cls_aa_EMB_concat	0.902	0.910	0.908	**0.913**	0.912	**0.916**	0.972	**0.972**

- ✓ cls_t_EMB performs poorly in prediction tasks
- ✓ ProtBERT embeddings (except cls_t_EMB) outperform OHE at test stage
- ✓ Performance improves: cls_raw_EMB → aa_EMB → cls_aa_EMB_concat
- ✓ Combining amino acid–level and sequence-level embeddings into a single representation yields the best overall results
- ✓ Same conclusions for LogR (data not shown)

Conclusions

- ➢ Learning functional protein properties requires supervised learning to uncover biologically relevant patterns
- ➢ ProtBERT embeddings provide better generalization and prediction performance than OHE at the test stage

References:

[1]. A. Elnaggar, M. Heinzinger, C. Dallago, G. Rihawi, Y. Wang, L. Jones, T. Gibbs, T. Feher, C. Angerer, M. Steinegger, D. Bhowmik, B. Rost, ProtTrans: Towards Cracking the Language of Life's Code Through Self-Supervised Deep Learning and High Performance Computing, (2020). https://doi.org/10.48550/ARXIV.2007.06225.

[2]. D.H. Bryant, A. Bashir, S. Sinai, N.K. Jain, P.J. Ogden, P.F. Riley, G.M. Church, L.J. Colwell, E.D. Kelsic, Deep diversification of an AAV capsid protein by machine learning, Nat Biotechnol 39 (2021) 691–696

Acknowledgements: This work was supported by the LASIGE Research Unit, ref. UID/00408/2025. It was also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453. This work was also supported partially by project 41, HfPT: Health from Portugal, funded by the Portuguese Plano de Recuperação e Resiliência

Qualweb powered Web Accessibility Dashboard

Tomás Barreto, Letícia Pereira, Carlos Duarte

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE inclusive human-computer interaction

- Nearly one-sixth of the world's population has disabilities.
- Web accessibility is essential for inclusivity, yet 95.6% of the top million websites have Accessibility standards failures.
- Regulations like EAA and Section 508 require accessibility, creating ethical and legal obligations.

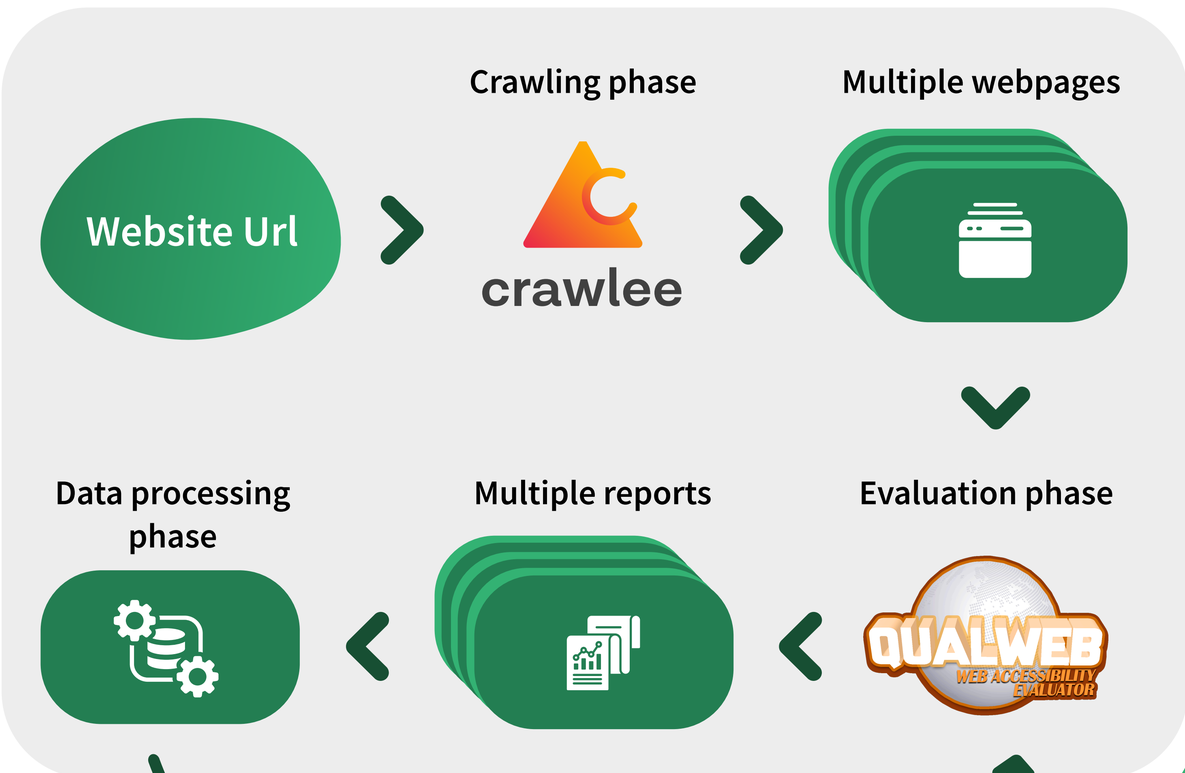
How do we score Web Accessibility?

$$A3 = 1 - \prod_{b} (1 - F_b)^{\frac{B_{pb}}{N_{pb}} + \frac{B_{pb}}{B_p}}$$

- B_{pb}: Number of failed instances of a barrier in a webpage
- B_p: Number of failed instances in a webpage
- N_{pb}: Number of testable instances of a barrier in a webpage
- F_b: Barrier score
- b: Accessibility barrier

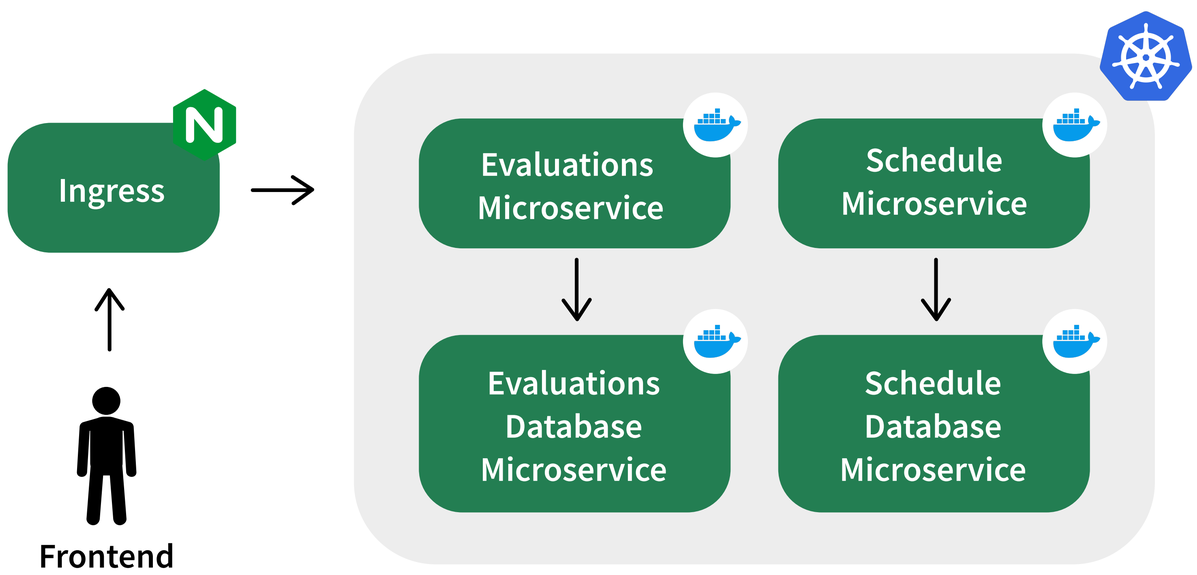
How do we keep track of it?

By continuously assessing websites over time.



The Dashboard

System Architecture



*each database microservice has its own postgres database

What's next?

- The next immediate focus will be implementing the comparison of evaluations from different dates to enable users to track accessibility improvements over time.
- We will then introduce filtering issues by guideline, login bypass for password protected content evaluation, and scan scheduling capabilities.

Acknowledgements: This work was supported by Fundação para a Ciência e Tecnologia (FCT) in the LASIGE Research Unit under the ref. (UID/00408/2025)

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

rospec: A Domain-Specific Language for ROS-based Robot Software

Paulo Canelas[†,‡], Bradley Schmerl[‡], Alcides Fonseca[†], and Christopher S. Timperley[‡]

[†] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[‡] Software and Societal Systems Department, Carnegie Mellon University, US

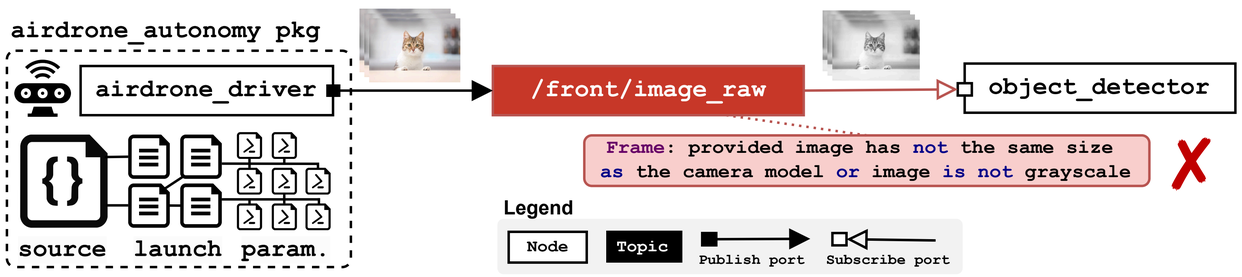
LASIGE reliable software systems

Motivation

The **Robot Operating System (ROS)** is the de facto open-source framework for building complex robot software offering reusable and configurable off-the-shelf components. ROS components often **lack proper documentation**, making system configuration challenging and forcing developers to rely on **unverified assumptions**, leading to errors.

In this work, we propose a ROS-tailored **domain-specific language** to specify component **configuration** and **integration** using domain concepts.

Misconfigurations arise from mismatched expectations between components, potentially causing dangerous robot behaviors in physical environments.



By specifying properties over component configurations and their integration, we can detect misconfigurations prior to execution

AMCL Component Specification

```
type alias AfterHumbleVersion: Enum[…, Humble, Iron, Jazzy, Kilted] where {_ >= Humble};
type alias LaserModelType: Enum[Beam, LikelihoodField, LikelihoodFieldProb];

type alias Meter: float32;

message alias RestrictedLoadMap: nav2_msgs/LoadMap {
    request  field map_url: string;
    response field map: nav2_msgs/OccupancyGrid;
    response field result: uint8 where {(_ >= 0 and _ <= 3) or _ == 255};
}

node type amcl_type {
```

Type aliases act as documentation providing semantic information about the purpose of each configuration

Restrictions over types → Z3 Solver = (SAT, SAT, UNSAT, UNSAT)

(0, 255, 4, -1)

Typing

Bounds

Presence

```
context distribution: AfterHumbleVersion;

param laser_model_type: LaserModelType;

optional param z_hit: double = 0.5;
optional param z_max: double = 0.05;
optional param z_rand: double = 0.5;
optional param z_short: double = 0.005;
optional param always_reset_initial_pose: bool = false;
```

QoS

Cardinality

Connection

```
@qos{sensor_qos}
publishes to particle_cloud: nav2_msgs/ParticleCloud;

@qos{default_qos}
subscribes to /initialpose: geometry_msgs/PoseWithCovarianceStamped
                                    where {count(publishers(_)) == 1};

provides service set_initial_pose: nav2_msgs/SetInitialPose;
```

```
} where {
    laser_model_type == Beam -> z_hit + z_max + z_rand + z_short == 1;
    laser_model_type == LikelihoodField -> z_hit + z_rand == 1;
    always_reset_initial_pose -> exists(initial_pose);
}
```

Dependency

Conditionals

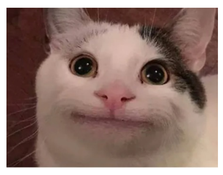
Contextual information allows developers to specify deployment-specific requirements and dependencies

Sensor Quality of Service (QoS)

```
policy instance sensor_qos: qos {
    setting history = KeepAll;
    setting depth = 1;
    setting reliability = BestEffort;
}
```

+

@color_format{RGB8}



@color_format{Grayscale}

Case Study

We evaluated rospec on **Neobotix MP-400** robot navigating a warehouse environment through specification and configuration of configurations across **18 components** including navigation, localization, and obstacle avoidance.

Evaluation

Analyzed **182 questions** from Stack Exchange, from which,

61	23	28 + 39	31
Detectable	Documentation	Missing Info / Out of Scope	Not Supported

Case Study System Specification

```
system {

✗   node instance amcl: amcl_type {

        context distribution = Jazzy;

        param laser_model_type = Beam;

        param z_hit = 0.5;
        param z_max = 0.0;
        param z_rand = 0.5;

        param always_reset_initial_pose = false;
    }
}
```

rospec verifier

D-System

$$\Gamma \vdash \overline{D} \dashv \Gamma' \qquad \frac{\overline{s \mapsto \textbf{publishes to}(x, T_1) \in \Gamma' \wedge \Gamma' \vdash T_1 <: T_2}^{\,s \mapsto \textbf{subscribes to}(x,\, T_2) \in \Gamma'} \quad \overline{\textbf{check_qos}(q1,\ q2)}^{\,s \mapsto \textbf{publishes to}(x) \text{ with } qos(q1) \in \Gamma',\ s' \mapsto \textbf{subscribes to}(x) \text{ with } qos(q2) \in \Gamma'} \quad \frac{\overline{s \mapsto \textbf{broadcasts}(x_1, x_2) \in \Gamma'}^{\,s \mapsto \textbf{listens}(x_1, x_2) \in \Gamma'} \quad \overline{x_1 = x_2}^{\,s \mapsto \textbf{broadcasts}(x_1, x_2) \in \Gamma',\ s' \mapsto \textbf{broadcasts}(x_2, x_3) \in \Gamma'}}{\Gamma \vdash \textbf{system}\ \{\ \overline{D}\ \} \dashv \Gamma'}$$

ERROR: Dependency laser_model_type == Beam -> z_hit + z_max + z_rand + z_short == 1 in amcl is not respected.

ERROR: Publisher not found for the subscriber amcl to the topic /initialpose.

Acknowledgements:

This work was supported by Fundação para a Ciência e Tecnologia (FCT) in the LASIGE Research Unit under the ref. (UID/00408/2025 and EXPL/CCI-COM/1306/2021), and the CMU Portugal Dual PhD program (SFRH/BD/151469/2021).

LLM-based Framework for Email Classification and Phishing Detection

Nuno Costa[1,2], José Cecilio[1]
Ruben Salgueiro[2], Mauricio Rosa[2], Dulce Domingos[1,2]

[1]LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2]Serviços Centrais, Universidade de Lisboa, Portugal

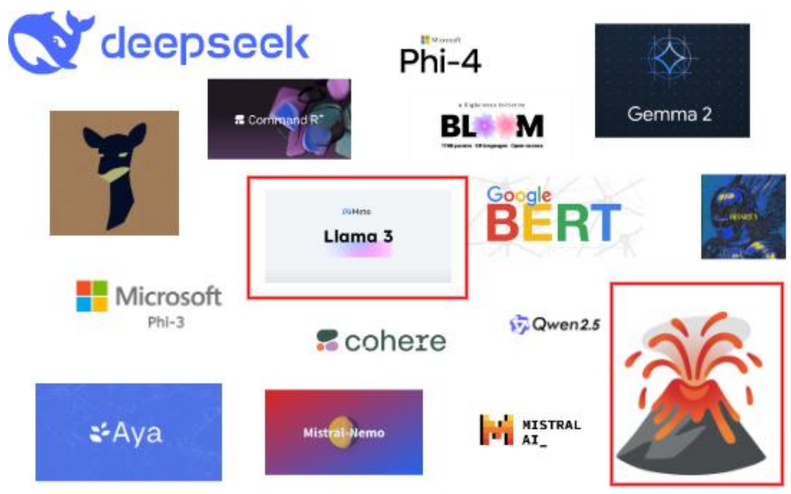
LASIGE cyber-physical systems

Context and Motivation

Phishing and spam remain major security threats as they evolve beyond traditional email filters. This project introduces a framework using Large Language Models (LLMs) to improve email classification and phishing detection. When a user flags an email through an inbox plugin, it is stored, cleaned, and analyzed using tailored prompts to classify it as "normal," "unwanted," or "malicious."

Model Selection

Selection criteria: Open-source, relatively lightweight, compatible with Ollama. After applying this criteria and testing the LLMs with a set of emails the final LLMs were Llama3.1 and Llava1.8.



Model Training

QLoRa was selected due to its significantly higher efficiency and lower computational costs.

Feature	Full Tuning	LoRA	QLoRA
Memory Use	High	Low	Very Low
Speed	Slow	Fast	Faster
VRAM Need	Very High	Medium	Low
Training Cost	Expensive	Affordable	Very cheap
Performance	Best	Very good	Very good

Selection of platform and datsets:

Platform: unsloth

Datasets:

- Kaggle Open Source kaggle
- Custom made with Ulisboa Anubis anubisnetworks
- Data Augmentation with books on email security Google Books

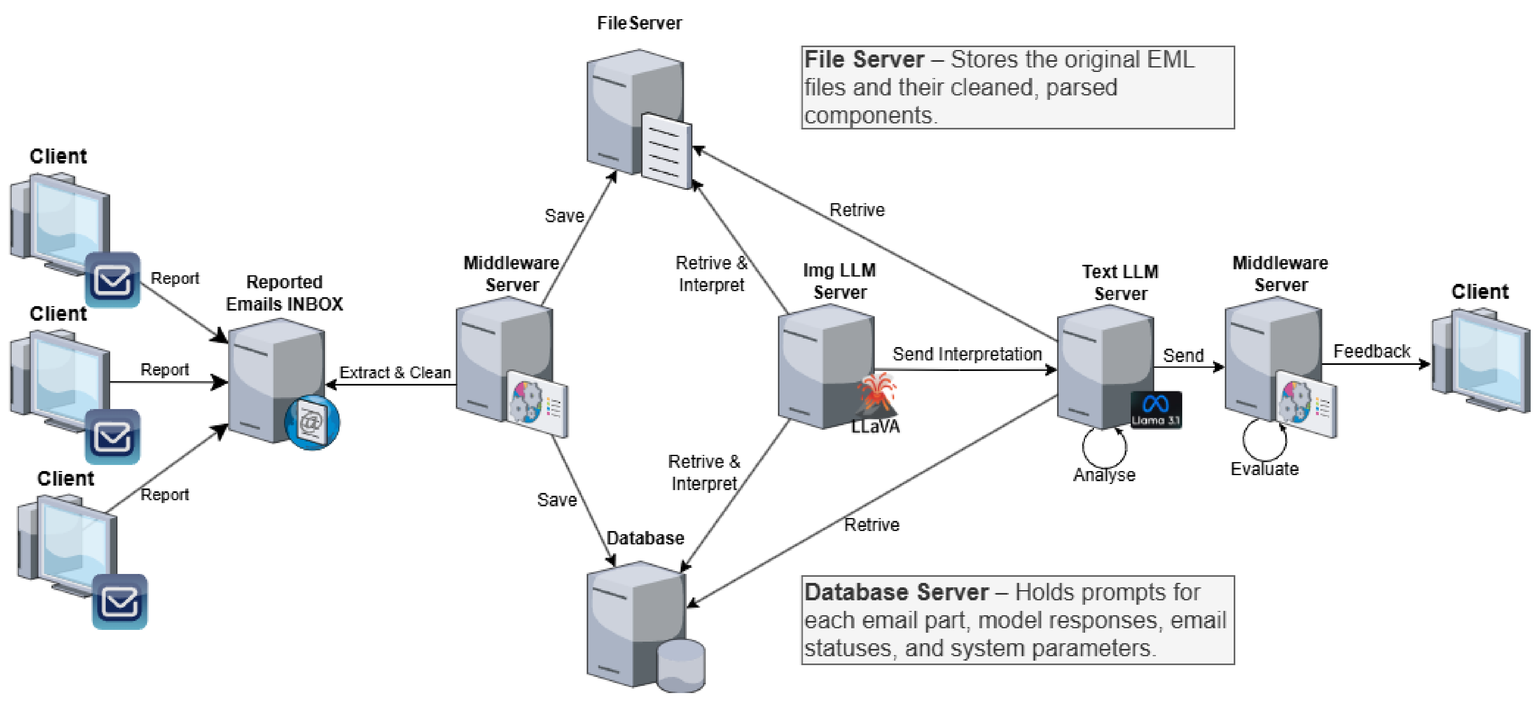
The Framework

There are three main servers that operate independently, using the database and file servers as shared support components.

Middleware Server – Extracts, cleans, parses, and stores emails; evaluates model responses, determines the final classification, and sends feedback to the user.

Image LLM Server – Uses LLaVa to analyze email images and generate descriptive text for each.

LLM Server – Uses a fine-tuned LLaMA 3.1 model to process all email components (including image interpretations) with tailored prompts for classification.



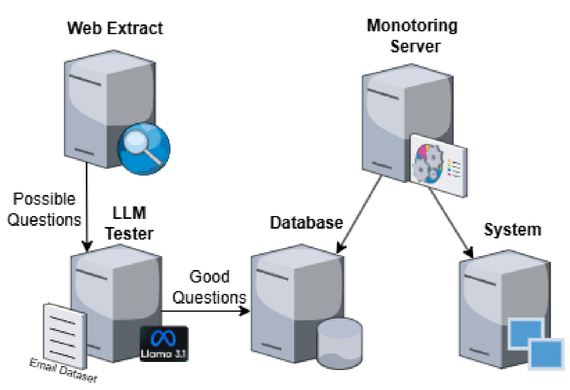
In Progress

Several features are still under development or refinement, including:

- Creation of the custom ULisboa Anubis dataset
- Fine-tuning LLaMA 3.1 using all available datasets
- Implementation of system logging
- Enhancement of the current prompt set

Future Work

- Improving LLM training for better performance.
- Automating question generation via web scraping and testing.
- Implementing a monitoring system to handle ULisboa's volume of over 150,000 daily emails.



U LISBOA | Ciências ULisboa | FCT Fundação para a Ciência e a Tecnologia | LASIGE driven by excellence

Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UID/00408/2025.

AI-Powered Multi-Target Antiviral Discovery

Maryam Lotfi , Francisco M Couto, Marcia Barros

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE data and systems intelligence

WHY: The Challenge

Urgent Need for Broad Spectrum Antivirals such As SARS-COVID

HOW: Our Approach

Combine ML + Proteins/Ligands Features + Docking/MD/MMPBSA Validation

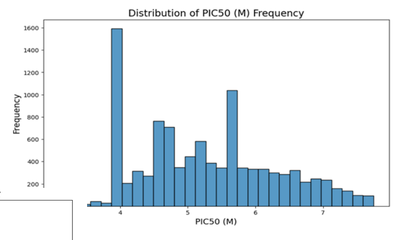
WHAT: Our Innovach

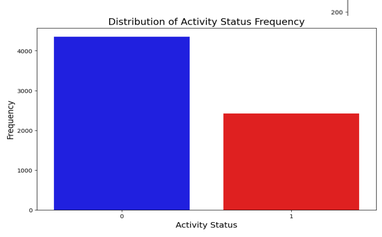
Multi-Target Screening in Action& Real World Virtual Screen

Integrated Workflow& Key Outcomes

Dataset Preparation

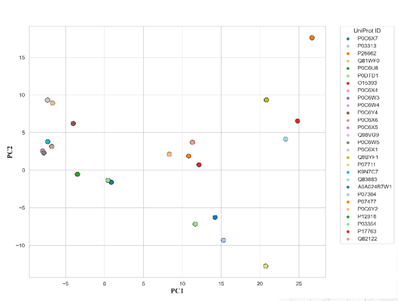
- Data Extraction from BindingDB
- Excluding Non-Human Targets
- Normalization





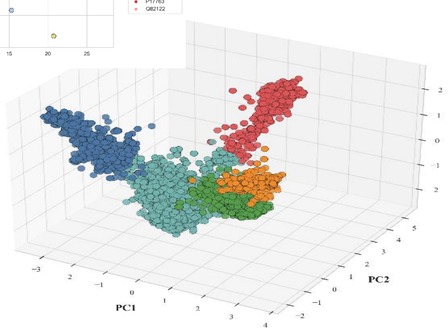
Feature Extraction

- Ligand SMILES Encoding
- Protein Binding site Sequences Encoding



Assessing Dataset Diversity

- PCA for Proteins
- PCA for Ligands



Model Preparation & Optimization

- Feature Selection
- Model Training
- Hyper parameter Optimization

Model	Accuracy	Precision	Recall	Macro F1-Score	AUC-ROC
Random Forest	0.8671	0.8602	0.8547	0.8571	0.9338
SVC	0.9050	0.9030	0.9012	0.9017	0.9482
XGBoost	0.8956	0.8921	0.8841	0.8875	0.9496
LightGBM	0.8857	0.8826	0.8719	0.8765	0.9419

Performance comparison of machine learning models on training, validation, and test sets.

MM/PBSA Calculations

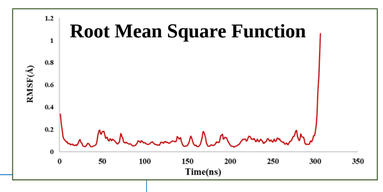
- Snapshot Extraction
- Energy Calculation

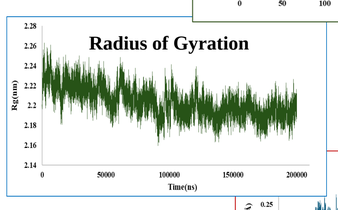
Energy component	Van der Waals	Electrostatic	Polar Solvation	SASA(non-polar solvation)	Total-Binding Energy
Value	−190.48 ± 3.12	−47.31 ± 5.73	+164.06 ± 10.89	−19.75 ± 1.75	−93.48 ± 11.74

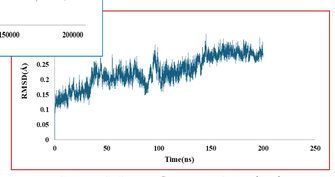
MM/PBSA binding free energy components for the ligand–target complex.

Molecular Dynamics Simulation

- Force Field Selection
- System Solvation
- Trajectory Analysis





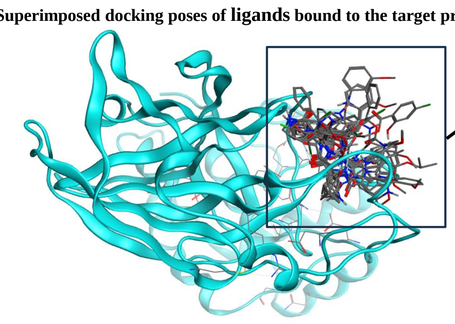


Root Mean Square Deviation

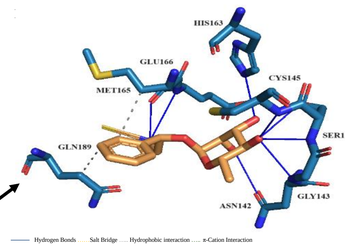
Molecular Docking Simulation

- Protein Preparation
- Ligand Preparation
- Docking Runs

Superimposed docking poses of ligands bound to the target protein



Top active ligand interactions with



Conclusion

- ML pipeline for multi-target antiviral prediction
- Combined Morgan fingerprints + ProtBERT embeddings
- XGBoost model: F1 = 0.89, AUC-ROC = 0.95
- Generalized across 28 SARS-CoV-2 targets
- Validated via : MD& MM/PBSA
- Enables rapid, AI-driven antiviral screening

Future Work

- Add 3D interaction maps + docking features to enrich protein representation
- Embed docking scores into training for structure-aware predictions
- Build probabilistic recommender for ranking multi-target compounds

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

Acknowledgements: This work was supported by FCT(Fundação para a Ciência e a Tecnologia) through project 2023.03164.BD, and the LASIGE Research Unit, ref. UID/00408/2025.

Secure Execution of Automated Penetration Tests

Tomás Páscoa, Alan Oliveira, José Cecílio

LASIGE, Faculdade de ciências, Universidade de Lisboa, Portugal

LASIGE dependable and secure decentralized systems

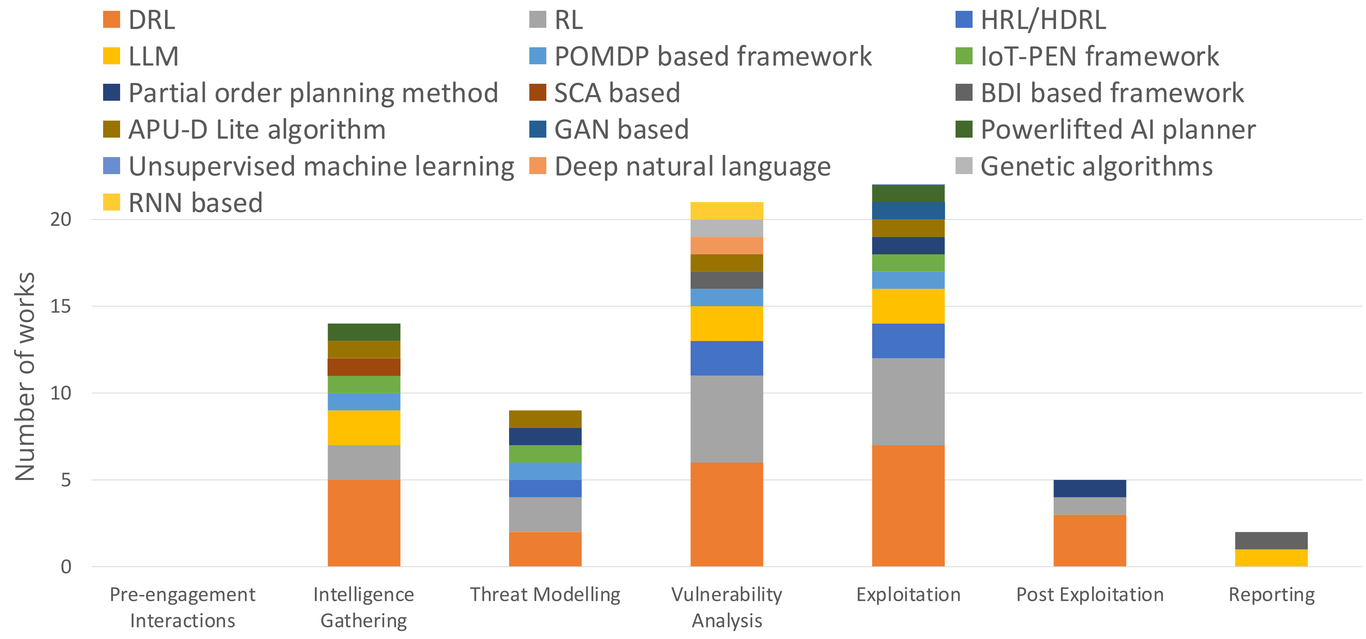
INTRODUCTION

One way to improve cybersecurity is through **penetration tests** (i.e., controlled attacks to identify system vulnerabilities).

