Coordinating and Regulating the IoT Mess with SIRE

Tiago Carvalho, Alysson Bessani, António Casimiro

LASIGE, Faculdade de Ciências, Universidade de Lisboa

LASIGE resilient distributed and networked systems

MOTIVATION

Characteristics of the Internet of Things

Critical Applications

Highly Dynamic

Highly Heterogenous

Problems

- Keep an up-to-date list of participating devices, i.e., the membership
- Ensure devices satisfy the security requirements
- Support multiple applications with different contexts

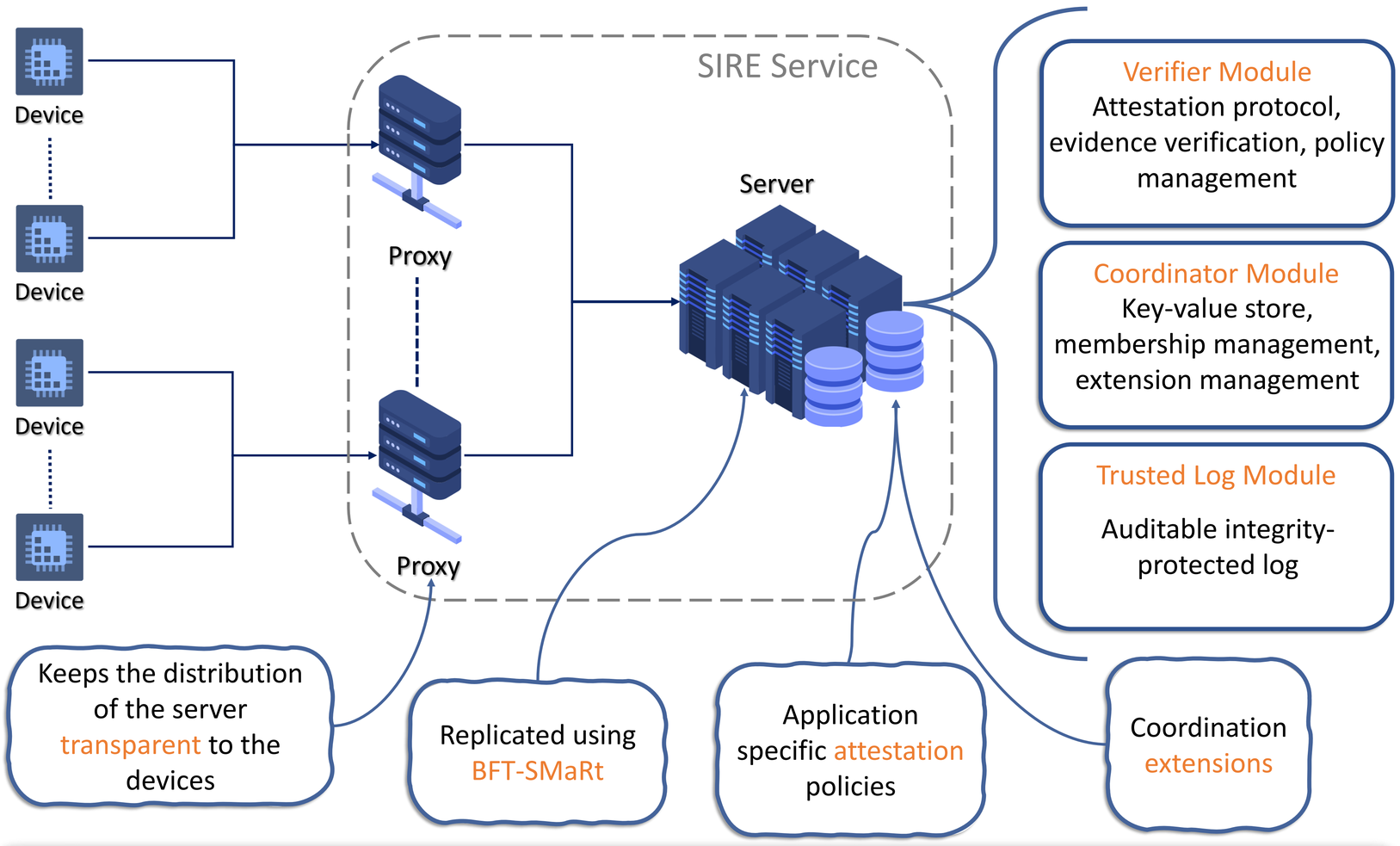
GOALS

System with these features:

- Remote Attestation
- Coordination Primitives
- Membership Management
- Byzantine Fault Tolerance

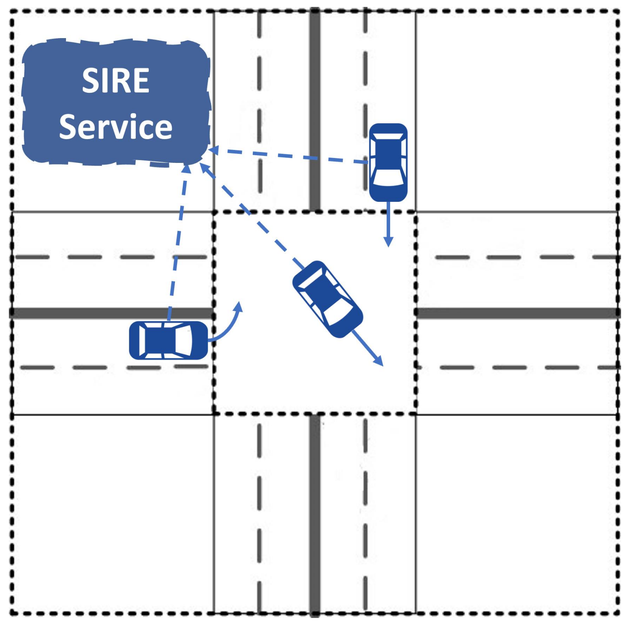
- Provide a reliable source of information about which devices can participate in an application
- Increase the trust between interacting IoT devices

OUR SOLUTION – SIRE

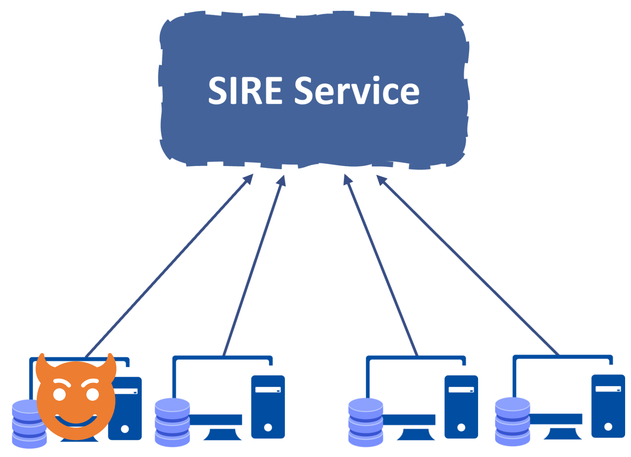


USE CASES

Example applications for SIRE:



1. Decide the order of passage at intersections for driverless vehicles



2. Perform robust distributed machine learning

CONCLUSION & FUTURE WORK

SIRE provides services beyond basic attestation to distributed, dynamic IoT applications operating in untrusted environments

Provide confidentiality

Integrate trusted hardware component

U LISBOA Ciências ULisboa FCT Fundação para a Ciência e a Tecnologia LASIGE driven by excellence VEDLIoT Very Efficient Deep Learning in IoT

Acknowledgements

This work was supported by FCT through the research unit LASIGE (ref. UIDB/00408/2020 and ref. UIDP/00408/2020) and by European Union's Horizon 2020 Research and Innovation Programme through the VEDLIoT project (agreement 957197).

Happy or Sad, Smiling or Drawing:
Multimodal Search and Visualisation of Movies Based on Emotions Along Time

Francisco Caldeira, João Lourenço, Teresa Chambel

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE health and biomedical informatics

Multimodal Search

Results match queries' modalities

by Dominant Emotions

Homepage & Random movie

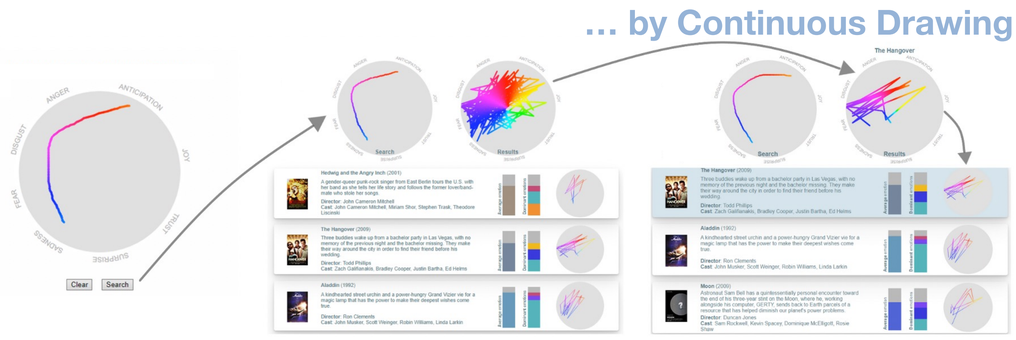
... by Categorical emotions in Words

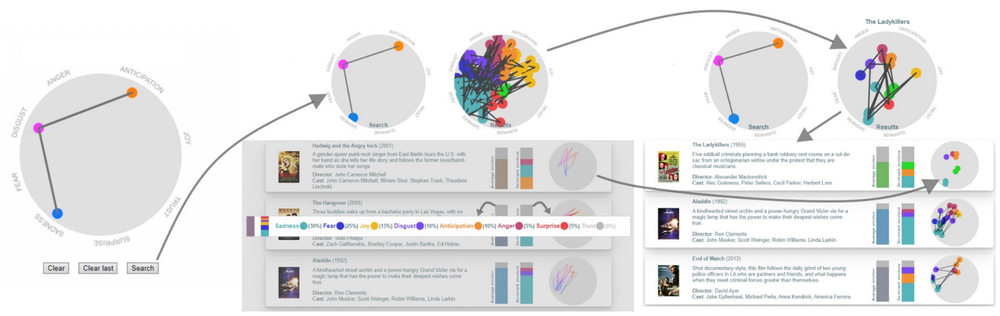
... by Emojis and Camera

... by Colors in Wheel

by Trajectory

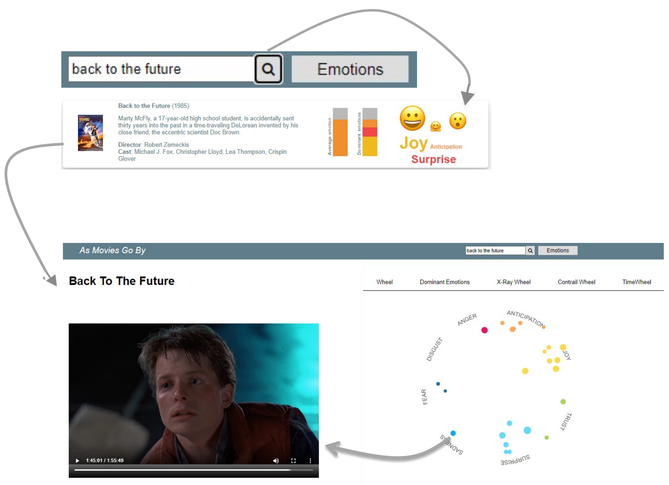
... by Continuous Drawing





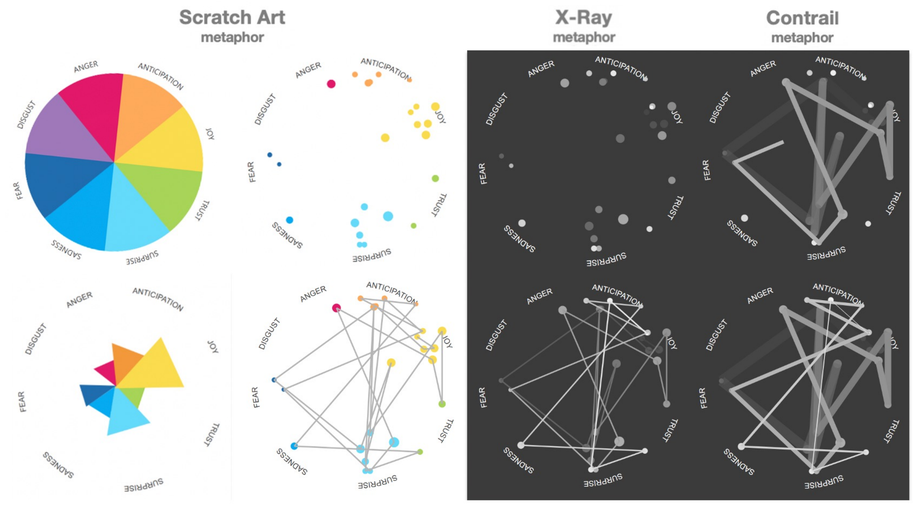
... by Discrete Points

by Title



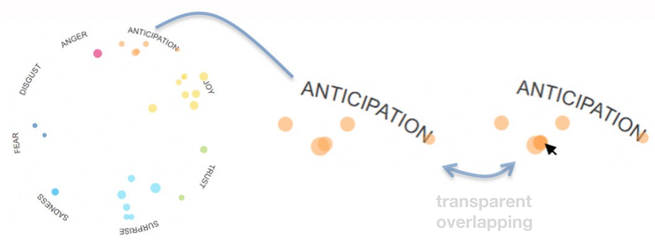
Movie Visualization

Visualization of Emotional Impact along Time

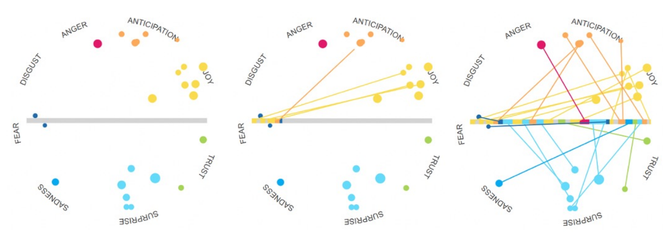


Dominant Emotions | Stationary or Overlapping + Trajectories

Emotion Wheel



Time Wheel



Highlights

Searching & Visualizing of Previous Annotated Movies Emotions . Multimodal . along Time
extend As Music Goes By, in the AWESOME Project

Flexibility in Emotion Models

Valence & Arousal Wheel + Categorical + Colors + Emojis
Personalization

Main Challenges addressed

1) timing of emotional trajectories

on the wheel: how to represent time direction and speed?
(visible in real-time or replay then static + sync with timeline)

2) stationary or overlapping spots

how to distinguish when the same emotions are felt for a long time or again at different times

User Evaluation

Task / # Feature	U M	U SD	S M	S SD	E M	E SD
1 Home: random movie	**4.3**	1.3	**4.2**	1.1	**5.0**	0.0
2 Search and Results *(mean)*	*3.7*	*1.0*	*3.8*	*0.8*	*3.9*	*0.7*
2.1 **by Emotions** in **Words**	**4.2**	1.1	**4.3**	0.8	**4.4**	0.7
Results	**4.1**	1.3	**4.1**	1.1	3.9	0.9
2.2 **by Emojis** and Camera	3.9	1.2	**4.6**	0.7	3.7	0.7
Results	3.2	0.8	**3.5**	0.5	3.8	0.6
2.3 **by Colors** in Wheel	3.3	1.2	3.2	0.6	3.3	0.7
Results	3.5	1.4	3.5	1.2	**4.4**	0.7
2.4 **by Trajectory Free** Drawing	2.8	0.9	3.0	0.9	3.4	0.8
Results	3.2	0.9	3.2	1.1	2.7	0.8
2.5 by Trajectory **Discrete** Points	**4.3**	0.8	**4.3**	0.9	**4.7**	0.5
Results	3.5	1.0	3.5	1.0	3.5	0.7
2.6 **by Title**	**4.8**	0.6	**4.7**	0.5	**4.9**	0.3
3 Movie Visualization *(mean)*	*3.7*	*1.1*	*3.7*	*0.7*	*3.7*	*0.6*
3.1 **Wheel** View	3.6	1.0	3.5	0.7	3.2	0.6
3.1.1 Overlays	3.8	1.4	3.4	0.7	3.4	0.7
3.1.2 Time and Trajectory	3.6	1.3	3.9	0.9	**4.1**	0.7
3.2 **Dominant Emotions** View	**4.0**	1.3	**4.2**	0.8	**4.5**	0.7
3.3 **X-ray** View	2.9	0.9	2.9	0.3	3.0	0.5
3.4 **Contrail** View	3.4	0.7	3.6	0.7	3.8	0.6
3.5 **TimeWheel** View	**4.4**	0.8	**4.1**	0.7	**4.0**	0.5
Global Evaluation	3.7	0.7	3.9	0.6	3.8	0.4
Total per Task (mean)	*3.9*	*1.1*	*3.9*	*0.9*	*4.2*	*0.4*

(1-5) (N=10)

Most chosen quality terms

Terms	type	#
Interesting	E	10
Comprehensible	H	8
Pleasant	A	8
Clear	E	6
Trustworthy	E	6
Innovative	H	5
Original	H	5
Good	A	4
Motivating	A	4
Sympathetic	A	4
Inviting	A	3
Controllable	E	3
Exciting	H	3
Exclusive	H	3
Impressive	H	3
Aesthetic	A	3
Attractive	A	3
Complex	*E*	*3*
Simple	E	2
Desirable	A	2
Familiar	A	1

(E)rgonomic; (H)edonic; (A)ppeal

U LISBOA | Ciências ULisboa | FCT Fundação para a Ciência e a Tecnologia

check out the paper

Best Paper Award @ HUCAPP 2023

Acknowledgements: This work was supported by Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020, and FCT grant PTDC/CCI/ 29234/2017 funding the AWESOME Project.

A Comparative Study of Automated Deep Learning Segmentation Models for Prostate MRI

Nuno M. Rodrigues[1,2], Sara Silva[1], Leonardo Vannechi[3], Nickolas Papanikolaou[2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Champalimaud Foundation
[3] NOVA Information Management School (NOVA IMS),
Universidade Nova de Lisboa, Campus de Campolide, Lisboa, 1070-312, Portugal

LASIGE
data and systems intelligence

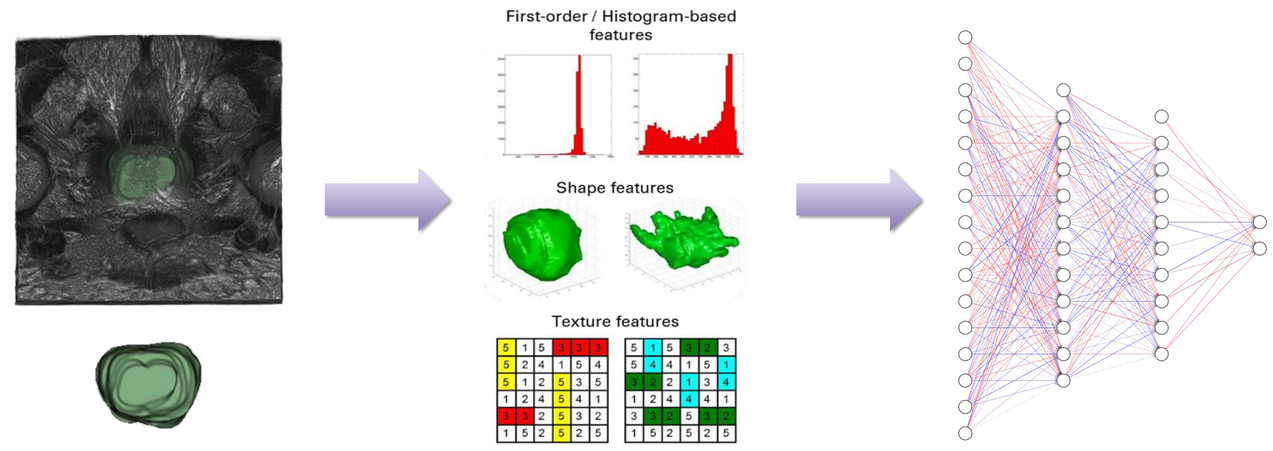
Full paper

Motivation

Radiomics is the process of turning medical images into mineable high-dimensional data, which allows for the extraction of quantitative high-dimensional features to be used in different clinical applications.

Rq. 1
Do significant differences exist between current popular segmentation architectures in a controlled experimental setting?

Rq. 2
Can an object detection approach be employed to locate the prostate and facilitate supervised reduction of image dimensions, thus enhancing the speed and quality of segmentation?

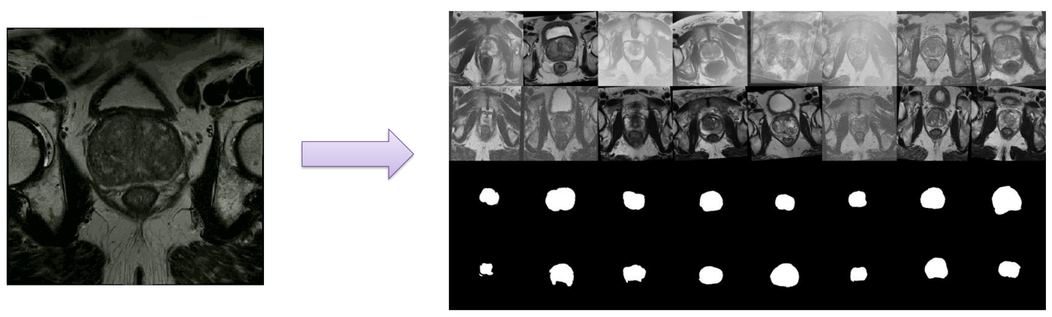


Segmentation and volume extraction → **Feature extraction and selection** → **Model evaluation**

Methodology

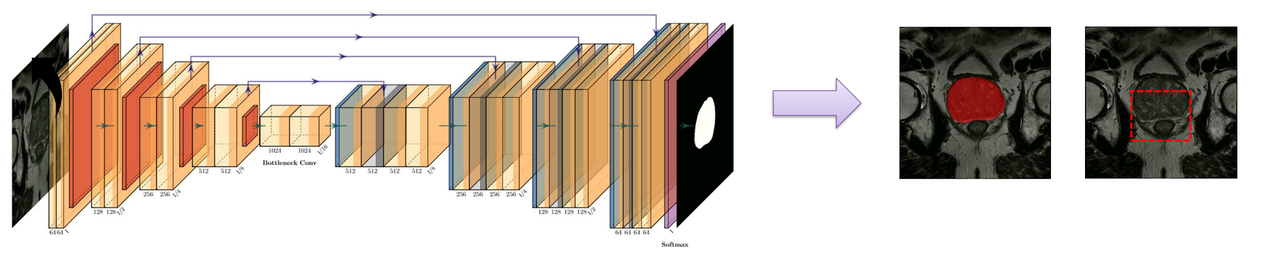
Segmentation
Pixel-level classification of an image. In medical imaging, segmentation can be used to isolate specific organs or tissues for further analysis or diagnosis.

Augmentation & pre-processing



Object Detection
Identification and classification of objects within images or videos. In medical imaging, it can be used to identify anatomical structures and lesions.

Prostate segmentation & detection



14 different encoder-decoder convolution models

- ❑ Attention
- ❑ Recurrence
- ❑ Dense connections
- ❑ Skip connections
- ❑ nnUnet

Results

(A)

	Gland				Cropped Gland			
	MDS	CI	MHD (mm)	ASD (mm)	MDS	CI	MHD (mm)	ASD (mm)
unet	0.9082	0.0154	11.5690	0.8732	0.9094	0.0064	7.0381	0.6990
unet++	0.9066	0.0108	9.7262	0.9289	0.9081	0.0037	7.9955	0.7533
runet	0.9081	0.0078	8.8370	0.7610	0.9083	0.0062	7.3480	0.7186
aunet	0.9052	0.0151	11.5549	1.0364	0.9105	0.0079	7.0760	0.7047
daunet	0.9023	0.0203	13.5009	1.0892	0.9123	0.0076	7.1538	0.7265
d2unet	0.9078	0.0170	12.5156	1.2563	0.9145	0.0047	6.6846	0.6664
d2aunet	0.9110	0.0115	9.9763	0.7997	0.9158	0.0068	7.6787	0.6715
r2unet	0.9084	0.0087	9.6704	0.8259	0.9117	0.0081	7.0649	0.6994
r2aunet	0.9069	0.0110	10.1465	0.8990	0.9130	0.0036	7.1827	0.7096
vnet	0.8991	0.0062	12.5448	1.0621	0.9067	0.0070	7.2024	0.7220
segresnet	0.8960	0.0117	9.3760	0.8540	0.9048	0.0051	6.9018	0.7490
highresnet	0.8980	0.0130	31.1897	2.1118	0.9029	0.0084	25.8385	0.7533
nnunet	0.9289	0.0046	5.7155	0.9218	0.9139	0.0044	6.2565	0.7406

(B)

	Gland			Cropped Gland		
	MDS	MHD (mm)	ASD (mm)	MDS	MHD (mm)	ASD (mm)
unet	0.8474	16.9045	2.3177	0.8595	9.9819	1.2044
unet++	0.8544	25.7024	2.0092	0.8472	9.0465	1.3394
runet	0.8501	19.8740	1.9686	0.8547	9.9693	1.2591
aunet	0.8291	25.3664	2.5814	0.8521	9.9022	1.3618
daunet	0.8277	21.4430	2.1456	0.8579	10.6797	1.2818
d2unet	0.8528	15.2098	1.7646	0.8567	9.1642	1.3421
d2aunet	0.8464	13.6260	1.9367	0.8562	8.2389	1.3334
r2unet	0.8517	14.9619	1.7727	0.8546	8.7872	1.3819
r2aunet	0.8514	16.2630	1.9737	0.8573	10.3648	1.7792
vnet	0.8368	21.3305	2.7043	0.8548	9.6195	1.3882
segresnet	0.8479	17.4913	1.7528	0.8468	9.7496	1.3192
highresnet	0.7448	69.6783	11.5506	0.8092	35.0037	3.6740
nnunet	0.8678	10.0231	3.3704	0.8558	9.2565	1.6221

(A)

	Tz				Cropped Tz			
	MDS	CI	MHD (mm)	ASD (mm)	MDS	CI	MHD (mm)	ASD (mm)
unet	0.8423	0.0243	14.0153	1.0092	0.7790	0.0250	14.0153	1.3668
unet++	0.8457	0.0254	14.0649	1.0042	0.7715	0.0278	14.0649	1.3840
runet	0.8469	0.0186	12.9054	0.9450	0.7895	0.0432	12.9054	1.3410
Aunet	0.8509	0.0164	14.8667	0.9109	0.7543	0.1045	14.8667	1.4419
daunet	0.8481	0.0155	14.6603	0.9105	0.7703	0.0311	14.6603	1.3604
d2unet	0.8447	0.0144	13.9448	0.9676	0.7790	0.0483	13.9448	1.4247
d2Aunet	0.8328	0.0456	14.5523	1.9611	0.7724	0.0351	14.5523	1.3411
r2unet	0.8390	0.0366	14.4749	1.0748	0.7788	0.0579	14.1360	1.3977
r2Aunet	0.8456	0.0152	13.1392	1.0938	0.7799	0.0365	13.1392	1.3406
vnet	0.8389	0.0172	12.4145	1.0663	0.7998	0.0282	12.4145	1.2589
segresnet	0.8234	0.0199	13.7308	1.2383	0.7729	0.0347	13.7308	1.3995
highresnet	0.8457	0.0186	15.0525	1.5373	0.8083	0.0212	15.0525	1.3386
nnunet	0.8760	0.0099	8.7049	1.2928	0.8561	0.0133	10.2390	1.1226

(B)

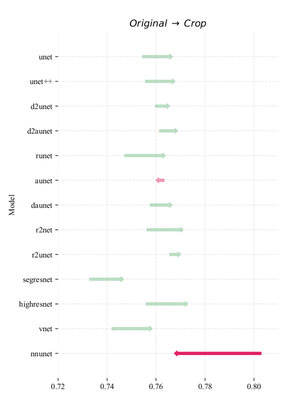
	Tz			Cropped Tz		
	MDS	MHD (mm)	ASD (mm)	MDS	MHD (mm)	ASD (mm)
unet	0.7700	22.5822	2.4379	0.6424	16.8615	2.6838
unet++	0.7681	19.8427	2.3727	0.6536	16.9354	2.4165
runet	0.7671	19.9993	2.0027	0.6477	17.0520	2.5169
Aunet	0.7680	21.7833	2.4354	0.6577	16.0525	2.5444
daunet	0.7688	19.6406	2.4507	0.6471	17.4325	2.5915
d2unet	0.7627	21.2574	2.4447	0.6581	17.3590	2.7397
d2Aunet	0.7638	22.9369	2.3869	0.6581	16.7426	2.5895
r2unet	0.7631	19.8119	2.5560	0.6523	17.3374	2.8883
r2Aunet	0.7672	21.2282	2.1350	0.6553	14.3813	2.6186
vnet	0.7643	20.3670	2.8237	0.6603	18.4141	2.8833
segresnet	0.7381	19.6969	2.6089	0.6482	18.0439	2.7591
highresnet	0.7209	46.6032	6.4525	0.6381	22.3236	3.6603
nnunet	0.7300	16.9107	7.4144	0.7540	14.4592	2.1459

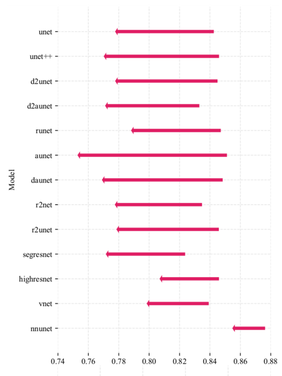
(A)

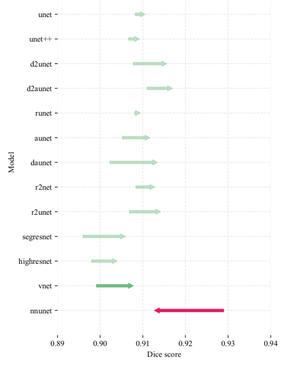
	Pz				Cropped Pz			
	MDS	CI	MHD (mm)	ASD (mm)	MDS	CI	MHD (mm)	ASD (mm)
unet	0.7545	0.0300	22.1193	1.5215	0.7656	0.0326	15.8711	1.1443
unet++	0.7557	0.0363	21.6545	1.4484	0.7666	0.0328	15.8122	1.1031
runet	0.7472	0.0296	23.0029	1.4938	0.7627	0.0243	15.8278	1.1482
aunet	0.7632	0.0280	19.6521	1.2725	0.7611	0.0360	15.0016	1.1684
daunet	0.7576	0.0384	20.1854	1.3109	0.7655	0.0349	15.1258	1.1131
d2unet	0.7597	0.0336	22.9887	1.4304	0.7644	0.0327	15.5292	1.0531
d2aunet	0.7614	0.0219	21.2204	1.5646	0.7677	0.0258	15.0999	1.1449
r2unet	0.7563	0.0280	24.6377	1.2007	0.7700	0.0307	15.4505	1.1609
r2aunet	0.7656	0.0294	19.6629	1.2778	0.7689	0.0343	14.8106	1.1195
vnet	0.7420	0.0315	22.0089	1.7818	0.7573	0.0297	15.9604	1.1659
segresnet	0.7329	0.0285	21.4516	1.4579	0.7456	0.0380	15.9265	1.2347
highresnet	0.7560	0.0301	35.7066	2.2781	0.7719	0.0203	18.4239	1.1279
nnunet	0.8029	0.0063	9.8693	1.0210	0.7686	0.0110	14.4054	1.1633

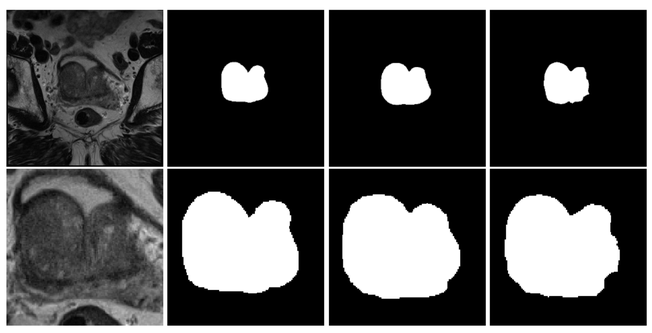
(B)

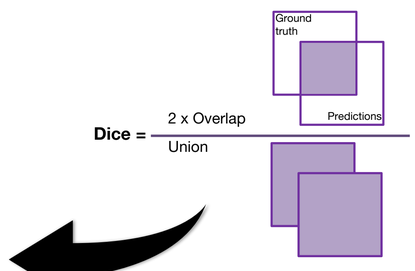
	Pz			Cropped Pz		
	MDS	MHD (mm)	ASD (mm)	MDS	MHD (mm)	ASD (mm)
unet	0.6160	26.9297	2.7719	0.6217	21.3145	2.0339
unet++	0.6129	37.0787	3.5482	0.6064	19.8930	2.3919
runet	0.5828	19.5994	2.8543	0.6104	22.1387	2.4339
aunet	0.6224	20.8184	2.5188	0.6141	20.3642	2.5048
daunet	0.6282	23.0450	3.0015	0.6252	20.2531	2.0680
d2unet	0.6222	18.9975	2.5521	0.6387	16.3358	1.9674
d2aunet	0.6218	26.7974	2.6237	0.6372	17.8484	2.1356
r2unet	0.6273	27.9654	2.5934	0.6308	21.7236	2.2581
r2aunet	0.6228	21.5421	2.7400	0.6300	17.1794	2.0054
vnet	0.5848	40.8916	4.7720	0.5853	23.1373	2.5561
segresnet	0.6106	21.5214	2.5409	0.5339	41.5031	4.4359
highresnet	0.5669	60.2008	5.0504	0.6075	19.7673	2.8573
nnunet	0.6835	13.4527	1.8224	0.6038	17.6085	2.5589



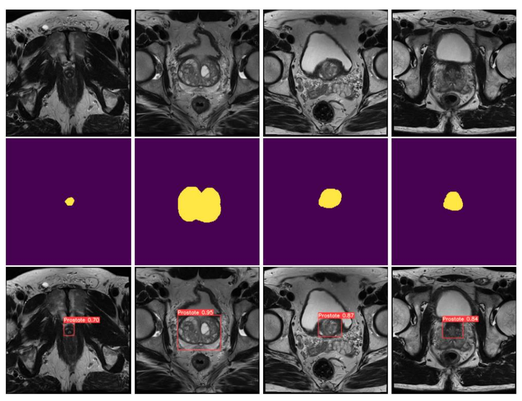




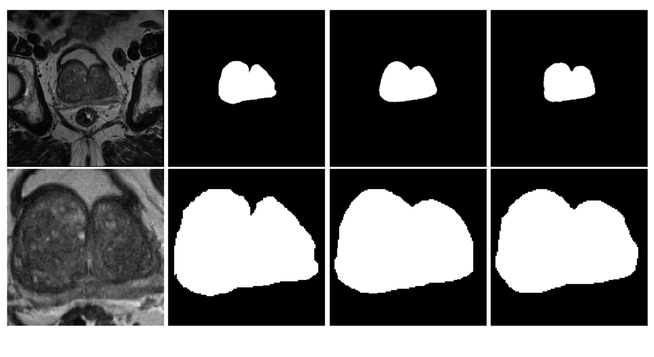


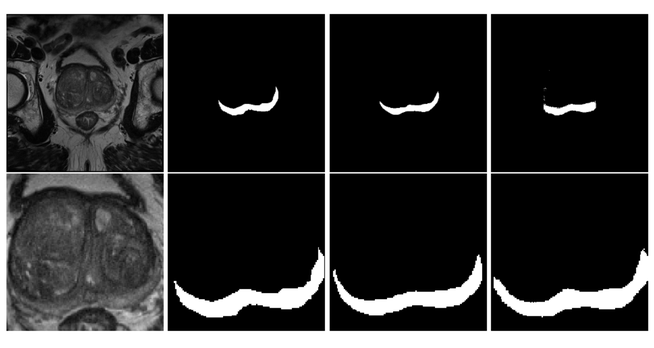


Evaluation



Prostate Detector performance





Conclusions

Rq. 1
The majority of the models exhibit insignificant differences. Nevertheless, **nnUnet** stands out with superior performance across all tasks (full volumes).

Rq. 2
While training the prostate detection model was seamless and yielded outstanding results, it provided no significant aid to the models at the additional cost of fully training another model.
In fact, in one of the tasks all models had a significant drop in performance when using the cropped data.

ULISBOA · Ciências ULisboa · FCT Fundação para a Ciência e a Tecnologia · LASIGE · NOVA IMS Information Management School · Fundação Champalimaud

Acknowledgements: The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952159 (ProCAncer-I). This work was partially supported by FCT, Portugal, through funding of the LASIGE Research Unit (UIDB/00408/2020 and UIDP/00408/2020), and under the project UIDB/04152/2020—Centro de Investigação em Gestão de Informação (MagIC)/NOVA IMS. Nuno Rodrigues was supported by PhD Grant 2021/05322/BD.

LWSEE – Lightweight Secured Software-Based Execution Environment

José Ferreira, Alan Oliveira, Georg Jäger
André Souto, José Cecílio

LASIGE, Faculdade de Ciências da Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

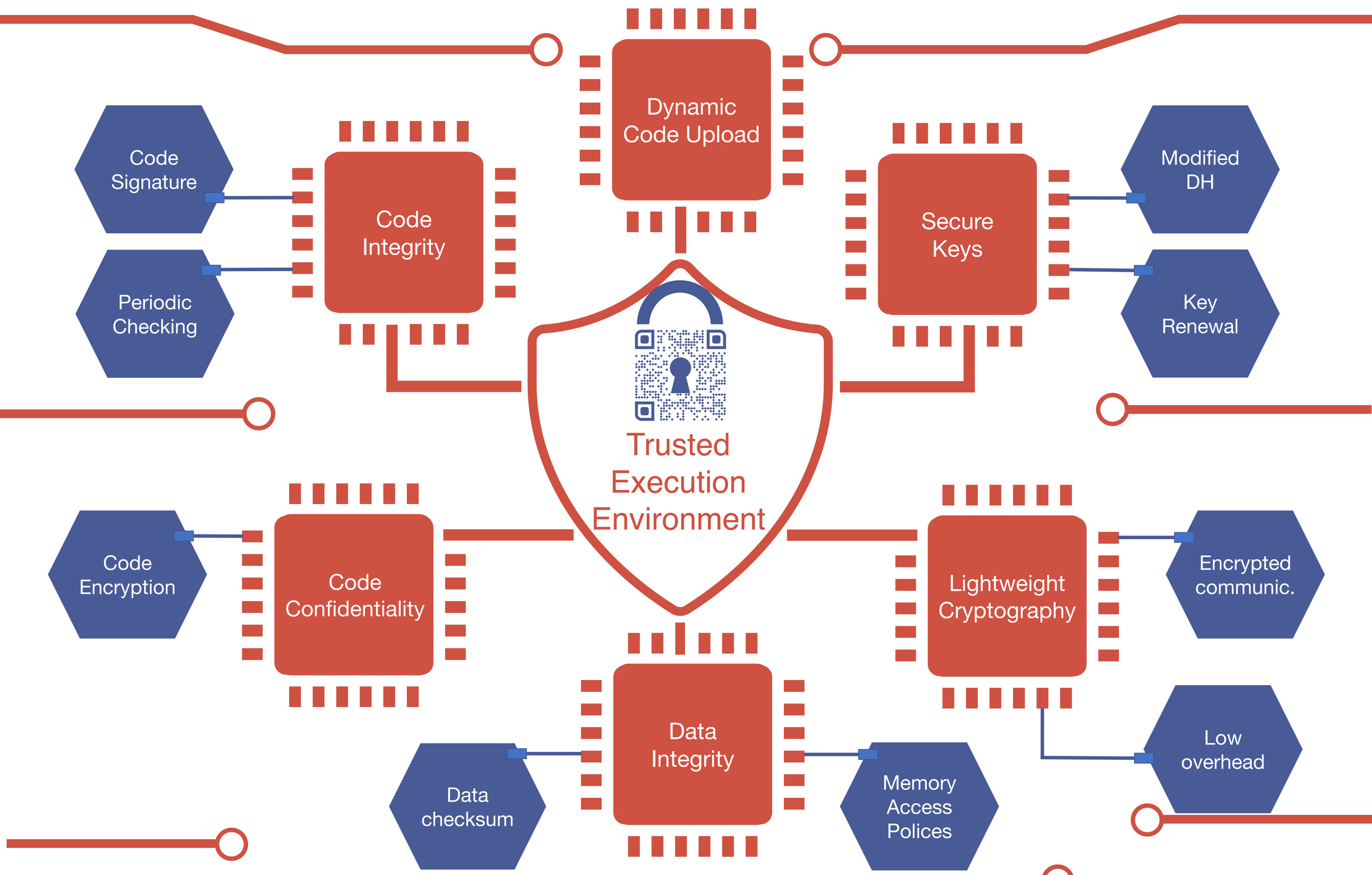
Motivation

- Lack of hardware security features on cost-effective and low-end embedded devices to support hardware-assisted Trusted Executions Environments
- Portably of software-based solutions
- Supported by legacy devices

Objectives

- Develop a Software-based Trusted Execution Environment architecture that:
  - Provides lightweight mechanisms to protect application and data integrity
  - Offers correction in case of unexpected changes
- Validate and assess the performance of the developed architecture

LWSEE Features



LWSEE Attacks protection

- ✓ Man-in-the-middle
- ✓ Replay
- ✓ Integrity of data and code
- ✓ Reverse engineering, code corruption and code reuse
- ✓ Modification of encryption key
- ✓ Data access to I/O operations

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

Acknowledgments: This work was supported by the LASIGE Research Unit (ref.UIDB/00408/2020 and ref. UIDP/00408/2020), and by the European Union's Horizon 2020 research and innovation programme under grant agreement No 871259 (ADMORPH project).