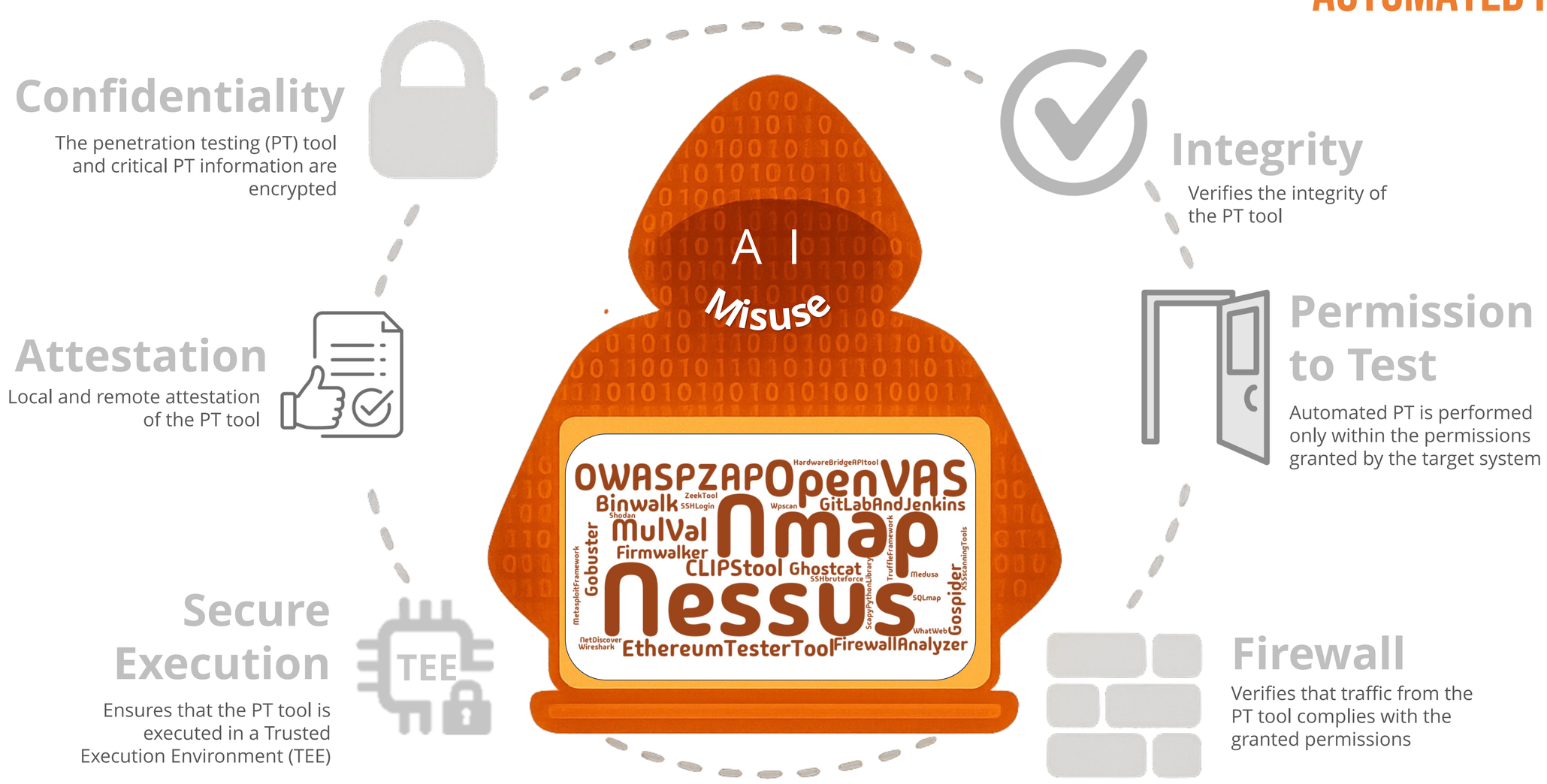
Penetration tests require highly skilled and often costly professionals, which motivates penetration testing (PT) automation.

Automated PT tools are handy but can be **misused for attack purposes**.

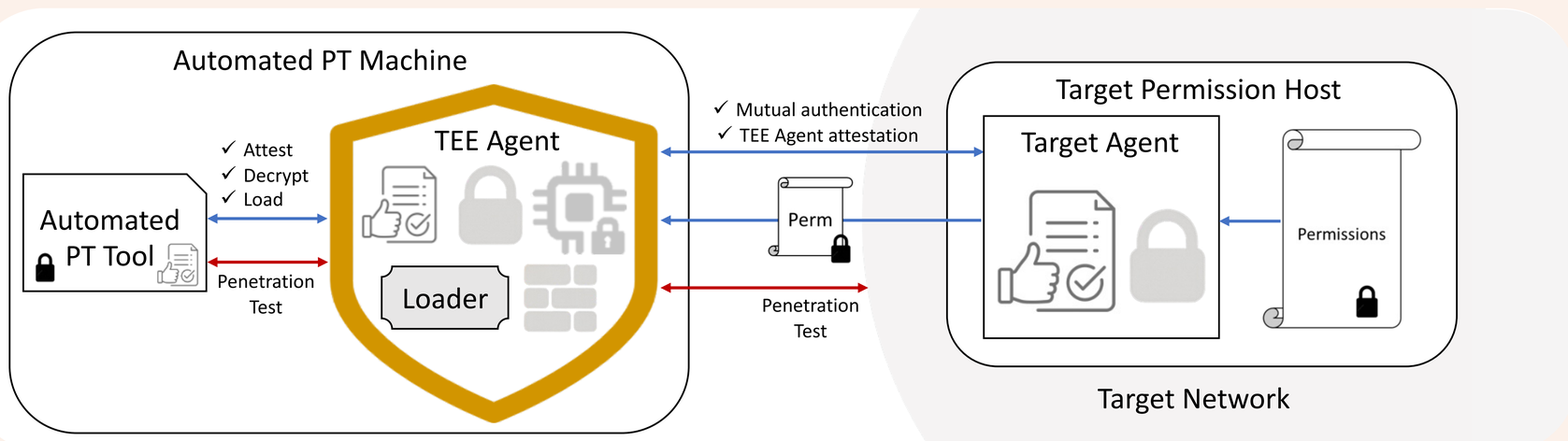
AI IN EACH PENETRATION TESTING PHASE



ETHICAL DEVELOPMENT OF AUTOMATED PT TOOLS



KEEP AI-BASED PENETRATION TESTS UNDER CONTROL



Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UID/000408/2025.

Explainable and Personalized Drug Recommendations via Metapaths and LLMs

Donato Aveiro[1], Francisco Couto[1], Matilde Pato[1,2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Instituto Superior de Engenharia de Lisboa, Portugal

LASIGE health and biomedical informatics

Recommendation systems have demonstrated great success across various domains, but their application in **healthcare** presents unique challenges that demand special consideration.

Got a bad movie?
2h wasted

★☆☆☆☆

SAFENESS

TRANSPARENCY

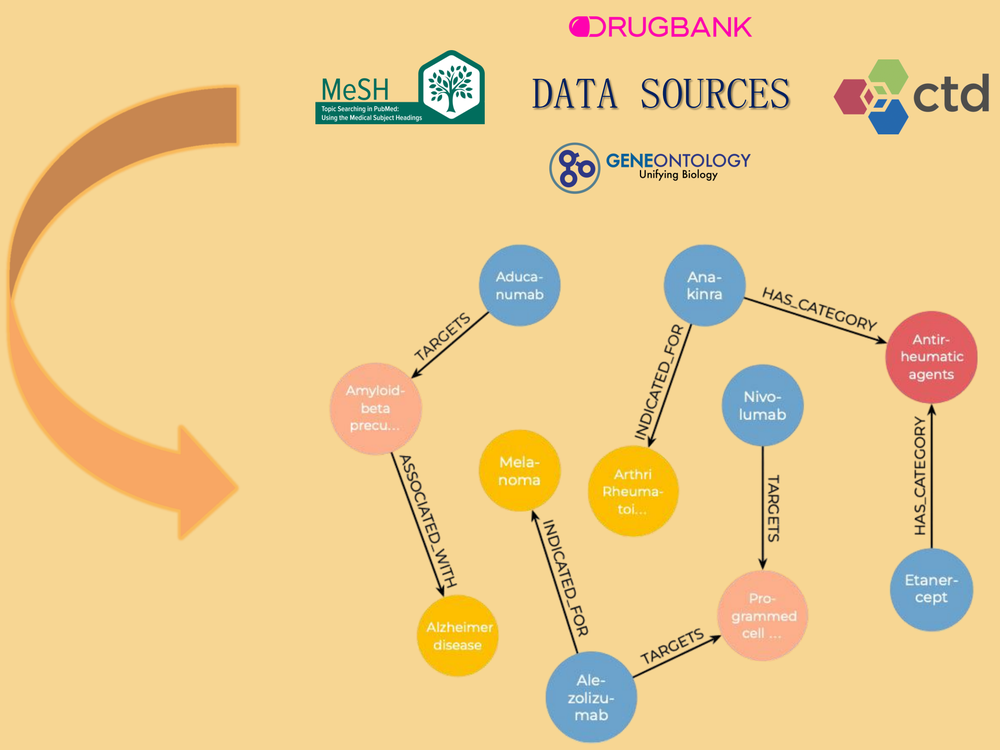
TRUSTWORTHINESS

Got a bad drug?
LIFE AT RISK

DON'T TAKE THE SUBSTANCE!

1. Why was I recommended this drug?
2. Has my doctor considered other drugs that I am already taking?

1. Leverage **metapaths** in a **knowledge graph** to capture rich, multi-hop relations between drugs and medical conditions.
1 2. Fine-tune a **large language model (LLM)** to provide **natural language explanations** and tailor recommendations to individual patient profiles.



*Nivolumab is indicated for **melanoma** because it is an inhibitory antibody of **programmed cell death protein 1**, like **atezolizumab**, which also treats that condition. However, you are taking **estradiol**, which may increase the thrombogenic activities of **nivolumab**, so be careful!*

Which metapaths are the most important?
Hetnet connectivity search identifies types of paths that occur more frequently than would be expected by chance, based on node degree alone.

+ **Drug-Drug Interaction** information

Acknowledgements: This work was supported by the LASIGE Research Unit, ref. UID/00408/2025.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Exploring the Uncanny Valley Effect: How Visual Perception-Related Autistic Traits Shape Human-Robot Interaction

Isabel D. Trancoso[1], Tomás Palma[2], Soraia M. Alarcão[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] CICPSI, Faculdade de Psicologia, Universidade de Lisboa, Portugal

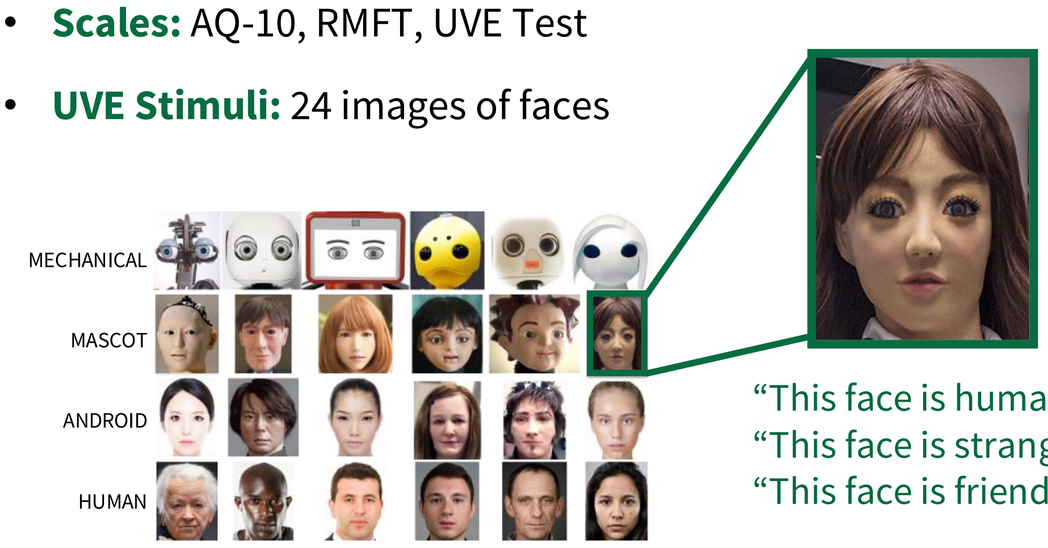
LASIGE inclusive human-computer interaction

Motivation

- Autistic individuals commonly present visual perception-related traits, such as difficulties in interpreting facial expressions, and hyper attention to detail, hindering their adaptation to the environment
- Robots have been used as Assistive Technologies (ATs) by the autistic population to foster their cognitive, social, and communication skills. However, their degree of realism can lead to negative emotional responses - the so-called Uncanny Valley Effect (UVE) - which may harm acceptance and limit their effectiveness as ATs
- While the UVE has been widely studied for the general population, little is known about its onset in individuals with visual perception-related autistic traits

Within-subjects Design Study

- **Goal:** Investigate whether visual perception-related autistic traits impact the emergence of the UVE, while studying the role of gender in moderating this effect
- **Sample:**
  - 287 participants (M_{age} = 21.8 ± 4.65; 174 females, 107 males)
  - 59 participants with severe autistic traits, and 209 with at least one autistic trait related to visual perception
- **Scales:** AQ-10, RMFT, UVE Test
- **UVE Stimuli:** 24 images of faces

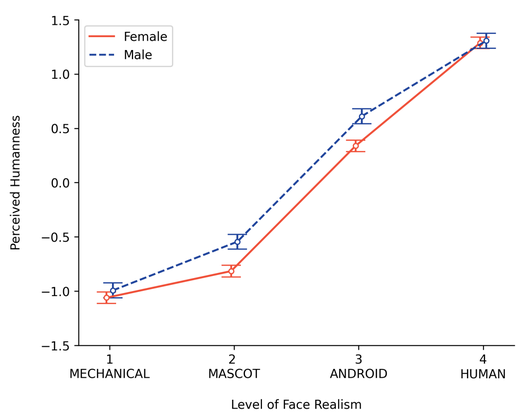


"This face is human."
"This face is strange."
"This face is friendly."

Results

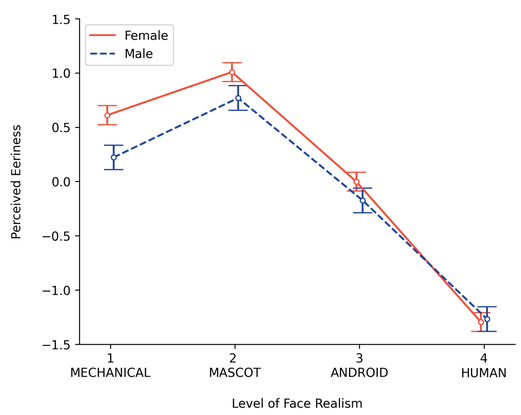
- **Emotion Recognition:** The greater the autistic traits, the lower the ability to recognize emotions from facial expressions

- **Humanness Ratings:**



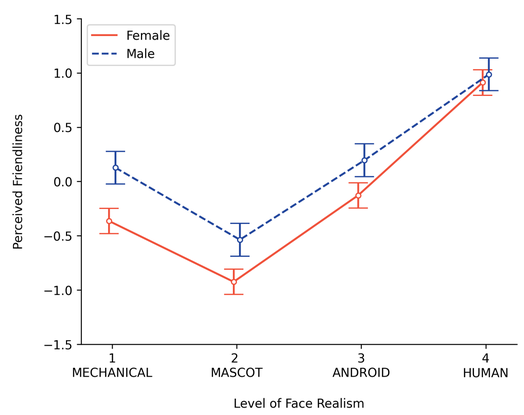
The greater the autistic traits, the higher the perceived humanness. Higher ratings in males than females

- **Eeriness Ratings:**



Only gender had a significant effect, with males attributing lower eeriness ratings than females

- **Friendliness Ratings:**



The greater the global autistic traits, the higher the perceived friendliness. Higher ratings in males than in females

Conclusions

- Despite difficulties in recognizing emotions and perceiving faces, individuals with autistic traits related to visual perception experienced a UVE similar to neurotypical individuals
- Interestingly, gender was the factor that most influenced how faces in the robot-human spectrum were perceived, with UVE being less potentiated in males

Main Design Implications

- We suggest greater care in using highly ambiguous features (robotic *vs.* humanoid), such as Mascot robots, as these were the eeriest and least friendly rated faces
- We reinforce the importance of analyzing each individual's preferences and needs to enhance the adequacy and acceptance of ATs developed in the future

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · Psicologia ULISBOA

Acknowledgments: This work was supported by FCT through LASIGE Research Unit, ref. UID/00408/2025, as well as the Research Center for Psychological Science of the Faculty of Psychology, University of Lisbon (UIDB/04527/2020; UIDP/04527/2020)

Safe Programming Language Interoperability with Session Types

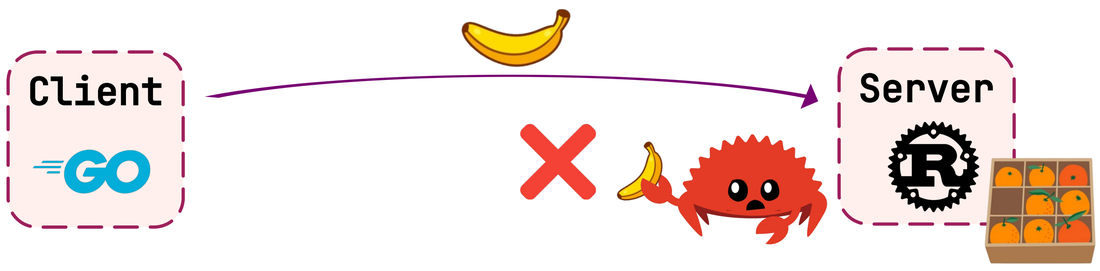
João Magalhães, Bernardo Almeida, Alcides Fonseca, Vasco Vasconcelos

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE reliable software systems

Motivation

In the microservices architectural approach, issues such as message mismatches, unexpected sequences of interactions, and subtle integration bugs can lead to system failures that are difficult to detect and debug



Microservices are often written in different languages and rely on communication mechanisms that provide limited guarantees about the structure or ordering of exchanged data

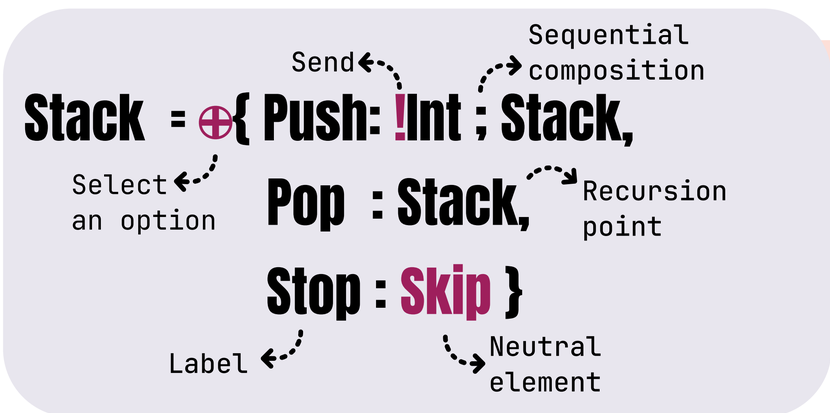
Session Types

Upgrade communication channels from simple message-passing mechanisms to structured conduits that capture and enforce the rules of interaction between distributed components. Session types can avoid a wide range of coordination and integration errors in complex systems

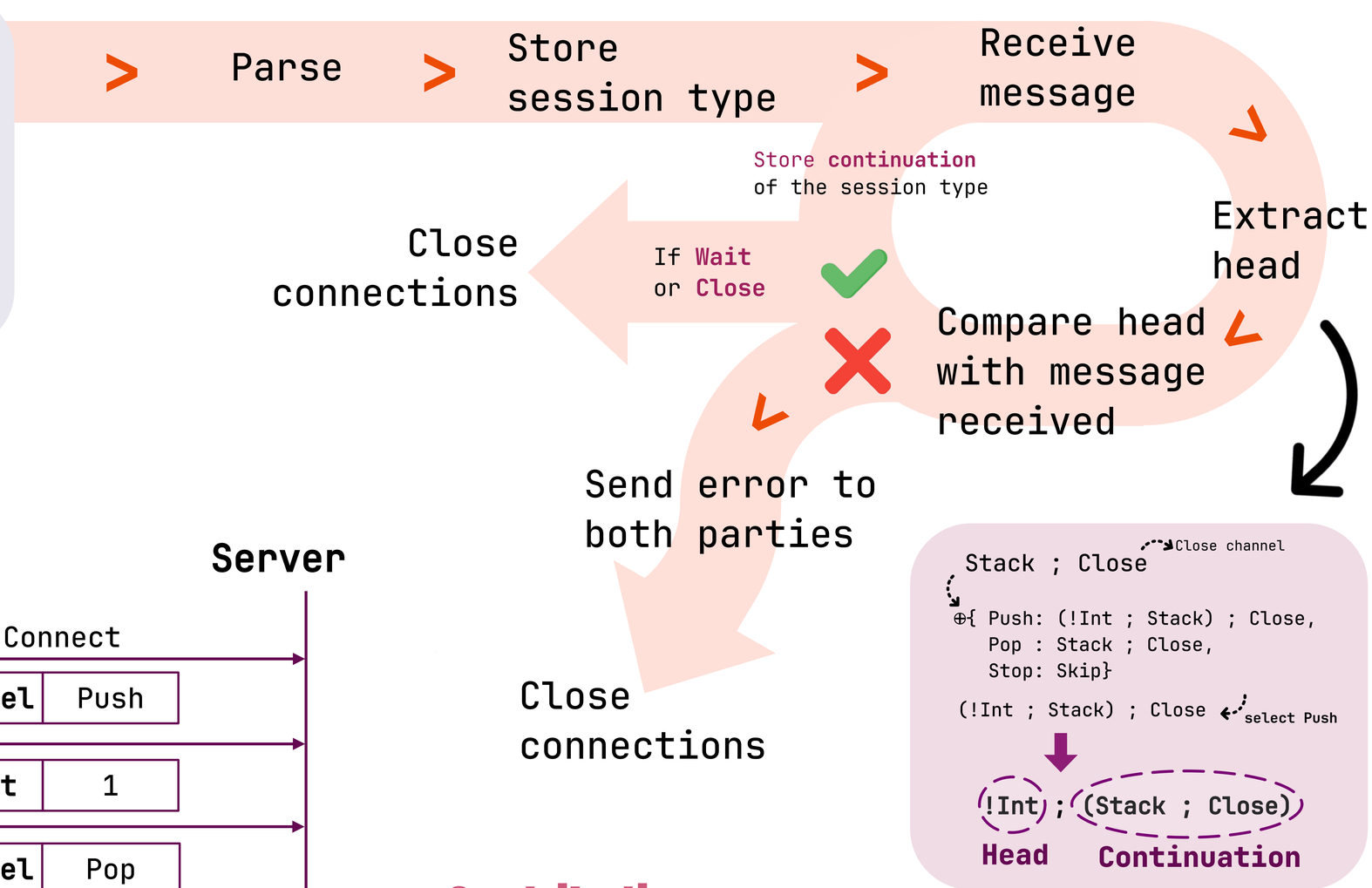
FreeST supports session types natively. But it is not enough when communicating with other languages that does not support session types

How can cross-language communication reliably enforce a common protocol?

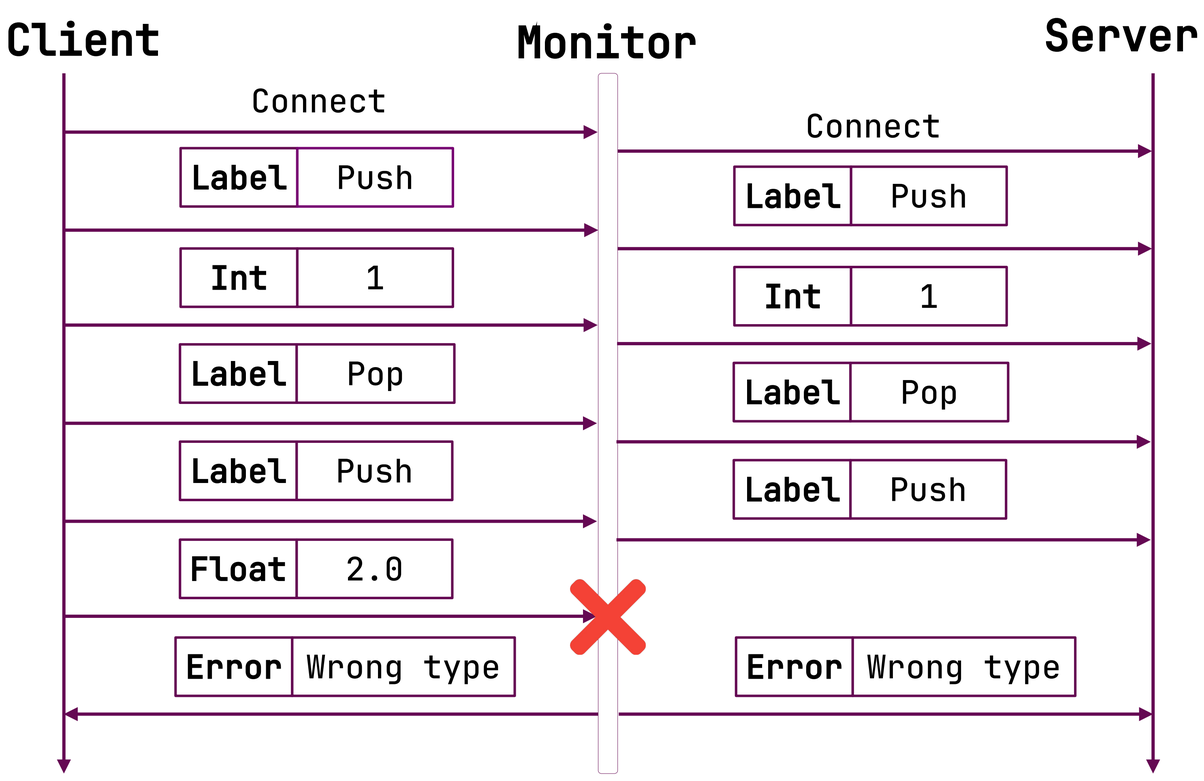
Session Type Example



State Machine Flow



Communication Example



Contributions

- FreeST now enables communication with external programs, extending its capabilities beyond isolated systems
- Programs written in different languages can now interact safely, with the guarantee that all participants adhere to a clearly defined communication protocol

Acknowledgements: This work was supported by Fundação para a Ciência e Tecnologia (FCT) in the LASIGE Research Unit under the ref. (UID/00408/2025).

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

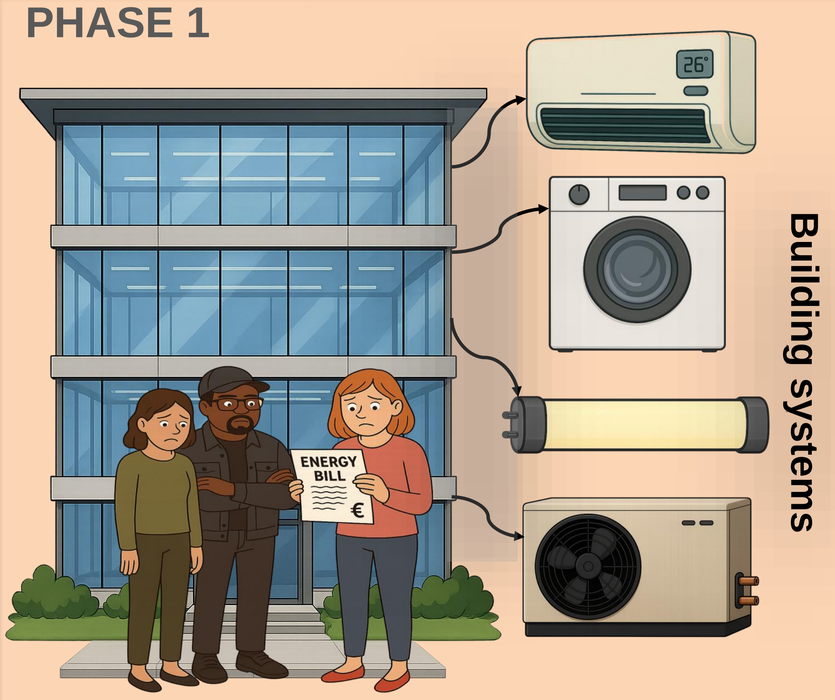
Context-Aware Machine Learning-Based Predictive Framework For Building Energy Management

Miracle Aniakor, Vinicius V. Cogo, Pedro M. Ferreira

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

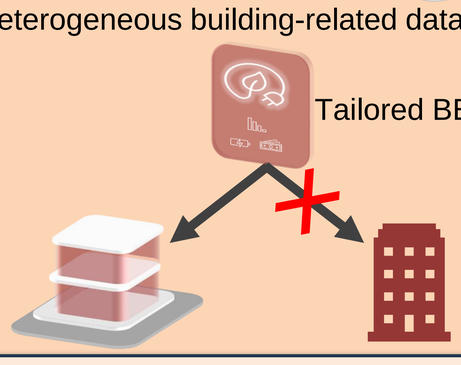
PHASE 1



- Buildings account for about 30% of global energy consumption.
- Lack of transparency on building actors' energy usage and impact on utility bills.

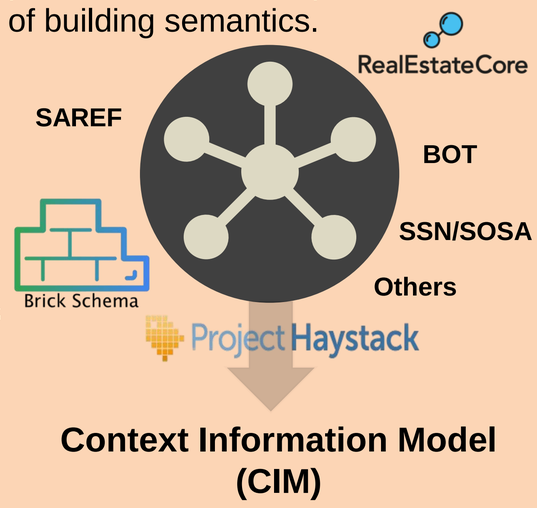
- A digital building ecosystem is initiated to support Building Energy Management systems (BEMS) and applications.

Heterogeneous building-related data and contexts



Tailored BEM Applications

We employ semantic models to provide a common representation of building semantics.



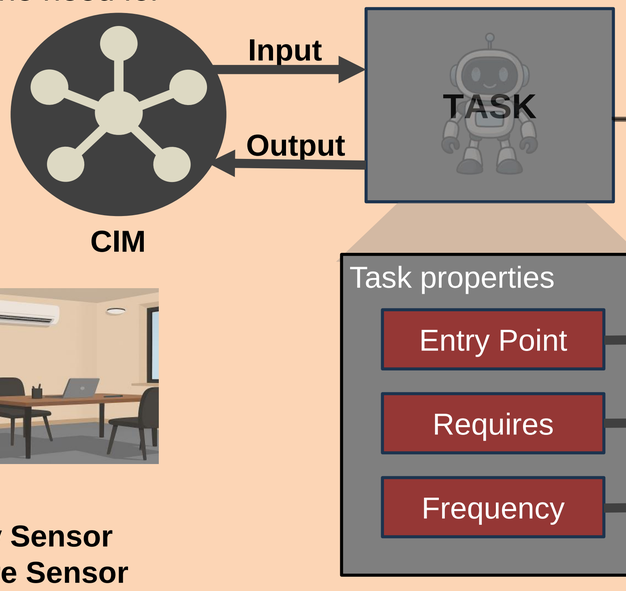
GAP 1: Lack of semantic representation of computational tasks associated with buildings

PHASE 2

Computational tasks impacting buildings are often difficult to identify.

- These tasks are usually conducted frequently necessitating the need for automation.

We **propose a novel** approach **to extend pre-existing ontologies** to model computational tasks that can be instantiated with specific details.



Task categories
- Virtual sensors
- KPIs
- Assessments
- Services

Triggers

Points to a **function** for the execution of a task

What **input** is needed to the task to be executed

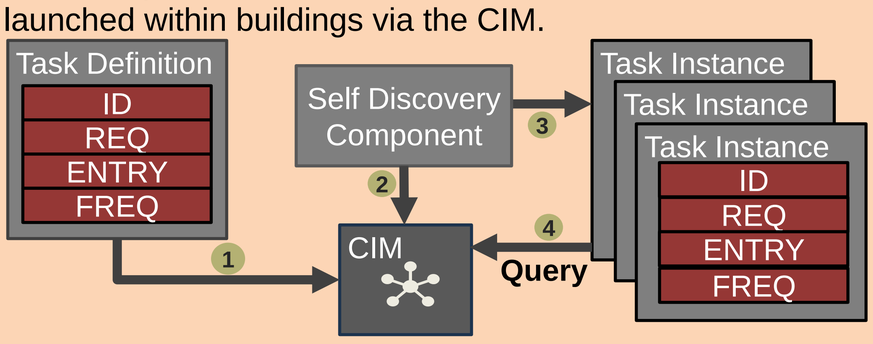
Specified **periodic** interval of execution

Thermal comfort task

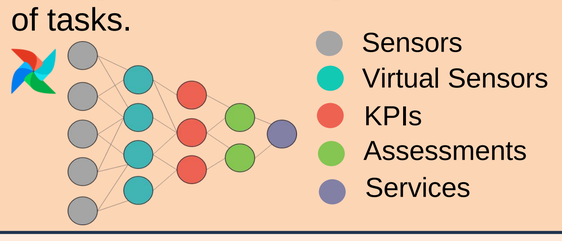
	Building Spaces	
1	T1 O1	✓
2	O3	✗
3	T2	✗

Requirements
- **O - Occupancy Sensor**
- **T - Temperature Sensor**

Self-discovery: Intelligent mechanism to find all possible tasks that can be launched within buildings via the CIM.



Task automation: employs a task orchestrator such as Apache Airflow to manage the scheduling and execution of tasks.

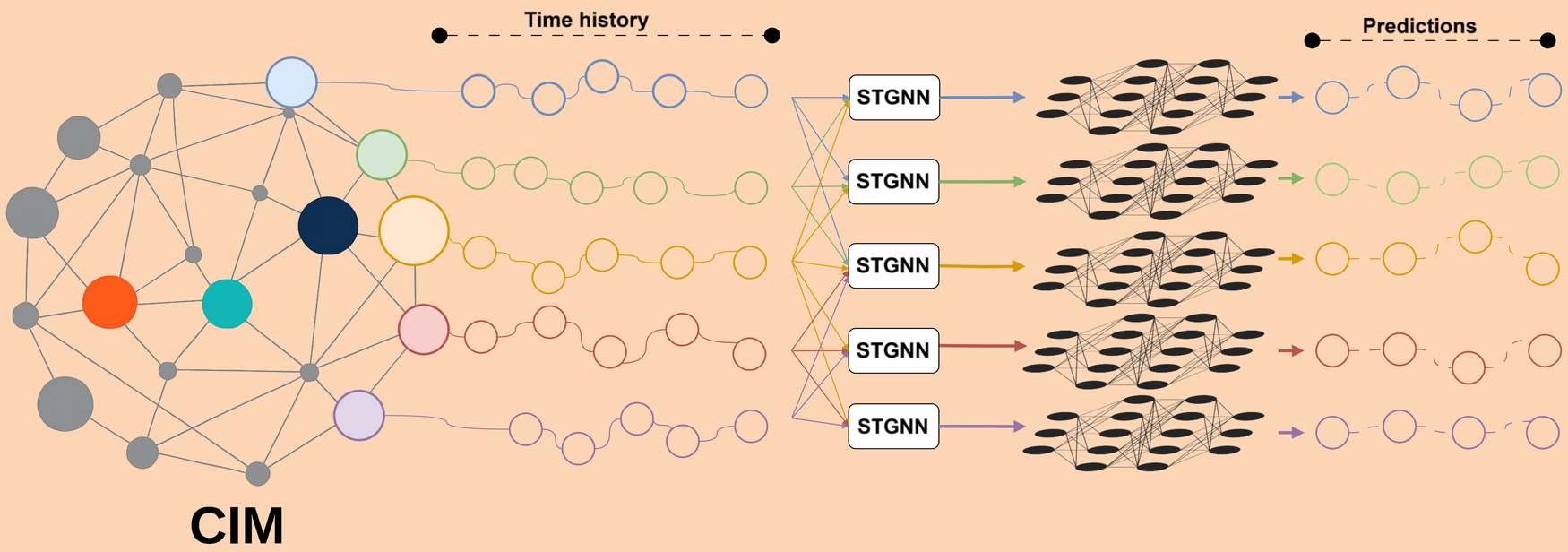


- Autonomous discovery of tasks that can be computed for a given building.
- Automate repetitive computational tasks in buildings.

GAP 2: Lack of context-aware adaptive predictive models to support BEM applications

PHASE 3

- Develop a CAML-based predictive framework using graph-based models such as Spatio-Temporal Graph Neural Networks (STGNN) and graph transformer networks.
- STGNN design exploration of architectural components for embedding building semantics.
- Dynamic creation of the predictive network model architecture.



U LISBOA | Ciências ULisboa | LASIGE | SATO OF BUILDING ENERGY

Paper: A Survey on Semantic Modeling for Building Energy

This work was partially supported by the European Commission through the SATO project,ref. H2020-IA-957128 (https://doi.org/10.3030/957128) and by the Fundação para a Ciência e Tecnologia (FCT) through the LASIGE Research Unit, ref. UIDB/00408/2025 and ref. UIDP/00408/2025

Beyond Unimodal Understanding: A Unified Multimodal Breast Exam Dataset and Small Vision-Language Model for Semi-Structured Report Generation

José Domingues[1,2], Hugo Figueiras[1,2], Nuno Matela[2], Nuno Garcia[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] IBEB, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE data and systems intelligence

Problem Statement

Screening results may take 1+ month since exam. Cancer doubles size every 6 months

Breast cancer research traditionally relies on single imaging modalities, **limiting integration of contextual information from multiple sources**

Standard auto-regressive next token prediction **does not ensure context and consistency**

AI Clinical deployment needs transparency and interpretability, but **current trends favor large, complex models that sacrifice efficiency and explainability for better accuracy**

Dataset

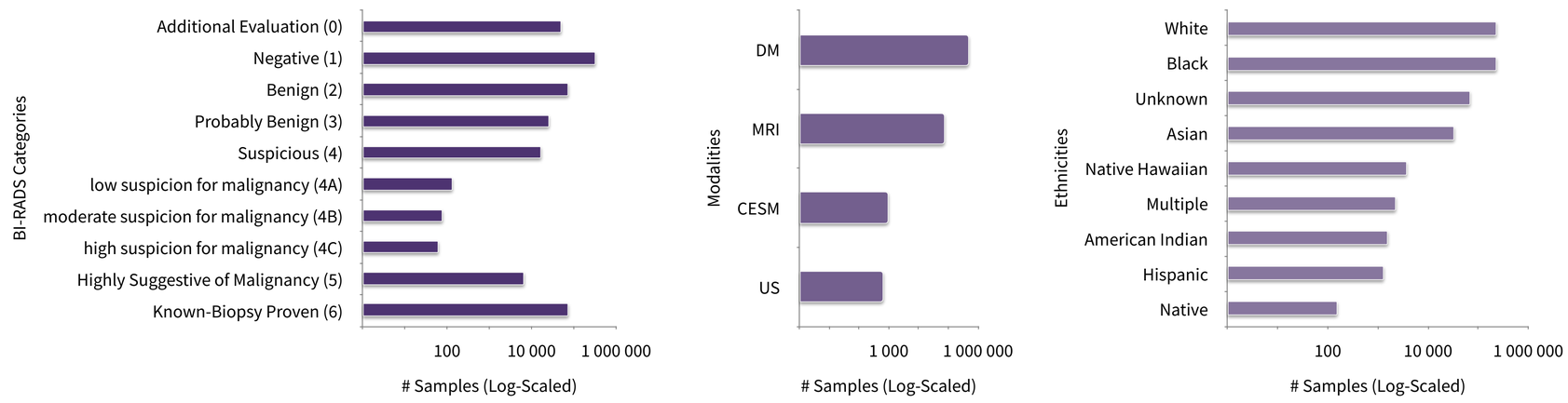
First open-source, unified dataset for breast cancer imaging, **composed of Digital Mammography, Magnetic Resonance Imaging, and Ultrasound**

Imbalance hack: Weighted Random Sampler + Heavy Augmentation

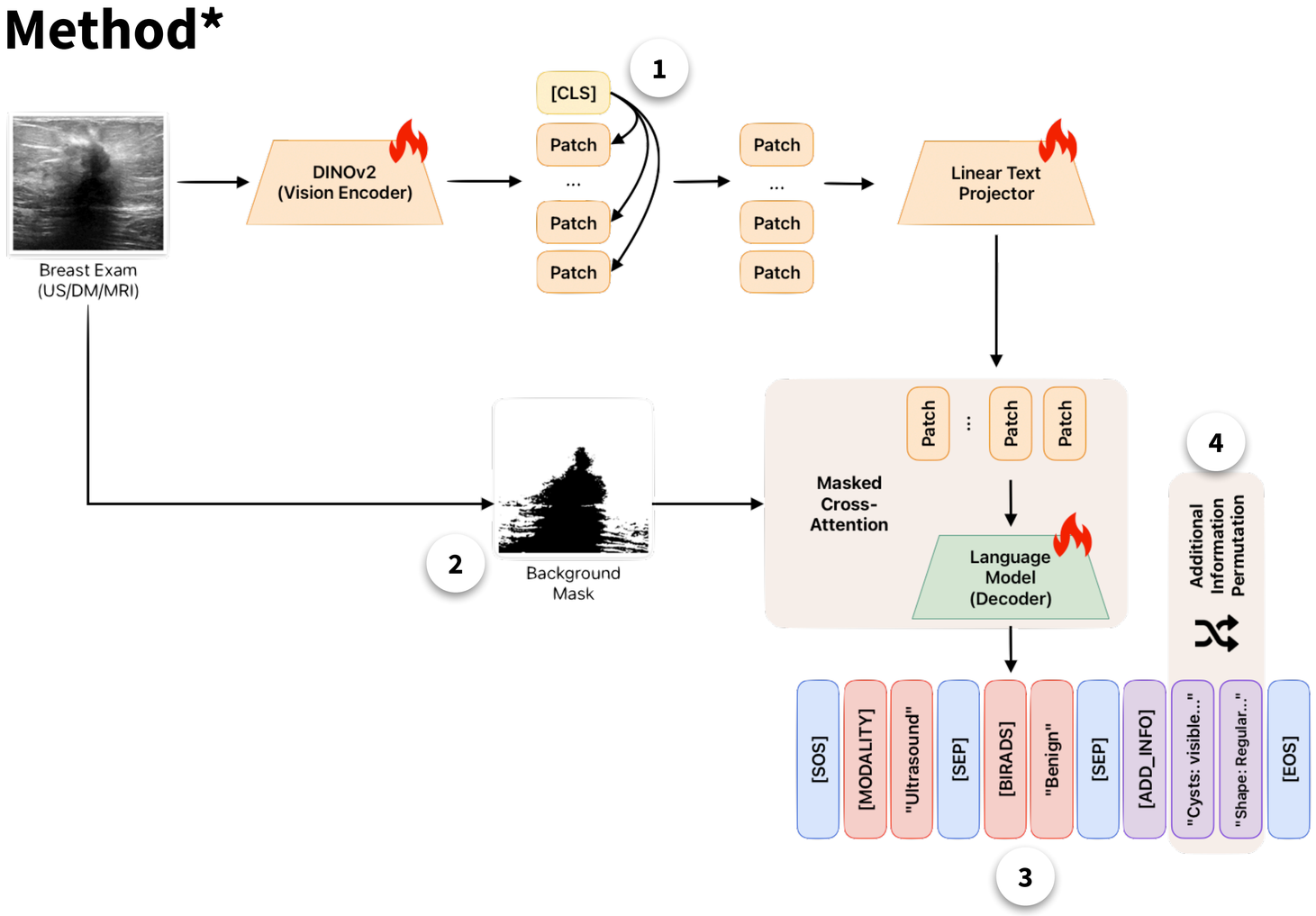
$$w_i = \frac{1}{N_{C_i}}$$

N_C – Number of samples in class $c \in C$

W_i – Weight w_i for a sample i in class C_i



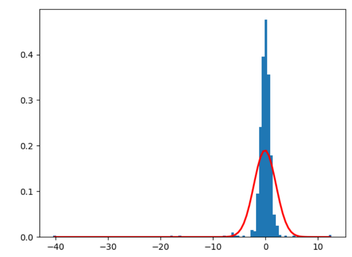
Method*



1 Adjusted DINOv2 Patch Representations

Long tailed distribution helps shifting features from ImageNet Pre-trained DINOv2

$$F = LN(\sum_{-1} P(CP))$$



2 Masked Cross-Attention

Mask access to background regions based on pixel values distribution

$$\tilde{I} = \mathrm{Norm}(I)$$

$$I'(x,y) = \begin{cases} 1, & \text{if } \tilde{I}(x,y) > T \\ 0, & \text{otherwise} \end{cases} \quad \text{where } T = Q_1(\tilde{I}) + 0.1 \cdot \mathrm{IQR}$$

$$M = \begin{cases} 1, & \text{if } \sum_{(x,y)\in N} I'(x,y) > 0 \\ 0, & \text{otherwise} \end{cases}$$

3 Ordered Conditional Auto-Regressive Semi-Structured Generation

Modality and BI-RADS enforce context when predicting fine-grained findings

$$\text{Modality} = \arg\max_{y \in M} \log p_\theta(y \mid I)$$

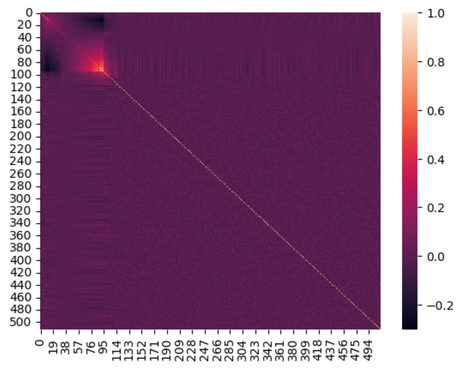
$$\text{BIRADS} = \arg\max_{y \in B} \log p_\theta(y \mid I, \text{Modality})$$

$$\text{Additional Information} = \arg\max_{y \in V} \log p_\theta(y \mid I, \text{Modality}, \text{BIRADS}, y_{<t})$$

4 Permutation Invariance Findings

Shuffled findings to avoid bias towards order

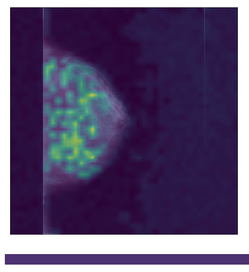
Unobserved pattern after ~150 tokens is due to short findings dataset



Results

Eval	MRI Set	US Set	DM Set
# Exams	3497	61	5568
# BI-RADS Cases	4	7	10
Unimodal Model			
— Add. Info BLEU-4 (SE)	0.6729 (0.0062)	0.2569 (0.0113)	**0.4044 (0.0017)**
— BI-RADS F1-Score (%) (SE)	25.57 (1.06)	12.86 (4.59)	21.78 (0.66)
— Modality F1-Score (%) (SE)	98.98 (0.29)	97.14 (2.30)	**99.71 (0.10)**
Multimodal Model			
— Add. Info BLEU-4 (SE)	**0.7084 (0.0060)**	**0.4143 (0.0161)**	0.3754 (0.0017)
— BI-RADS F1-Score (%) (SE)	**64.57 (0.70)**	**28.57 (6.09)**	**30.25 (0.54)**
— Modality F1-Score (%) (SE)	**100.00 (0.00)**	**98.57 (1.64)**	90.75 (0.42)

Attention (BI-RADS Prediction)



Target Report

Modality: dm
Birads: suspicious
Additional info:
- image view: cranial caudal
- breast density: scattered fibroglandular densities
- abnormality type: calcification
- calcification type: dystrophic
- calcification distribution: clustered
- pathology: benign

Inferred Report

Modality: dm
Birads: suspicious
Additional info:
- image view: mediolateral oblique
- breast density: scattered fibroglandular densities
- abnormality type: calcification
- calcification type: pleomorphic
- calcification distribution: clustered
- pathology: malignant

Future Work

Better context enforcement to avoid diagnosing out of scope and maintain consistency

Improve alignment of attention with real-world expectations

Acknowledgements: This work was supported by FCT through project 2023 Call for PhD Studentships — Regular Line of Application, ref. 2023.00273.BD, and the LASIGE Research Unit, ref. UID/00408/2025.

* Submitted to MICCAI '25 and IEEE TPAMI

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · IBEB · MICCAI · IEEE

Improving Oblivious BFT State Machine Replication

David Dantas[1], Bernardo Ferreira[1], Alysson Bessani[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

dependable and secure decentralized systems

MOTIVATION

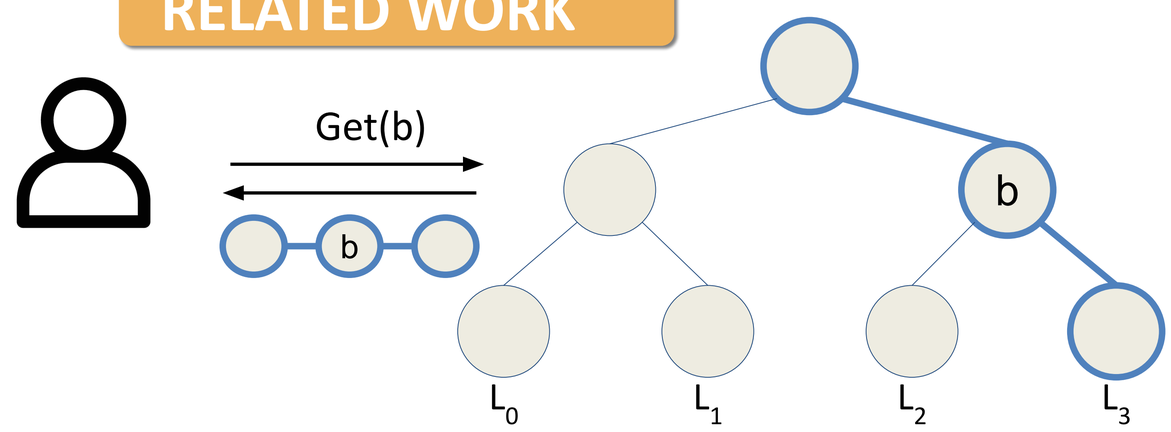
Oblivious RAM (ORAM) is a cryptographic algorithm that ensures a server learns nothing about a client's access in a remote database, thus hiding not only the data, but access patterns too.

Remote databases, however, must account for **malicious clients** wishing to alter unauthorized data.

PROBLEM

How can an oblivious server validate client accesses without learning anything?

RELATED WORK



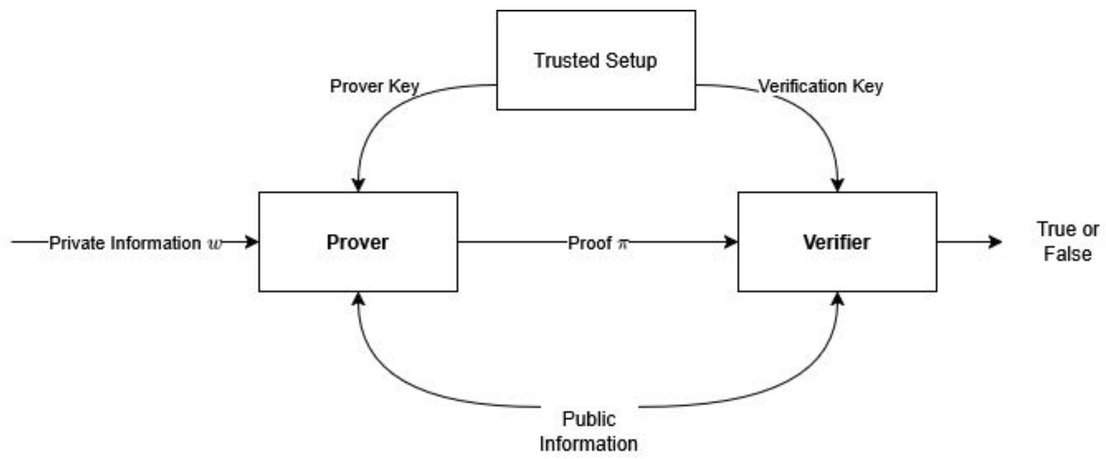
Path ORAM hides access patterns by storing data in a binary tree, and having clients read and write a full path. Blocks are distributed randomly through the tree according to a position map

GOALS & APPROACH

- **Secrecy:** Clients can only read entries if they have *read* permission
- **Integrity:** Clients can only write entries if they have *write* permission
- **Tamper-Resistance:** A adversary can't modify an entry without eventual detection
- **Obliviousness:** The server can learn no access patterns

Our implementation combines zk-SNARKs with the Path ORAM protocol to build a multi-client ORAM, enabling the data owners to create an oblivious database with customizable read/write access controls.

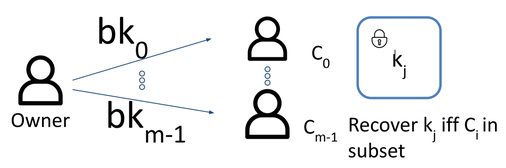
To enforce these controls efficiently, we introduced several modifications that enable zero-knowledge proofs with practical overhead.



Zero-**K**nowledge **S**uccinct **N**on-Interactive **Ar**gument of **K**nowledge (zk-SNARK)

Access Control

Read:
- Broadcast Encryption



Write:
- Write Access Merkle List

m	$M_{root}(wb_0)$	…	$M_{root}(wb_{m-1})$

- Server stores a list of size m (One entry per client) with merkle roots of a list wb_i
- Items in wb_i are either:
  - $1 \| b_j$ if client C_i has write access to block b_j
  - $0 \| b_j$ if client C_i does not have write access

Obliviousness

Elliptic Curve ElGamal Encryption:
- Keygen: Generate random private key x, and public key Y
- Encrypt(m, Y) ➡ C
- Decrypt(C, x) ➡ m

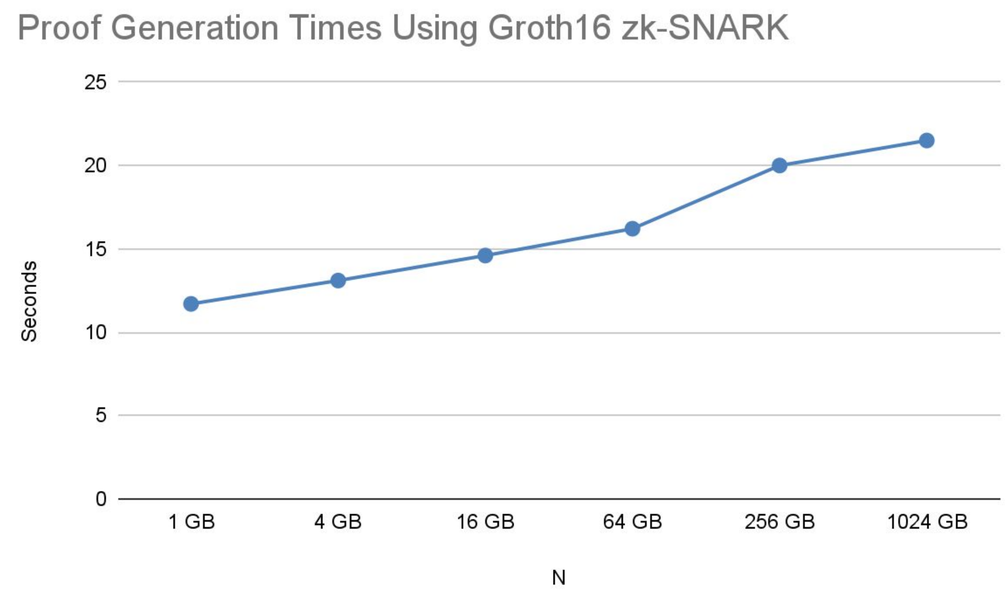
- **Rerand(C, *Y*) ➡ *C'***
- ***Benefits:***
  - Rerand allows clients to refresh the cipher without learning it's content, preserving obliviousness to the server and other clients.
  - Elliptic curve elgamal is highly efficient in zk-SNARKs.

Protocol Correctness

Zero-knowledge proof is generated during each eviction to:
- Prove client C_i only modified a block b_j they have write access to ($1 \| b_j \in wb_i$)
- Ensure all other blocks are re-randomized
- Confirm blocks are evicted to correct paths per the position map
- Verify the position map is correctly updated

This proof:
- Enforces the data owner's **access control policy**
- **Prevents malicious behavior** that could force honest clients to download large parts of the database

RESULTS



Proving times for a DB of varying size with blocks of 16kB. Proof times largely depend on the height of the tree which is *log N*. Verification takes constant time (≈130ms)

References:
[1] Emil Stefanov, Marten Van Dijk, Elaine Shi, T.-H. Hubert Chan, Christopher Fletcher, Ling Ren, Xiangyao Yu, and Srinivas Devadas. 2018. Path ORAM: An Extremely Simple Oblivious RAM Protocol. J. ACM 65, 4 12(2018), 18:1–18:26.
[2] Jens Groth. 2016. On the Size of Pairing-Based Non-interactive Arguments. In Advances in Cryptology – EUROCRYPT 2016, Marc Fischlin and Jean-Sébastien Coron (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 305–326.

This work was supported by FCT through the SMaRtChain project (2022.08431.PTDC) and the LASIGE Research Unit (UIDB/00408/2020, UIDP/00408/2020, and UID/00408/2025).