GitUI: A Community-Based Platform to Democratize User Interfaces

Sérgio Alves[1], Ricardo Costa[1], Kyle Montague[2], Tiago Guerreiro[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Northumbria University, United Kingdom

LASIGE accessibility and ageing

Motivation

UI customization can improve usability and user experience.

However, it often requires significant effort and technical skills, beyond the reach of most users.

A Community-Based Approach to Customization

Concept

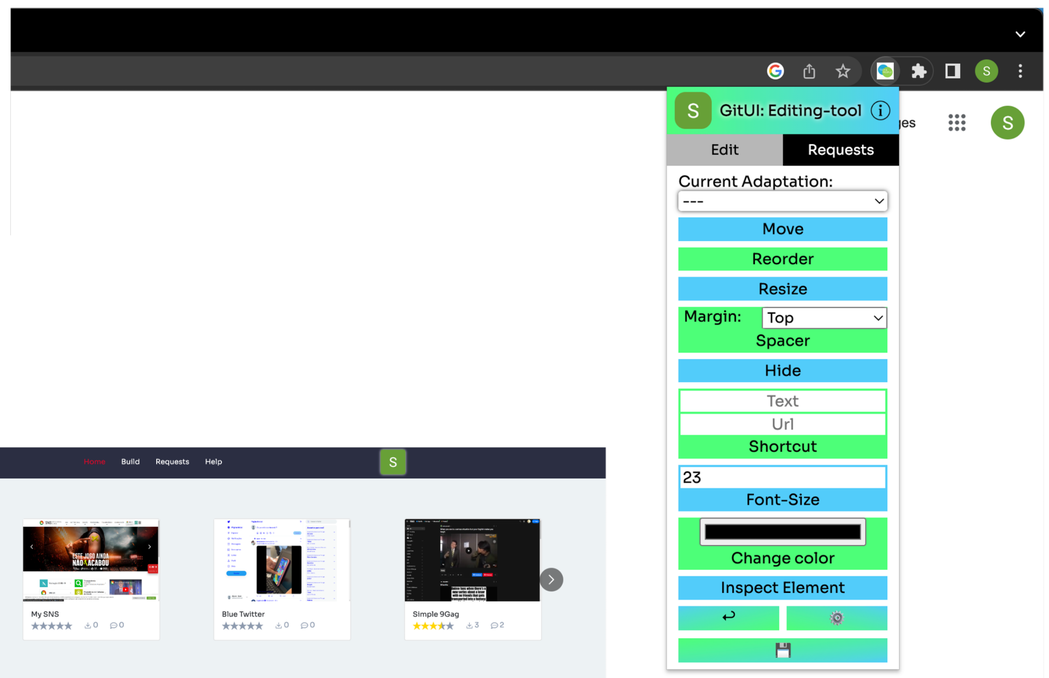
The community collaborates to create customized UIs.

- People can customize themselves, ask for assistance, or search for publicly available customization templates.
- Templates are built atop existing ones.

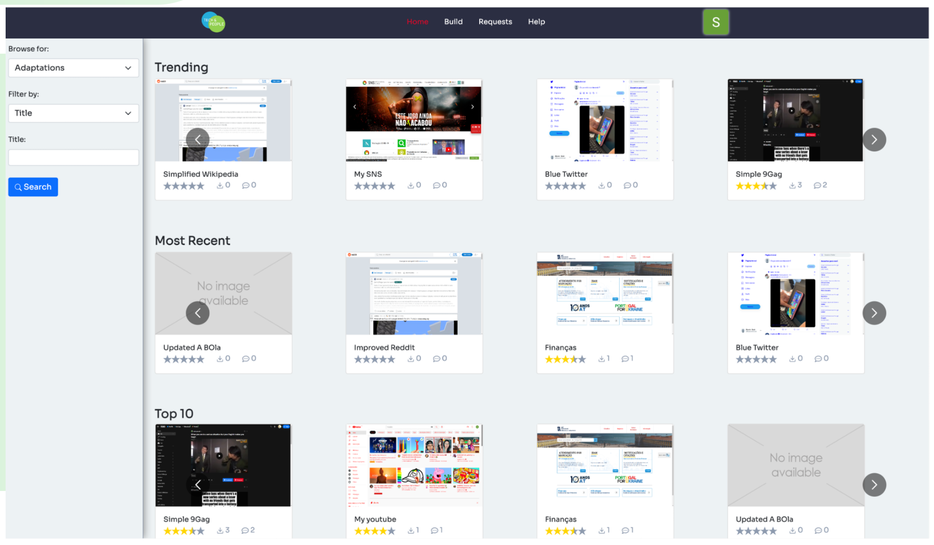
Prototype

GitUI allows website customization. Users can:

- Create, share, and apply templates on any website.
- Create and reply to assistance requests.



a) Customization toolkit



b) Customization repository

Preliminary Study

Procedure

- Two-phase study with 9 participants: people customized for the self and others (two weeks) and explored the Web repository.

Results

- People enjoy customizing for others — motivated by the challenge it represents and the positive feeling of helping others.
- The *shortcut*, *move*, and *change color* operations are particularly valued.
- Future work should leverage the motivation factors (e.g., recognition and gamification) and improve the feedback and communication between users.

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) through LASIGE Research Unit funding refs. UIDB/00408/2020, UIDP/00408/2020 and SFRH/BD/146847/2019.

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

Following the Facts:

Explaining Knowledge Graph Embeddings-based Node Classification

Filipe Paulino, Cátia Pesquita

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE health and biomedical informatics

Background

Explainable AI makes AI systems' results more understandable to humans, critical in high-risk applications

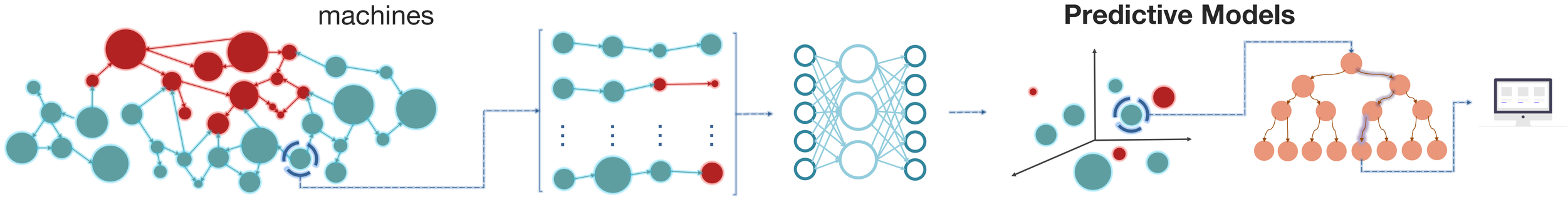
Perturbation-based methods analyse how small changes in the input data impact the outcome of the system

The **Knowledge Graph** (KG) combines **ontologies** of concepts and relations with **data** on real world entities and relations

The KG is easily **understandable by humans** and machines

Knowledge Graph Embeddings (KGE) are **Machine Learning** models that project graph nodes into a **dense vector format**

KGE is frequently used as input for **Node Classification Predictive Models**



Problem

KGE solutions are **black-box** because their meaning and inner workings are too complex to be understood by humans

For Transductive KGE, pertubation-based explanations are hindered by the **costly retrain of Out-of-KG Entities**

Objective

Provide **insight about AI models based on Transductive KGE and downstream Node Classification Models**. With:

- **Local** explanations more suitable for healthcare applications
- **Model-agnostic** approach for a wider range of applications
- **Time efficient** explanations useful in the decision support context

Data

- KG in **RDF graph** format
- Smaller and **general purpose benchmarks**
- Larger and **biomedical domain datasets**

Main principles:

- **Example-based** explanation finds the most relevant **KG facts**
- KGE + Node Classification Model in a **single prediction step**

Methodology

Search space reduction by **pre-selection** of most promising facts considering:

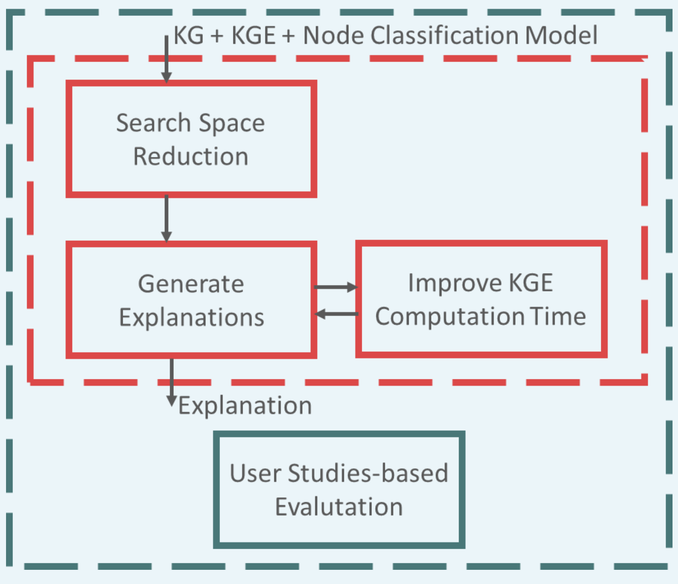
- **Distance, information content** and **KGE additional information**

Improve KGE computation times by testing different approaches:

- **Simplified retraining**, **solutions for evolving networks**, or lightweigtht algorithms

Generate explanations focusing on **necessary and sufficient explanations**

Validation through user studies including metrics for **user satisfaction, understanding and trust**



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

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020, and by the KATY Project.

The KATY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017453.

Inclusive Social Virtual Environments: Exploring the Acceptability of Different Navigation and Awareness Techniques

Inês Gonçalves[1], André Rodrigues[1], Tiago Guerreiro[1], João Guerreiro[1]

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

LASIGE accessibility and ageing

Motivation

Social Virtual Environments are becoming ubiquitous in modern society.

But they are not designed to welcome everyone, with many currently disregarding accessibility.

However, these environments have potential to surpass the accessibility of the physical world, as elements and their locations are known at all times.

- How are different navigation methods in social spaces perceived by people navigating and in conversations?
- Which information should be conveyed in groups and outside?

Our Social Virtual Environment

We developed a Social Virtual Environment composed of 6 scenarios to explore 4 navigation techniques and 2 awareness of group dynamics:

Navigation-Based Scenarios

Free Exploration: Freely move the avatar around the environment, by pressing the arrow keys.

Teleport: Navigate a menu of options to go through the different groups and select the desired group to teleport instantly.

Auto-Walk: Very similar to Teleport, but instead of reaching the desired group instantly, the avatar moves autonomously through the environment.

Co-Pilot: Join in with a second person (i.e. co-pilot represented by an autonomous avatar) who guides the user through the environment.

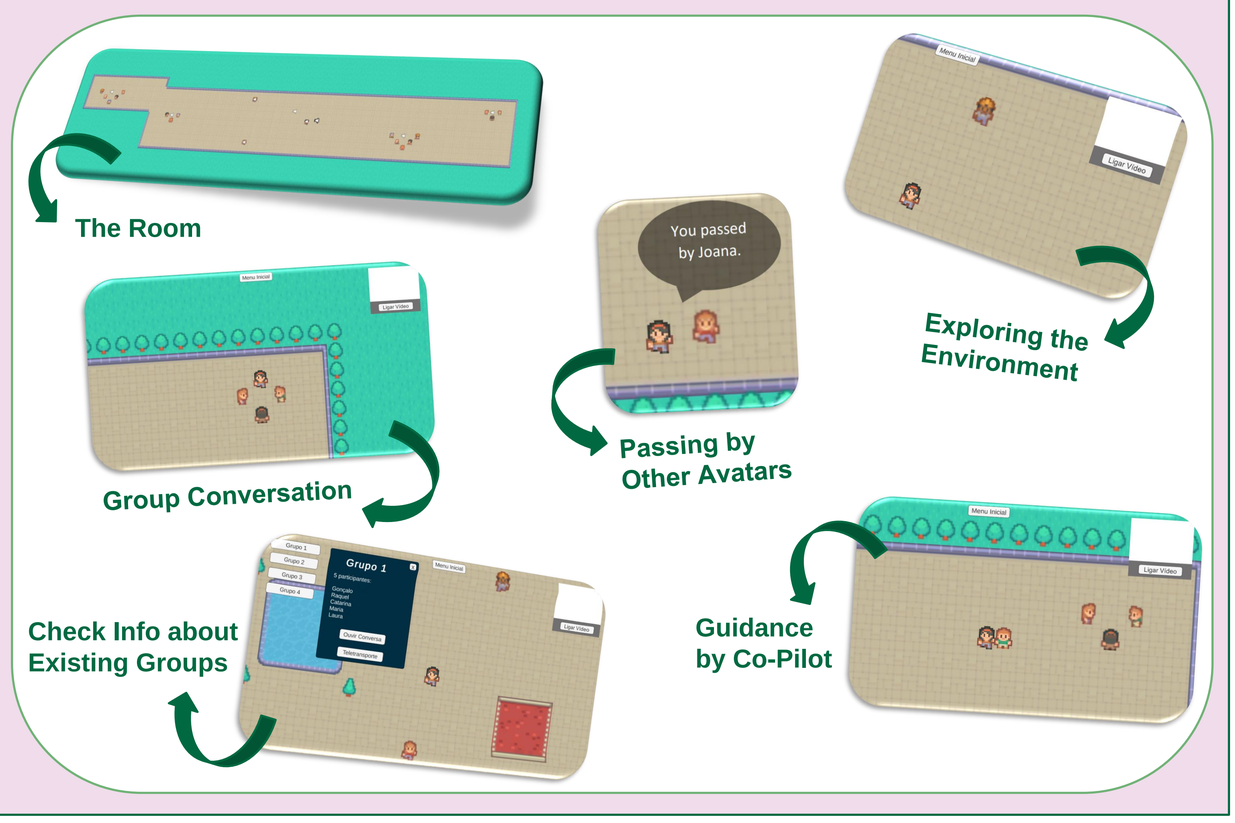
Group Awareness-Based Scenarios

In-Group Footsteps: Perceive other avatars who join the group through the sound of footsteps.

In-Group Teleport: Earcons indicate that someone is teleporting to the group.

It is possible to experience 8 auditory awareness cues that are typically conveyed only visually:

- **Movement:** Footsteps Sounds
- **Collision:** Bumping Sounds
- **Information:** Passer by
- **Notification:** Corresponding keys
- **Environment:** Crowd Noise
- **Distance:** Conversation Previews
- **Presence:** Group Conversations
- **Action:** Joining Group Sounds



User Study

Participants

16 Participants (8 Blind + 8 Sighted): aged 19-63 (M=37.25,SD=13.75), with varied levels of experience with computers and virtual environments.

Procedure

Complete tasks:

Part One: experience different navigation techniques and auditory cues to find a specific group.

Part Two: while in a group, experience a sequence of events to illustrate different navigation and awareness techniques, executed by others.

Findings

- Priviledge autonomy over efficiency
- Trade-offs between difficulty, efficiency, and room awareness
- Flexible navigation, no clear preferred method
- Audio feedback is essential for blind, but redundant for sighted people
- Important to access information, even from afar

Participant's Feedback

- *You can truly feel the sound getting closer and farther away!*
- *How fun!*
- *It really seems like a person is walking near me!*
- *Free Exploration has the charm of being a challange, allowing you to know the configuration of the room!*
- *Teleport is my dream!*
- *This is so cool!*
- *I want to have this on my computer!*
- *This is really interactive!*

Scan me to access the paper!

Acknowledgements: This work was supported in part by FCT through the LASIGE Research Unit (UIDB/00408/2020, UIDP/00408/2020), the AccessVR project (2022.08286.PTDC), and the PlayFam project (2022.08895.PTDC).

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

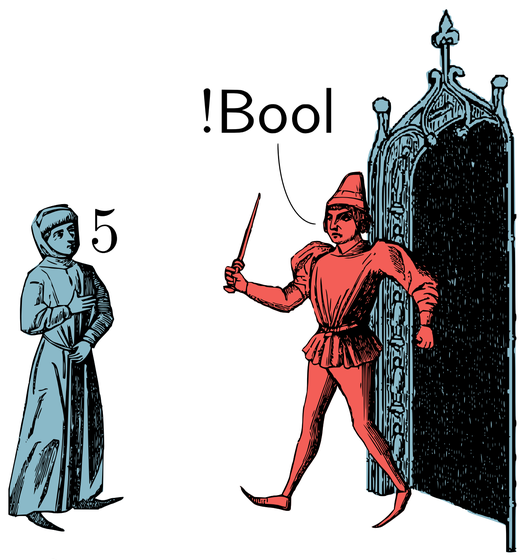
Subtyping Context-Free Session Types

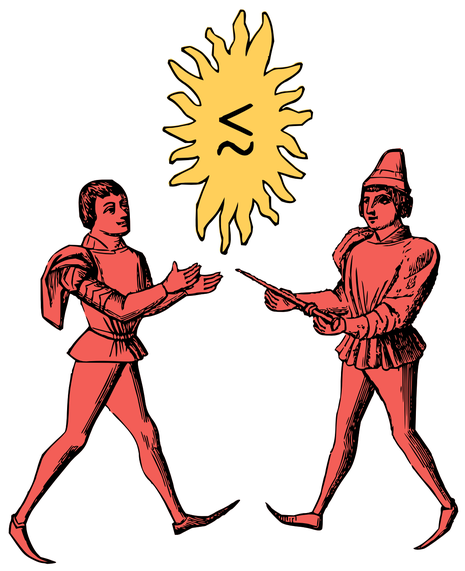
Gil Silva, Andreia Mordido, Vasco T. Vasconcelos

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

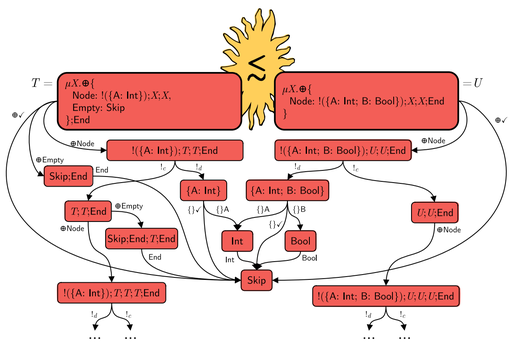
LASIGE reliable software systems

Types are a valuable tool for the static verification of programs, allowing the detection of many errors at compile-time. *Session types* specialize in the verification of message-passing concurrent programs. They are attached to communication channels, and are used to specify and enforce messaging protocols. *Context-free session types* are their recent, most expressive incarnation yet. Their protocols correspond to *simple deterministic context-free languages*.