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

Exploring Physiological Indicators of Gaming Disorder

Marco Viana[1], Ana Paula Afonso[1], Manuel J. Fonseca[1], Joana Cardoso[2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Center for Psychology at the University of Porto, Porto, Portugal

LASIGE health and biomedical informatics

Motivation

- Video games are rapidly growing, raising concerns
- The excessive use of video games can escalate to pathological and even addictive levels
- **Internet Gaming Disorder** (IGD) is recognized by WHO and APA
- Diagnosis relies on self-reports, often inaccurate

Goals

- Investigate correlations between physiological responses and IGD symptoms in Counter-Strike 2 players during gameplay
- Analyze physiological signals (EEG, PPG) to identify potential patterns and biomarkers
- Distinguish amateur or professional players with pathological symptoms from those without
- Create a dataset

Methodology

- Data collection
- Analysis of the collected data
- Create a dataset
- Identify correlations and patterns between physiological biomarkers and IGD

Related work

P300
- Amplitude ↓ in IGD
- Latency ↑ (slower processing)

N100
- Amplitude ↓ (weaker early sensory response)

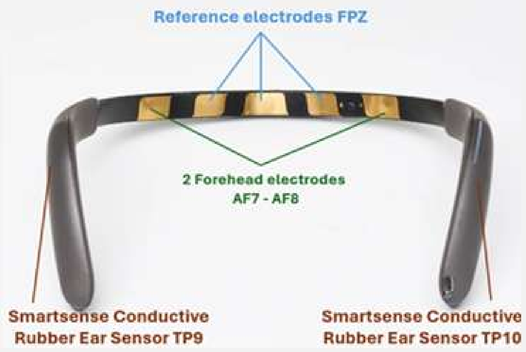
Delta / Theta waves
- Power ↑ in IGD (indicates lower alertness)

Alpha wave
- Tends to ↓ with stronger game engagement

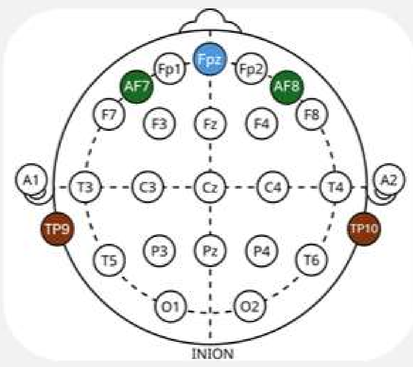
Beta / Gamma waves
- ↑ in some studies; ↓ in others

Data collection

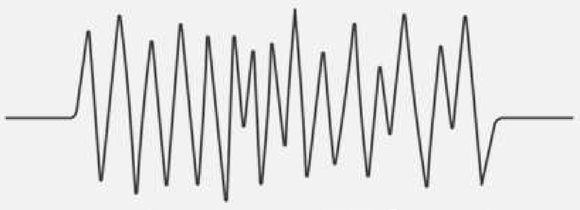
a. Participant fills 4 questionnaires (sociodemographic, IGD9-SF, BSI, SWLS)



b. Fit Muse 2



Position of the EEG electrodes captured by the Muse 2 device according to the international 10-20 system



c. 2-minute EEG and PPG recording

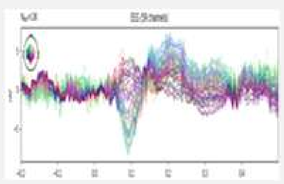
d. Play a game with the assigned team while recording the signals

e. 5–10 minute break; record game results and participant performance

f. Repeat d) to e) with the participant playing with the other team

g. Participant completes a questionnaire on experience and emotions

h. Analysis of the Data



i. Create a dataset and identify Patterns

Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UID/000408/2025

U LISBOA · Ciências ULisboa · FCT Fundação para a Ciência e a Tecnologia · LASIGE

HexGaze
Virtual Keyboard for Text Input Using Gaze

Robert Cachapa, Manuel J. Fonseca

LASIGE, Faculdade de Ciências da Faculdade de Lisboa, Portugal

LASIGE inclusive human-computer interaction

MOTIVATION

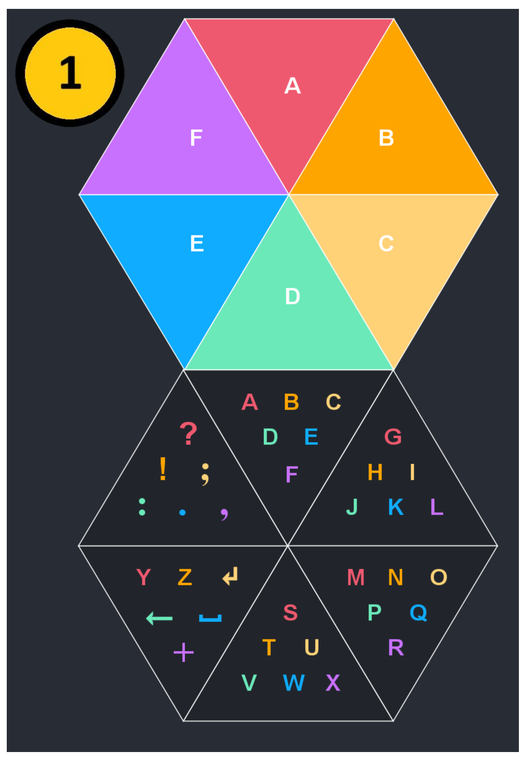
Eye typing is a technique that allows users to interact with computers and type sentences using only gaze

Current eye typing techniques lack speed and comfort
Some are not viable with low precision eye trackers

GOALS

Gaze only interaction technique
Support efficient, continuous and fluid typing
Avoid Midas touch problem
Maximize prolonged use comfort and minimize eye strain

HEXGAZE TECHNIQUE



Initial State

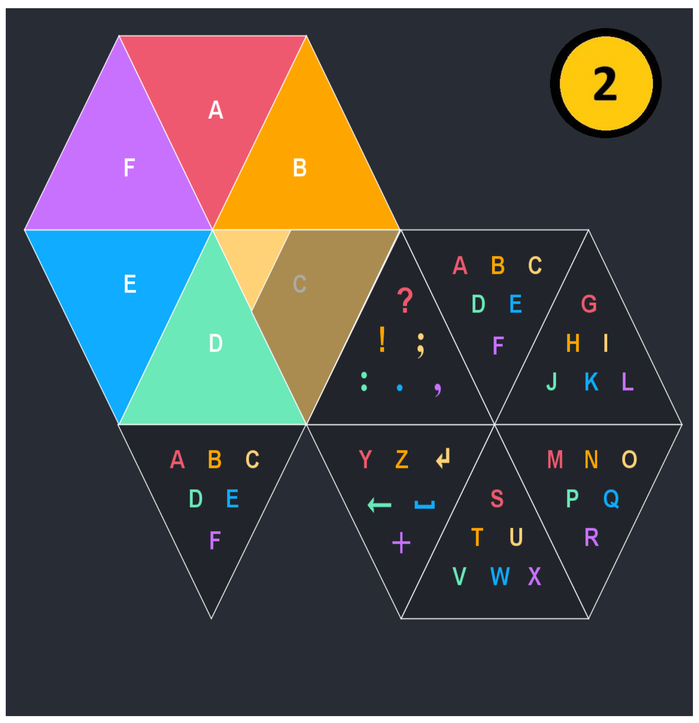
Group Hexagon: Hexagon with 6 segments containing 6 characters each

Letter Hexagon: Hexagon with 6 segments containing 1 character each (from the previously gazed group)

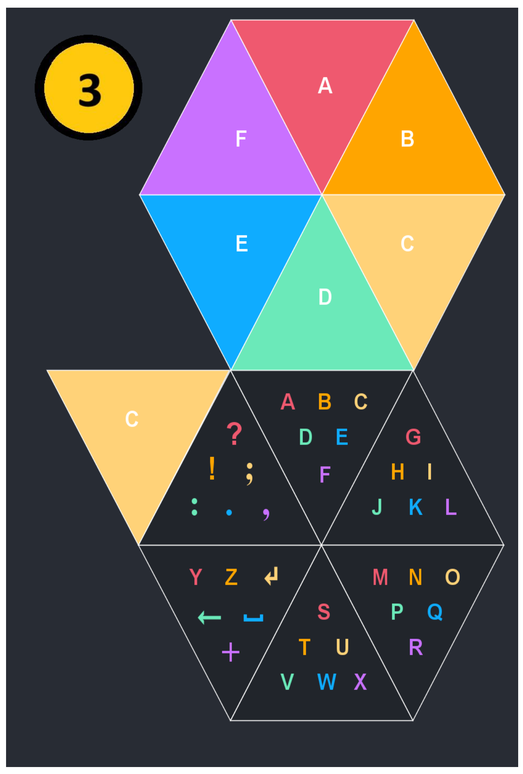
Letter Selection: Gaze at the next Group Hexagon or by gazing at the letter for 800 ms

Example: Typing “CAT”

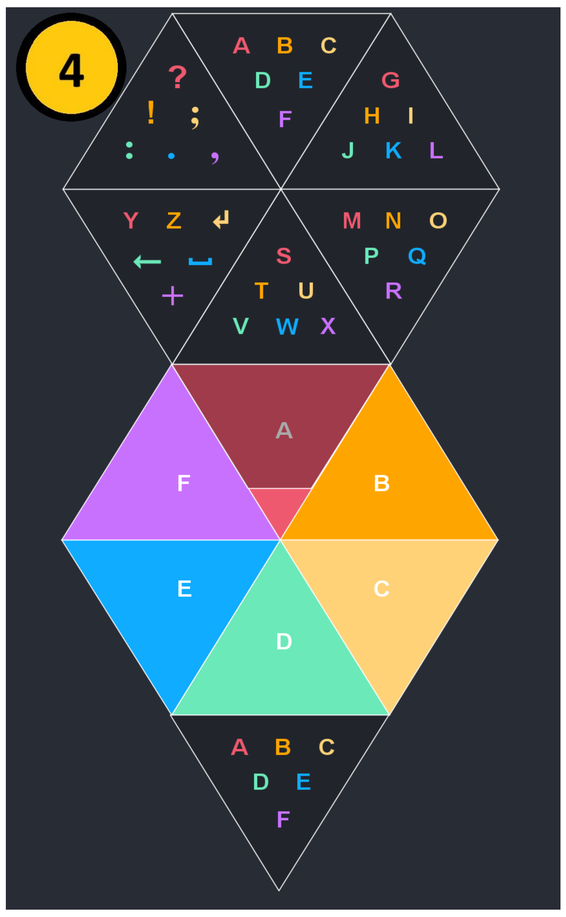
Word Prediction: Display complete words according to typed letters



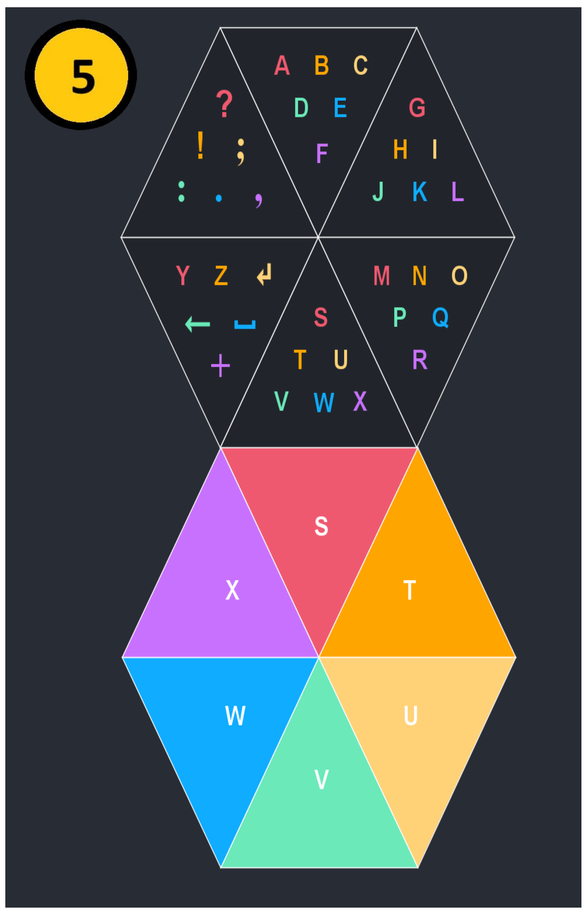
User Looks at C



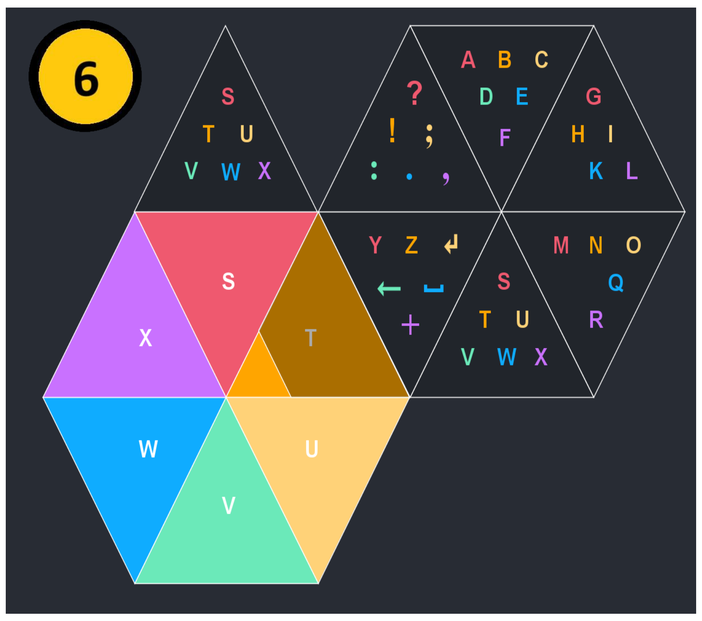
User Looks at Segment with A



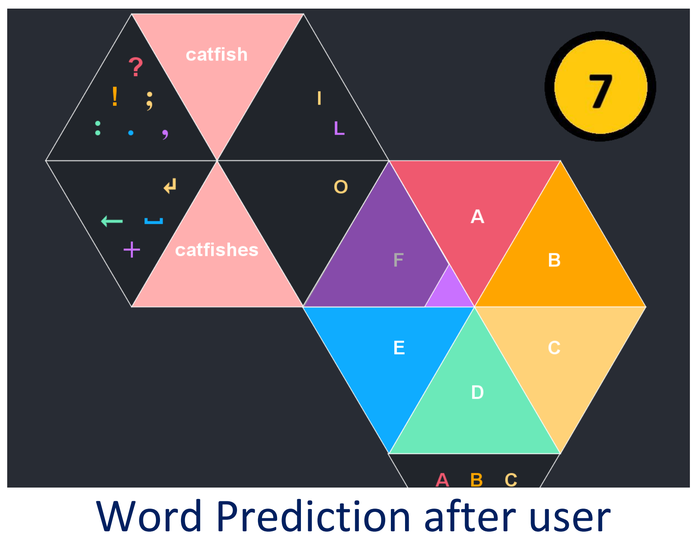
User Looks at A



User Looks at Segment with T



User Looks at T



Word Prediction after user looks at F

NEXT STEPS

User Testing: Tests with users to identify usability issues and validate design decisions
Comparison Test: Comparison of HexGaze with two existing techniques
Learning Evolution: Eight sessions over time (one on each day)
Metrics: Words per Minute (WPM), Minimum String Distance (MSD), Comfort

Acknowledgements: This work was supported by FCT through project <title>, ref.\ <project reference> (<project DOI>), and the LASIGE Research Unit, ref. UID/00408/2025.

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

Deadlock Freedom in FreeST

Luís Santos, Andreia Mordido, Vasco T. Vasconcelos

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

reliable software systems

FREEST & DEADLOCKS

FreeST is an experimental strictly typed concurrent programming language in which message-passing communication protocols are defined through context-free session types. **Session types describe structured dyadic interactions**, and FreeST's context-free variant removes their original tail-recursive limitation, making them even more expressive.

A **deadlock** is a state in which all members of a given set of processes are **stuck and cannot proceed** to their next instruction. In message-passing contexts, they commonly occur as a result of a state of **circular wait**.

FreeST is a powerful tool that can be leveraged to develop complex concurrent systems, and **deadlocks stand in the way of** the level of **reliability and availability** that these systems should guarantee. As such, we aim to create a new, more robust version of FreeST that can guarantee that **only deadlock-free programs are compiled.**

DESIGNING A SOLUTION

```
type A = !String;Close
type B = !String;Close
```

We can model the interaction between Alice and Bob illustrated above in FreeST, using a pair of identical session types describing the same structured interaction: **sending a message and then closing the channel**. Channels have two endpoints, with one endpoint having the other's **dual operations**.

```
alice : A ->[T,⊥] dualof B 1->[a,d] ()
alice a b =
    let (_,b) = receive b in
    let a = send "hello" a in
    close a;
    wait b
```

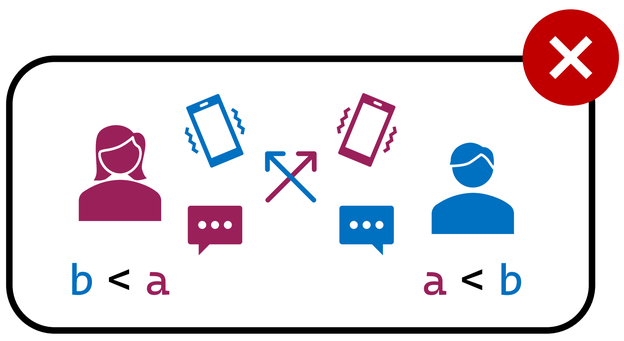
```
bob : dualof A ->[T,⊥] B 1->[a,d] ()
bob a b =
    let (_,a) = receive a in
    let b = send "hello" b in
    wait a;
    close b
```

However, both Alice and Bob's **first instruction is to receive a message** through their dual endpoint. This is obviously a deadlock scenario, because they will only send their messages after receiving one, so **both processes are stuck**.

Our proposed solution is to **annotate** our session types. These annotations are used within FreeST's typechecker to establish a **set of priorities**, which are then submitted to an **SMT Solver**. The solver determines whether our constraints are **satisfiable or not**. If no valid order of instructions is found, the program is considered **deadlock prone** and does not compile.

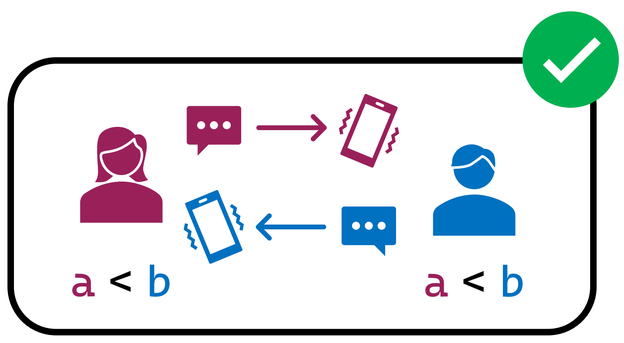
```
type A = !aString;Close c
type B = !bString;Close d
```

RESULTS



Using our new solver, we can determine that our program will result in a deadlock, because we have **two conflicting constraints**.

However, if we alter Alice's behavior, making her **send her message before receiving** Bob's, there is no conflict and our new deadlock-free program compiles.



Acknowledgements: This work was supported by FCT through project Reliable and Expressive Concurrent Systems through Advanced Session Types ref. 2023.13752.PEX (https://doi.org/10.54499/2023.13752.PEX), and the LASIGE Research Unit, ref. UIDB/00408/2025

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

A soft sensor to assess the energy performance of laundry washing machines

Zygimantas Jasiunas, José Cecílio, Pedro M. Ferreira

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

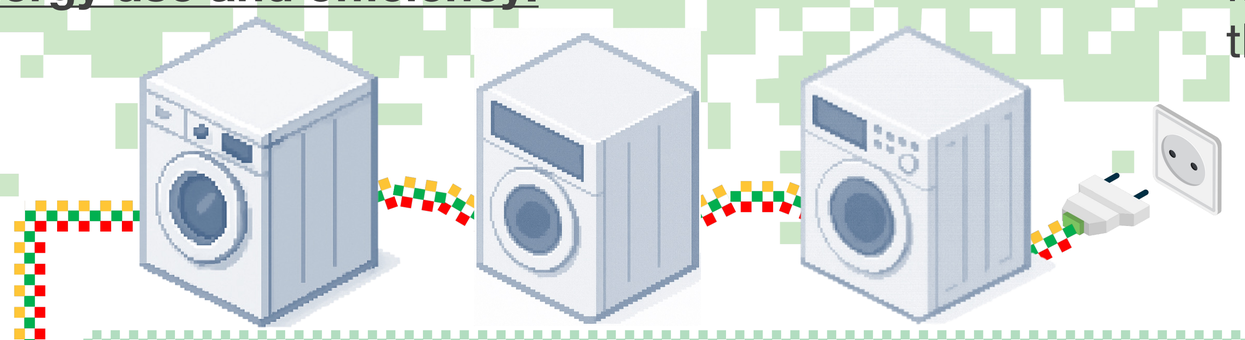
LASIGE cyber-physical systems

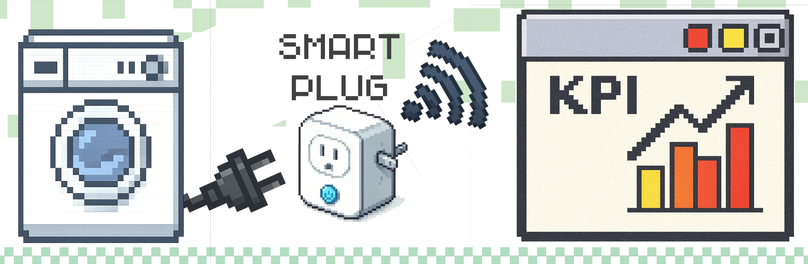
Household energy consumption increased to 27% of global demand, due to appliances.

ECO design and energy labelling **overlook real-life energy use and efficiency.**

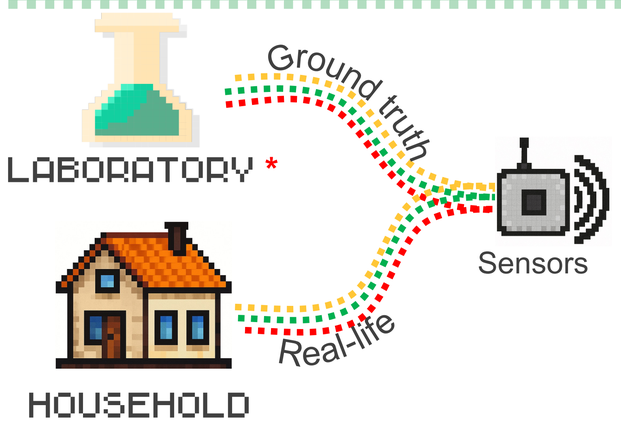
We propose using off-the-shelf smart-plug and machine learning to estimate real-life appliance efficiency via **key performance indicators**.

It considers real-life use and provides actionable insights that promote efficiency by optimising user practice.





DATA COLLECTION



ENERGY
- Power
- Current
- Voltage
- Frequency
- Power factor

OPERATION
- Water inlet / outlet*
- Water inlet / outlet temp.
- Vibrations X,Y,Z axes
- Drum temperature*
- Door temperature*
- Acoustic sound*

ENVIRONMENTAL
- Ambient temperature
- Ambient humidity
- Outside / Inside

CONFIGURATION*
- Varying laundry load
- Controlled ambient temp.
- Operational setting selection
  - Temperature
  - Program

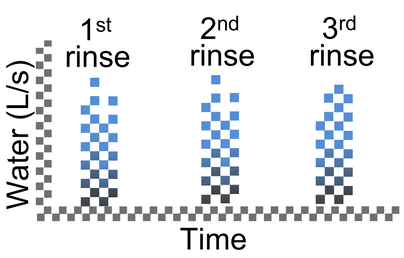
OBSERVATIONS*
- Laundry load weight before/after.
- Dryer tank water quantity

OUTPUT
Dataset over 800 time series records of washing machines and dryers

SOFT SENSOR DEVELOPMENT

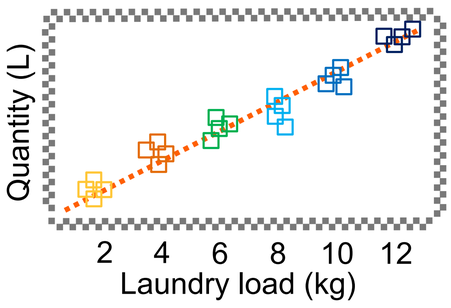
WATER BASED SOFT SENSOR

Washing machine water inflow during the washing cycle



Extracting 1st rinse quantity

Relationship between 1st rinse water and laundry load



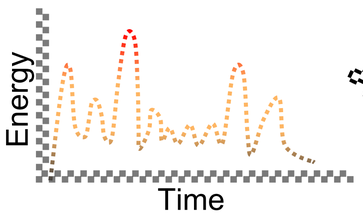
Input: 1st rinse water quantity

Training: AdaBoost, KNN – Regression, SVR, XGBoost

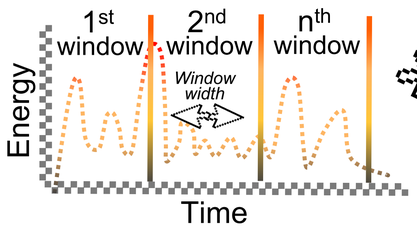
Output: Laundry Load (kg)

ENERGY BASED SOFT SENSOR

Instantaneous washing machine energy measurements during washing cycle



Selecting: window width, # of windows



Time series feature extraction

Temporal
Statistical
Spectral

Feature importance selection

Training: AdaBoost, KNN – Regression, SVR, XGBoost

Output: Laundry Load (kg)

APPLIANCE ASSESSMENT

KEY PERFORMANCE INDICATORS

Energy consumption (kw/h)

Laundry Load (kg)

Water consumption (L)

Energy per unit of laundry

Water per unit of laundry load

KEY PERFORMANCE INDICATORS

Energy consumption (kw/h)

Laundry Load (kg)

Energy per unit of laundry

A water KPI is traded for a convenience and **realistic installation**

A **smart plug** is enough to perform **real-life assessment**

RESULTS

REFERENCE: AN ERROR OF TWO MEDIUM-SIZED COTTON T-SHIRTS

SOFT SENSOR

	Laundry load (kg) MAE
AdaBoost	**310g**
KNN	320g
SVR	330g

SOFT SENSOR

	Laundry load (kg) MAE
AdaBoost	**467g**

The Energy-based prediction is slightly worse than water-based soft sensor. Nevertheless, method provides high enough accuracy and can easily be applied using a **smart plug in real-life conditions**.

EFFICIENCY RANKING PRESERVATION

In 10 washing machines with 260 samples, **80% could preserve correct efficiency prediction**.

It ensures that the majority of **recommendations align** with the **optimal energy** usage.

U LISBOA | Ciências ULisboa | FCT Fundação para a Ciência e a Tecnologia | LASIGE | SATO OF BUILDING ENERGY

This project receives funding in the European Commission's Horizon 2020 Research Programme under Grant Agreement Number 957128

Acknowledgements: This work was supported by FCT grant 2023.00280.BD, SATO H2020 project and the LASIGE Research Unit, ref. UID/00408/2025.

A Low-Cost Acoustic Bird Identification and Classification System for Biodiversity Monitoring

Diogo Anjos[1], Duarte Gonçalves[1], Wellington Oliveira[1], Fernando Ascensão[2], Sara Silva[1], José Cecílio[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] CE3C, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE
data and systems intelligence

1 Motivation

1. Birds are crucial indicators of ecosystem health and biodiversity.
2. Tracking bird populations helps assess environmental health.
3. Our work aims to automate this process, providing a low-cost, easily scalable solution.

2 Monitoring Station

- Records audio
- Computes spectrogram
- Tiny machine learning detects whether the audio has birds
- Forwards the bird audios

3 Processing Station

- Receives audio data flagged with bird presence
- Evaluates the data using advanced machine learning models
- Identifies the bird species and adds it to the population count

Summary

This project focuses on developing a robust, machine learning-based automated bird monitoring system that continuously records and analyses acoustic data using IoT devices.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · ce3c centre for ecology, evolution and environmental changes

Acknowledgements:
This work was supported by FCT through the LASIGE Research Unit, ref. UID/00408/2025.

False Alarms, Real Damage: Adversarial Attacks on CTI Systems

Samaneh Shafee, Alysson Bessani, Pedro M.Ferreira

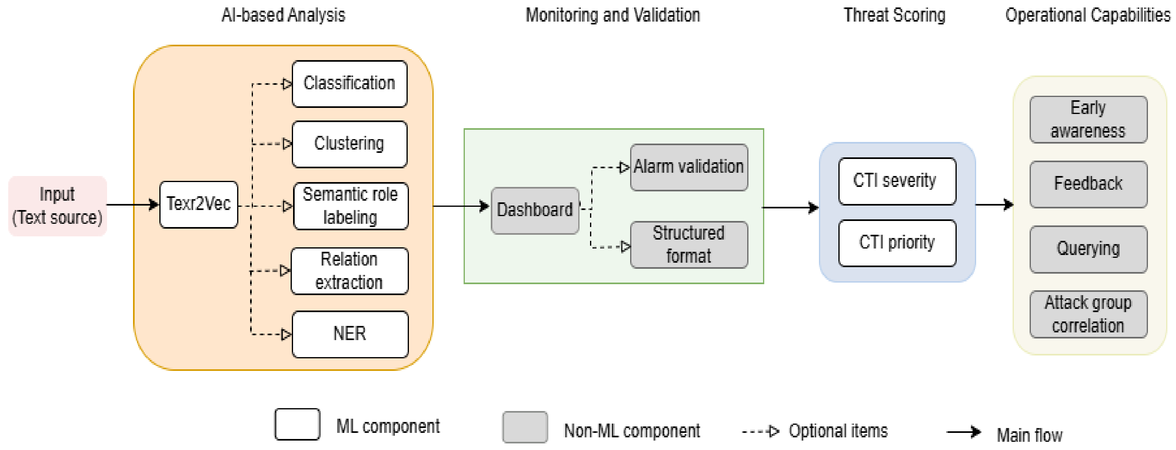
LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

Dependable and Secure Decentralized Systems

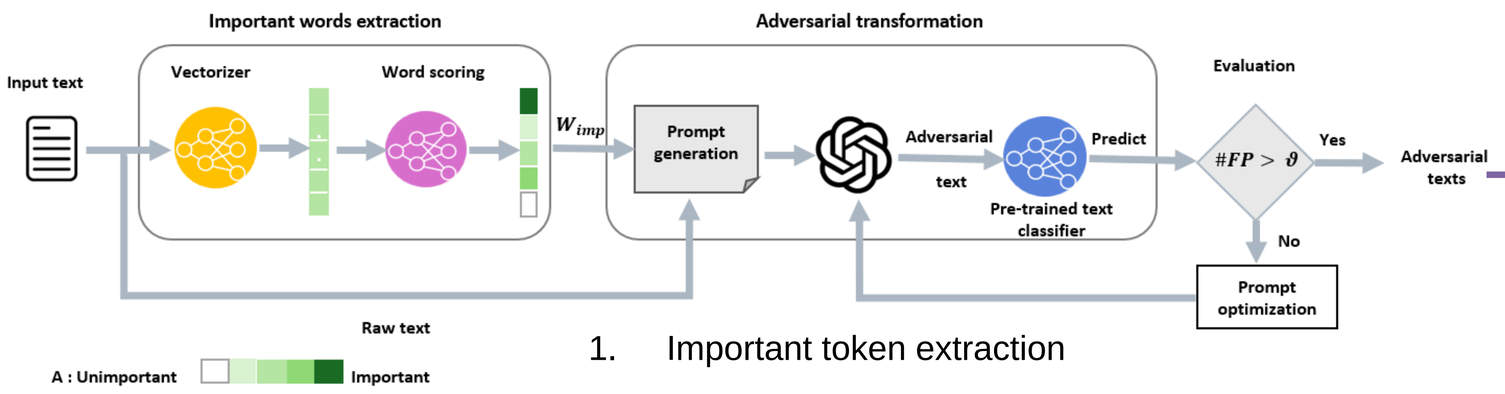
Motivation

- Using even untrusted OSINTs for extracting CTIs in proposed integrated CTI pipeline.
- Generating and applying fake TN text for adversarial evasion, poisoning and flooding attacks on integrated CTI pipeline.

Integrated CTI pipeline



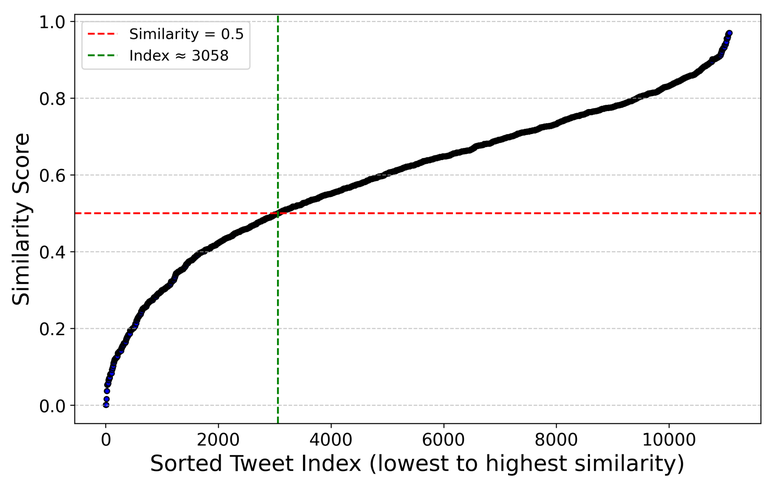
Method: generating fake TN texts



1. Important token extraction
2. Adversarial transformation
3. Human evaluation

Adversarial texts → Evaluation

Similarity scores result



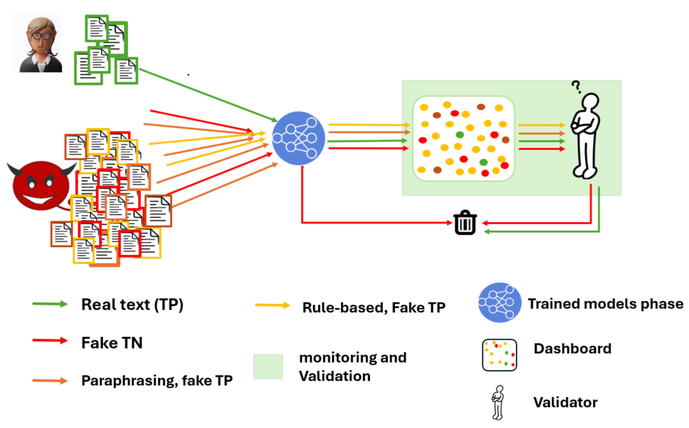
Attack Categorization

Evasion

- Applying fake TN texts on target models as input
- Deceiving target models

Flooding

- Data preparation including fake TP texts, TNs text.
- Bombarding Dashboard component in Monitoring and Validation stage of integrated CTI pipeline.



Poisoning

- Passing fake TN texts from Data validation component in the Monitoring and Validation stage of proposed integrated CTI system.
- Adding fake TN texts to training set and retrain target model.

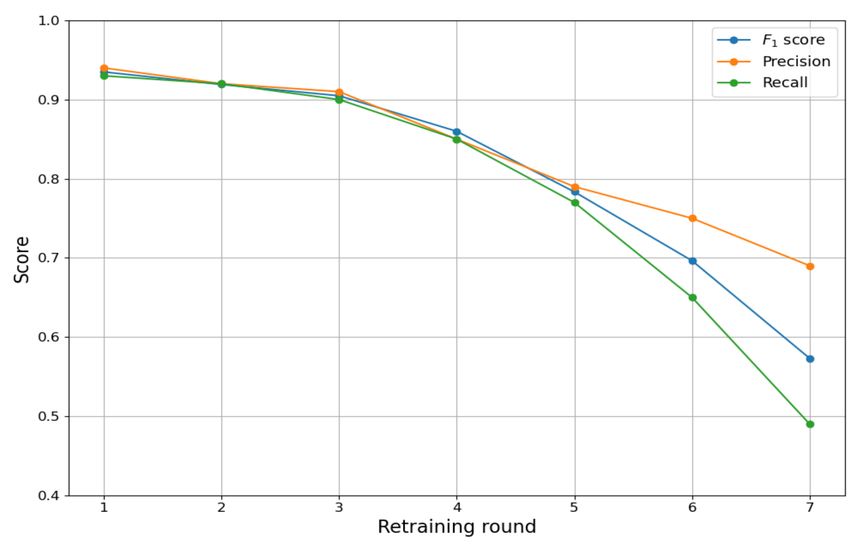
E

Models	Original text	Optimised text
ChatGPT-4o	0.51	0.75
Binary classifier	0.87	0.97

F

Text generation techniques	Fake Text Generator	Text	TP	FP	TN	FN	FPR	TPR
Adversarial texts generation	GPT-4o	11,074 (FaN)	0	9,402	332	0	0.97	0
Paraphrasing	GPT-3.5-Turbo	110,740 (FaP)	97,865	0	0	12875	0	0.88
Rule-based	semantic grouping	23,547 (FaP)	19,266	0	0	4,282	0	0.81

P



Conclusion

- Untrusted source-based CTI systems are vulnerable to evasion, flooding, and poisoning.
- Early-stage fact-checking is critical to prevent fake texts entrance to pipeline.

This work was partially supported by the European Commission through the SATO project,ref. H2020-IA-957128 (https://doi.org/10.3030/957128) and by the Fundação para a Ciência e Tecnologia (FCT) through the LASIGE Research Unit, ref. UIDB/00408/2025 and ref. UIDP/00408/2025

GDHelper Plus: Tool to Help the Diagnosis of Gaming Disorder for Multiple Games

Filipe Lopes[1], Ana Paula Afonso[1], Manuel J. Fonseca[1], Joana Cardoso[2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Center for Psychology at University of Porto, Porto, Portugal

LASIGE health and biomedical informatics

Motivation and Problem

- Video games are becoming more popular
- This raises mental health concerns, such as **Internet Gaming Disorder (IGD)**
- Self-report questionnaires provide inaccurate diagnoses because players' subjective responses rarely reflect their actual behavior

- **GDHelper** is a web tool that tracks gaming behavior over time using CS2 data, providing metrics like played hours, session frequency, and match count
- But addicted players may favor one genre but often play multiple types of games.

Objective

Develop **GDHelper Plus**:

- Extension of GDHelper covering players data from other video games, such as **League of Legends (LoL)**
- Evaluate the usability based on feedback from mental health professionals

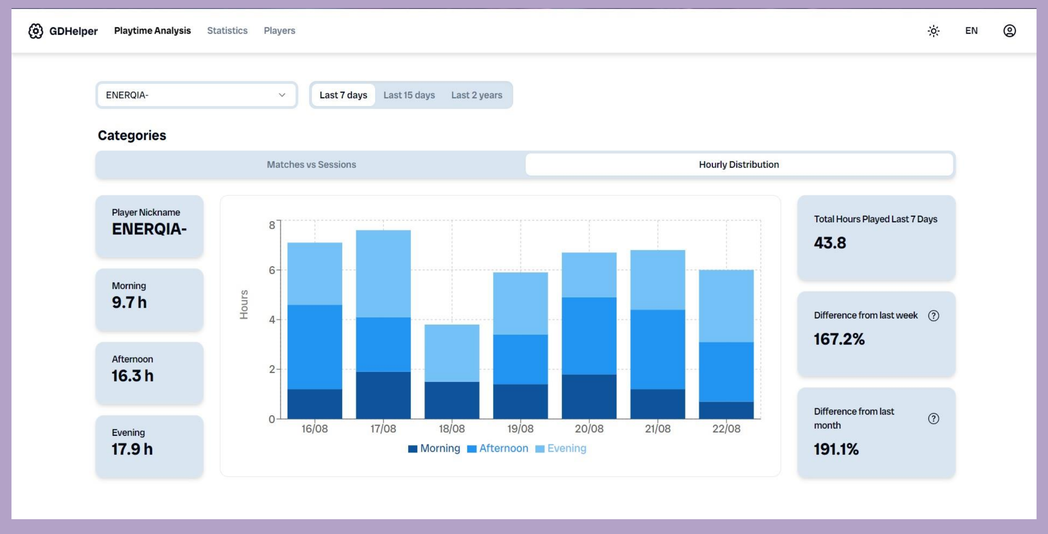
Expanding coverage across games allows professionals to monitor player progress more accurately

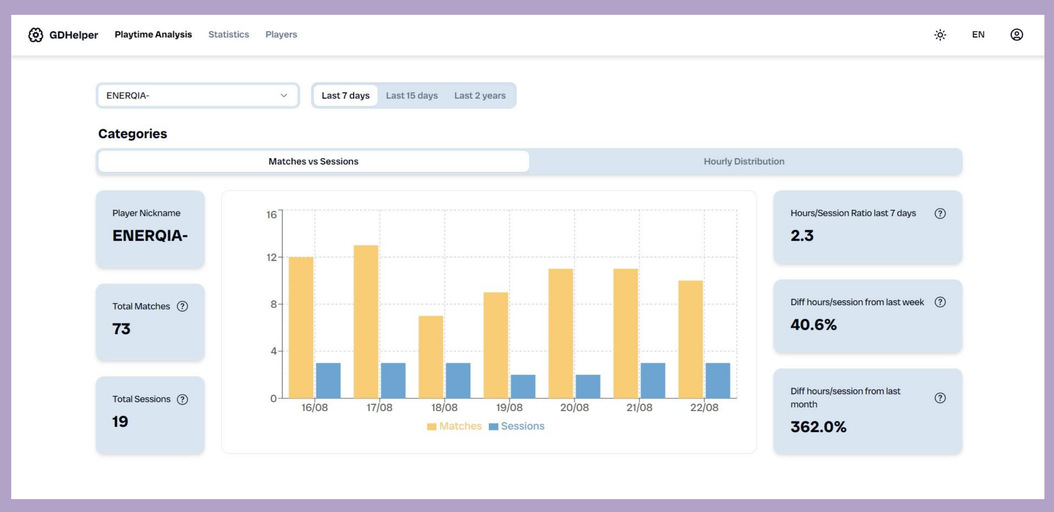
Data Collection

- Telemetry data is collected from League of Legends using **Riot API**
- Access to two endpoints to gather players data:
  - Accessing the **Account-V1 endpoint** provides player account details, including the account identifier
  - Accessing the **Match-V5 endpoint** retrieves match data, including the desired metrics

Visualization

The visualization of the data is made for all the games (CS2 and LoL):





- Metrics include **total matches, session frequency and playtime** over certain periods of time
- It helps identify patterns of excessive gaming
- The first page shows hours played in the last 7 days
- The second page shows matches and sessions from the last 7 days

Ongoing Steps

Iterative Design and Development of the application:

- GDHelper Plus development
- Data Collection
- User Feedback

Acknowledgements: This work was supported by FCT through project <title>, ref.\, <project reference> (*<project DOI>*), and the LASIGE Research Unit, ref. UID/00408/2025.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Development and Enhancement of the Portuguese Accessibility Monitoring Ecosystem

Filipe Egipto[1], Letícia Pereira[1], Carlos Duarte[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE inclusive human-computer interaction

What's the problem?

- Despite several guidelines, standards and directives, the web remains an inaccessible place.
- Moreover, there is a lack of tools that allow the monitoring of mobile applications accessibility.

Context

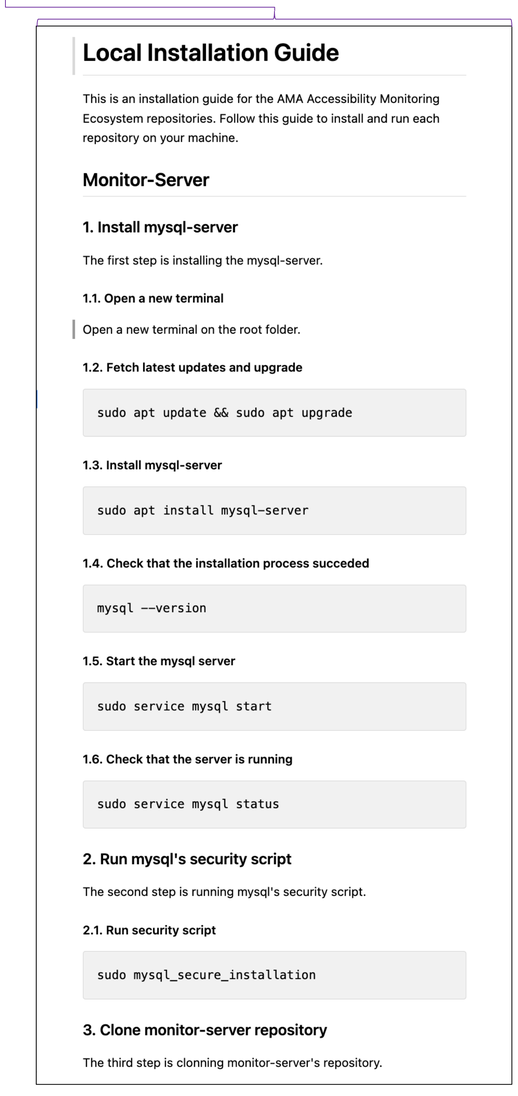
AMA and FCUL have a partnership that has existed for over a decade. This partnership made this thesis possible, for which AMA made nine requirements.

Who's AMA?

- In Portugal, the *Agência para a Modernização Administrativa* (AMA) is the entity responsible for promoting and advancing administrative modernization, offering multiple accessibility-related tools.
- AMA is also responsible for monitoring the portuguese state of compliance with the Web Accessibility Directive.

Requirements

- Create Guide to Replicate AMA's Environment (September 3rd to November 11th, 2024)
- Create Mobile Applications Observatory (November 21st, 2024 to January 23rd, 2025)
- Create AccessMonitor's Chrome Extension (January 24th to May 2nd, 2025)
- Create Documentation on Accessibility Problems (January 24th to May 2nd, 2025)
- Integrate Extension's Results in the Observatory's DB
- Register Manual Evaluations on the Observatory
- Elaborate Rankings Based on Manual and Automatic Evaluations
- Create Archives of Temporal Datasets from the Observatory's Data
- Allow the Visualization of the Archives of Temporal Datasets from the Observatory's Data



Local Installation Guide

This is an installation guide for the AMA Accessibility Monitoring Ecosystem repositories. Follow this guide to install and run each repository on your machine.

Monitor-Server

1. Install mysql-server

The first step is installing the mysql-server.

1.1. Open a new terminal

Open a new terminal on the root folder.

1.2. Fetch latest updates and upgrade

```
sudo apt update && sudo apt upgrade
```

1.3. Install mysql-server

```
sudo apt install mysql-server
```

1.4. Check that the installation process succeded

```
mysql --version
```

1.5. Start the mysql server

```
sudo service mysql start
```

1.6. Check that the server is running

```
sudo service mysql status
```

2. Run mysql's security script

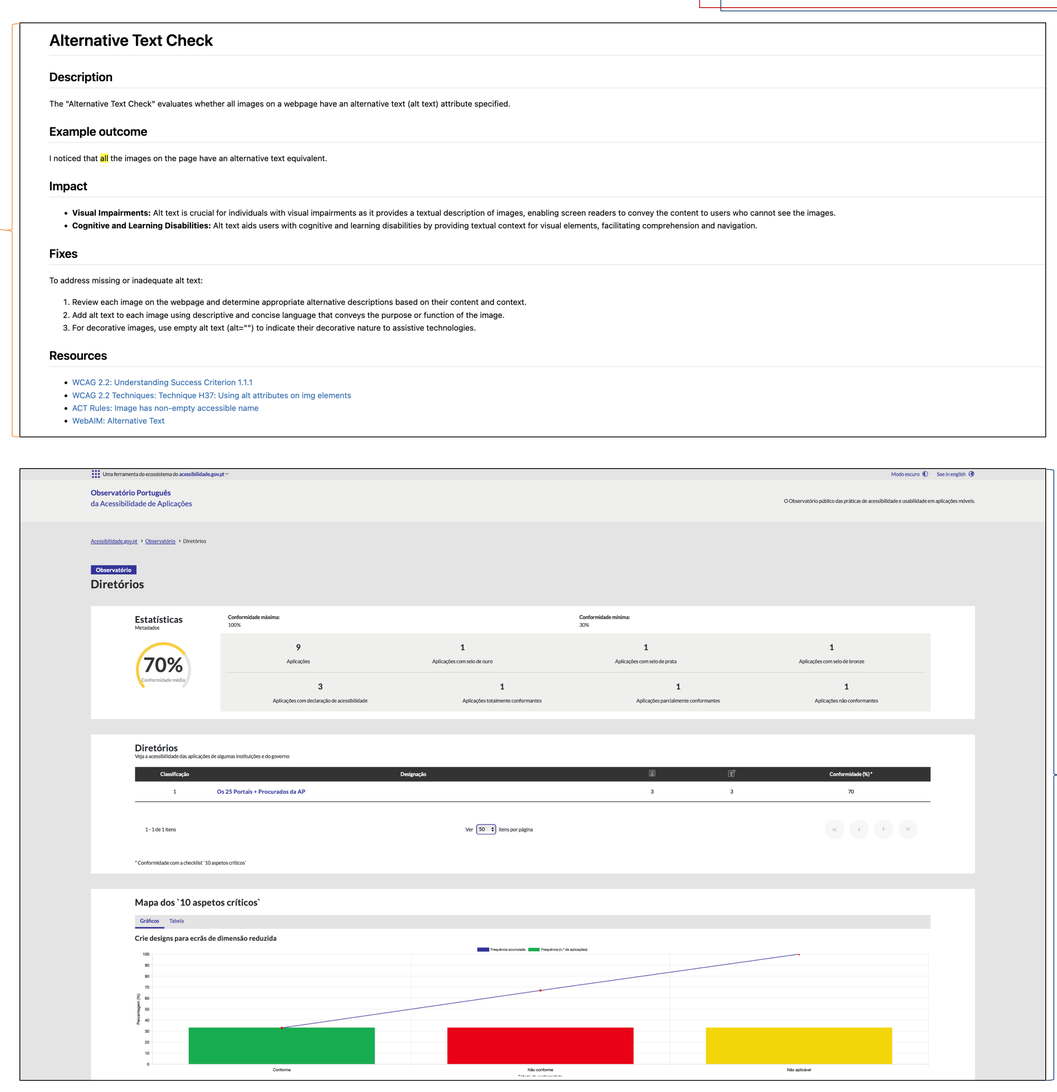
The second step is running mysql's security script.

2.1. Run security script

```
sudo mysql_secure_installation
```

3. Clone monitor-server repository

The third step is clonning monitor-server's repository.



Alternative Text Check

Description

The "Alternative Text Check" evaluates whether all images on a webpage have an alternative text (alt text) attribute specified.

Example outcome

I noticed that all the images on the page have an alternative text equivalent.

Impact

- **Visual Impairments:** Alt text is crucial for individuals with visual impairments as it provides a textual description of images, enabling screen readers to convey the content to users who cannot see the images.
- **Cognitive and Learning Disabilities:** Alt text aids users with cognitive and learning disabilities by providing textual context for visual elements, facilitating comprehension and navigation.

Fixes

To address missing or inadequate alt text:

1. Review each image on the webpage and determine appropriate alternative descriptions based on their content and context.
2. Add alt text to each image using descriptive and concise language that conveys the purpose or function of the image.
3. For decorative images, use empty alt text (alt="") to indicate their decorative nature to assistive technologies.

Resources

- WCAG 2.2: Understanding Success Criterion 1.1.1
- WCAG 2.2 Techniques: Technique H37: Using alt attributes on img elements
- ACT Rules: Image has non-empty accessible name
- WebAIM: Alternative Text

Uma ferramenta do ecossistema do acessibilidade.gov.pt

Modo escuro | See in english

Observatório Português da Acessibilidade de Aplicações

O Observatório público das práticas de acessibilidade e usabilidade em aplicações móveis.

Acessibilidade.gov.pt > Observatório > Diretórios

Observatório

Diretórios

Estatísticas
Metadados

70% Conformidade média

Conformidade máxima: 100%

Conformidade mínima: 30%

9 Aplicações	1 Aplicações com selo de ouro	1 Aplicações com selo de prata	1 Aplicações com selo de bronze
3 Aplicações com declaração de acessibilidade	1 Aplicações totalmente conformantes	1 Aplicações parcialmente conformantes	1 Aplicações não conformantes

Diretórios
Veja a acessibilidade das aplicações de algumas instituições e do governo

Classificação	Designação	A	P	Conformidade (%)*
1	Os 25 Portais + Procurados da AP	3	3	70

1 - 1 de 1 Itens

Ver 50 Itens por página

* Conformidade com a checklist '10 aspetos críticos'

Mapa dos `10 aspetos críticos`

Gráficos | Tabela

Crie designs para ecrãs de dimensão reduzida

Frequência acumulada | Frequência (n.º de aplicações)

Percentagem (%)

Conforme | Não conforme | Não aplicável

Access Monitor Extension

Relatório de práticas de acessibilidade Web (WCAG 2.1 do W3C)

Re-avaliar a página | Descarregar dados em CSV

Sumário

7.7 Pontuação

Título
A Investigação na Faculdade de Ciências | Faculdade de Ciências da Universidade de Lisboa

31 práticas encontradas

Tipo de prática	A	AA	AAA
Aceitáveis	9	10	0
Para ver manualmente	4	1	2
Não aceitáveis	4	1	0
Total	17	12	2

Acknowledgements: This work was supported by FCT through project <title>, ref.\ <project reference> (*<project DOI>*), and the LASIGE Research Unit, ref. UID/00408/2025.

ULISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

Secure Communication in FreeST

Martim Baptista, Andreia Mordido, Vasco T. Vasconcelos

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

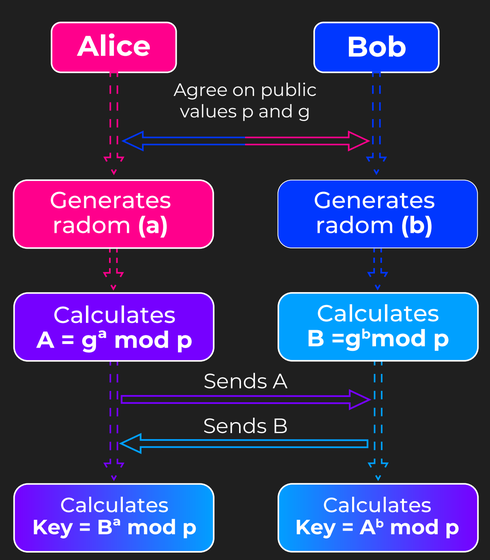
LASIGE reliable software systems

INTRODUCTION

FreeST is a strongly typed functional programming language with a focus on Session Types. The objective of this work is to bring secure communication to FreeST. To achieve this goal various problems need to be solved, such as key exchange, confidentiality and authenticity.

1. Establishing Keys

In order provide confidentiality to FreeST's messages we must encrypt them, and in order to encrypt data we need a shared key between the two sides of the communication. Solution? Diffie-Hellman Key Exchange Protocol, which allows two clients to establish a secrete shared key through an unprotected/public connection.



Here Alice and Bob manage to reach the same key only by sharing the values A and B. Note that it is not possible to find the key knowing only A, B, p and g, since the functions used to obtain A/B are one-way functions (known as the discrete logarithmic problem) and as such one can not find a/b through A/B.

2. Encrypting Messages

To keep messages private and thus provide confidentially to the connection, messages need to be encrypted and decrypted. Solution? Chacha20 Stream Cypher, which deterministically creates a bit stream that is XORed with our data resulting in the encrypted message, and similarly XORed with said encrypted message to obtain the original message.

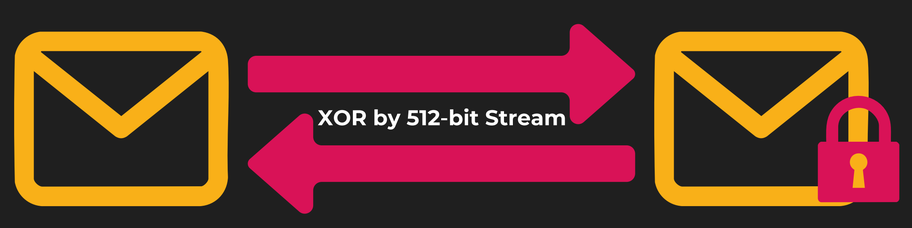
"expa"	"nd 3"	"n2-by"	"te k"
Key	Key	Key	Key
Key	Key	Key	Key
Counter	Nonce	Nonce	Nonce

Chacha20 starts by creating a 4 by 4 array of 32-bits, initially composed by a constant ("expand 32-byte k"), the given key, a counter and a Nonce.

It then applies a total of 20 rounds of shuffling which include sum, XOR and bitrotation operations.

1001...	1110...	1000...	1010...
1100...	1110...	1010...	1010...
1101...	1010...	0010...	0110...
1010...	1000...	1110...	1101...

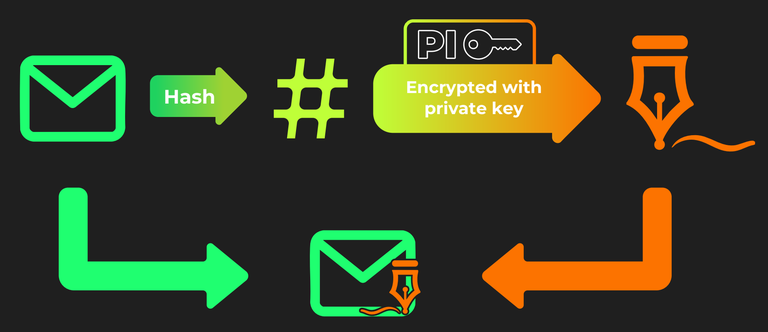
The resulting array is then serialized into a stream which can then applied with and XOR operation to the plain message in order to encrypt it or to the encrypted message in order to decrypt it.



3. Providing Authenticity

Diffie-Hellman by itself has a vulnerability, an attacker can intercept the messages sent by the clients when establishing a key and in turn establish a Diffie-Hellman connection with each client, without them noticing. Essentially the clients need a way to make sure that the messages that they receive during Diffie-Hellman are from who they expect, authenticity is required. Solution? Use RSA and Asymmetric Encryption to authenticate the shared values (A and B).

When using asymmetric keys, data can be encrypted with either the private or the public key and subsequently decrypted with the opposite key. Each approach has its own advantages, but in this case, we encrypt a hash of the shared value with the sender's private key (this is called a signature).



This assures the receiver that the value received was indeed sent by the sender whose public key is known. Through this public key the receiver can decrypt the signature and compare it to the hash of the received value.

4. Abstraction and Modular Design

These features were implemented in modules which are changeable by other equivalent algorithms. Using these modules and their features was made as simple as possible, abstracting their complexity.

```
type Session = !Int : ?Int : Close
alice : Session -> Int
alice c =
  let c = send 21 c in
  let (msg, c) = receive c in
  close c;
  msg
```

Besides having to establish the secure channel, the structure remains identical, only having to change the send, receive and close functions to their secure counter-parts.

```
type SecureSession = EstablishSecureChannel ; SecureSendInt ; SecureReceiveInt ; SecureClose
alice : SecureSession -> Int
alice c =
  let piAlice = getKeyFromFile "AliceKey.priv" in
  let puBob = getKeyFromFile "BobKey.pub" in
  let sc = establishSecureAuthenticatedChannelA piAlice puBob c in
  let sc = secureSendInt 21 sc in
  let (msg, sc) = secureReceiveInt sc in
  secureClose sc;
  msg
```

Additional Features:

To implement these algorithms many other features, had to be added to FreeST. Features such as Arbitrary precision Integers, Random Number Generators and seeding sources, among others.

Acknowledgements: This work was supported by FCT through project Reliable and Expressive Concurrent Systems through Advanced Session Types ref. 2023.13752.PEX (https://doi.org/10.54499/2023.13752.PEX), and the LASIGE Research Unit, ref. UIDB/00408/2025

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

A Resilient Framework for Automated Monitoring and Security Verification for Cloud Environments

Bruno Lima, Wellington Oliveira, José Cecílio

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

MOTIVATION

Microservices and containerized architectures have fundamentally transformed the design and deployment of scalable systems.

However, these modern paradigms introduce new challenges in **security**, **resource management**, and **energy efficiency**.

OBJECTIVES

This study aims to strengthen the **resilience** and **efficiency** of containerized microservices by:

- Integrating **AI-driven** security protocols
- Applying energy-aware **optimization**
- Delivering **real-time operational insights** for informed resource decisions

SYSTEM ARCHITECTURE

The solution is **system agnostic** and may be easily **adapted to different contexts**

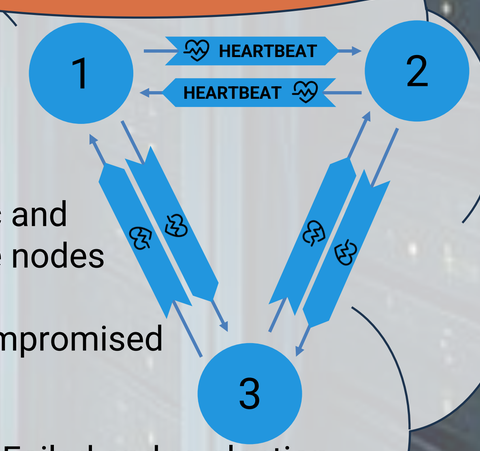
Nodes are **independently and distributed**

Nodes **cooperate** to find the most suitable **Leader**

One of the nodes will be automatically selected as **Leader**, using context-aware metrics

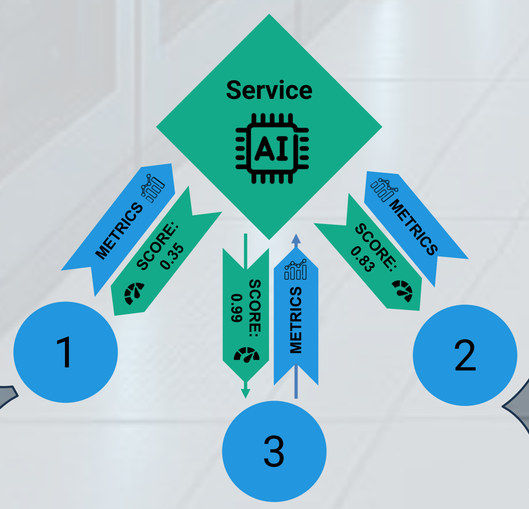
The **LEADER** will coordinate the nodes to manage security

Heartbeat Protocol



- It will sync and connect the nodes
- Detect compromised nodes
- Master Failed and re-election triggered

AI Service



The **AI SERVICE** collects metrics and provides a confidence score for each node

CONCLUSIONS

Microservice architectures provide scalability and flexibility but also pose growing challenges in security, resource management, and operational efficiency. These complexities intensify as adoption increases.

This study contributes to the ongoing discourse by taking a ***holistic approach***, addressing multiple critical aspects of microservice security and performance, rather than treating them in isolation.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE driven by excellence

Acknowledgements: This work was supported by the LASIGE Research Unit, ref. UID/000408/2025.

Exploring Grammatical Constraints with Large Language Models

David Silva[1,2], Vânia Mendonça[1], João Veríssimo[2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] CLUL, Faculdade de Letras, Universidade de Lisboa, Portugal

LASIGE data and systems intelligence

Motivation

- **How do we acquire language?**
- Are there innate grammatical constraints that direct this cognitive process?
- Transformer-based Large Language Models (LLMs) present new ways to explore such questions.
- They test the limit of data-driven learning, allowing us to revisit discussions focused on the Argument for the Poverty of the Stimulus (APS).

Plurals-in-Compounds

- A well-known morphological constraint previously unexplored in this context and used as a case study.
- English speakers tend to avoid regular plurals in compound words while accepting irregular plurals:

  rat-catcher vs **rats-catcher** (regular)

  duck-feeder vs **ducks-feeder** (regular)

  mouse-catcher vs **mice-catcher** (irregular)

  goose-feeder vs **geese-feeder** (irregular)
- It has been argued that this pattern reveals innate grammatical constraints.

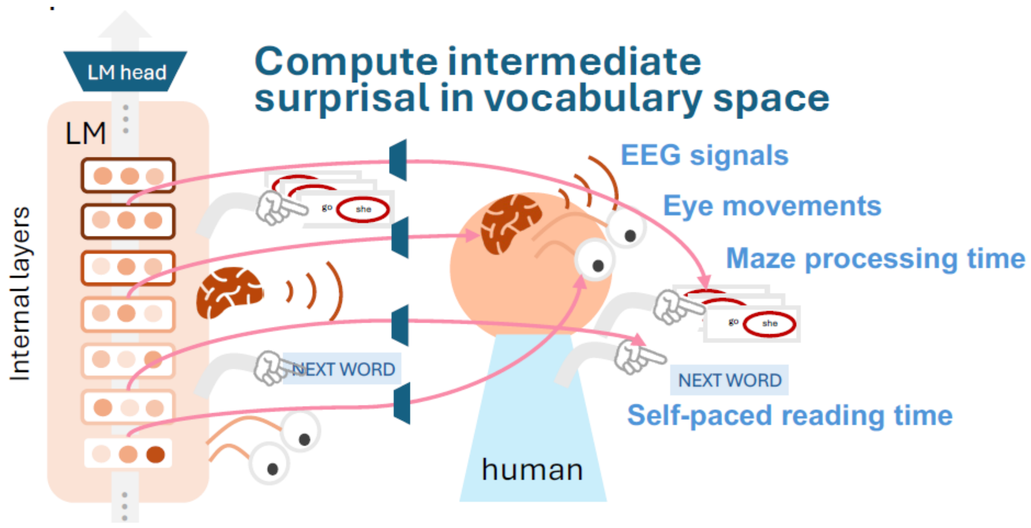
Research Questions

Comparison with Human Linguistic Knowledge: Does the behavior of a model align with the results of human judgements?