Equivalence is not trivial for context-free session types. It relies on *observational equivalence*, or *bisimulation*: two types are considered equal if they exhibit exactly the same communication behavior (even if they are syntactically different). The behavior of types is specified by a *labelled transition system*.

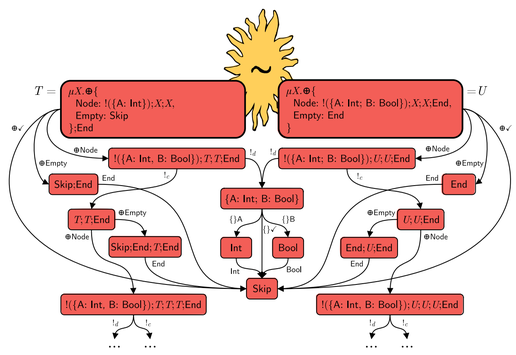


By adapting an existing type bisimilarity algorithm to our new kind of simulation, we obtain a subtyping algorithm capable of handling both functional and higher-order context-free session types. The algorithm can be described in three main steps:

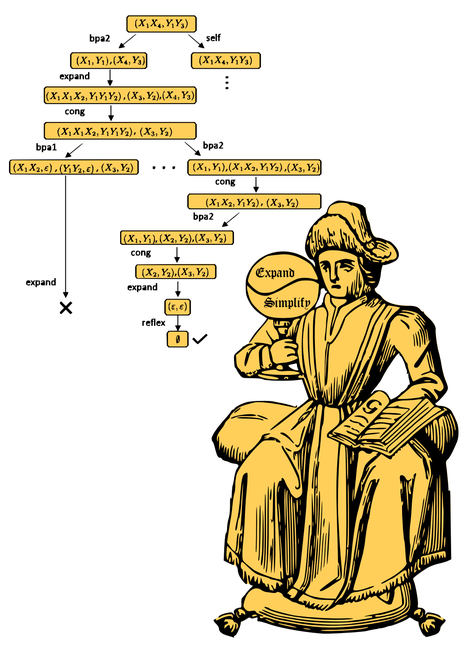
1. Convert the given types to words from a simple deterministic context-free grammar in Greibach Normal Form;
2. Prune unreachable symbols from the grammar;
3. Explore an expansion tree, alternating between expansion and simplification steps, until an empty node is reached (**True**) or expansion fails (**False**).

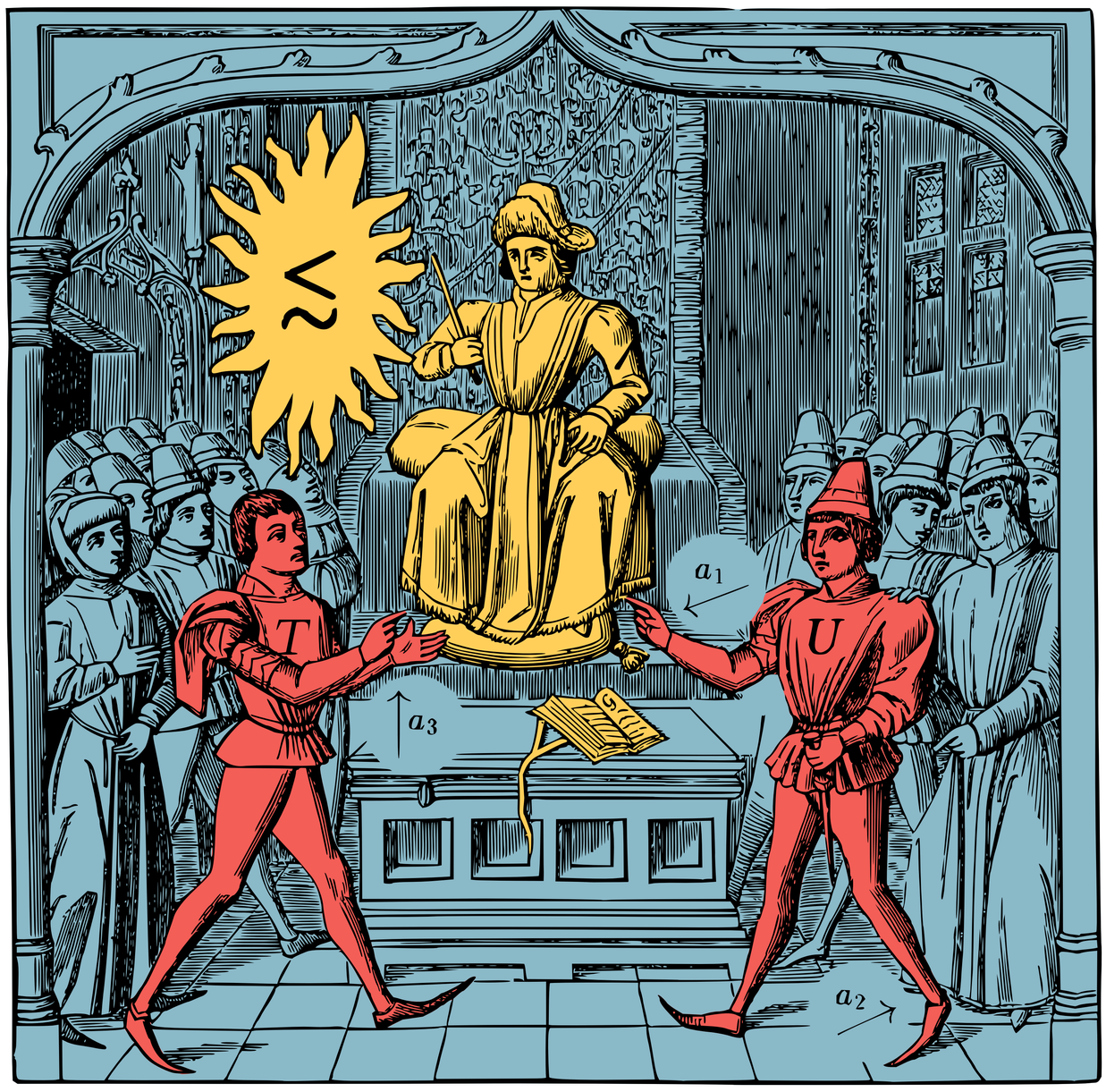
Subtyping formalizes the intuition that some types may be safely used in place of others, making type systems more flexible and type specifications more precise. When applied to session types, subtyping promotes *modular development* by allowing compatible changes in communicating agents' views of protocols. In this project we:

1. Look for a notion of subtyping suitable for both functional and context-free session types;
2. Use it to develop an algorithm;
3. Implement the algorithm in the FreeST programming language.



We follow a similar approach to define our notion of subtyping, relying instead on a novel form of *observational preorder* we call *XYZW-simulation*, which allows us to specify precisely how the behavior of subtypes may differ from that of supertypes.





FreeST

is a functional, concurrent programming language where processes communicate via message-passing on channels governed by context-free session types.

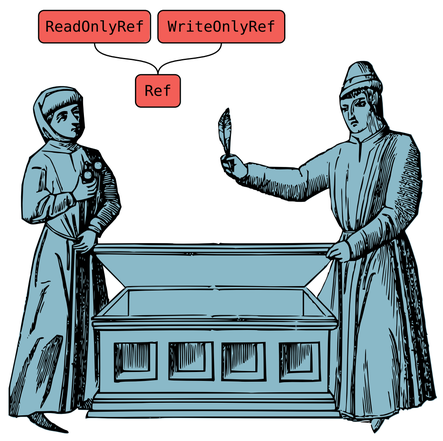
Ref.fst

```
-- Type of the reference cell
type RefSession = +{Write: !Int, Read: ?Int};End
type Ref = *?(RefSession)

-- Types with downgraded capabilities
type ReadOnlyRef  = *?(+{Read : ?Int};End)
type WriteOnlyRef = *?(+{Write: !Int};End)

-- Constructs a reference cell.
ref : Int -> Ref
ref n = forkWith (runServer refServer n)

-- Handles cell state.
refServer : Int -> dualof RefSession 1-> Int
refServer v (Write c) = receiveAndClose c
refServer v (Read  c) = c |> send v |> close; v

-- Stores a value.
write : Int -> WriteOnlyRef -> ()
write v r =
  r |> receive_ |> select Write |> send v |> close

-- Returns the stored value.
read : ReadOnlyRef -> Int
read r =
  r |> receive_ |> select Read |> receiveAndClose

-- Examples
main : Int
main = let x = ref 0 in
       let y = ref 5 in
       let z = x     in
       fork (\_ -> write 2 z);
       fork (\_ -> write 3 x);
       write 6 y;
       read x + read y -- what's the result?

-- Error! Function has write permission only.
f : WriteOnlyRef -> ()
f r = write (read r + 1) r
```

Try FreeST on your browser!
rss.di.fc.ul.pt/tryit/FreeST
Downloads, tutorials and more
freest-lang.github.io

FreeST is not purely functional, as channel-based communication requires side-effects. Another effectful feature present in several functional languages is *mutable storage* (a.k.a. *references* in OCaml). FreeST does not come with this feature out of the box, but we can simulate it using threads and channels, as the program on the left shows.

This program also shows how subtyping can be useful in FreeST. In this case, the subtype hierarchy formed by the reference types naturally acts as a mechanism to control read and write permissions.



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

Acknowledgements
This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020 and the SafeSessions project (PTDC/CCI-CIF/6453/2020).

Illustrations adapted from "Manners, Customs, and Dress During the Middle Ages, and During the Renaissance Period" by Paul Lacroix and "Structure and Interpretation of Computer Programs" by Harold Abelson and Gerald Sussman.

WildKey Suite: a Privacy-Aware Toolkit for Data Collection In-the-Wild

André Rodrigues, André Santos, Diogo Branco, Tiago Guerreiro
LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

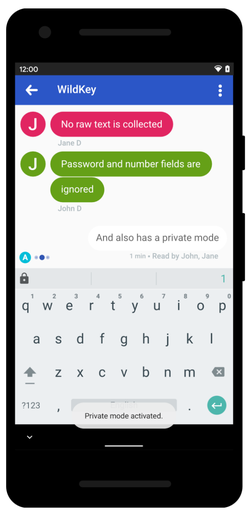
LASIGE health and biomedical informatics

Smartphones are becoming **an extension of oneself, and data privacy is of utmost importance**.

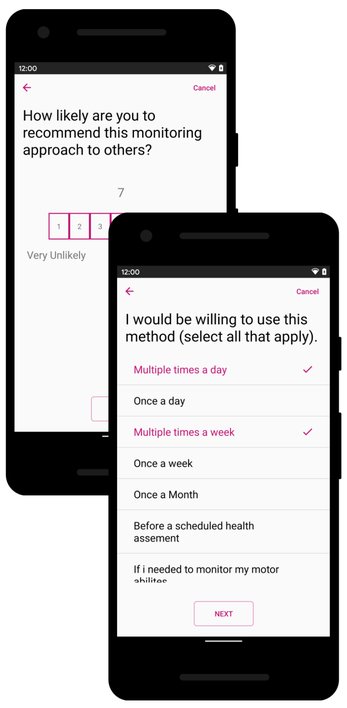
Keyboard

Collect everyday typing data, enabling continuous and spontaneous assessments, without collecting any sensitive information.

Data	Implicit	Transcription	Composition
Speed			
Words per Minute	✔	✔	✔
Time per Word	✔	✔	✔
Errors			
Corrected Error Rate	✔	✔	✔
Uncorrected Error Rate	✔	✔	✔
Total Error Rate	✔	✔	✔
Insertion Error Rate		✔	✔
Omission Error Rate		✔	✔
Substitution Error Rate		✔	✔
Error Correction Attempts	✔	✔	✔
Touch Dynamics			
Flight Time	✔	✔	✔
Hold Time	✔	✔	✔
Touch Major/Minor	✔	✔	✔
Touch Offset	✔	✔	✔
Key Selected		✔	✔
Motion Info		✔	✔
Timestamp	✔	✔	✔
Action and Character Counts			
Action Count	✔	✔	✔
Correction Action Count	✔	✔	✔
Entry Action Count	✔	✔	✔
Number of Auto Corrects	✔	✔	✔
Number of Changed Characters	✔	✔	✔
Number of Selected Suggestions	✔	✔	✔
Number of Written Characters	✔	✔	✔
Number of Written Numbers	✔	✔	✔
Number of Written Special Characters	✔	✔	✔



Embed questionnaires



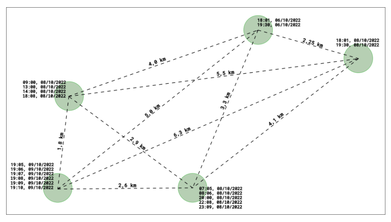
Social

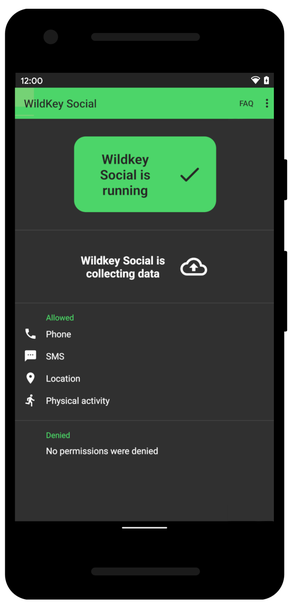
Collect location, used app categories, physical activity, ambient light, screen states, call and SMS events.

Location Data

What you do

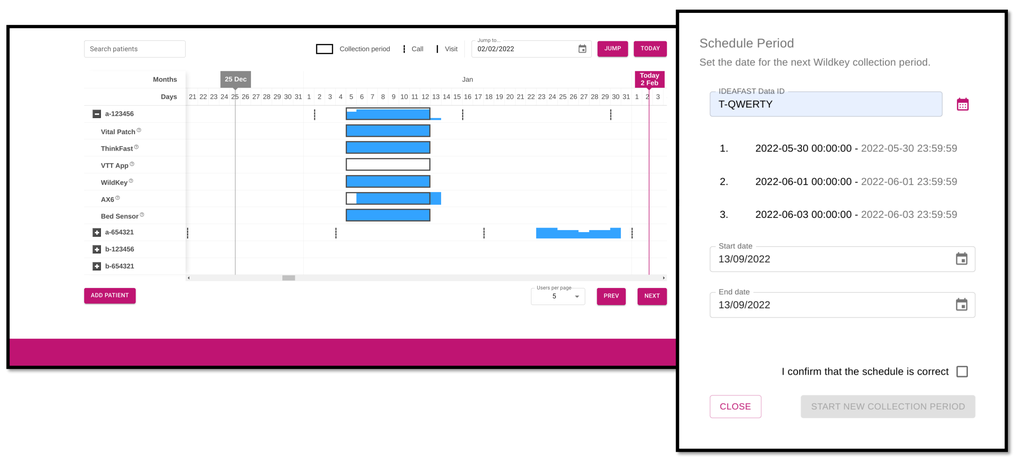
What we get





Dashboard

Manage your study, check data coverage and schedule technology wear periods



Anonymized Data, Right to be Forgotten

WildKey analyzes all data locally on the device before it synchronizes computed metrics.

Track Record

Used by over *250 participants* with a variety of neurodegenerative diseases.

Two completed studies with two ongoing, used in 10+ countries, 19 clinical sites

Best paper award for its debut paper *"Investigating the Tradeoffs of Everyday Text-Entry Collection Methods"*

André Rodrigues, Hugo Nicolau, André Santos, Diogo Branco, Jay Rainey, David Verweij, Jan Smeddinck, Kyle Montague, Tiago Guerreiro. "Investigating the Tradeoffs of Everyday Text-Entry Collection Methods" CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems.

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

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020, by 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.

Comparative analysis of oral fibroblasts from Human and Mouse by scRNA-seq

Alexandre Fernandes[1], Diana Pereira[2], Inês Sequeira[2], Márcia Barros[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Institute of Dentistry, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, 4 Newark Street, London E1 2AT, United Kingdom

LASIGE health and biomedical informatics

Background

The wound healing process has differences between tissues. The oral tissue is characterized for scarless and rapid wound healing. Several studies have shown that fibroblast heterogeneity is the principal factor for these characteristics.

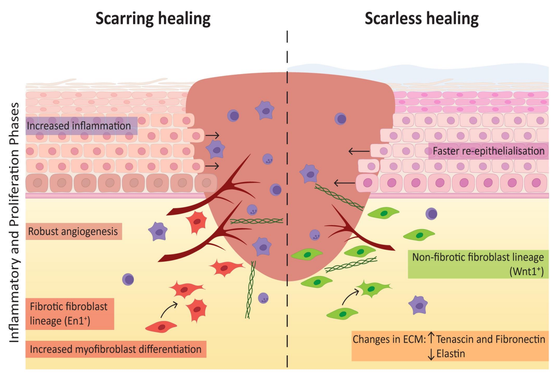


Figure 1 – Skin and oral mucosal healing comparison. Pereira & Sequeira, *Front. Cell Dev.* Biol. (2021)

Objectives

- Use bioinformatics tools applied to omics approaches to understand the differences between subpopulations of fibroblasts.
- Characterization of fibroblast subpopulations in each specie.
- Comparation of fibroblast subpopulations in human and mouse.

Methodologies

Step 1 – Data acquisition

Human Dataset

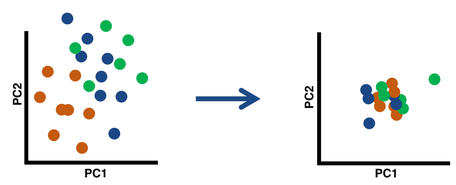
Download of Single Cell RNA sequencing dataset (10x Genomics) deposited on GEO (Accession number GEO: GSE164241

Mouse Dataset

Single Cell RNA sequencing dataset (10x Genomics) generated in our lab from *Mus musculus* buccal tissues.

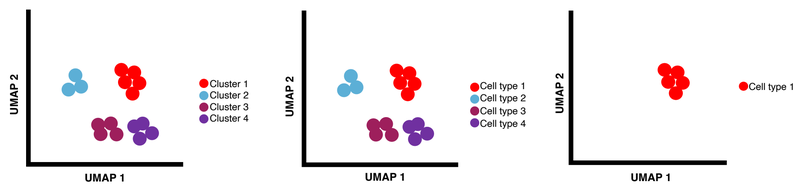
The CellRanger step was performed by the dataset authors

Step 2 – Data Cleaning



Filtered all cells with low counts, doublets removed, regressed out, normalized and logarithmized

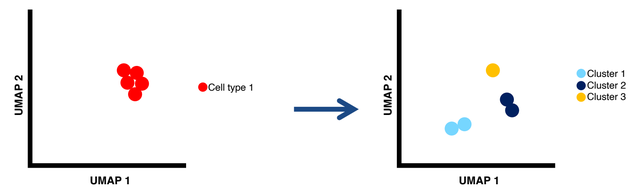
Step 3 – Data annotation and subsetting



Clustering: Leiden Algorithm → Manual annotation → Subsetting

Using highly expressed genes, rank genes groups and specific cell type markers

Step 4 – Data re-clustering



After performed the subsetting, was performed a new clustering using the Leiden algorithm

Further steps

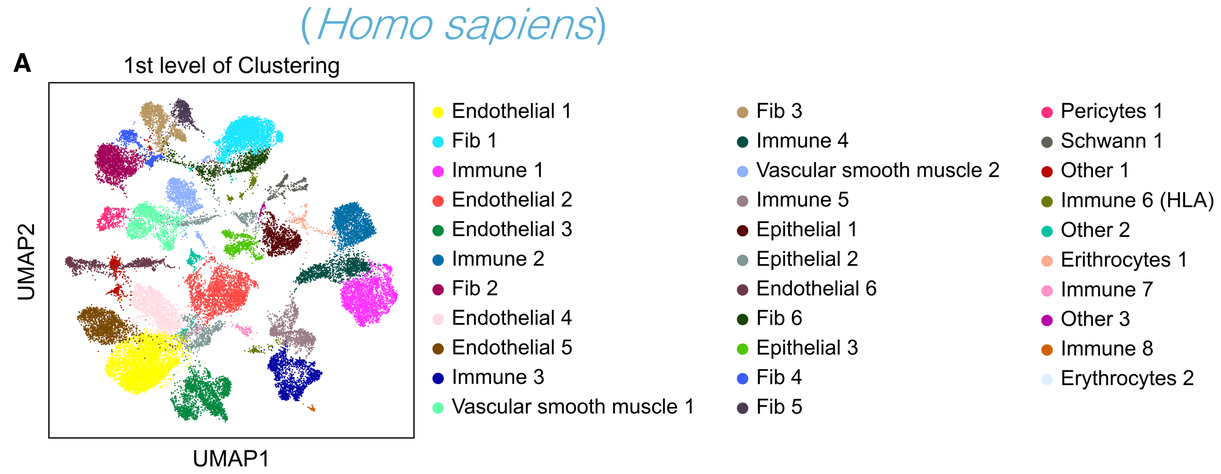
We will analyse each cluster in detail, using different fibroblast markers and comparing they expression between clusters in the same specie. Therefore, we expect to characterize new subpopulation of oral fibroblast. At the end of this work, we will try to understand the difference between species, trying to find homologous subpopulations. This knowledge is important to improve new approaches in the wound healing process.