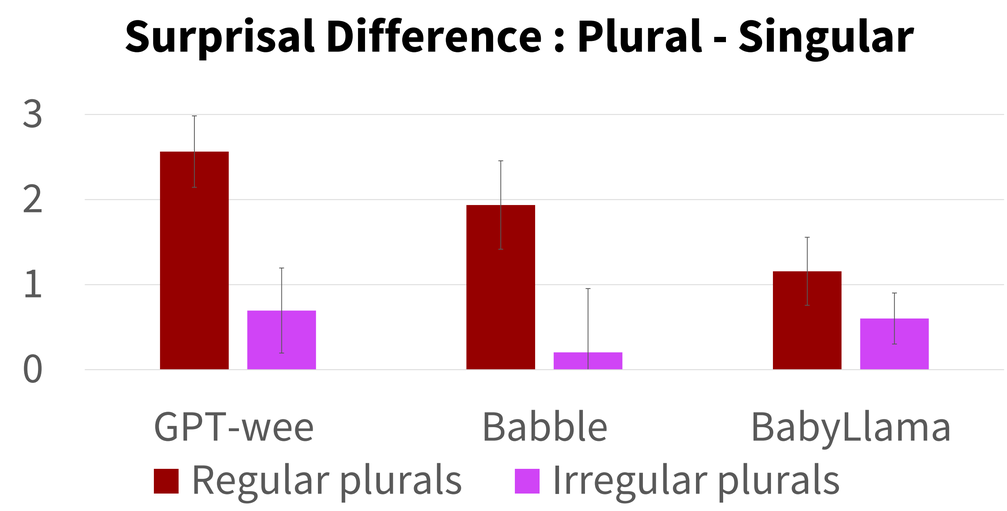
Nativism vs. Empiricism and APS: Can LLMs learn such constraints through their data-driven processes alone?

Methodology

- **Models trained on plausible amounts of data**: BabyLlama, GPT-wee and Babble, amongst others.
- **Linguistic stimuli:** examples of plurals in compounds matched for meaning, length and frequency.
- **Metrics:** Surprisal.



Preliminary Results



Preliminary results show an acceptability pattern aligned with human behavior.

Future Research

- *that-trace* effect
- Syntactic Islands

Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UID/00408/2025 and Center of Linguistics of the University of Lisbon, ref. UID/00214.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · FLUL FACULDADE DE LETRAS

Building a permissionless blockchain based on proof-of-space for archiving the world wide web

Bruno Cotrim[1], Miguel Matos[2], Bernardo Ferreira[1]

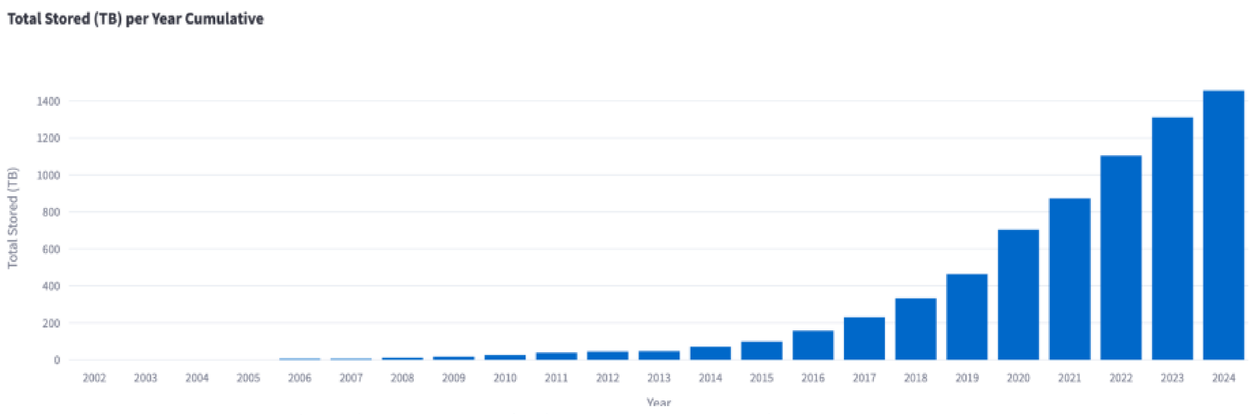
[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] INESC-ID , Instituto Superior Técnico, Universidade de Lisboa, Portugal

LASIGE dependable and secure decentralized systems

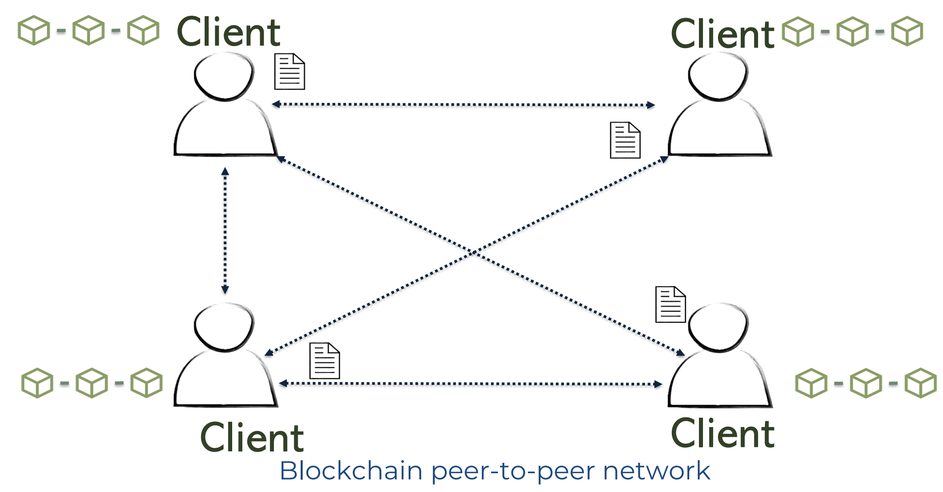
1 Context

The rapid growth of digital content has significantly increased the storage and operational costs for web archives like Arquivo.pt, managing 1 PB of data across 77 servers with 5.18 PB of storage. This creates an **urgent need for a scalable solution** that **reduces this costs**.

Permissionless blockchains enable decentralized storage, incentivizing participants with cryptocurrency rewards to reduce costs and enhance scalability for Arquivo.pt.



Arquivo.pt cumulative storage over the years

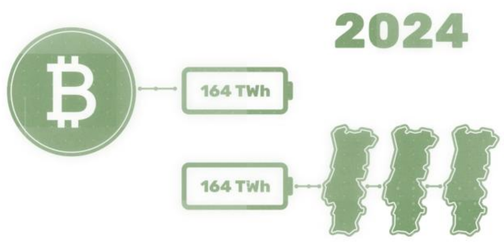


Blockchain peer-to-peer network

2 Problem

In current blockchain solutions, **resources expended** by users to prevent Sybil Attacks, are **used inefficiently**:

Bitcoin's [1] Proof-of-Work (PoW) requires **massive amounts of energy**



Chia's [2] Proof-of-Space(PoS) **fills disk space with useless random data** just to maintain the blockchain.

3 Challenges

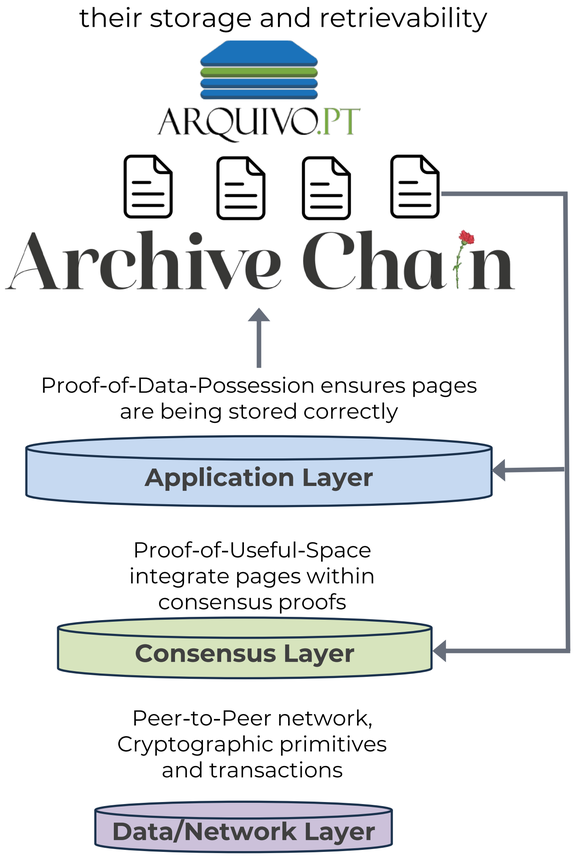
Creating a blockchain network where the consensus relies on **resource efficient primitives**?

Scale the blockchain into a fully distributed storage able to securely replicate Arquivo.pt files, with **file storage** and **integrity guarantees**

Incentive participants to commit storage space to the network

4 Solution

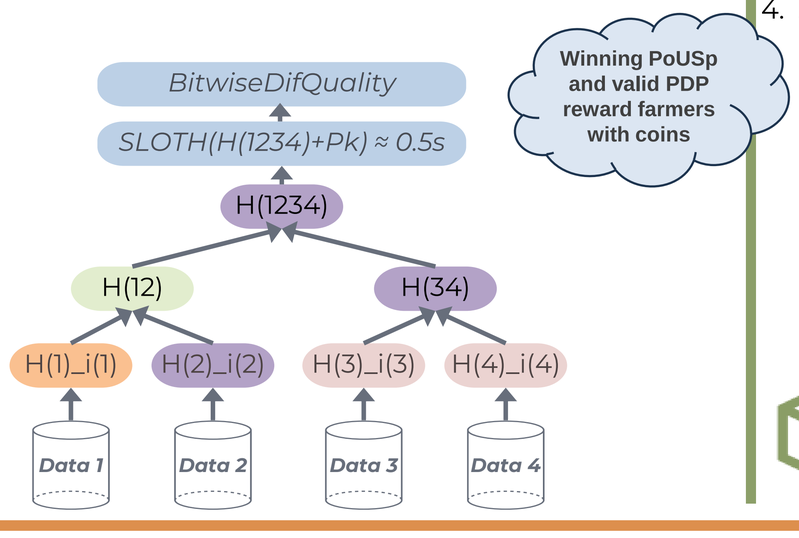
Use **Arquivo.pt pages** as non-forgeable tokens in proof of useful space consensus while ensuring their storage and retrievability



Proof-of-Useful-Space (PoUSp): Adapting Chia Bread Pudding[3], PoUSp integrates Arquivo.pt data into consensus proofs transforming **storage into a useful resource**:

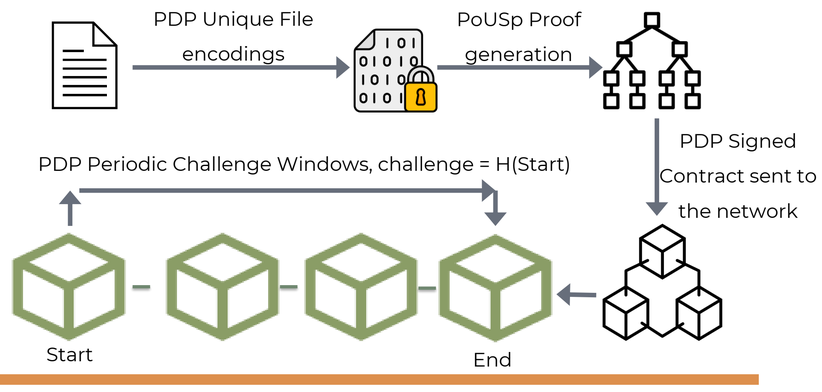
1. **Arquivo.pt data** is encoded into Merkle Trees
2. Slow Timed Hash Function (SLOTH) is computed on the Merkle Root, ensuring a **time-intensive generation** that deters Hellman Attacks
3. This **incentives storage of proofs** rather than computation on-the-fly

Achieves at least 89% useful space ratio, since only data and SLOTH proofs are stored



Proof-of-Data-Possession(PDP): Complements PoUSp by providing it with verifiable storage and integrity:

1. **Files are encoded** with AES-256/verifiable delay encoder (VDE), using unique salts (Hash(farmerPk + fileId)) and encoded into Merkle Trees
2. FCCN issues signed storage contracts with Merkle Roots for storage proofs, and metadata for reward and integrity verification
3. **Periodic challenge windows** defined by block heights, require farmers to generate Merkle Proofs for the file.
4. All this happens automatically in the application layer via Smart Contracts, preventing outsourcing attacks and ensuring verifiable storage.



5 Results

ArchiveChain is now live! Visit! **https://archivechain.pt**

References:
[1] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. Satoshi Nakamoto, 2008

[2] Bram Cohen and Krzysztof Pietrzak. The chia network blockchain. White Paper, Chia. net, 9, 2019

[3] Monica Chen Jin. Chia bread pudding: A blockchain of useful storage. Master of science thesis, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal, December 2024.

Acknowledgements: This work was supported by FCT through project APOSTLE , ref. 2023.12254.PEX(https://doi.org/10.54499/2023.12254.PEX), and the LASIGE Research Unit, ref. UID/00408/2025

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Sharing and Accessing Autobiographical Memories Tied with Media

Francisco Ludovico, Teresa Chambel

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE health and biomedical informatics

Motivation

Autobiographical Memory

vital **cognitive** tool;
crucial role for **identity**
sense of self and **continuity** time across
+ social bonding
sharing personal narratives.

Movies, Music & other Media

always present in our lives
among most relevant and **emotional**
forms of entertainment and education.

we can easily make **connections**
between media & memorable
moments & **places** in our lives
that tend to be **shared** with others.

Background

Autobiographical Memory
Collective Memory
Personal Computing
Emotions

Related Work
media platforms
personal and social media
Family Stories, EmoJar
As Music Goes By…

User Survey

Demographic Information
diverse background and age, Portuguese.
questionnaire (N=45) + Interviews (N=8)

Accessing & sharing media content
whatsapp. messenger, social net, music,
movies, videos… to relax, motivate, share

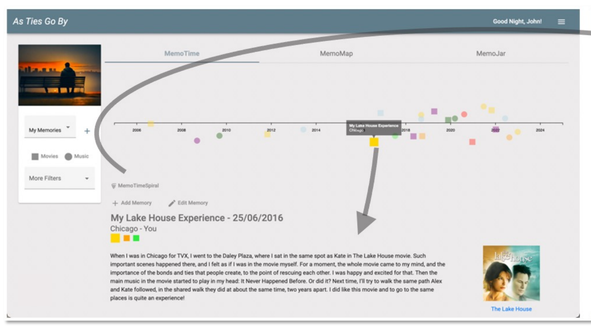
Memory register and sharing
(very) important (79%); relate to media (70%)

Time perception
events (66%), timeline (36%), cyclic (17%)

Platforms connect' memo to media
(very) interested (92%); media, time, emo
personal journal, share, filters, notifications…

*"… involves **retaining** and **retrieving** specific **events**, **emotions**, and **contextual** details **tied** to **personal experiences**"* Byrne, 2017

memotime

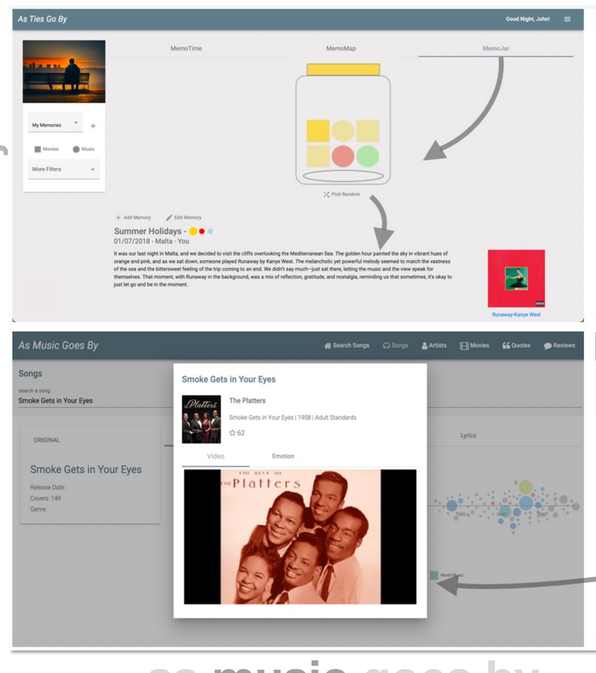


memory & movie

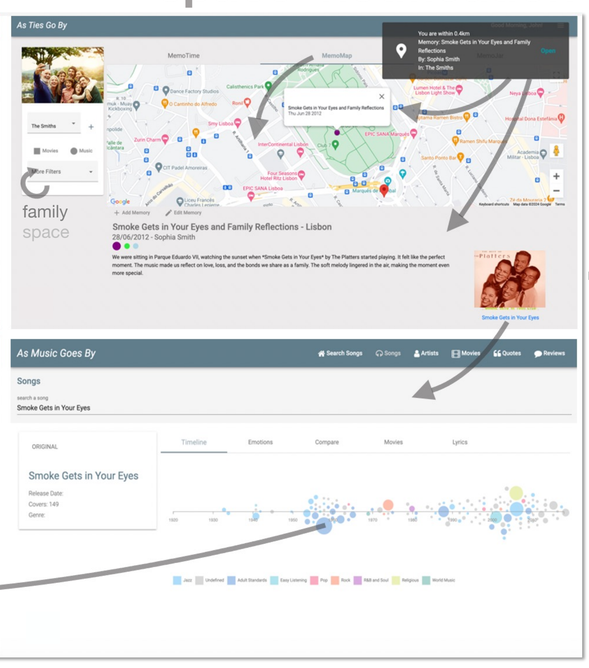
timeSpiral

space notification

memojar



as music goes by …



memomap

As Ties Go By

Memory Spaces

my memories,
Smith **family**, **friends**, etc.

Memory Views

MemoTime & **TimeSpiral**
MemoMap
MemoJar

Memories

text + **music** & **movies**
linking to As Music Goes By

Filters

by media & hashtag

Notifications

by time & space …& emo

User Evaluation

11 tasks | features , Interviews (N=10)

USE mean (1-5): Usefulness: **4.6**
Satisfaction: **4.3**, Ease of use: **4.7**

Overall: **preferred** features are specific to As Ties Go By: **MemoTime**, **MemoMap**, **time** and **location**-based **Notifications**, and **conutextalized Access** to **Music** & **Movies** from memories.

Most frequent terms: **Original** *(H)*
Understandable *(E)* **Simple** *(E)*
Interesting *(H)* **Sympathetic** *(A)*
(H)edonic; (E)rgonomic; (A)ppeal

Future Work

Refining based on user feedback
Other representations of **time**
Dedicated **mobile app**

User generated media content
Larger scale media platforms

Richer user **expression**; media **linking trajectories** in maps; group perspectives and **storytelling**

Keeping privacy & significant ties in larger scale.

© book by X. González & A.M. Garcia

the paper

Acknowledgements: This work was partially supported by FCT through the AWESOME project, ref. PTDC/CCIINF/29234/2017 and the LASIGE Research Unit, ref. UID/00408/2025.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Real-Time Accessibility Dashboards for Higher Education

Beatriz Patatas, Carlos Duarte, Letícia Seixas Pereira

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE inclusive human-computer interaction

Context

LMS platforms host educational content in various formats, including files (e.g., PDFs) and content embedded directly into course pages (e.g., text, images, videos).

LMS (Learning Management System) are platforms used to manage and deliver online learning content and assessments

Main challenges in PDF Accessibility

- Lack of proper structure;
- Images without alternative descriptions;

Many accessibility issues arise not only from uploaded documents but also from the way **content is structured** within course pages.

Problem

Teachers are often unaware that the content they create may not be accessible

Institutions face challenges in effectively monitoring digital accessibility

Proposal

Development of a **real-time**, interactive **dashboard** integrated with **Moodle** and tested on **FCUL's Moodle** instance to support both professors and institutions in evaluating digital accessibility.

An interactive panel that presents information visually and intuitively.

- Centralizes key data about digital accessibility issues in one location.
- Facilitates analysis and decision-making with charts and metrics.
- Enables real-time monitoring of accessibility status.

Features

For Professors

- Real-time content evaluation
- Alerts for non-compliant content
- Accessibility reports by course
- Guidance to fix accessibility issues

For Institutions

- Detailed accessibility reporting
- Continuous monitoring of accessibility scores
- Filters by course or department
- Set accessibility goals

Flow Diagram

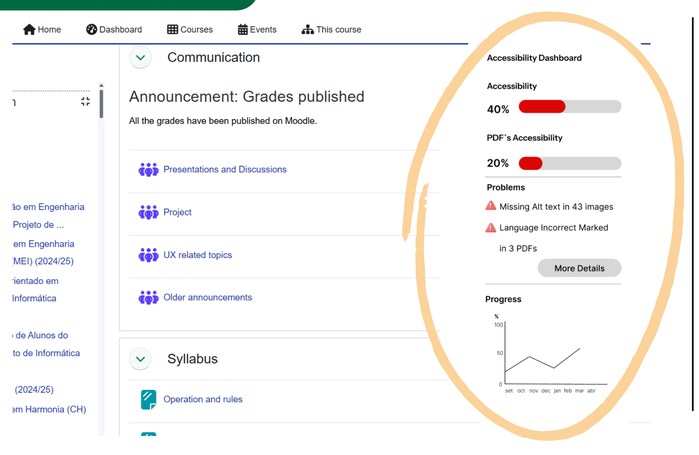
1. **PDF Upload to Moodle** — The professor uploads a PDF file to the Moodle course page.
2. **Automated Accessibility Analysis** — The system checks PDF accessibility and generates a list of accessibility issues with guidance for correction.
3. **Real-Time Dashboard Update** — The dashboard is updated with accessibility data from the course page and the uploaded PDF, showing common issues and progress over time.
4. **Monitoring and Feedback** — Professors receive actionable feedback. Institutions monitor accessibility trends by course or department.

Check PDF Accessibility

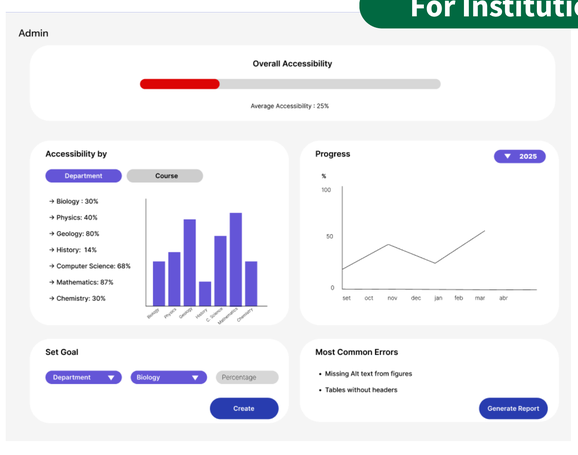
- Valid Links
- Marked Lists
- Compare Languages
- Figures with alt text
- Tables with headers
- OCR on a scanned PDF
- Verify Title

Dashboard Mockup

For Professors



For Institutions



Acknowledgements: This work was supported by FCT through the LASIGE Research Unit, ref. UID/000408/2025.

Classification and scoring approach for IoT threats

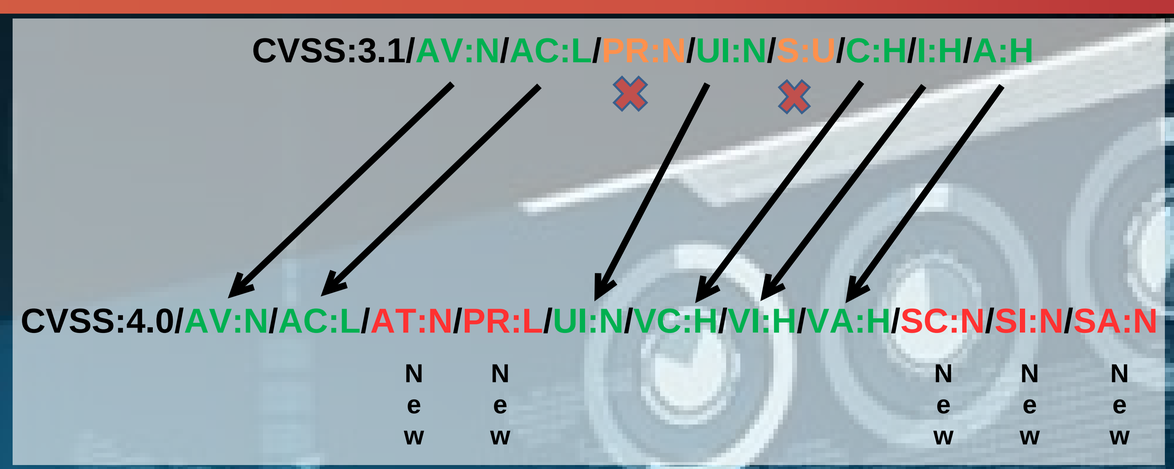
Nuno Vieira, André Souto, José Cecílio

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

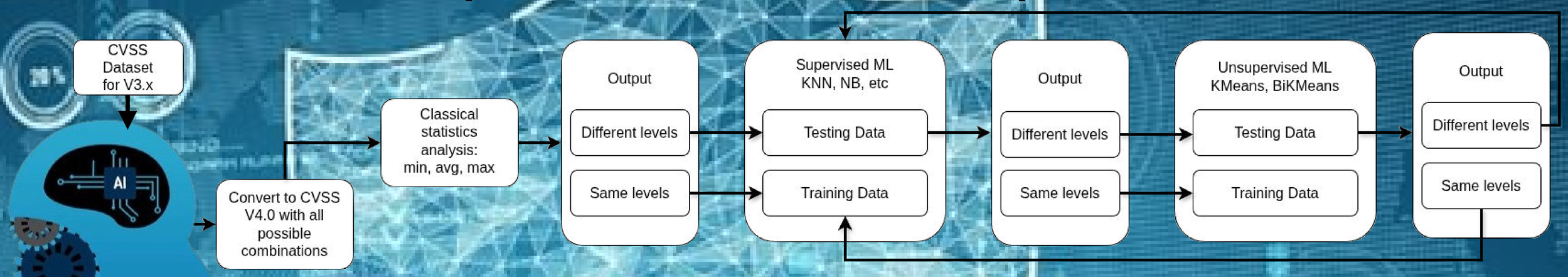
LASIGE cyber-physical systems

The Common Vulnerability Scoring System (CVSS) is evolving from version 3.x to version 4.0.

Current systems that track existing attacks and their severity do not account for the new version.

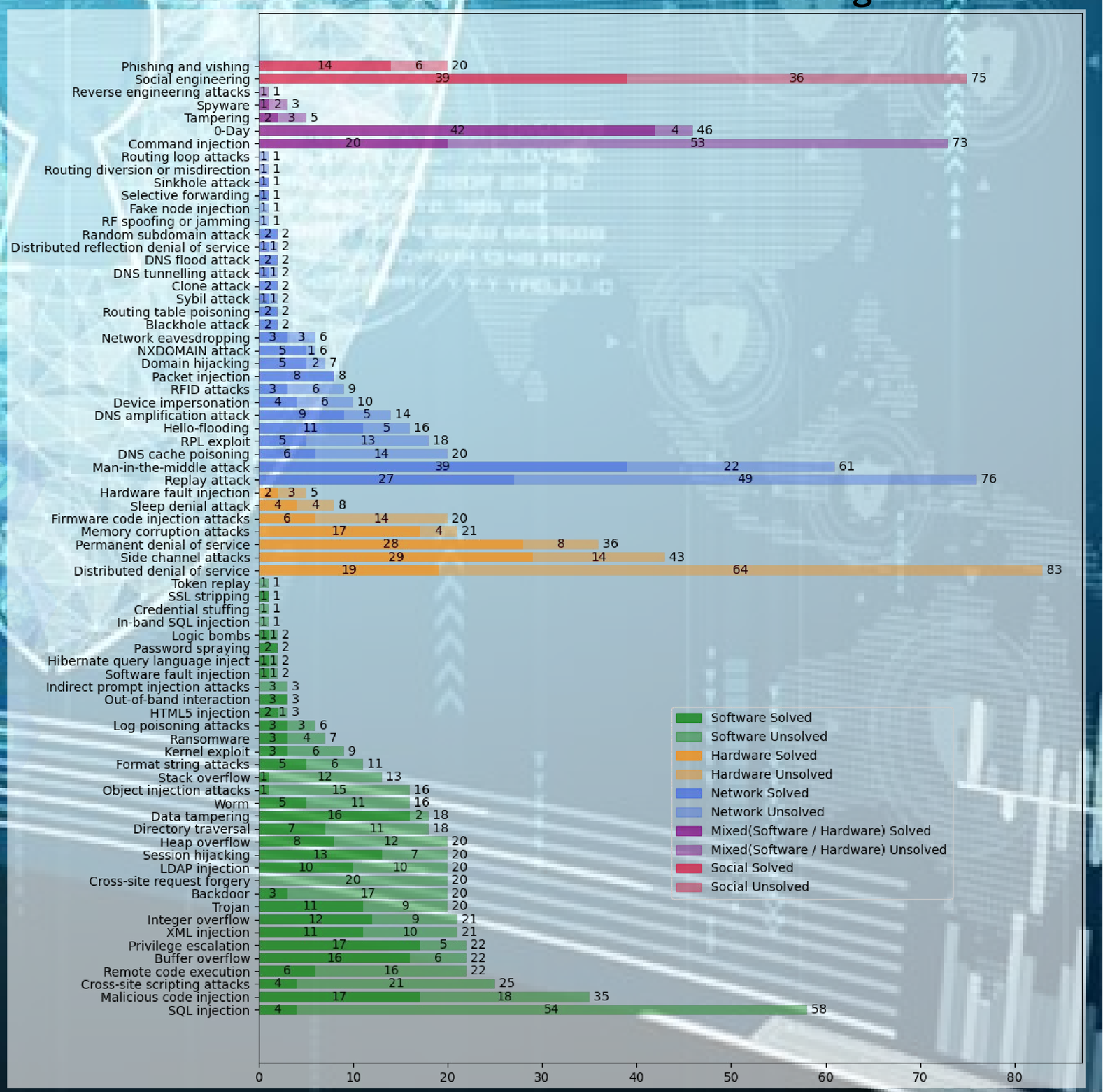


Proposed conversion process



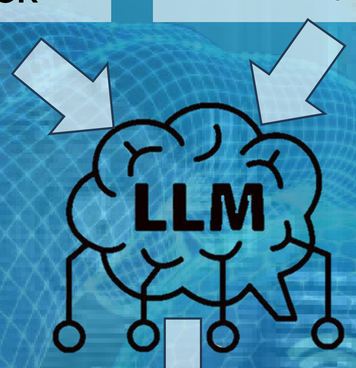
Results:
- 98.90% w/cross-validation

Considered attack categories



Description of the attack

IoT scenarios in which the devices are deployed



Get an adjusted severity score for the vulnerability

Preliminary Results:
- 87,81% accuracy

Recommendations on how to improve Device's security

The approach enables a user to better prioritize patching vulnerabilities when dealing with multiple attacks on the same network of devices.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE

Acknowedgements: This work was supported by the LASIGE Research Unit, ref. UID/00408/2025.

Deep Learning-based Digital Twin of Manufacturing Systems

Nessrine FARHAT[1,2], Pierre DAVID[2], Nuno CRUZ GARCIA[1], Leah RIFI [2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] G-SCOP, Institute of Engineering Univ. Grenoble Alpes, Grenoble, France

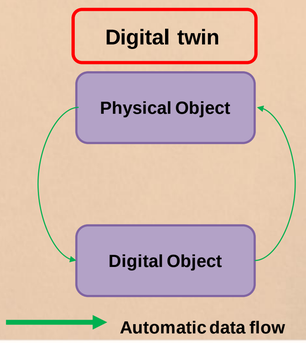
LASIGE data and systems intelligence

Context & Problematic

Context: The rise of **Industry 4.0** drives the need for efficient, data-driven production using **Digital Twins** to optimize real-time operations.