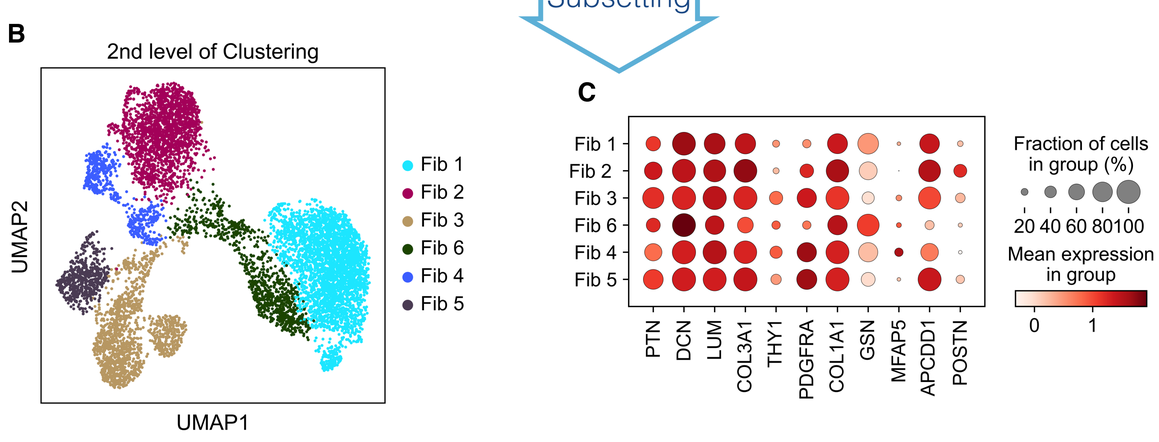
scanpy

For all steps was used Scanpy (v.1.9.1) package for Python environment (v.3.10.9)

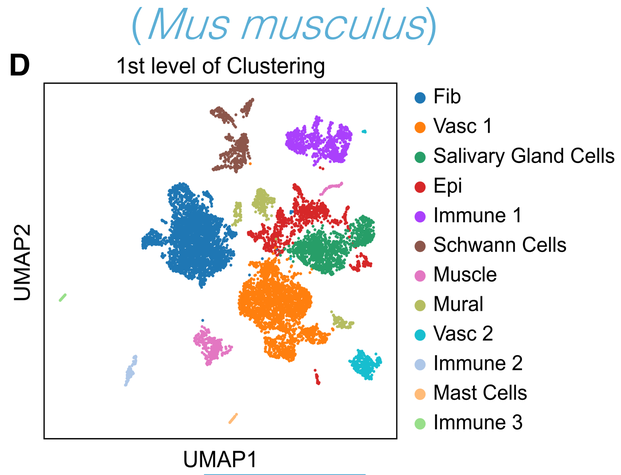
Results

Human (*Homo sapiens*)





Mouse (*Mus musculus*)



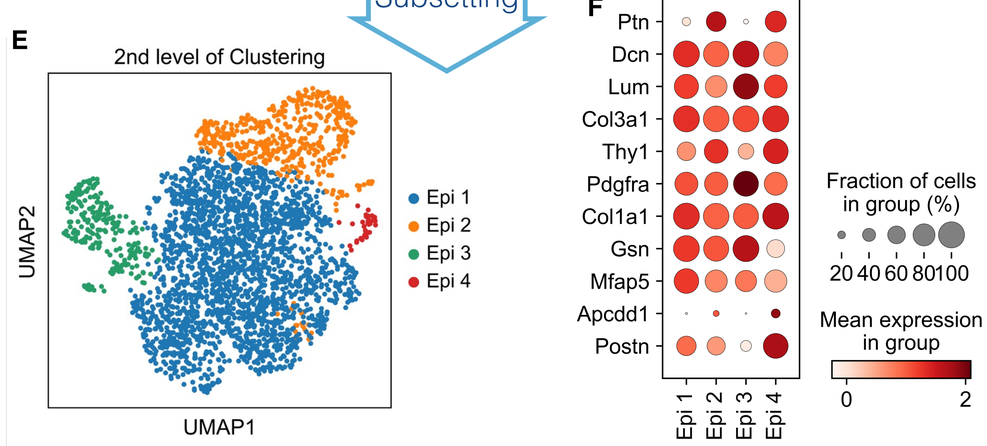


Figure 2 – **(A)** Human and **(D)** Mouse UMAP visualization of oral buccal scRNA-seq datasets. **(B)** Human and **(E)** Mouse UMAP visualization of the fibroblast second level of clustering, resolution = 0.15. **(C)** Human and **(F)** Mouse dotplot of specific fibroblats gene markers

1. Pereira D, Sequeira I. A Scarless Healing Tale: Comparing Homeostasis and Wound Healing of Oral Mucosa With Skin and Oesophagus. Front Cell Dev Biol. 2021 Jul 26;9:682143. doi: 10.3389/fcell.2021.682143. PMID: 34381771; PMCID: PMC8350526.
2. Ascensión AM, Fuertes-Alvarez S, Ibañez-Solé O, Izeta A, Araúzo-Bravo MJ. Human Dermal Fibroblast Subpopulations Are Conserved across Single-Cell RNA Sequencing Studies. J Invest Dermatol. 2021 Jul;141(7):1735-1744.e35. doi: 10.1016/j.jid.2020.11.028. Epub 2020 Dec 30. PMID: 33385399.
3. Williams DW, Greenwell-Wild T, Brenchley L, Dutzan N, Overmiller A, Sawaya AP, Webb S, Martin D; NIDCD/NIDCR Genomics and Computational Biology Core; Hajishengallis G, Divaris K, Morasso M, Haniffa M, Moutsopoulos NM. Human oral mucosa cell atlas reveals a stromal-neutrophil axis regulating tissue immunity. Cell. 2021 Jul 22;184(15):4090-4104.e15. doi: 10.1016/j.cell.2021.05.013. Epub 2021 Jun 14. PMID: 34129837; PMCID: PMC8359928.

ULISBOA | Ciências ULisboa | FCT Fundação para a Ciência e a Tecnologia | LASIGE | Barts and The London Queen Mary's School of Medicine and Dentistry

Acknowledgements:
This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020, by the European Commission.

A consumer-oriented mobile application for assessment of household appliances

Catarina Moita, Zygimantas Jasiunas
Ana Paula Afonso, José Cecílio, Pedro M. Ferreira

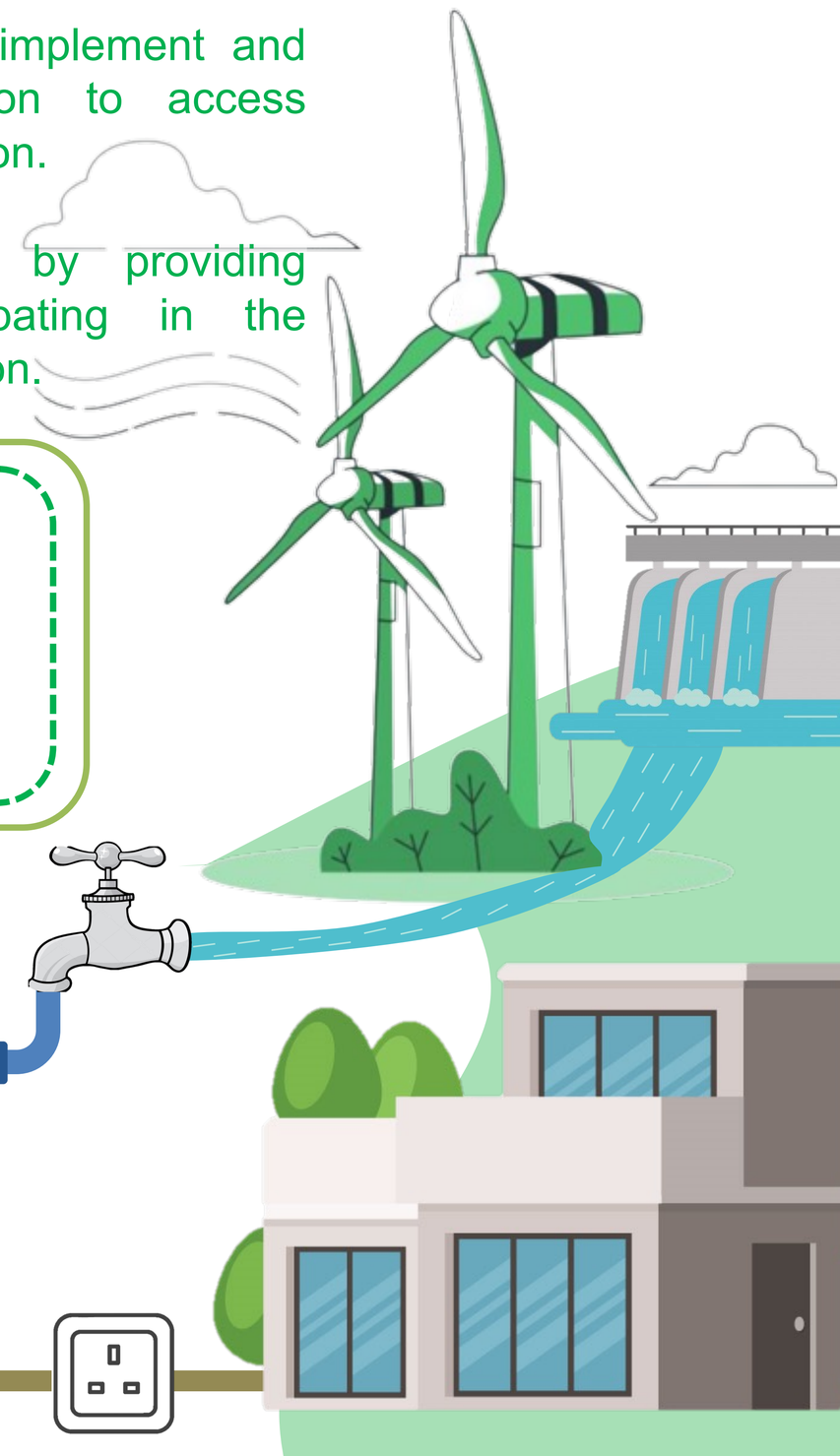
LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

Energy efficiency and performance indicators in European Union energy labelling regulations do not reflect real-life utilization of appliances. The SATO project is developing energy efficiency and performance assessments capable of expressing real-life utilization and life-cycle evolution.

This project aims to design, implement and evaluate a mobile application to access household appliances information.

Worten supports this work by providing appliances, and by participating in the application design and evaluation.



SATO platform

- Contextual information
- Assessments
- Worten API
- Mobile app API
- Data storage

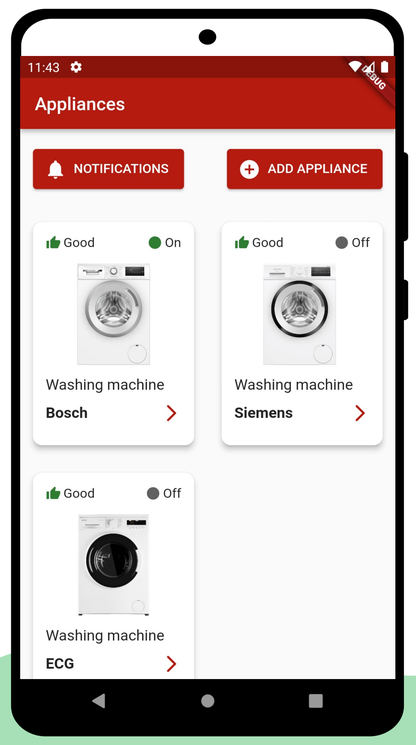
Data acquisition

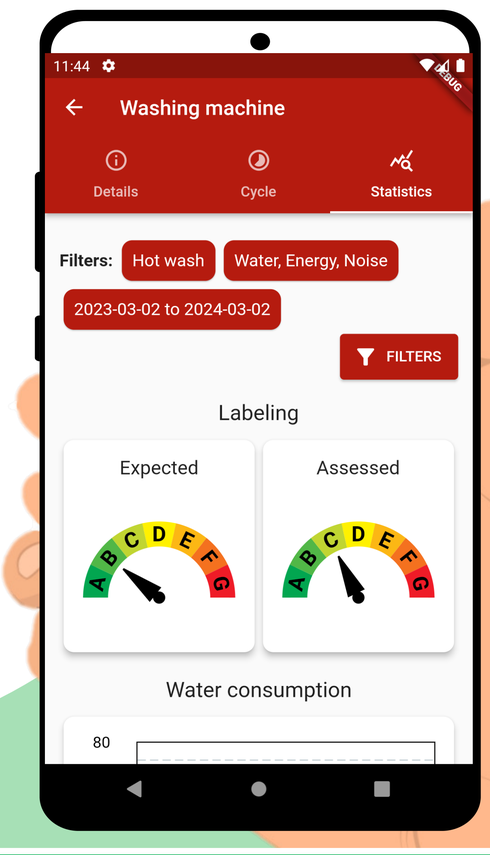
Temperature, vibration, water flow, energy consumption

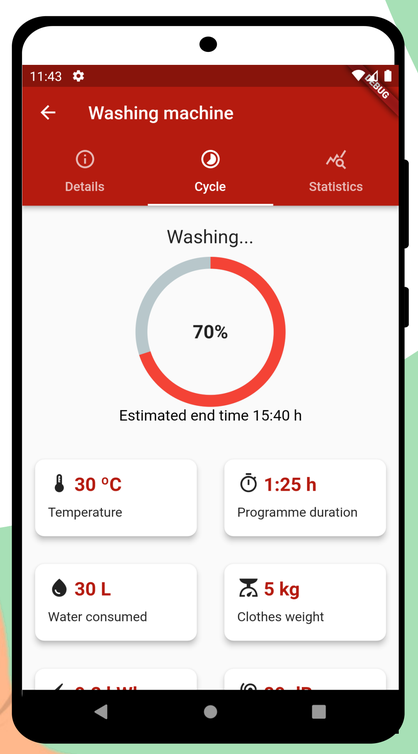
worten

- Data analysis
- Notifications

- Add a new appliance
- Visualization of appliance's statistics
- Data filtering
- Comparison between real-life energy label and EPREL
- Notifications







- UI simple and intuitive
- UI responsive and adaptable
- Support for multiple platforms (Android and iOS)
- Security and privacy respected

Performance assessment

Weekly, monthly and yearly energy consumption. Utilization patterns and statistics

Real-life energy label

Real-life energy efficiency label and indicators based on real-life utilization of appliance

Fault / anomaly detection

Detection of abnormal efficiency and performance indicators (values out of range, life-cycle drift)

U LISBOA · Ciências ULisboa · FCT Fundação para a Ciência e a Tecnologia · LASIGE driven by excellence · SATO OF BUILDING ENERGY

Acknowledgements: This work was supported by the European Commission through the Sato project (Grant agreement ID: 957128) and the LASIGE Research Unit, ref. UIDB/00408/2020 and ref. UIDP/00408/2020.

Designing a theory-driven, evidence-based and user-centred web application to simulate behaviour change support in persons with chronic disease

Diogo Branco[1], **Duarte Taranta**[1], **Ricardo Santos**[1], **Ana Paula Cláudio**[1], **Beatriz Carmo**[1], Gregor Stiglic[2], **Luís Correia**[1], Nino Fijacko[2], Isa Brito Félix[3,4], Anastasiya Barabash[5], Margarida Henriques[5], Mustafa Abacioglu[6], Mara Pereira Guerreiro[3,4]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Univ. of Maribor, Slovenia, [3]Egas Moniz School of Health & Science, [4] Egas Moniz Center for Interdisciplinary Research, [5]Escola Sup de Enfermagem de Lisboa, [6]Marmara University, Turkey

LASIGE data and systems intelligence

1. Background and aim

Healthcare and other professionals are expected to support behaviour change for the self-management of chronic disease

The Train4Health Project addresses this challenge through an innovative educational package

Undergraduate students present a skills gap in behaviour change support, resulting from curricular insufficiencies and limited training opportunities

Aim
Reporting the design of a web-based application for training behaviour change support

2. Design

Literature search, co-creation with users and requirements prioritisation

- Evidence on design elements and components were retrieved from the literature
- Twelve international focus groups were conducted with students and educators from pharmacy, nursing and sport sciences
- Requirements were prioritised through the Moscow approach

Prior project work informing software design

- Interprofessional European competency framework to support behaviour change in chronic disease
- List of core behaviour change techniques from an established taxonomy (BCTTv.1)1
- Learning outcomes-based curriculum derived from the competency framework.
- Persons' profiles developed as part of case studies on behaviour change in chronic disease

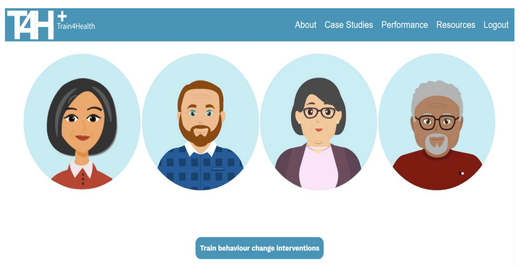
3. Results

The use of virtual patient simulation and gamification appear promising in health professionals' education[2, 3]. Selected gamification features include points, levels, stats, acknowledgement and progression.

The self-determination theory enables an understanding of how gamification may enhance engagement and motivation, whilst helping avoid pitfalls in implementation.

Requirements elicited from focus groups (e.g. gamification, feedback) and team discussion were prioritised into "must have", "should have" and "won't have".

1

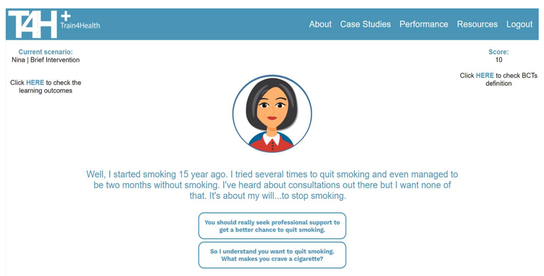


Web-based application depicting four persons with chronic disease as 2D virtual humans (VHs), presented in **levels** with increasing complexity. Users interact with the VHs playing the role of a professional supporting behaviour change.

2

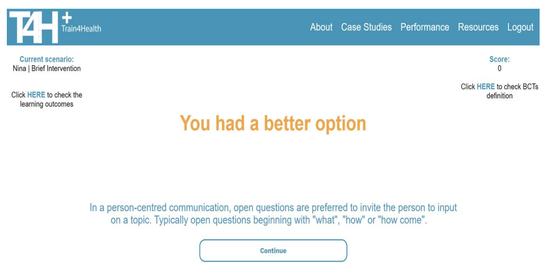
Each person profile offers training in long and brief behaviour change interventions. Learning outcomes are presented for each session.

3



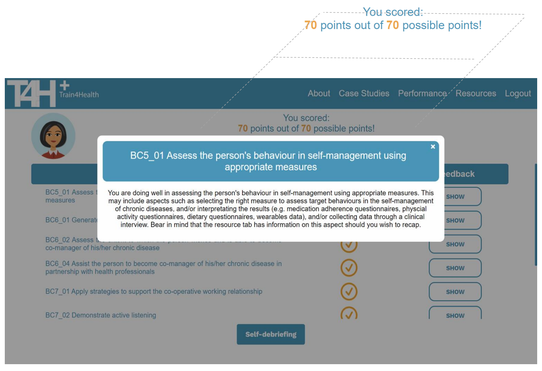
In each session, VHs communicate with users through a synthetic voice and facial expressions; input from users consists of buttons depicting two options.

4



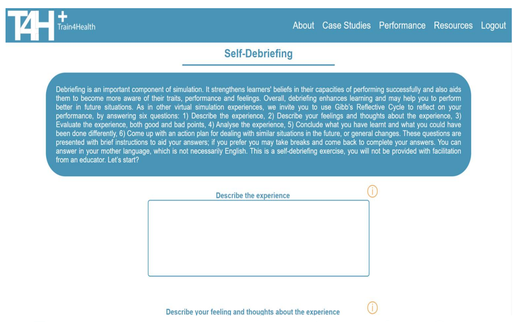
The choice of the less correct option prompts immediate feedback within each session.

5



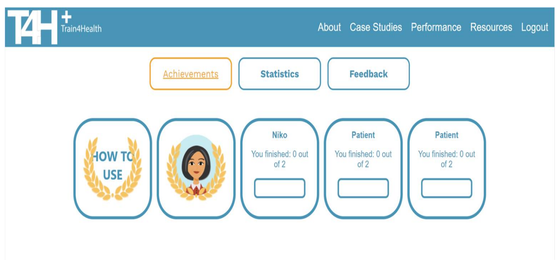
Once the session is finished, users firstly receive feedback for each learning outcome, plus **points** providing information about the user's performance.

6



Then users are prompted to self-debrief, via Gibbs' reflective cycle.

7



The "performance" menu deploys gamification features, such as **stats**, **acknowledgement** (badges) and **progression**.

4. Conclusion

The design of the web-based application includes simulation with 2D virtual humans and gamification, underpinned by the self-determination theory, plus components such as feedback and self-debriefing. Development is complete and the application is being tested iteratively with students and educators.

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020. This project has also received funding from the Erasmus+ Programme of the European Union under the grant agreement no. 2019-1-PT01-KA203-061389. The European Commission's support for the production of this publication does not constitute an endorsement of the contents, which reflect the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

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

Deep learning interpretation to understand RNA splicing

Pedro Barbosa[1,2], Rosina Savisaar[3], Maria Carmo-Fonseca[2], Alcides Fonseca[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Instituto de Medicina Molecular João Lobo Antunes, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal
[3] Mondego Science , Coimbra, Portugal

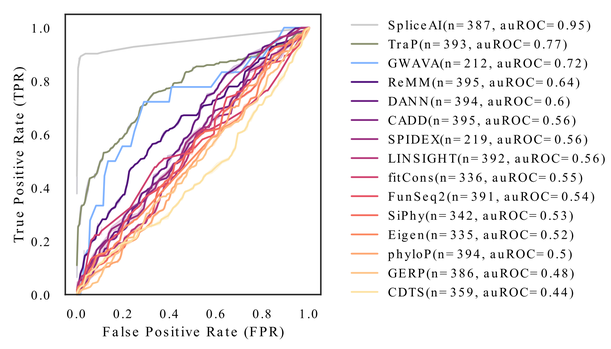
LASIGE data and systems intelligence

Instituto de Medicina Molecular João Lobo Antunes

Motivation

RNA Splicing involves the removal of introns from pre-mRNA molecules, which allows exons to be joined and functional proteins to be produced. It is an essential process for all living cells, and mutations in DNA affecting this process frequently cause human disease.

SpliceAI, an artificial neural network, has achieved impressive accuracy on identifying splicing-associated disease-causing mutations.



Complete benchmark paper:

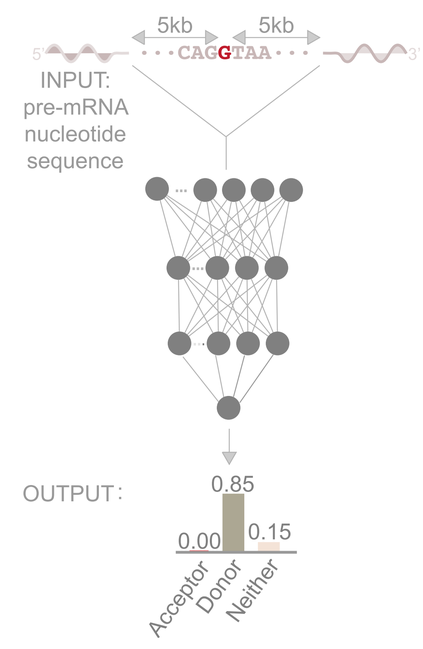
Research question

SpliceAI was trained using only raw mRNA sequences. Did this model learn combinatorial rules of splicing regulation?

Hypothesis

Combining of knowledge about splicing regulators and explainable AI techniques enhances the comprehension of the splicing logic learned by SpliceAI.

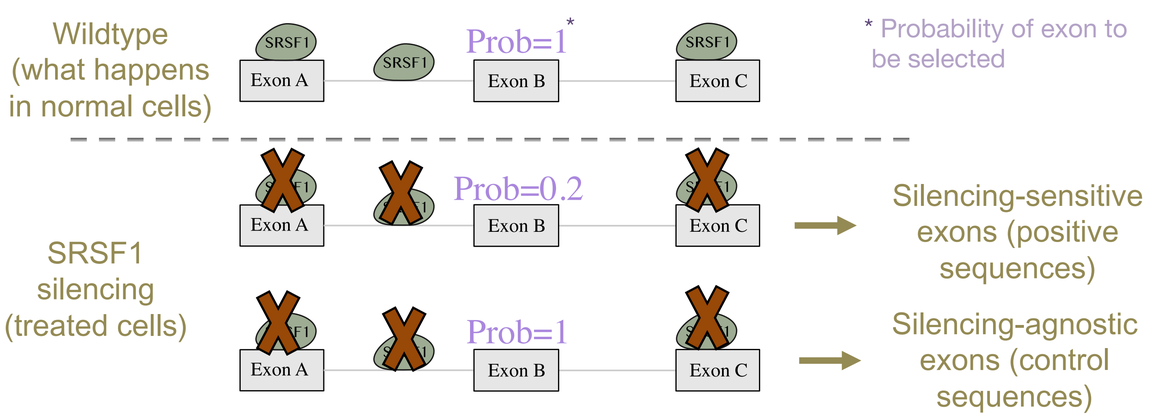
Ultimate goal
Use SpliceAI as a source of knowledge to generate new biological insights to be tested experimentally.



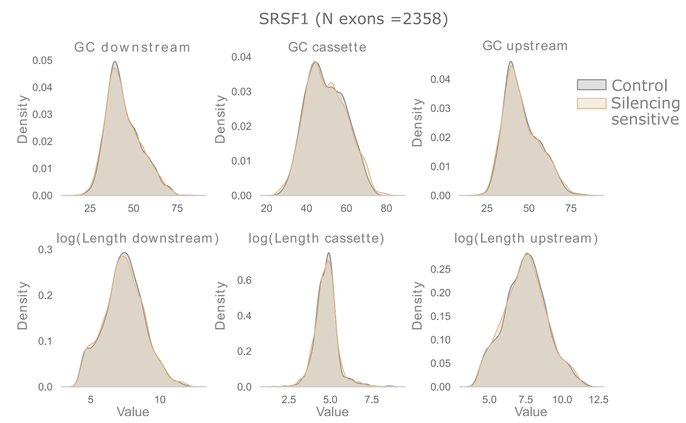
Generation of high-quality datasets to study the sensitivity of SpliceAI to specific motifs

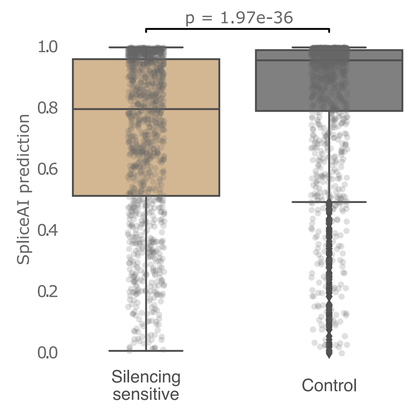
We re-analyzed publicly available RNA sequencing datasets to identify regions of the genome (exons) that exhibited differential selection for protein production when a single gene was silenced experimentally.

The idea was to generate datasets of genomic sequences that were either sensitive or insensitive to the silencing of other genes. To illustrate this approach, we used as example the SRSF1 gene:



To ensure that further analysis of SpliceAI predictions on these sequences was not confounded by genome architecture or nucleotide composition features (which are known to influence the splicing outcome), paired-control exons were extracted.



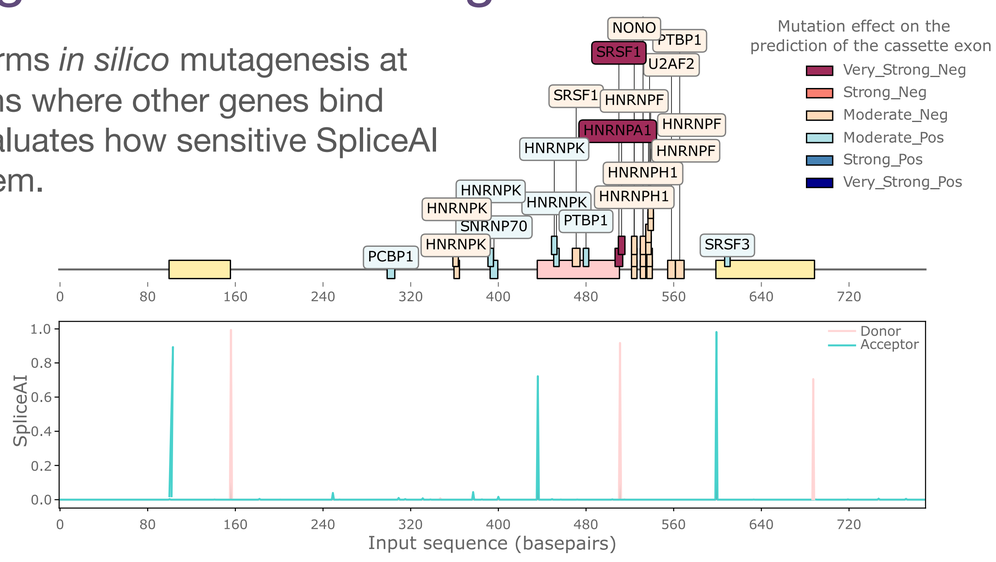


Yet, SpliceAI predictions are significantly different between the two exon groups, demonstrating the ability of the model to predict what was observed in the RNA-sequencing data, without being confounded by unwanted features.

Prediction difference between exon groups	Number of gene silencings
significant	82
non-significant	3

MutSplice: performing model perturbations using domain knowledge

It performs *in silico* mutagenesis at locations where other genes bind and evaluates how sensitive SpliceAI is to them.



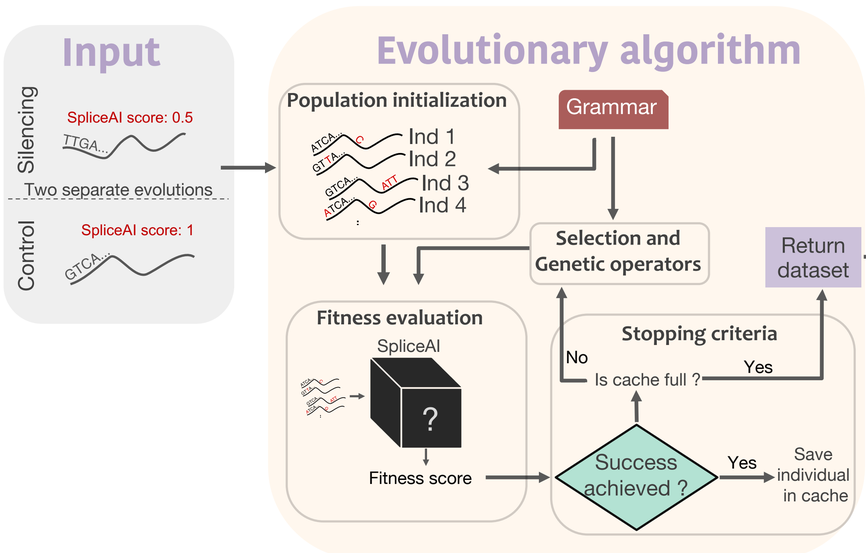
We observed that mutating specific motif locations drive changes in SpliceAI predictions that are in agreement with what it known from biology. Good!

However, this approach was not sufficient to explain differences observed between the two groups of exons in actual cells. **Synergistic interactions between motifs of different genes are still lacking with this approach.**

Ongoing: Generative model of RNA sequences to produce local SpliceAI explanations

The goal is to find a set of mutations that maximizes a custom fitness function (e.g. how much do SpliceAI predictions change after employing the set of mutations present in the individual genotype ?)

The key point is that we design a grammar to restrict the list of mutations based on our domain knowledge about RNA splicing.



Interpretability

- **Directly inspect each sequence phenotype**, which is composed by mutations applied to the original sequence, according to the syntax defined in the grammar).
- **Develop a parser of the phenotype** to generate more biology-driven characterization of the sequences in the dataset.
- **Use the generated dataset and train a decision tree** to distinguish between the two sequence groups based on motif occurrences (features composing the tabular dataset).

LISBOA · Ciências ULisboa · FCT Fundação para a Ciência e a Tecnologia · LASIGE · Instituto de Medicina Molecular João Lobo Antunes

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020, and individual FCT grant SFRH/BD/137062/2018

Co-designing ecologically-valid platforms for neuropsychological rehabilitation

Filipa Ferreira-Brito[1,2], Sérgio Alves[1], Osvaldo Santos[2], Tiago Guerreiro[1], Cátia Caneiras[3,4], Luís Carriço[1], Ana Verdelho[1,5,6]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal, [2] Instituto de Saúde Ambiental, Faculdade de Medicina, Universidade de Lisboa, Portugal, [3]Laboratório de Investigação em Microbiologia na Saúde Ambiental, Faculdade de Medicina, Universidade de Lisboa, Portugal, [4] Healthcare Department, Nippon Gases Portugal, [5] Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Portugal, [6]Neurology Service, Department of Neurosciences and Mental Health, Hospital de Santa Maria, Centro Hospitalar Universitário Lisboa Norte, Portugal.

LASIGE accessibility and ageing

Motivation and Objective

Gamified digital cognitive platforms (GDCP) are a mainstream medium to deliver health interventions[1] in patients with mild cognitive impairment (MCI); however, the impact of such training only rarely impacts patients' daily life functioning[2,3]. We aimed to develop an ecologically-valid platform to train cognition and functional capacity to execute instrumental activities daily living (i.e., grocery shopping).

Methods

Step 1 – Literature Review and older adults' shopping habits characterization

- Which game elements (GE) are used in gamified digital cognitive platforms?
  → We identified that score systems and narrative context are the two most frequently used GE[4]
- Are video games effective to promote cognition and functional capacity in MCI patients?
  → We found a significant large effect in MMSE scores favouring the VG group (MD= 1.64, 95% CI [0.60 - 2.69])[5]
- What are older adults' shopping habits and routines?
  → 53.5% go to supermarkets → 56.2% usually make shopping lists[6]

Step 2 – Development of NeuroVRehab.PT[6]

360º photos of a real-world supermarket

Gamification layer

Neuropsychological rehabilitation principles

NeuroVRehab.PT

An image-based full interactive virtual supermarket with three game modes: shopping list, recipes, and challenges

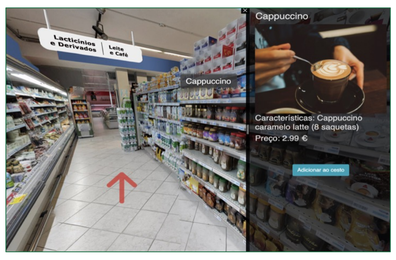
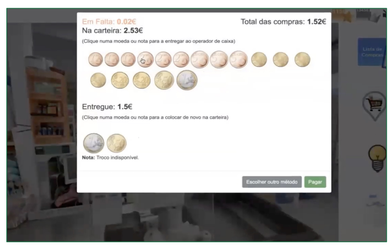


Image-based virtual supermarket

Challenges game mode



Groceries payment activity

Step 3 – NeuroVRehab.PT evaluation

Rehabilitation Potential[6]

Interviews with health professionals

- Valid approach to develop GDCP
- Ecologically valid

Digital Literacy

Interviews with older adults

- ↓ perceived computer confidence negatively impacts the interaction experience
- Necessity of a training period

Feasibility in MCI

Interviews with MCI patients

- Feasible in MCI patients
- Therapist presence is crucial for long-term adherence to GDCP

Conclusion

NeuroVRehab.PT is an ecologically-valid GDCP adequate for neuropsychological rehabilitation in MCI patients. User-perceived computer confidence and therapist presence seemed to be crucial factors for intention-to-use and long-term adherence to GDCP. Including a training period that ensures patients have the necessary digital skills to use GDCP is paramount for the effective and efficient use of GDCP.

U LISBOA | Ciências ULisboa | FCT Fundação para a Ciência e a Tecnologia | LASIGE driven by excellence | NIPPON GASES | FACULDADE DE MEDICINA LISBOA

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) and Nippon Nippon Gases Portugal under the program European Social Fund and Human Capital Operational Programme, co-financed by Portugal 2020 and European Union (Ref: PDE/BDE/127784/2016) and LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020.

References & Credits

Investigating Game Mechanics for Real-World Interactions

Daniel Reis[1], Pedro Pais[1], David Gonçalves[1], João Guerreiro[1], Tiago Guerreiro[1], André Rodrigues[1]

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

LASIGE accessibility and ageing

MOTIVATION

Digital gaming

Has the potential to prompt feelings of **togetherness** through a shared **challenging and immersive activity**, particularly in the real-world.

However

Most commercial games **do not incorporate game mechanics explicitly designed to induce real-world social interactions.**

Growing opportunity to expand existing game mechanics

OBJECTIVE

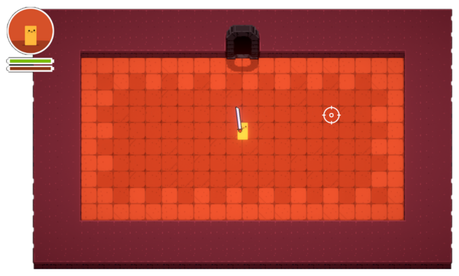
We aim to analyse gaming's **potential for real-world social interaction**, by attempting to:

Understand how games are able to **foster real-world interpersonal communication**

Explore concepts such as **asymmetry** and **interdependence** through a collaborative gaming experience

BREAKING DOWN SOCIAL INTERACTION

We identified and analysed pre-existing game mechanics that could be **sources of real-world social interaction** between players, such as:



Roguelite Multiplayer Game in Development

- Interaction through asymmetry of abilities.
- Interaction through decision-making.
- **Interaction through asymmetry of information.**

We also identified how we can design game mechanics that **force real-world interpersonal interaction**.

	Real-World Co-Location	Coordination	Sharing Personal Content	Sharing Audio	Device Sensors	Accessing Physical Locations
A	●				●	
B		●				
	...	...	...	...	...	...

GAME MECHANICS

A	**If the player is co-located in the real-world with their play partner**, they receive extra currency from completing quests **(Requires access to GPS)**.
B	**The players are able to coordinate and choose a happy hour for their play partner**, meaning that during that hour of the day, if their play partner is playing the game, they get a small percentage of the currency their partner makes.

Examples (work in progress)

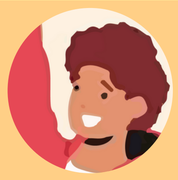
ASYMMETRY OF INFORMATION

Focus: Impacts of information asymmetry between players as a source of real-world social interactions.

INFORMATION

Piece of Information ⟷ **Purpose of the Piece of Information**

The game mechanics were designed to fit each possible combination of these categories for the two players.



	Has Information	Knows Purpose	Has Information	Knows Purpose
C		●	●	
D	●		●	
	...	...	...	...