Challenge: **Automatically** and **accurately** reflecting the complexities of real-world systems.

Solution: **A Digital twin** powered by **computer vision** to monitor, analyze, and optimize production **in real-time.**



Case study

The **Operations Management (OM) platform** at G-SCOP Lab simulates industrial environments to optimize production processes in Industry 4.0. It serves as the foundation for our case study on **Tidal Turbine Production**, where we focus on the assembly and disassembly processes.

Tidal Turbine Assembly

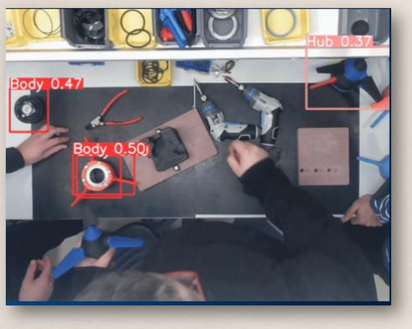
Tidal Turbine Disassembly

Sub-Assembly of the Tidal Turbine components

Preliminary Results

In the initial phase, we focused on an **object detection** task to recognize the parts of the tidal turbine using the **YOLOv8** model. We tested two approaches: **real data** and **synthetic data**.

Test of the model trained on real data



- Low precision
- Time-consuming and resource-intensive
- Privacy concerns

Test of the model trained on synthetic data

- Data gap
- Poor performance on real data
- Need for better data fidelity

Methodology

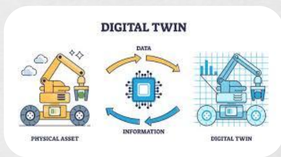
1 Data Collection

- Spontaneous real data
- Controlled real data
- Synthetic data

2 Development of CV Techniques

- Object detection
- Object tracking
- Action recognition
- Anomaly detection
- Object counting

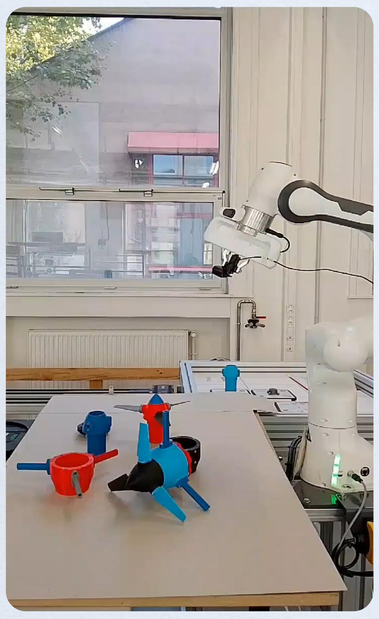
3 Integration of DT with the Model



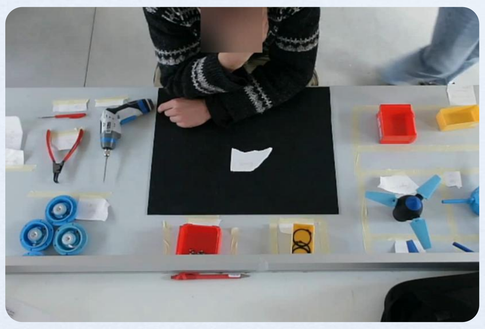
- Production Order Tracking
- Work in Progress Monitoring
- Operating Times

New Dataset

Controlled_Real_Data

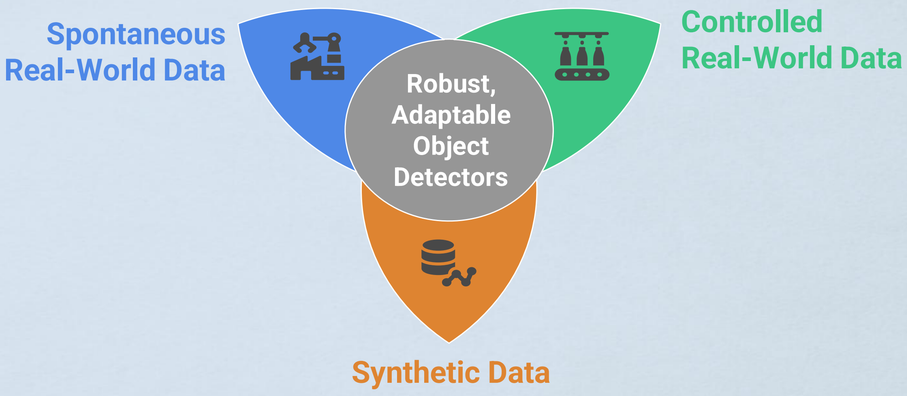


Spontaneous_Real_Data



Synthetic_Data

Future Perspectives



To enhance the robustness and generalizability of object, we propose a mixed approach that leverages a hybrid dataset composed of spontaneous real-world data, controlled real-world data, and synthetic data. This strategy aims to overcome limitations related to data scarcity, annotation cost, and variability in manufacturing environments.

Partners

This project is the result of a successful application to the UNITE!* PhD Fellowship Call between the Faculty of Sciences at the University of Lisbon and Grenoble INP of Grenoble Alpes University

LASIGE

unite! University Network for Innovation, Technology and Engineering

G-SCOP SCIENCES POUR LA CONCEPTION, L'OPTIMISATION ET LA PRODUCTION

Project Team

Nessrine FARHAT
PhD student

Nuno Garcia (LASIGE)
Thesis Co-director

Pierre DAVID (G-SCOP)
Thesis Director

Leah RIFI (G-SCOP)
Thesis co-supervisor

Keywords

- Digital Twin
- Industry 4.0
- Smart Manufacturing
- Machine Learning
- Computer Vision
- Simulation

Contacts

Nessrine.farhat@grenoble-inp.fr
Pierre.david@grenoble-inp.fr
ncgarcia@ciencias.ulisboa.pt
Leah.rifi@grenoble-inp.fr

Related Links

UNITE! * : https://www.unite-university.eu/about-us/unite-projects
G-SCOP[1] : https://g-scop.grenoble-inp.fr/
OM plateform[2]: https://g-scop.grenoble-inp.fr/fr/equipements/operations-management

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · UGA Université Grenoble Alpes · GRENOBLE INP UGA · unite! University Network for Innovation, Technology and Engineering · G-SCOP

Acknowledgements: This project was supported by the UNITE! PhD Fellowship Call, a collaborative initiative between the Faculty of Sciences at the University of Lisbon and Grenoble INP of Grenoble Alpes University, and by FCT through the LASIGE Research Unit, ref. UIDB/00408/2025 and ref. UIDP/00408/2025

Rewarding Explainability in Drug Repurposing with Knowledge Graphs

Susana Nunes[1,2], Samy Badreddine[2], Catia Pesquita[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Sony AI, Barcelona, Spain

Scan to check the paper!

Accepted at IJCAI 2025

LASIGE health and biomedical informatics

AI and Validating Scientific Explanations

Drug repurposing seeks new therapeutic uses for existing drugs. Knowledge graphs (KGs) can power hypothesis generation but often lack explainability. However, existing methods prioritize predictive accuracy, but not scientific insight.

Objective

Generate scientifically meaningful explanations for predictions over KGs using an RL-based framework **REx**.

Method

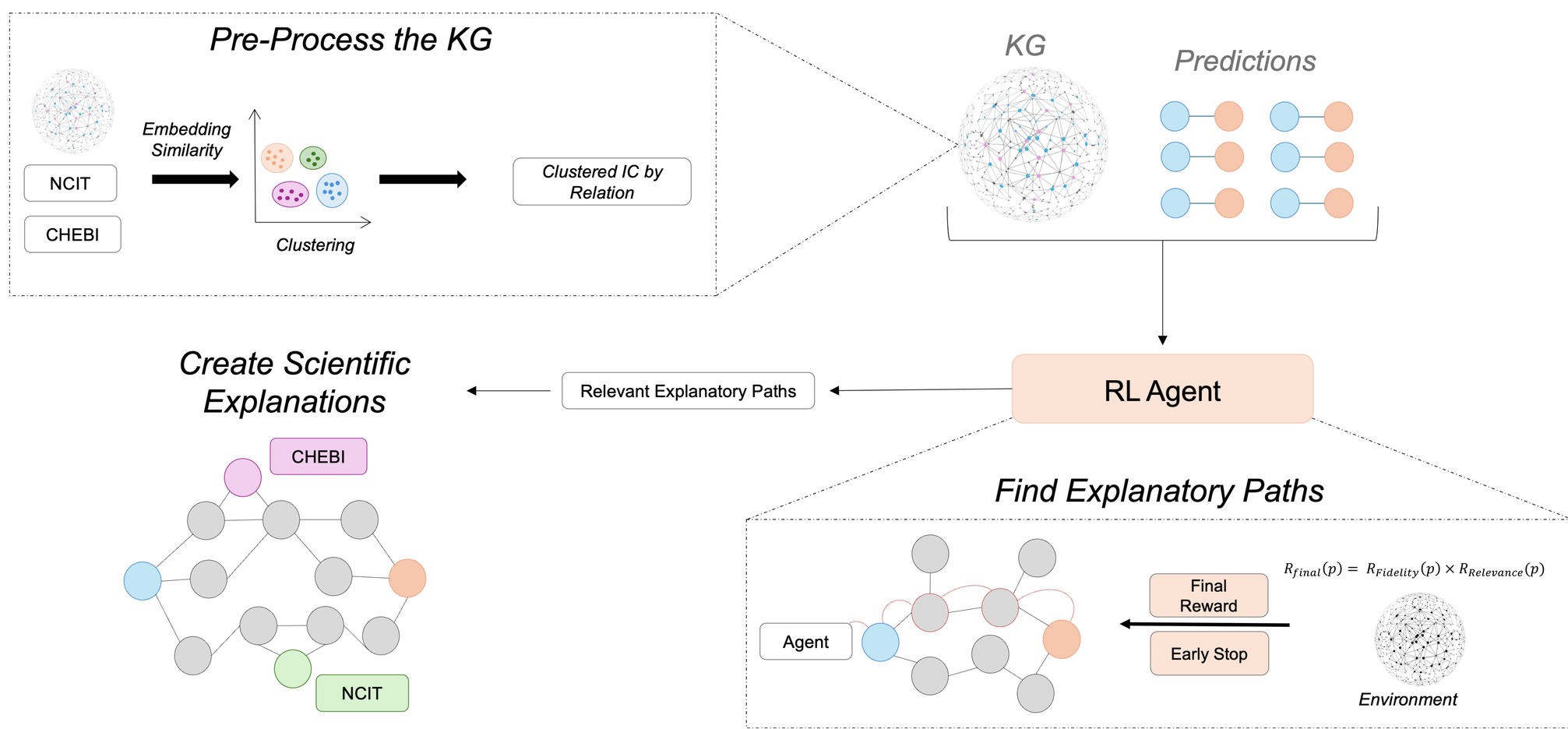
Rex finds explanatory paths in biomedical knowledge graphs. It employs reward and policy mechanisms that consider desirable properties of an scientific explanation to guide an RL agent in the identification of explanatory paths within a KG.

Fidelity

A scientific explanation should necessarily align with the hypothesis. Paths must connect the head and tail of the triple.

Relevance

A scientific explanation should provide detailed insights into the mechanisms underlying the hypothesis.

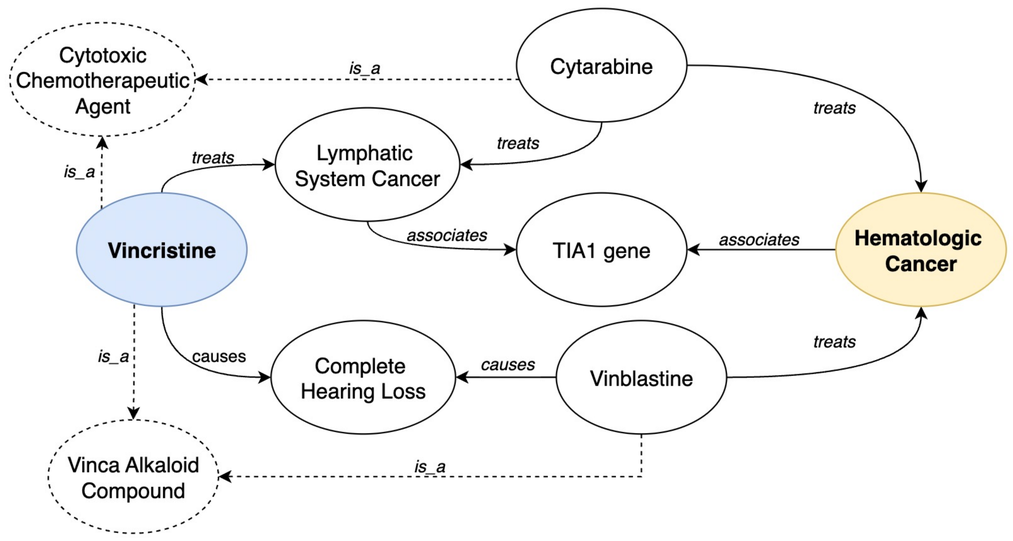


Results

Performance comparison for predictions based on Hits@1, Hits@3, Hits@10, and MRR metrics.

		Hits@1	Hits@3	Hits@10	MRR
Hetionet	Minerva	0.264	0.409	0.593	0.370
	PoLo	0.314	0.428	**0.609**	0.402
	REx	**0.338**	**0.461**	**0.609**	**0.427**
PrimeKG	Minerva	0.262	0.420	0.546	0.359
	PoLo	0.245	0.408	0.526	0.344
	REx	**0.286**	**0.429**	**0.544**	**0.376**
OREGANO	Minerva	0.133	0.200	0.489	0.220
	PoLo	**0.171**	0.292	0.473	0.259
	REx	**0.171**	**0.327**	**0.533**	**0.278**

Explanation between Vincristine and Hematologic Cancer from Hetionet dataset.



REx outperforms state-of-the-art methods in predictive accuracy, delivers more relevant explanatory paths, and produces expert-validated, higher-quality explanations across three drug repurposing benchmarks.

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

Acknowledgements: This work was supported by FCT through the fellowship 2023.00653.BD, and the LASIGE Research Unit, ref. UID/00408/2025.It was also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453. A part of this project was also developed during an internship at SONY AI.

Leveraging Large Language Models (LLMs) to Support Web Accessibility Evaluation

João Costa[1], Letícia Pereira[1], Carlos Duarte[1]

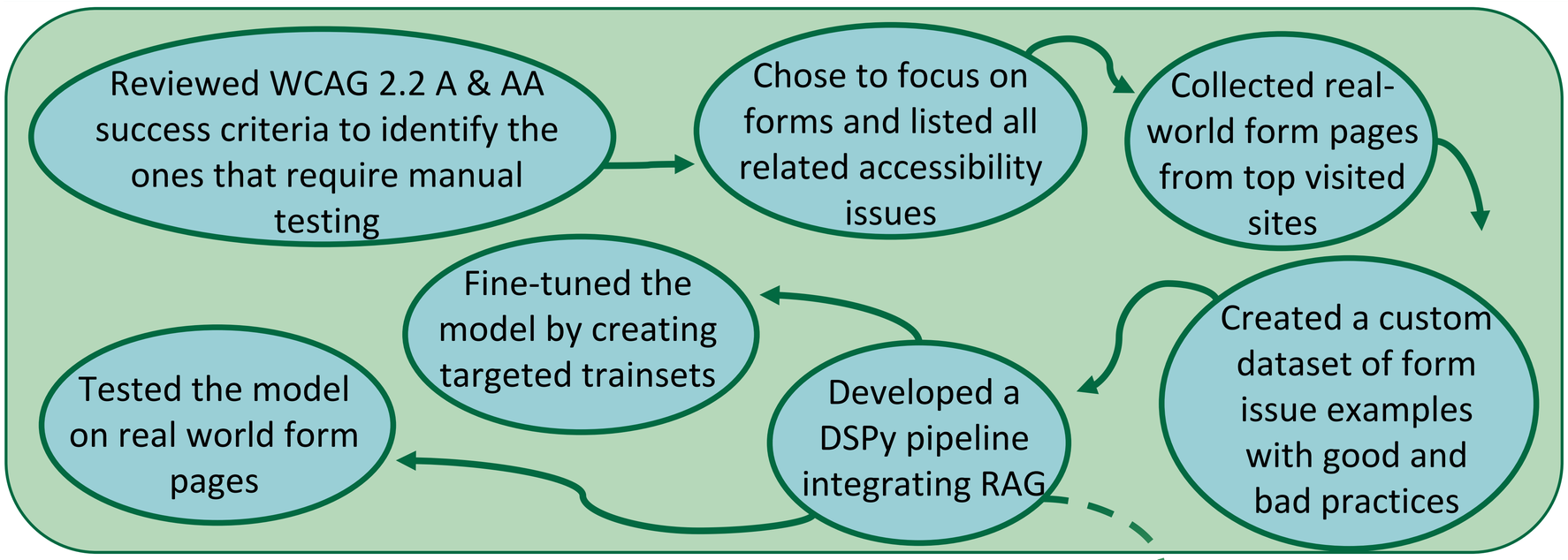
[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE inclusive human-computer interaction

MOTIVATION AND GOALS

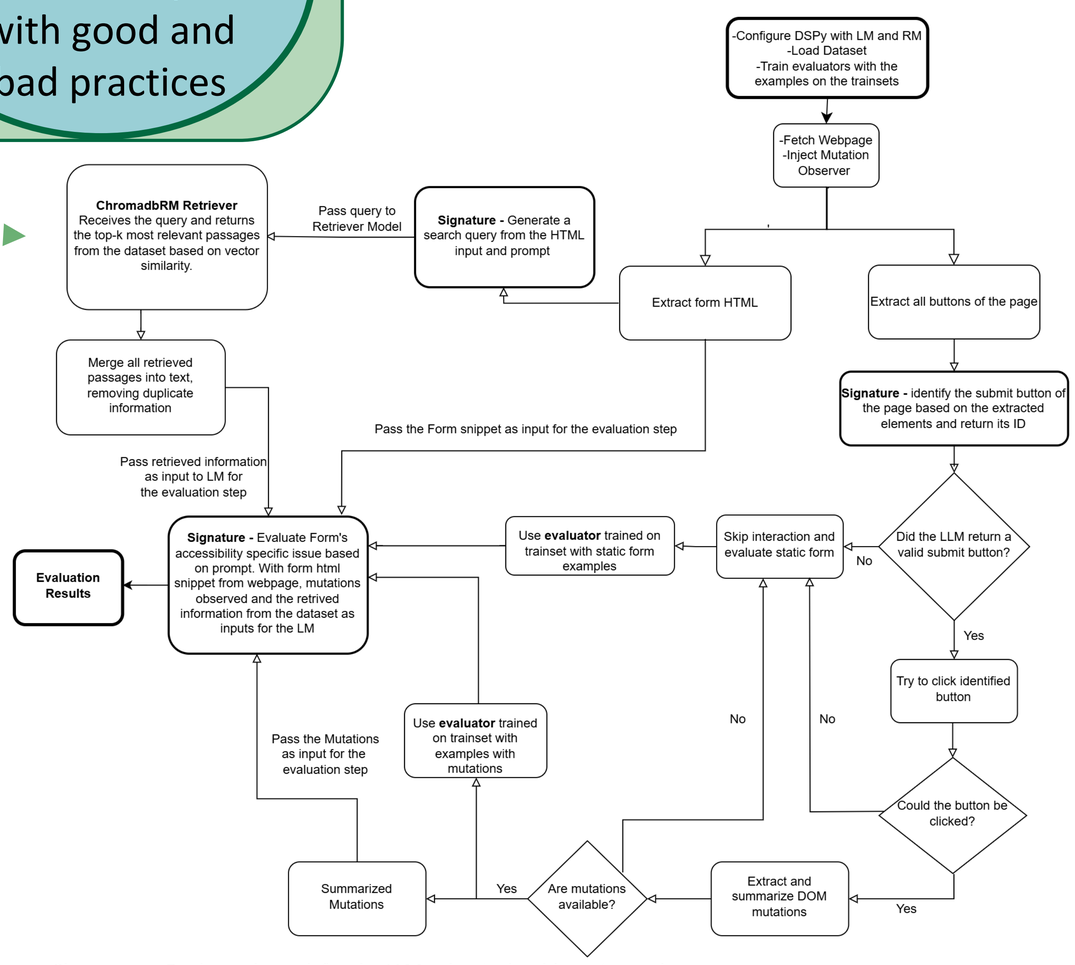
- Many Web Content Accessibility Guidelines (WCAG) success criteria still require manual inspection, limiting the scalability of current evaluation tools.
- Current web accessibility evaluation tools are effective for rule-based issues but struggle with more subjective checks.
- LLMs are capable of understanding context, making them well-suited to automate more complex accessibility checks.
- **By leveraging the capabilities of LLMs we aim to automate the accessibility checks that can't be covered by automated tools, reducing manual effort in accessibility evaluations.**

APPROACH



PIPELINE

- **Built with**: DSPy + Retriever Model (ChromaDB) + Language Model (Mistral Large)
- **Goal**: Evaluate web form accessibility issues with the aid of LLMs

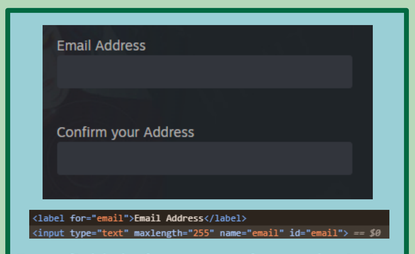


Signature -> Declares the task for the LLM, what it should receive and return.

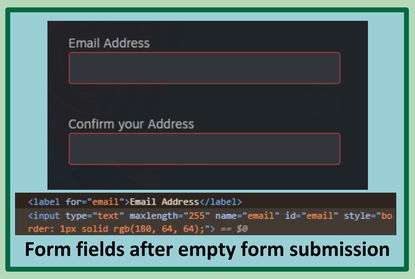
Evaluator -> DSPy module that wraps the Signature, prompt strategy, and uses trainsets to guide model behaviour.

EARLY RESULTS

- Focused on detecting correct use of **required** attribute in forms.
- Current approach attempts to submit empty forms on the page **capturing the HTML both before and after submission** when possible, passing it onto the LM.
- The model achieves around **90% accuracy** in evaluating correct use of the required attribute in form's fields, when empty form submission is possible.



Form fields before empty form submission



Form fields after empty form submission

Fields raise error messages when left empty but **do not have any required attribute** set in the HTML.
Also, these fields raise errors but **only use color changes** to indicate a problem.
These are examples of issues we are aiming to identify and evaluate.

FUTURE GOALS

- Test pipeline on other form-related criteria, beyond the required attribute check.
- Expand test set to validate scalability and robustness.
- Integrate this LLM-augmented evaluation module into QualWeb to complement existing rule-based checks.

Acknowledgements: This work was supported by FCT through project <title>, ref.\ <project reference> (*<project DOI>*), and the LASIGE Research Unit, ref. UID/00408/2025.

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

Do You Really Know How to Attack ML NIDS?

A Reality Check on Adversarial Machine Learning Research

Allan Espindola[1,2], Eduardo Viegas[2], António Casimiro[1], Pedro M. Ferreira[1], Altair Santin[2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] SeCPLab, Escola Politécnica, Pontificia Universidade Católica do Paraná

LASIGE cyber-physical systems

Many **adversarial attacks** against Machine Learning Network Intrusion Detection Systems (NIDS) show **high success**, but only **in controlled, unrealistic settings.**

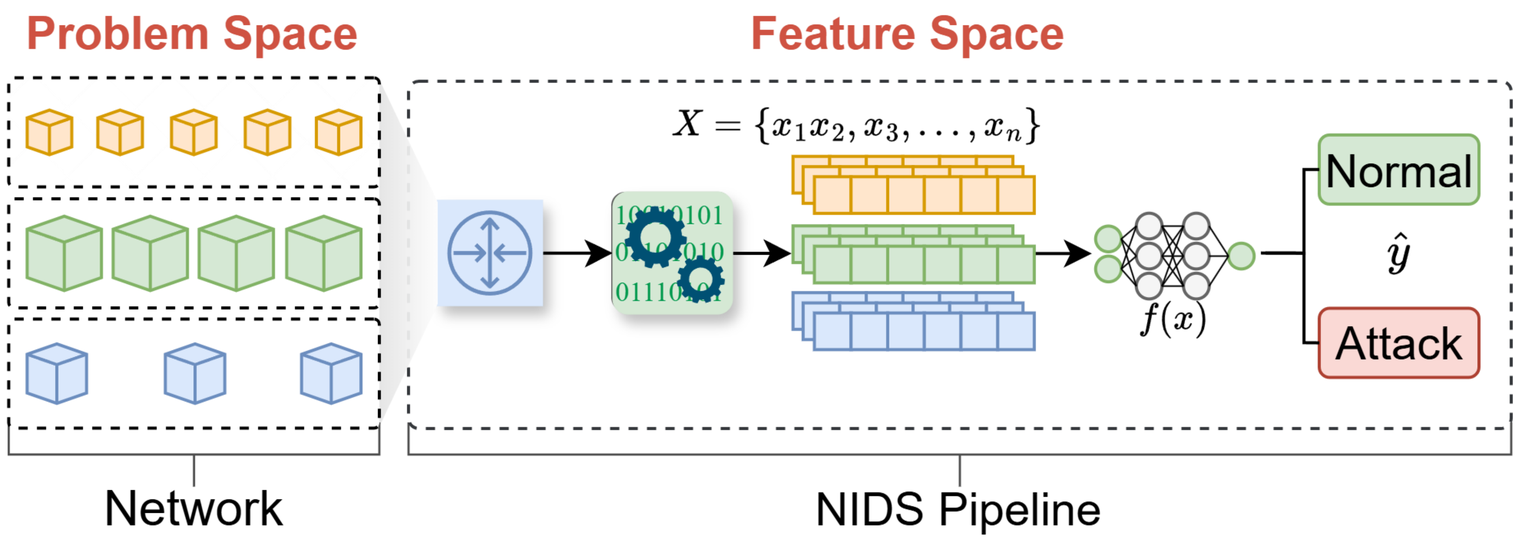
Most are performed **after feature extraction**, in the **feature space**. In the **real world**, attackers interact only with **network traffic**.

So: **Are these attacks even possible to execute?**

Our Study

Challenges the common **practice** of evaluating NIDS with attacks that **operate on internal** model inputs.

☠ **1. Feature-space attacks** (e.g., FGSM, PGD, C&W), which assume full knowledge of the model and its features



☠ **2. Problem-space attacks**, which modify **real network behavior**, such as timing or packet volume

The first are often **impossible**. The second are **realistic** — and effective.

FINDINGS

1. Feature-space attacks:

- ☠ Break natural correlations between features
- ☠ Can't be translated into valid traffic
- ☠ Require capabilities no attacker has

2.Problem-space attacks:

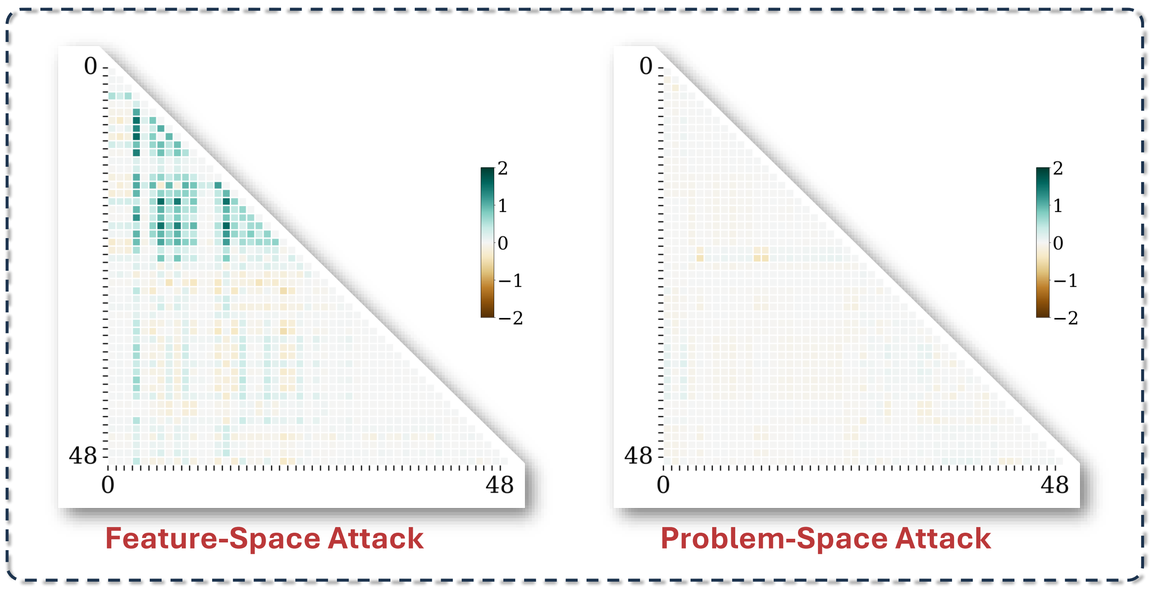
- ☠ Preserve network semantics
- ☠ Remain effective in real-world scenarios

Many published attacks succeed only on paper.

Attacks in the Feature Space ≠ Attacks in Reality

If your attacker needs access to your extracted features, you have bigger problems.

Feature Correlation Shift



THE TAKEAWAY

Evaluating NIDS with **unrealistic attacks** gives a **false sense of security.** Robustness must be measured with **feasible**, real-world attacks, not idealized scenarios.

U LISBOA · Ciências ULisboa · fct Fundação para a Ciência e a Tecnologia · LASIGE · PUCPR Grupo Marista Escola Politécnica · SecPLab Disclosing your future needs for Security & Privacy

Acknowledgements: LASIGE is supported by national funds through FCT within the framework of projects UIDB/00408/2025 and UIDP/00408/2025.

Source Code & Materials
Scan the QR code to access the GitHub repository.
GitHub

Online Clustering of News Articles

Duarte Raposo[1,2], Diogo Pernes[2], Luis Correia[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Priberam

LASIGE data and systems intelligence

Research Problem

Current news clustering approaches have high performance scores, but **perform badly in real-world Media Monitoring applications**.

Contributions

Enriched evaluation methodologies:

- Hard negatives evaluation dataset;
- New evaluation metrics.

New Model implementations:

- Usage of entities as new features;
- Neural Networks based on the use of Attention.

Evaluation Methods

These include both the datasets and metrics used:

Datasets

- **Rupnik Dataset –** standard for the news clustering;
- **Monitio Datasets –** imported from a Media Monitoring product (Monitio), three datasets were created – with 4k documents, 12k documents and 89k documents.

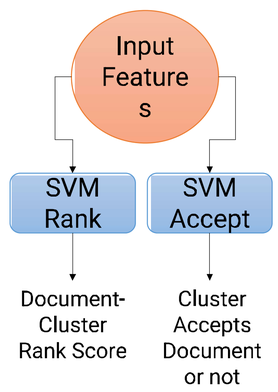
Metrics

- **BCubed Metrics –** standard for clustering problems;
- **Micro and Macro BCubed Metrics –** adapted from the regular BCubed metrics, instead of doing the mean over all items here the mean is done over the classes or clusters.

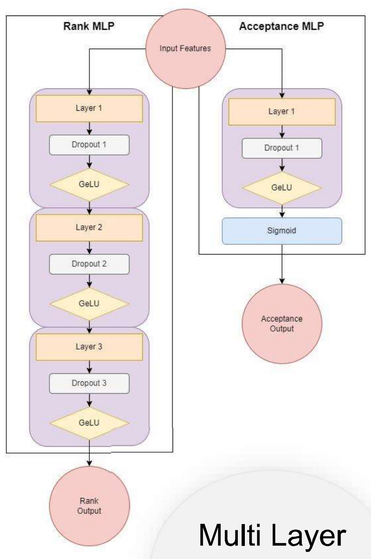
Model and Algorithm Implementations

Usage of the models:

- Document comes in;
- Document is used to rank each cluster;
- The best ranked cluster is selected and passed to the acceptance model together with the document;
- The acceptance model output will tell us if that document belongs to that cluster or not, and with that we decide to add the document and update the cluster's centroids or not.

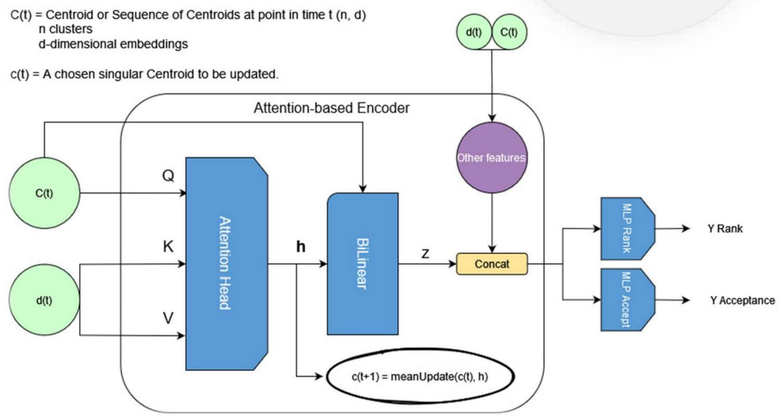


State of the Art Support Vector Machine Models



Multi Layer Perceptron Models

Attention Based Models



Results

		SOTA	MLP Base	MLP With Entities	Attention Based Model	Attention with Centroid Updates
Pure Rupnik	BCubed F1	**83.12**	81.33	79.63	81.22	80.23
	Macro Cluster BCubed F1	44.27	37.66	40.97	32.17	**51.78**
	Macro Class BCubed F1	**78.73**	74.84	78.38	71.32	75.83
	Ground truth Number of Clusters	**394**				
	Num of predicted clusters	1201	1470	1381	1719	**850**
89k Monitio	BCubed F1	**79.24**	67.42	76.14	69.32	73.68
	Macro Cluster BCubed F1	**42.02**	30.55	35.22	21.87	36.70
	Macro Class BCubed F1	71.87	69.60	**74.95**	59.96	66.75
	Num pred clusters	**1217**	1961	1677	2954	1320
	Clusters with monitio docs	286	367	**203**	242	963
	Num monitio docs	10577	3026	**1474**	3822	4568
	Mean percentage of rupnik docs per cluster	86.36	91.07	95.27	**97.01**	60.52

Conclusions

- The new testing methodologies proved efficient detecting centroid deviations and provide more strict evaluations of the problem;
- The model results aren't yet better than the ones tested with the SOTA model;
- The train dataset was proved insufficient, and the train method has its issues regarding the optimization for ranking and acceptance at the same time;
- The new contributions to the test methodologies and to the model provide us with more information on the problem and important baselines for future work.

Acknowledgements: This work was supported by Priberam and by FCT through the fellowship2024.01208.BD, and the LASIGE Research Unit, ref. UID/00408/2025.

priberam · U LISBOA · Ciências ULisboa · FCT Fundação para a Ciência e a Tecnologia · LASIGE driven by excellence

Learning Knowledge Graph Alignment

Pedro Giesteira Cotovio[1,2], Ernesto Jimenez-Ruiz[2], Catia Pesquita[1]

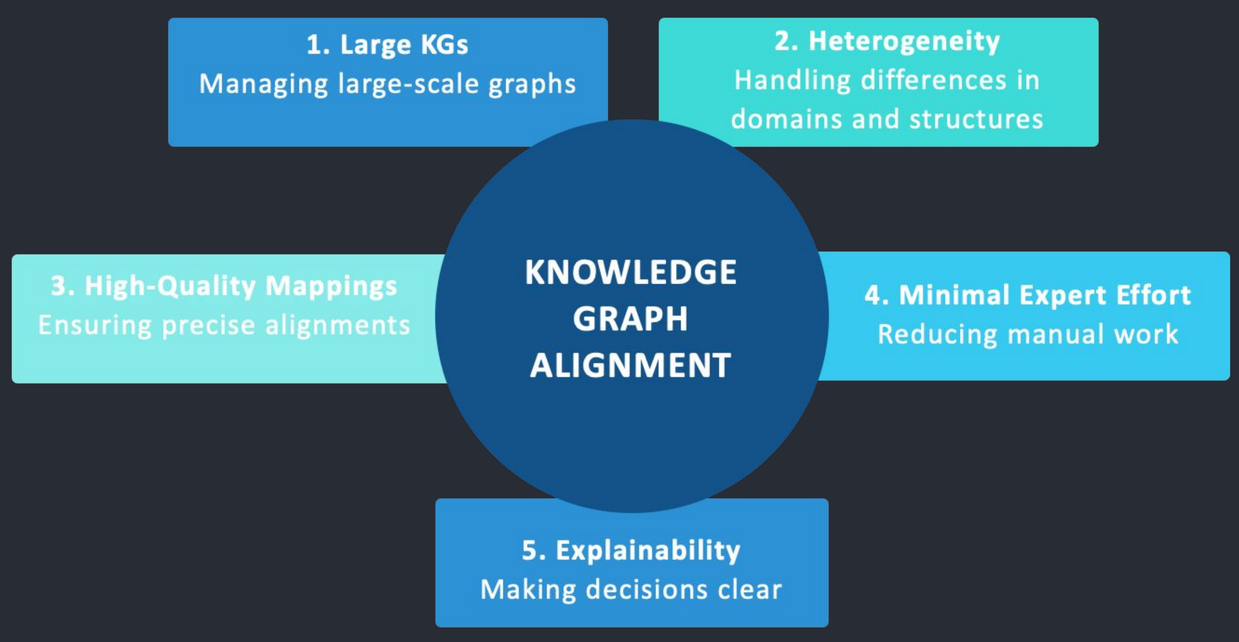
[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] City St George's, University of London

health and biomedical informatics

Knowledge Graph Alignment

Knowledge Graphs (KGs) effectively manage and integrate data, but individual KGs are often incomplete. Knowledge Graph Alignment (KGA) addresses this by identifying and merging equivalent entities and relationships across KGs, creating unified representations. This is essential for leveraging KGs in complex, multi-domain applications.

Challenges



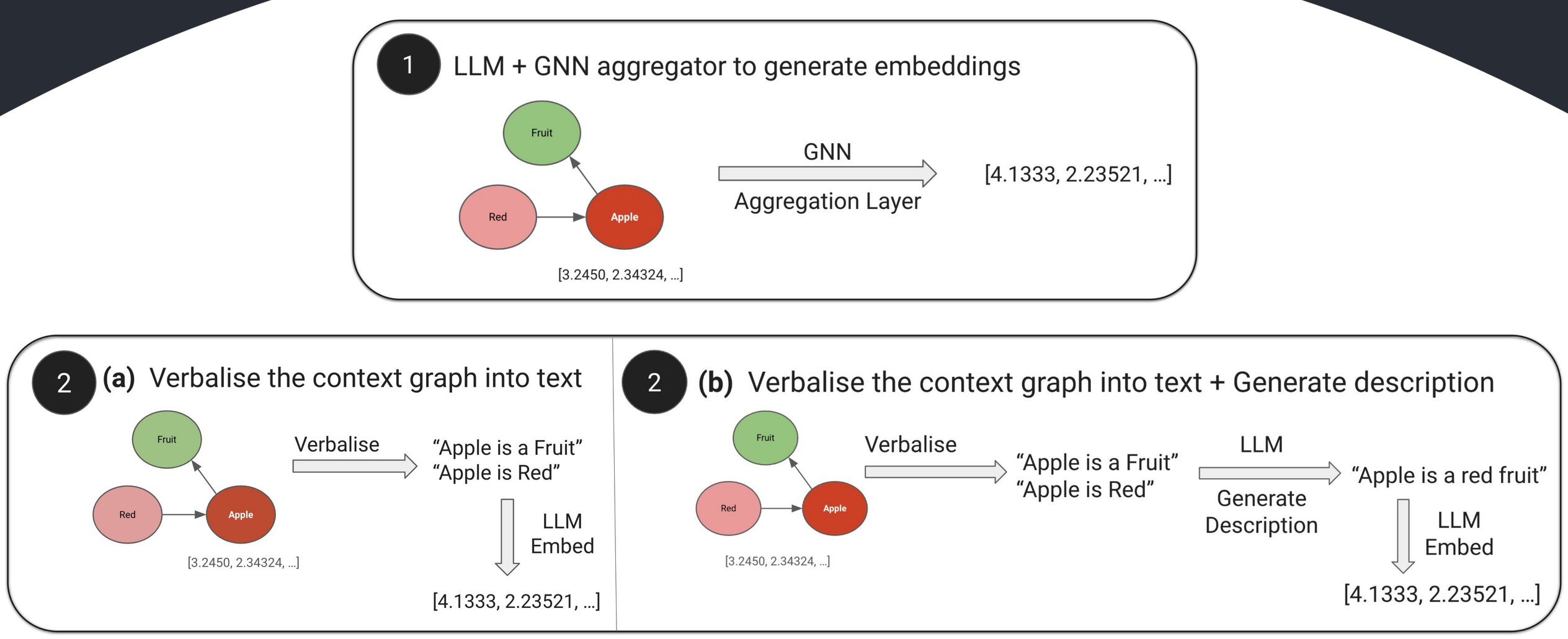
Hypothesis

Augmenting Large Language Models with Knowledge Graph semantics can boost Knowledge Graph Alignment

Objectives

RQ1: Explore strategies to integrate KG semantics into Large Language Models (LLMs) to improve KG alignment by leveraging LLMs' lexical and semantic capabilities while addressing challenges like hallucinations and logical inconsistencies.
RQ2: Maximize recall while minimizing logical errors by integrating logical repair into LLM training and leveraging expert feedback effectively to improve alignment quality with minimal manual effort.
RQ3: Develop approaches to generate interpretable explanations for KG alignment, enhancing human-feedback processes and supporting model transparency and usability.

Augmenting Large Language Models



Acknowledgements: This work was supported by FCT through the grant 2022.10557.BD, the LASIGE Research Unit, ref. UID/00408/2025, the European Union through ERASMUS+ and the Katy project funded with the Horizon2020 program under grant agreement 101017453 , and FLAD under grant 2024-0020.

P300.MM: Creating a Multimodal Dataset of the P300 Wave

João Martins, Manuel J. Fonseca

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE inclusive human-computer interaction

Problem and Motivation

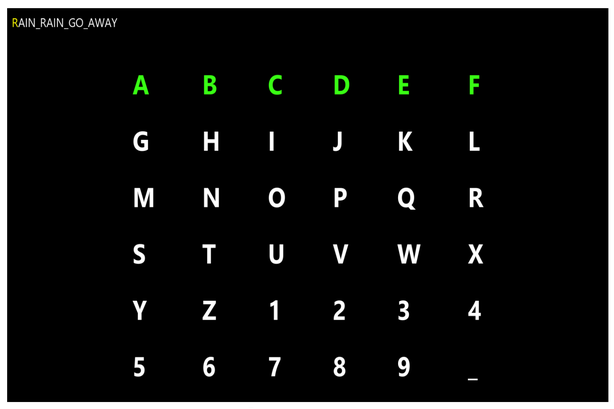
- BCI interfaces are typically based on three types of signals. The P300 wave is one of them
- Identification of a P300 wave typically done (only) by analyzing the EEG signal
- BCI interfaces based on P300 are often used in text input solutions

Objectives

- Create an annotated multimodal dataset containing various physiological signals (PPG, ACC, GYRO, and EEG)
- Use of two stimuli, a matrix keyboard and a rapid visual presentation (RSVP), to elicit P300 waves
- Investigate relationships between physiological signals and the presence of a P300 wave
- Identify which of the two stimuli is more effective in eliciting the P300

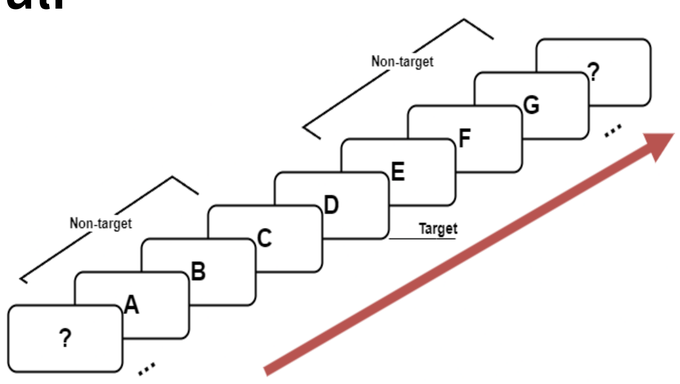
Description of the methodology

Stimuli



Matrix Keyboard

- Made up of 36 characters



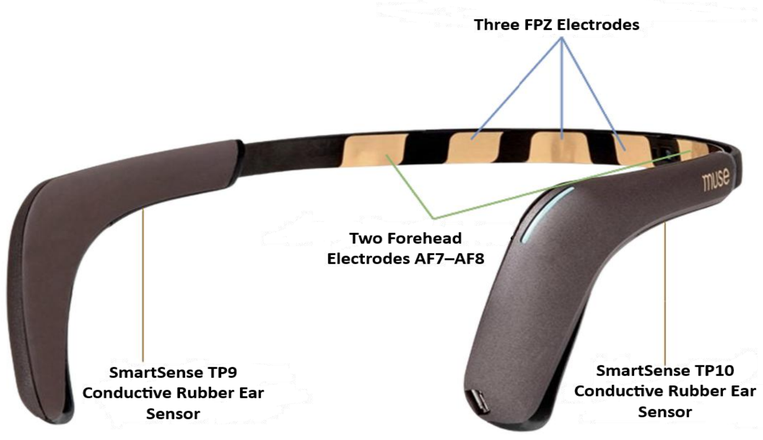
RSVP

Task and Procedure

- Task: Writing of a sentence with approximately 25 characters
- All 36 characters light up to elicit the P300 (1 trial)
- 15 trials for each letter to be written
- 250 ms between each stimulus

Collecting Physiological Signals

- Device: Muse 2
- Physiological signals: EEG (brain activity), PPG (heart rate), ACC + GYRO (movement)

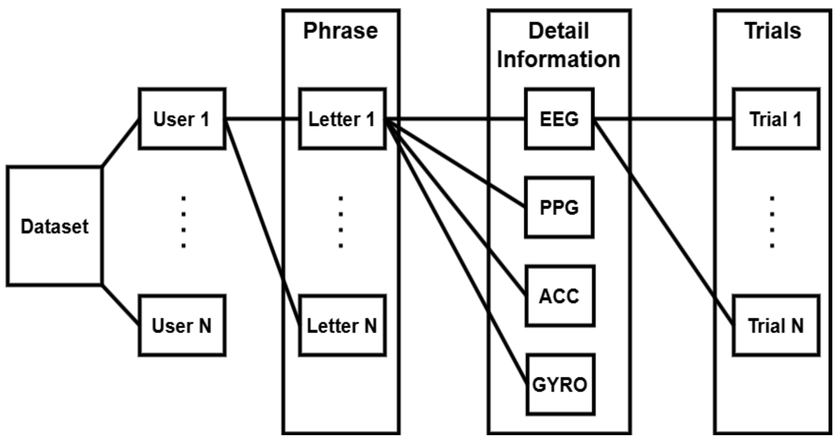


Data Cleaning

- Analyze each record for corrupted data (sudden movements, prolonged signal failures or synchronization problems)
- Format and synchronize valid signals with the relevant events

Dataset Organization

- Hierarchical Structure: Dataset -> User -> Letter -> Signals
- User Folder: Stimuli phases; Stimulus labeled as P300 or non-P300
- Synchronization: Synchronized recordings of EEG, PPG, ACC, and GYRO
- File Format: .CSV per modality, with shared timestamps



Acknowledgements: This work was supported by FCT through project <title>, ref.\ <project reference> (<*project DOI*>), and the LASIGE Research Unit, ref. UID/00408/2025.

U LISBOA Ciências ULisboa fct Fundação para a Ciência e a Tecnologia LASIGE

Personalized Image Recommendation based on Emotions

Ana Domingos[1], Soraia M. Alarcão[1], Márcia Barros[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE data and systems intelligence

Motivation & Objective

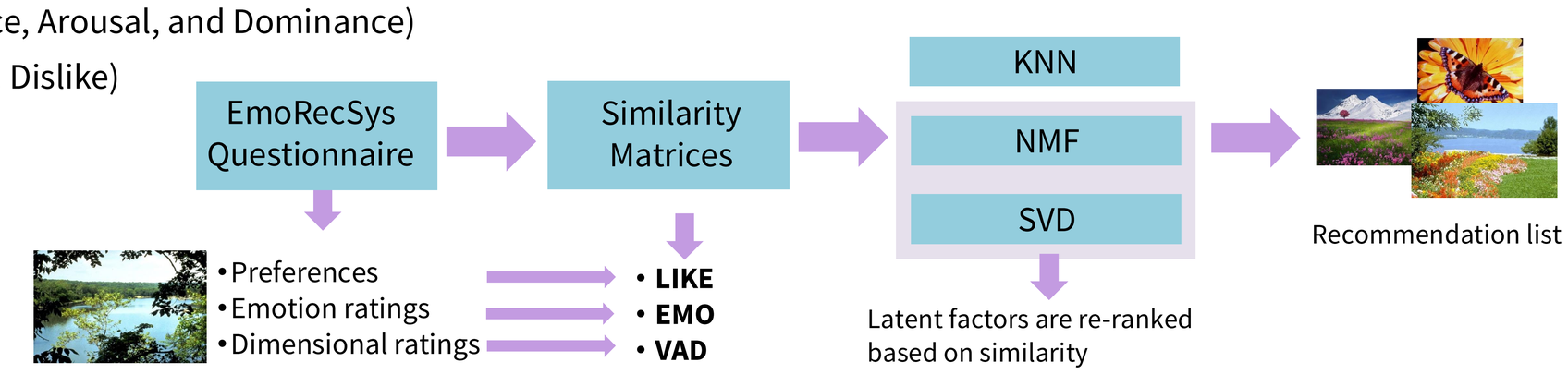
- Information overload on the web makes access to relevant content increasingly challenging
- Recommender systems (RS) attenuate this burden by suggesting personalized content to users
- One way to generate recommendations that deeply resonate with users' preferences is to consider the emotions elicited by content

Objective → Develop an RS that integrates the emotional responses of users to images

Methodology

Development of an emotion-based collaborative filtering RS comparing 3 different approaches to emotions:

- 1) Categorical model of emotions (Anger, Fear, Disgust, Sadness, Happiness, Surprise, and Neutral)
- 2) Dimensional (Valence, Arousal, and Dominance)
- 3) Preferences (Like *vs.* Dislike)



Experimental Setup

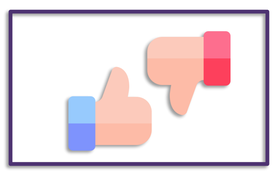
Dataset:

- 2500 images from Pexels
- 385 IAPS Mikels

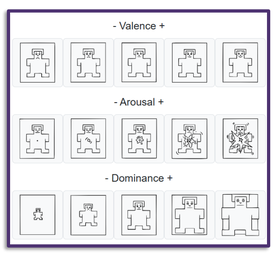
Number of Images	2885
Number of Users	163
Number of Images per User	15
Number of ratings	2445

EmoRecSys Questionnaire:

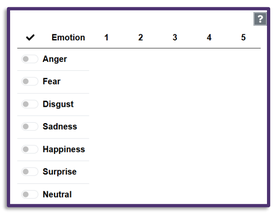
- Part I: Demographic information about the user
- Part II: Evaluation of 15 images



Preference:
Like *vs.* Dislike

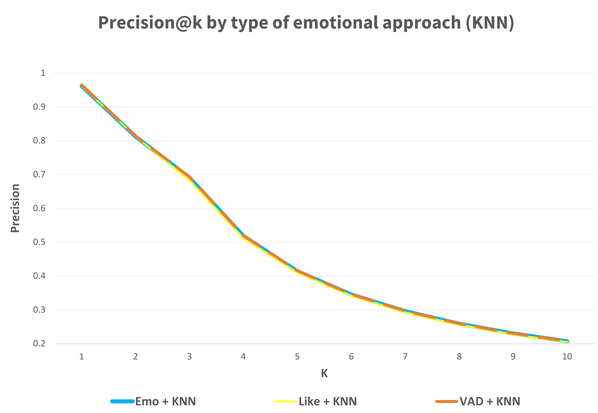


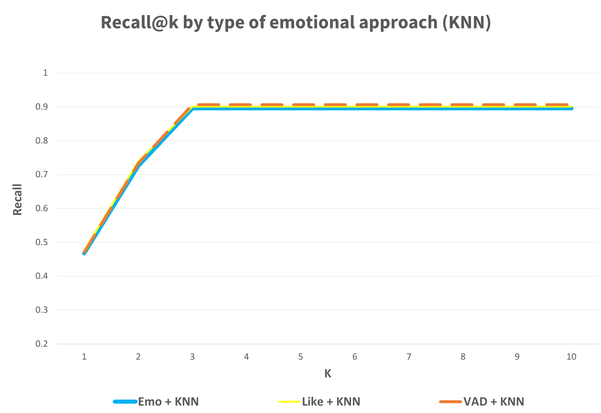
VAD Dimensions:
Self Assessment Manikin (SAM) on Valence-Arousal-Dominance



Emotions:
5-point Likert scale for 7 emotions
(1 - milder to 5 - intense)

Results & Discussion





- Best results achieved for **KNN** + **VAD** for Precision@1 (0.97) and Recall@3 and onward (0.91)
- Excessive granularity of **EMO** variant across RS algorithms may have had a detrimental effect, maybe due to a small sample size

Future Ideas:

- Integrate demographic profiles in RS, since emotions can be influenced by context information or individual experiences
- Use DL algorithms (e.g., Neural CF, Bert4rec, or SASRec)

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE

Acknowledgements: This work was supported by FCT through LASIGE Research Unit, ref. UID/00408/2025

From Time to Timelines: TKG Embeddings for Patient-Centric ICU Readmission Prediction

Ricardo M.S. Carvalho[1], Andreia Sofia Teixeira[1,2], Catia Pesquita[1]

[1]LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2]Network Science Institute, Northeastern University London, United Kingdom

LASIGE health and biomedical informatics

Introduction

In healthcare, data isn't just abundant—it's vital. Predicting ICU readmissions is a high-stakes challenge where every detail counts, yet most models miss the bigger picture by overlooking the semantic meaning and timing of clinical events.

This project explores KG embeddings and TKG Embeddings to see how well they can predict ICU readmissions by capturing the "who, what, and when" of patients' care. Our study focuses on two critical questions

- Does adding semantic knowledge (**ontologies**) enhance predictions?
- Does capturing the timing of events (**temporal data**) improve outcomes?

Armed with a comprehensive ICU dataset and a range of static and temporal embedding models, we offer fresh insights into using TKGs for clinical prediction.

Objective

Static-Simple
KGE's | Patient

Temporal-Simple
TKGE's | Patient

VS

Static-Semantic
KGE's | Patient + Ontology

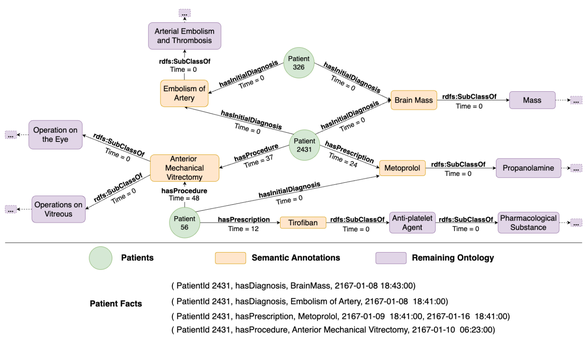
Temporal-Semantic
TKGE's | Patient + Ontology

Method

MIMIC-III » Construct Knowledge Graphs » Embedding » ML Prediction

Process	Simple KG	Semantic KG	Temporal KG	Semantic-Temporal KG
Data pre-processing	yes	yes	yes	yes
Semantic Annotation	yes	yes	yes	yes
Triple Generation	yes	yes	yes	yes
Semantic Enrichment	no	yes	no	yes
Temporal Annotation	no	no	yes	yes

Element	Simple	Semantic	Temporal	Semantic-Temporal
Patient Nodes	66,540	66,540	66,540	66,540
Ontology Class Nodes	0	232,157	0	232,157
Total Nodes	66,540	298,697	66,540	298,697
Patient-level properties	5	5	5	5
Ontology-level properties	0	2	0	2
Total properties	5	7	5	7
Patient Triples	3,110,248	3,110,248	11,207,832	11,207,832
Ontology Triples	0	226,326	0	226,326
Total Triples	3,110,248	3,336,574	11,207,832	11,434,158



I. **Temporal Knowledge Graph Construction:** Built four patient-centric TKGs from ICU data, distinguishing between simple and semantically enriched graphs. Time was encoded as hours since ICU.

II. **Semantic Enrichment with Ontologies:** Exclusively integrated NCIT ontology, capturing hierarchical relationships. Focus *SubClass* and *EquivalentClass* axioms.

III. **Embedding Generation:** Generated embeddings using three model families: translational (TransE, T-TransE, TA-TransE), bilinear (DistMult, TA-DistMult), and complex (ComplEx, T-ComplEx, TNT-ComplEx). Static embeddings represented patient states, while temporal variants captured time-evolving events.

IV. **Predictive Modeling and Evaluation:** Used Logistic Regression, Random Forest, and XGBoost for prediction. Models trained on embeddings were evaluated using Acc and ROC-AUC)

Results

Semantic Enrichment

- ✓ **ComplEx** family benefits
- ✓ **DistMult** fails to leverage semantic information due to its symmetric nature.
- ✓ **TransE paradox:** Semantic enrichment improves static TransE but harms its temporal variants.

Temporal Representation

- ✓ Temporal gains in **ComplEx family**
- ✓ **TransE** struggles with time, temporal TransE variants underperform, failing to capture time dynamics.
- ✓ **DistMult paradox:** Temporal DistMult performs better, but semantic enrichment has limited impact.

Baseline Comparison

- ✓ Baseline outperformed: models consistently exceed the performance of the numerical baseline.
- ✓ Embedding impact: The greatest performance gain comes from embedding clinical features, not raw data.

Model	Embedding	Simple-KG/TKG Accuracy Mean	Simple-KG/TKG Accuracy Std	Simple-KG/TKG ROC AUC Mean	Simple-KG/TKG ROC AUC Std	Semantic-KG/Semantic-TKG Accuracy Mean	Semantic-KG/Semantic-TKG Accuracy Std	Semantic-KG/Semantic-TKG ROC AUC Mean	Semantic-KG/Semantic-TKG ROC AUC Std
LR	Baseline	0.745	0.002	0.603	0.008	–	–	–	–
	complEx	0.794	0.003	0.756	0.008	0.791	0.004	0.760	0.008
	T-complEx	**0.802**	0.004	**0.775**	0.008	**0.806**	0.002	**0.776**	0.009
	TNTcomplEx	**0.807**	0.003	**0.782**	0.008	**0.810**	0.004	**0.786**	0.009
	transE	0.807	0.004	0.782	0.006	0.809	0.003	0.784	0.007
	T-transE	0.786	0.005	0.758	0.008	0.748	0.002	0.652	0.010
	TA-transE	0.765	0.003	0.722	0.007	0.762	0.003	0.701	0.008
	DistMult	0.798	0.002	0.767	0.010	0.795	0.003	0.760	0.008
	TA-DistMult	**0.807**	0.004	**0.779**	0.006	0.791	0.005	**0.760**	0.007
RF	Baseline	0.745	0.002	0.603	0.008	–	–	–	–
	ComplEx	0.805	0.003	0.733	0.008	0.794	0.003	0.734	0.009
	T-ComplEx	0.782	0.003	**0.749**	0.006	0.783	0.003	**0.751**	0.007
	TNTComplEx	0.785	0.003	**0.757**	0.007	0.791	0.004	**0.763**	0.005
	TransE	0.804	0.003	0.775	0.006	0.805	0.003	0.773	0.005
	T-TransE	0.748	0.002	0.689	0.006	0.745	0.002	0.607	0.009
	TA-TransE	0.755	0.002	0.692	0.005	0.747	0.003	0.653	0.007
	DistMult	0.791	0.003	0.749	0.006	0.775	0.003	0.730	0.007
	TA-DistMult	0.784	0.004	**0.762**	0.008	0.773	0.004	**0.743**	0.007
XGB	Baseline	0.384	0.072	0.602	0.009	–	–	–	–
	ComplEx	0.751	0.004	0.772	0.005	0.732	0.005	0.768	0.005
	T-ComplEx	**0.750**	0.005	**0.771**	0.007	**0.748**	0.007	**0.773**	0.010
	TNTComplEx	**0.753**	0.006	**0.781**	0.008	**0.758**	0.008	**0.785**	0.007
	TransE	0.766	0.009	0.795	0.008	0.769	0.009	0.793	0.006
	T-TransE	0.662	0.017	0.743	0.007	0.494	0.007	0.643	0.009
	TA-TransE	0.665	0.024	0.716	0.008	0.638	0.037	0.697	0.009
	DistMult	0.755	0.006	0.782	0.007	0.733	0.011	0.771	0.007
	TA-DistMult	**0.756**	0.008	**0.781**	0.006	0.726	0.004	0.765	0.007

Lessons Learned

KG enrichment did not mean improved performance!

- ❑ Semantic redundancy
- ❑ Embeddings can't leverage hierarchical information
- ❑ Ontology choice

Temporal modeling benefits depend on model design!

- ❑ Algorithmic choices
- ❑ Time zero for ontologies and temporal resolution
- ❑ Transductive TKGs are an issue

U LISBOA | Ciências ULisboa | fct Fundação para a Ciência e a Tecnologia | LASIGE | KATY Clinical Knowledge for Humanity

Acknowledgements: This work was supported by FCT through the fellowship 2022.10769.BD, and the LASIGE Research Unit, ref. UID/000408/2025. It was also partially supported by the KATY project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101017453.