GAME MECHANICS

C	The **player knows that a certain room should be accessed**, and the **play partner knows the map layout containing said location**.
D	The **player knows half of an item recipe**, while the **play partner knows the other half**, but none know that a combination of the two crafts an item.

Examples (work in progress)

USER STUDY

Target: Two-player dyads that live in the same household.

Requisite: Engage with the developed game through an in-the-wild study.

Evaluation: Player Experience Needs Satisfaction (PENS), Player Experience Inventory (PXI), Game Experience Questionnaire (GEQ).

Reflection: Benefits of explicitly designing for real-world interactions and how this impacts relationships and the players' outlooks on gaming as a whole.

Acknowledgements:
We gratefully acknowledge financial support by FCT through the project "Plug n' Play: Exploring Asymmetry and Modularity for Inclusive Game Design" ref. 2022.08895.PTDC, scholarships ref. UI/BD/151178/2021 and ref. 2022.12448.BD, and the LASIGE Research Unit, ref. UIDB/00408/2020 and ref. UIDP/00408/2020.

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

LayeredTypes – Combining Dependent and Independent Type Systems

Lukas Abelt [1,2], Alcides Fonseca [1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] Saarland University, Germany

reliable software systems

Motivation

A variety of different type systems that ensure specific properties of a program exist, such as resource usage (Linear Types), value predicates (Liquid Types) or correct following of a distributed protocol (Session Types). Some of these type systems can be dependent or independent of each other and different applications require different subsets of these type systems.

Current verification tools however are limited in combining different type systems and often require special approaches to create a single monolithic type system encompassing all the complexity of the different type systems.

A very similar issue can also be observed when operating on Liquid Types: In certain cases, we want to define predicates on orthogonal properties that could be verified independently from one another, which is not supported by current implementations of liquid types.

How to identify the exact cause of liquid type errors?

Before

```
-- Refinement type for a ordered and positive list
{-@ type OrderedList = { v:[Int] | len(v) > 0
                        && all(i,j | i<j => v[i]<v[j]) } @-}
{-@ type PositiveList = { v:[Int] | len(v) > 0
                          && all( i | v[i] > 0 ) } @-}

-- foo is a function from an OrderedList to an Integer
{-@ foo :: OrderedList -> Int @-}
foo :: [Int] -> Int

{-@ x :: PositiveList }
x :: [Int]

foo(x) = ... -- Perform some complicated operation
```

```
Error: type of 'x' is not compatible with liquid type:
   { v:[Int] | len(v) > 0 && all(i,j | i<j => v[i] < v[j]) }
```

After

```
-- Refinement types for individual properties
{-@ type NonEmpty = { v:[Int] | len(v) > 0 } @-}
{-@ type Ordered = { v:[Int] | all( i,j | i<j =>
                                    v[i] < v[j] ) } @-}
{-@ type Positive = { v:[Int] | all( i | v[i] > 0 ) } @-}

{-@ foo :: { NonEmpty && Ordered } -> Int @-}
foo :: [Int] -> Int

{-@ x :: { NonEmpty && Positive } @-}
x :: [Int]

foo(x) = ... -- Perform some complicated operation
```

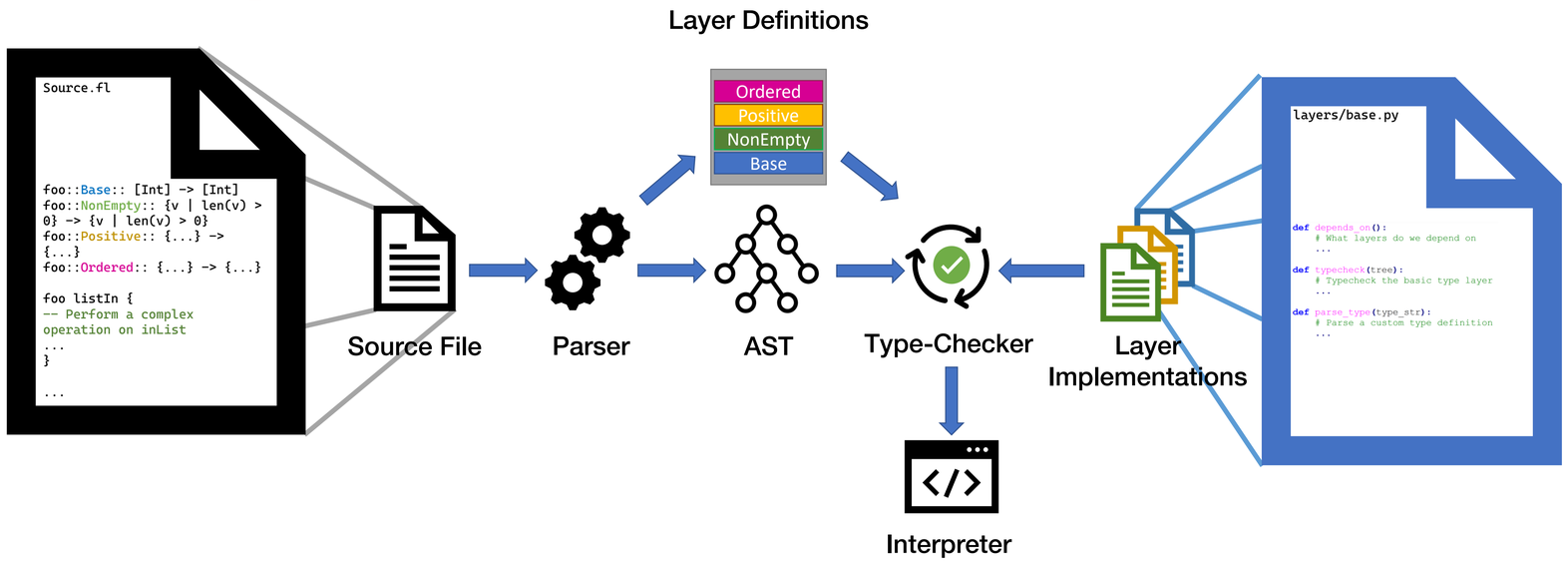
```
Error: type of 'x' is not compatible with layer Ordered
```

Approach

We present LayeredTypes: In LayeredTypes, developers can write programs, but can also define additional type systems as layers. Each layer can independently define its' basic types and typechecking rules. Additionally, layers can depend on one another and use information such as typing contexts from other layers if required.

Due to this separation of concerns, we can use refinements in a much more flexible way. Since each Layer can be verified independently, we can improve modularity, composability as well as error-detection.

Prototype Implementation Architecture



Acknowledgements: This work was supported by *Fundação para a Ciência e Tecnologia* (FCT) in the LASIGE Research Unit under the ref. UIDB/00408/2020 and UIDP/00408/2020, by the CMU – Portugal project CAMELOT (LISBOA-01-0247-FEDER-045915), and the RAP project under the reference (EXPL/CCI-COM/1306/2021.

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

MyHealthyTrack: Recommendation of Healthy Routes

João Lopes[1], Ana Paula Afonso[1], Ana Respício[1]

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

LASIGE cyber-physical systems

MOTIVATION

- Route recommendation systems (such as Google Maps or Waze) recommend the shortest, fastest and/or most economical route. In addition, they use voting systems to increase the competitiveness and performance of the service provided. However, these systems ignore other relevant user criteria.

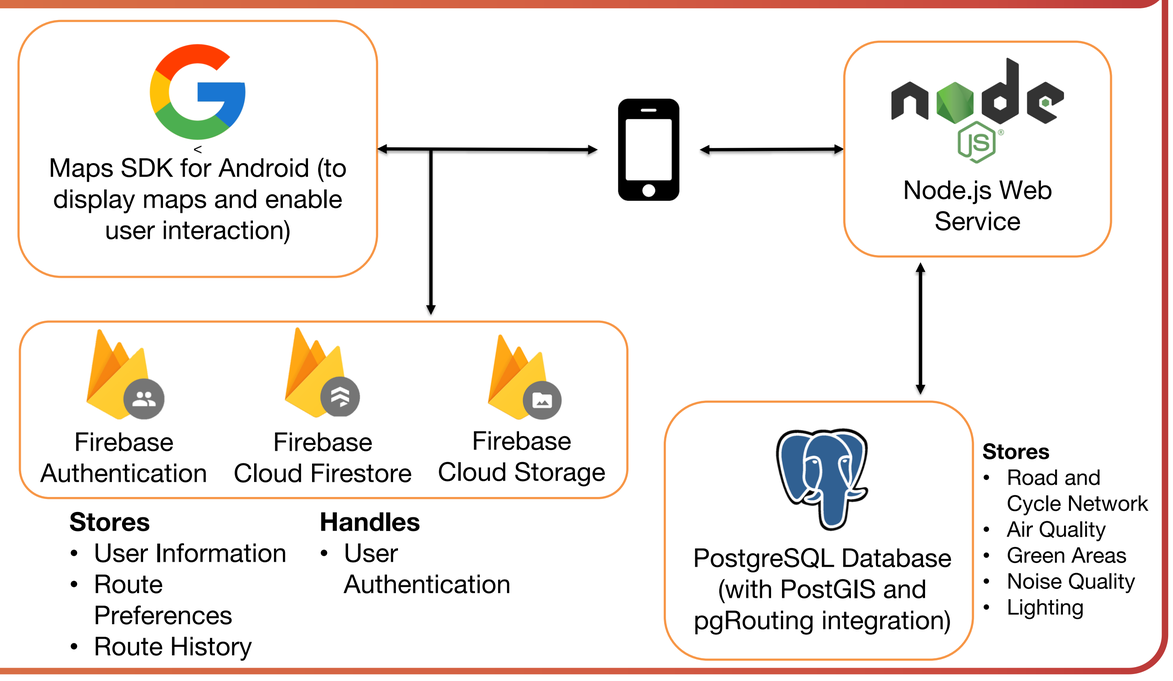
OBJECTIVE

- Develop a new version of the mobile application that recommends pedestrian and cycling routes considering criteria adjusted to the preferences of each user:

Distance | Air Quality | Green Proximity | Noise Quality | Lighting

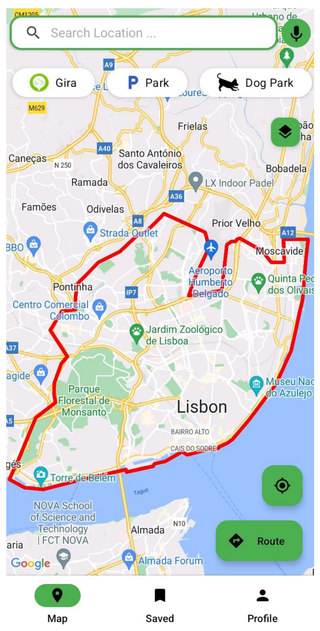
- Integrate new crowd voting mechanisms in order to improve recommended routes.

ARCHITECTURE

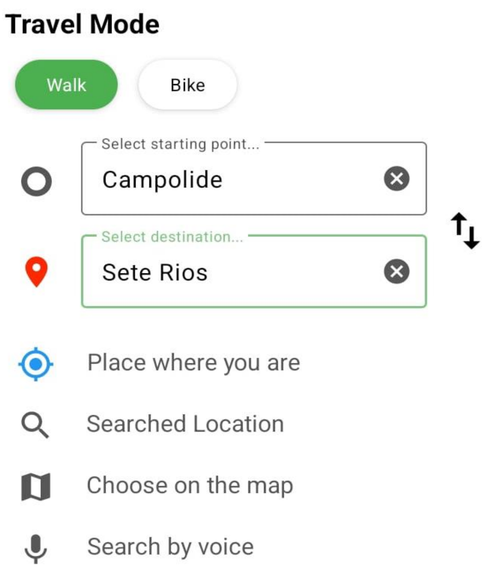


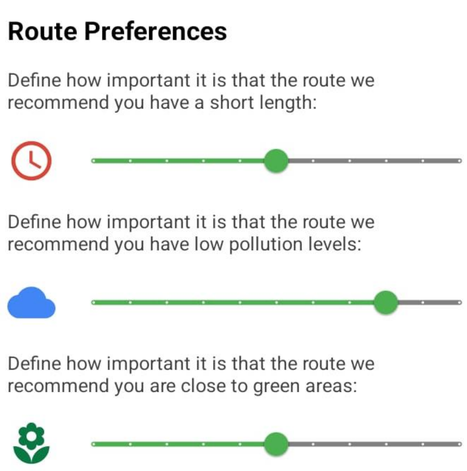
RECOMMENDATION AND CROWD VOTING PROCESS

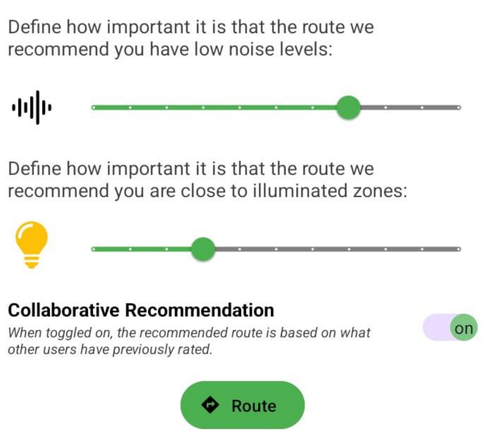
Home Screen



Set Origin, Destination and Route Preferences

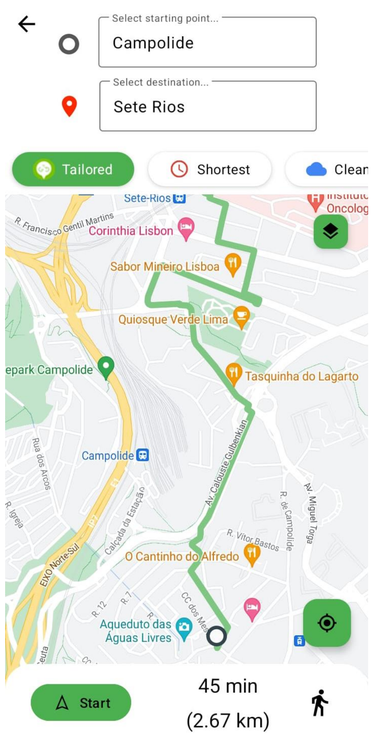




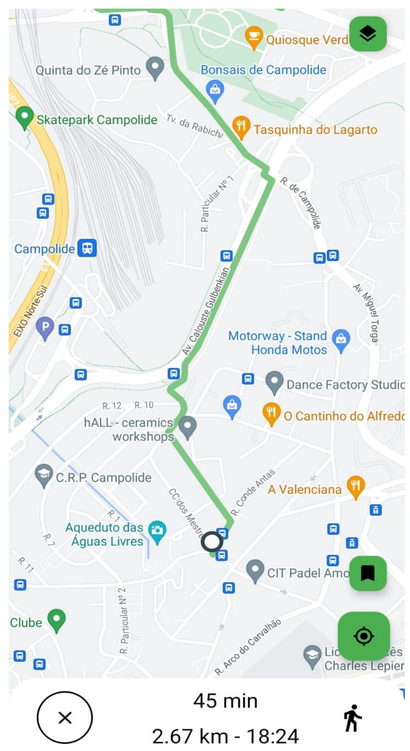


Each criterion is weighted according to the user preferences.

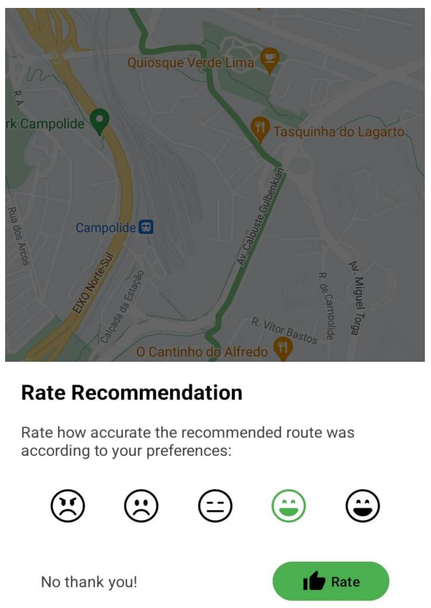
The tailored route optimizes the weighted average of the criteria.



Filter one of the six computed routes



Tailored Route



Rate Recommendation

Crowd Voting is based on a 5 star rating and the user experience (number of reviews).

Criteria values are updated considering the voting.

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020.

Social gaming:
A systematic review

David Gonçalves[1], Pedro Pais[1], Kathrin Gerling[2], Tiago Guerreiro[1], André Rodrigues[1]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal
[2] Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

LASIGE accessibility and ageing

MOTIVATION

A large body of research work focuses on better understanding **gaming as a social activity**.

However, it is not clear how research has approached this topic and what **outcomes** and experiences, in terms of **games**, **populations**, and **contexts** have been considered.

METHODS

We conducted a **systematic literature review** on social gaming, following PRISMA guidelines.

Identification phase
Search query in digital libraries (e.g., IEEE) based on keywords (e.g., "social", "games")
6315 **papers** returned

Screening phase
Filtering according to criteria (e.g., about digital games, focused on social outcomes)
263 **papers** included

Analysis
Mixed inductive and deductive coding
45 coding **fields**

FINDINGS

What makes it social?
Different definitions are presented but most works do not clarify

Social as non-solitary
Any experience with more than one person involved

Social in the outcomes
Any experience that fosters social outcomes (e.g., connectedness)

Social inherent to gaming
Any gaming experience has social outcomes, including single-player

What scenarios are considered?
Different ways to engage in social gaming

Co-playing

Spectating

Mediating

Playing with AI

What interactions are covered?
Research considers both the modality and nature of the interaction

In and out of the game
Interacting in the virtual (e.g., combat, trade) and physical world (e.g., events)

Communication
Both verbal (text, voice) and non-verbal (e.g., pointing)

What outcomes are assessed?

Player experience
In terms of immersion, autonomy, flow, etc.

Social outcomes
Mostly at a shallow level (e.g., co-presence)

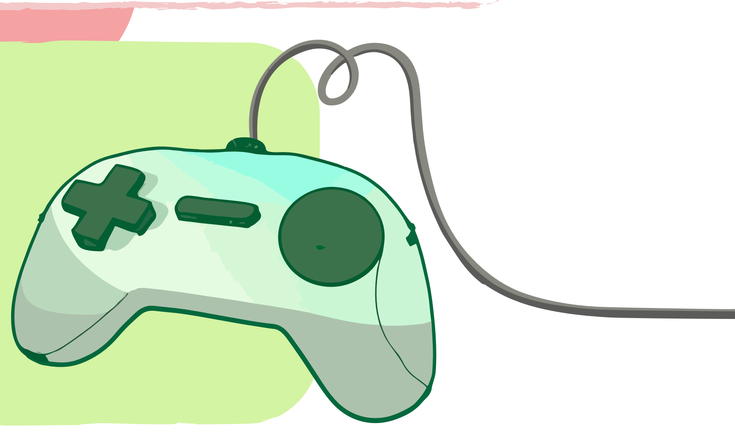
What determinant factors are acknowledged?

Game design and technology
Specific genres, mechanics, and devices

Players and setting
Few works consider the characteristics of the players (e.g., abilities) and the setting

HIGHLIGHTS & OUTLOOK

- Social gaming is **multifaceted**: varied roles (e.g., spectator), interactions, and outcomes.
- Tendency to explore **collaboration** and **novel elements** of design and technology.
- The need for more work looking at **barriers** to social gaming (e.g., inaccessibility).
- Lack of **in-the-wild studies** and established **constructs** to assess social outcomes.



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

Acknowledgements:
We gratefully acknowledge financial support by FCT through the project "Plug n' Play: Exploring Asymmetry and Modularity for Inclusive Game Design" ref. 2022.08895.PTDC, scholarships ref. UI/BD/151178/2021 and ref. 2022.12448.BD, and the LASIGE Research Unit, ref. UIDB/00408/2020 and ref. UIDP/00408/2020.

Images were generated with the assistance of AI (DALL-E).

AML-Robust Network Intrusion Detection based on diversity and Machine Learning

Allan Espindola[1, 2], António Casimiro[1], Pedro M. Ferreira[1], Altair Santin [2], Eduardo Viegas [2]

[1] LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal
[2] SeCPLab, Escola Politécnica, Pontificia Universidade Católica do Paraná

LASIGE cyber-physical systems

Motivation

Cyber attacks on Cyber-Physical Systems (CPS) pose a threat. Network intrusion detection systems (NIDS) detect malicious activities but are vulnerable to Adversarial Machine Learning (AML) attacks. Behavior-based ML techniques need to be strengthened against AML attacks for accurate results.

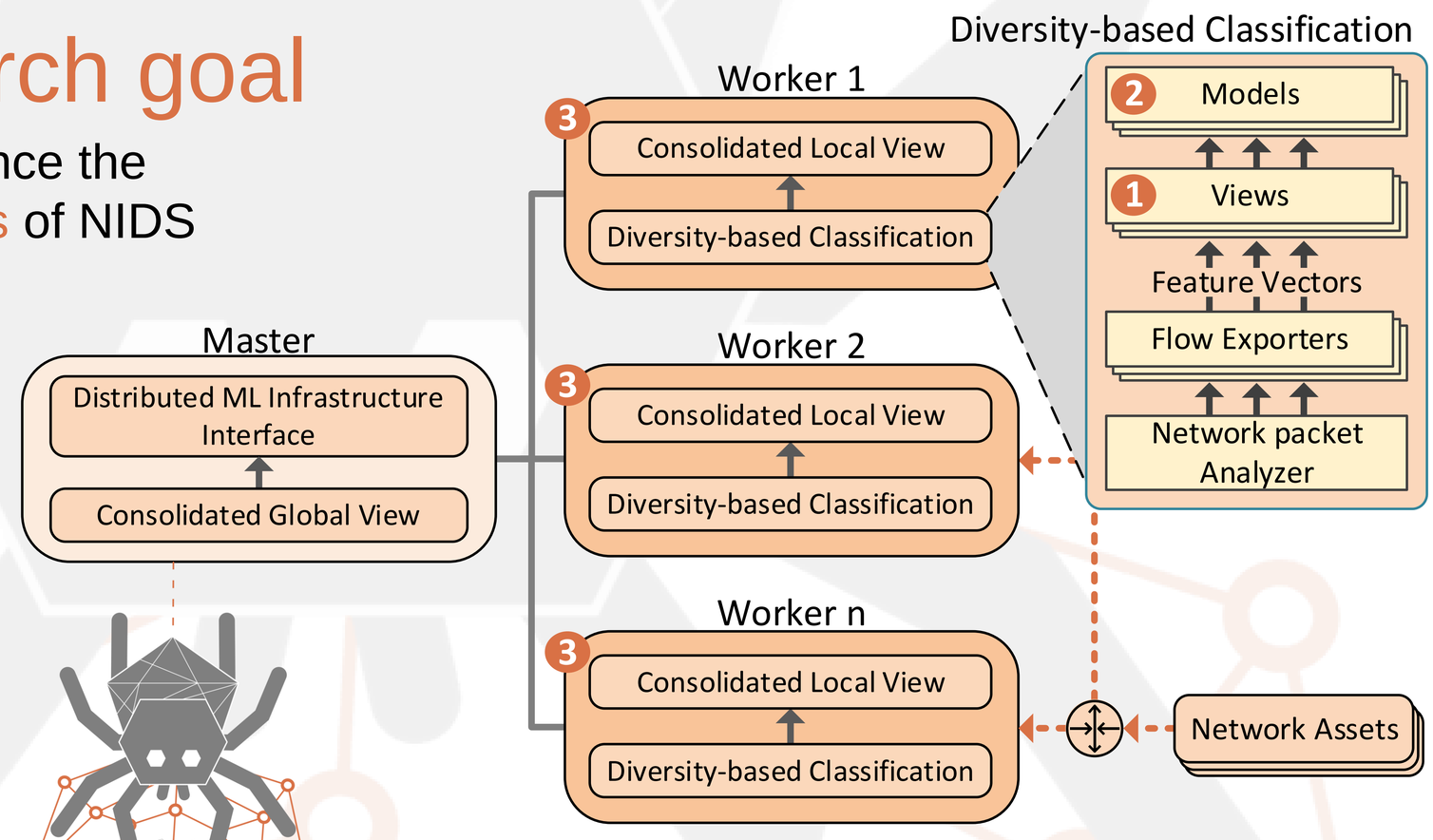
Main idea and research goal

Explore diversity and how it can enhance the detection performance and robustness of NIDS against AML attacks.

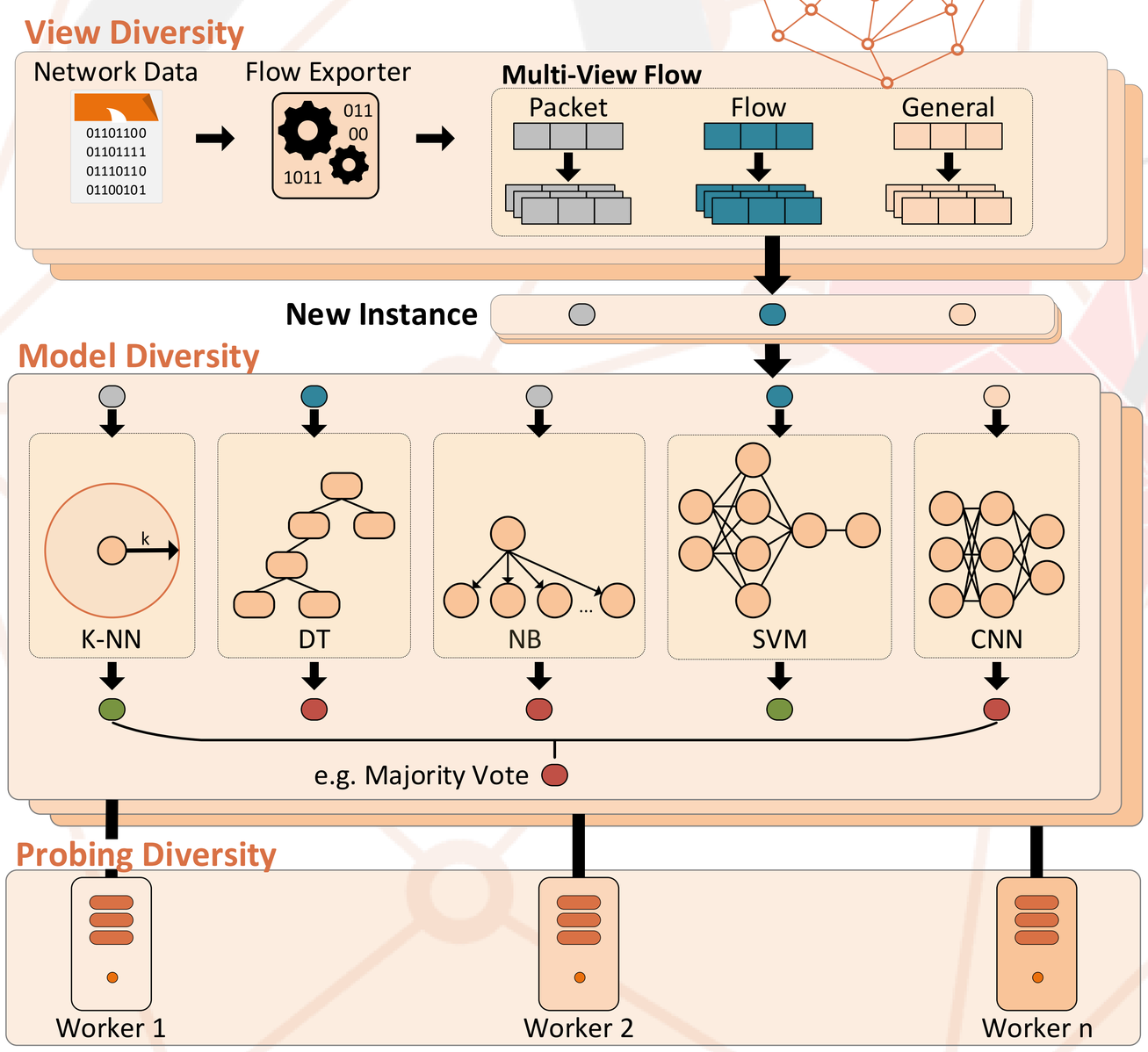
Main contribution

ARAchNID

AML-Robust Architecture for Network Intrusion Detection based on diversity and machine learning



Architecture



Diversity types

1. **View diversity** enables the examination of the same event from multiple perspectives, it can assist in detecting a wider range of attacks through complementary views.

2. **Model diversity** utilizes diverse ML models and hyper-parameters to boost detection performance and robustness.

3. **Probing diversity** creates unique workers with diverse views and models, mitigating common vulnerabilities. If some are compromised, others can continue detecting threats.

LASIGE is supported by national funds through FCT, within the framework of projects UIDB/00408/2020 and UIDP/00408/2020

Project funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 871259.

Quantum Specific Smells

Qihong Chen*, **Ruben Câmara, José Campos, André Souto**, Iftekhar Ahmed*

LASIGE reliable software systems

A **code smell** describes design and implementation flaws in source code that do not necessarily make the software behave incorrectly or crash but make it harder to understand and maintain. Since quantum is a different computing paradigm, quantum software has specific smells. The paper ***The Smelly Eight: An Empirical Study on the Prevalence of Code Smells in Quantum Computing*** defines some quantum-specific smells and proposes a novel tool Qsmell to detect those smells automatically.

Use of Customized Gates

All gates are written with built-in gates, but this process requires significantly more operators than using only built-in operators.

```
  qc = QuantumCircuit(3)
- qc.unitary([
- [1, 0, 0, 0, 0, 0, 0, 0], [0, 1, 0, 0, 0, 0, 0, 0],
- [0, 0, 1, 0, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0, 0],
- [0, 0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 0, 1, 0, 0],
- [0, 0, 0, 0, 0, 0, 0, 1], [0, 0, 0, 0, 0, 0, 1, 0]
- ], [0, 1, 2])
+ qc.ccx(0, 1, 2)
```

Repeated set of Operations on Circuit

Algorithms must be designed with minimal sequential repetition to avoid exceeding the limited operations to be passed to device.

```
  qc = QuantumCircuit(3, 3)
  hadamard = QuantumCircuit(1, name='H')
  hadamard.h(0)
  measureQubit = QuantumCircuit(1, 1, name='M')
  measureQubit.measure(0, 0)
- for i in range(3):
      for j in range(3):
          qc.append(hadamard, [j])
          qc.append(measureQubit, [j], [j])
+ qc.repeat(3)
```

Intermediate Measurements

Measurement affects the entire system and may introduce errors. They should be the last operation on the circuit.

```
  qc = QuantumCircuit(3, 2)
  qc.h(0)
- qc.measure(0, 0)
+ qc.cnot(0, 1)
+ qc.cnot(0, 2)
  qc.h(0)
- qc.measure(0, 1)
+ qc.measure(1, 0)
+ qc.measure(2, 1)
```

No-alignment between the Logical and Physical Qubits

Currently, quantum memories have asymmetries among the qubits. Then the geometry of physical qubits affect the results.

```
  backend = FakeVigo()
  qc = QuantumCircuit(3, 3)
  qc.h(0)
  qc.cx(0,range(1,3))
  qc.measure(range(3), range(3))
- qc = transpile(qc, backend)
+ qc = transpile(qc, backend,
+ initial_layout=[3, 4, 2])
```

Non-parameterized Circuit

Real devices work in a shared policy. Circuits should be parametrically designed to receive multiple initial values at once, minimizing communication payloads and avoiding queuing delays.

```
+ from qiskit import transpile
+ from qiskit.circuit import Parameter
+ theta = Parameter('0')
  def init_circuit(theta):
      qc = QuantumCircuit(5, 1)
      qc.h(0)
      for i in range(4):
          qc.cx(i, i+1)
    qc.rz(theta, range(5))
    return qc
  theta_range = [0.00, 0.25, 0.50, 0.75, 1.00]
  backend = Aer.get\_backend('qasm\_simulator')}
- for theta_val in theta_range:
-     qc = init_circuit(theta_val)
-     job = backend.run(qc)
-     job.result().get_counts()
+ qc = init_circuit(theta)
+ circuits = [qc.bind_parameters({theta: theta_val})
+             for theta_val in theta_range]
+ job = backend.run(transpile(circuits, backend))
+ job.result().get_counts()
```

Long circuits

Errors affect the performance of quantum circuits. The larger and deeper the circuit is, the worst it is its performance.

```
  qc = QuantumCircuit(1)
- qc.h(0)
- qc.z(0)
- qc.h(0)
+ qc.x(0)
```

Idle Qubits and its initialization differently from |0⟩

Having idle qubits for too long enhances wrong results since they decohere fast. Hence, initially, one should keep it in its ground state (i.e., in state |0>) as long as possible.

```
  qc = QuantumCircuit(3, 3)
- qc.h()
+ qc.h(0)
  qc.p(pi / 2, 0)
  qc.z(0)
+ qc.measure(0, 0)
+ qc.h(1)
  qc.p(pi / 4, 1)
  qc.z(1)
+ qc.measure(1, 1)
  qc.h(2)
  qc.p(pi / 8, 2)
  qc.z(2)
- qc.measure_all(add_bits=False)
+ qc.measure(2, 2)
```

Experiments on real-world QPs showed 73% of quantum programs contain at least one code smell, and, on average, programs have three code smells. The long circuit is the most prevalent code smell. The results reported lay the foundation for our future work. We plan to explore the evolution of the quantum-specific code smells and their effect on the quality of quantum programs. Furthermore, we plan to run a user study on the usage of QSmell and extend QSmell to support other quantum frameworks (e.g., Cirq).

The Smelly Eight: An Empirical Study on the Prevalence of Code Smells in Quantum Computing, Q. Chen, R. Câmara, J. Campos, A. Souto, I. Ahmed, IEEE/ACM International Conference on Software Engineering (ICSE) 2023.

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

Acknowledgments: This work was supported by the LASIGE Research Unit (ref. UIDB/00408/2020 and ref. UIDP/00408/2020).

*University of California, Irvine, USA

Evaluating Accessible Navigation for Blind People in Virtual Environments

Manuel Piçarra[1], André Rodrigues[1], João Guerreiro[1]

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

LASIGE accessibility and ageing

Prior work[1] allowed us to verify how blind people dealt with barriers to play mainstream games and when it was almost impossible to.

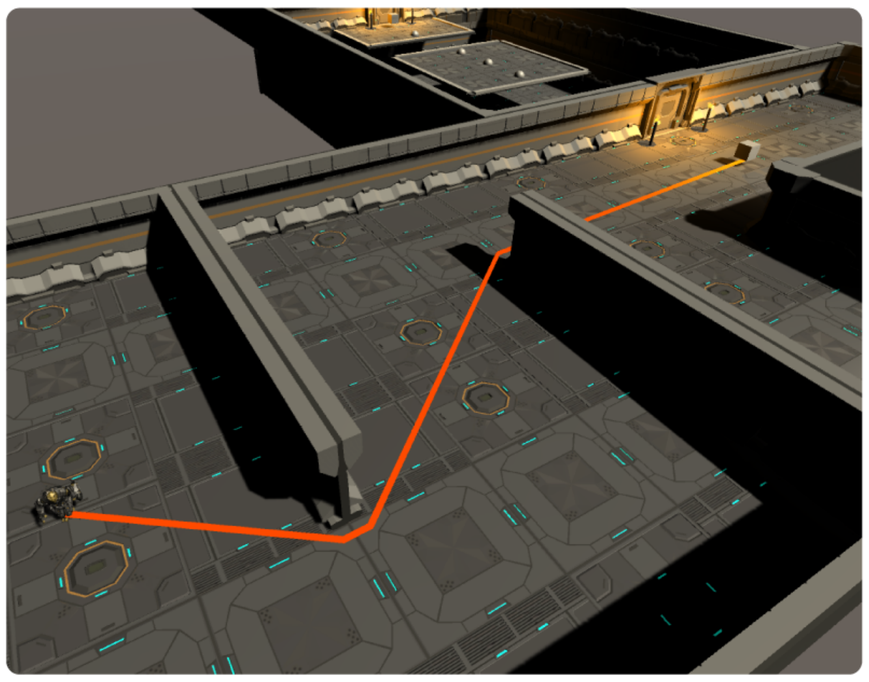
[1] Gonçalves, D., Piçarra, M., Pais, P., Guerreiro, J., & Rodrigues, A. (2023). My Zelda Cane: Strategies Used by Blind Players to Play Visual-Centric Digital Games. CHI '23, April 23–28, 2023, Hamburg, Germany.

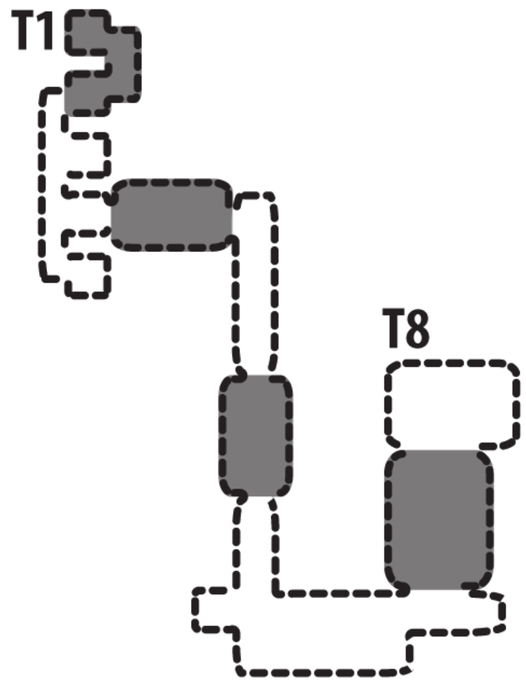
Contributions: Exploration of accessible techniques and design choices based on the requirements of blind players, sharing their perspectives on the experience.

We created a VE with most common barriers to blind people navigation and orientation.
We implemented eleven accessible tools to tackle such challenges.

→ Eight navigational tasks, 4 minutes to reach the objective's location.

→ After completing a task, a new accessibility technique is unlocked.



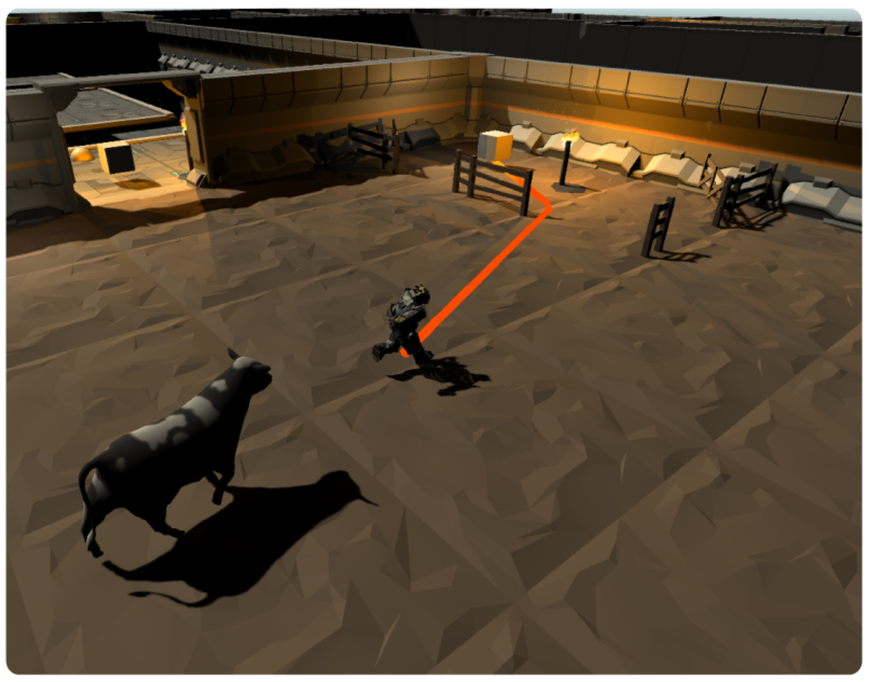


Tasks:

- **T1** Listen to **Footsteps** and **Collisions**;
- **T2** Automatic or Demanded **Voice Indication** of Current Occupied Area;
- **T3** **Sound Sources** as Reference Points;
- **T4** Interaction with **Occluded Objects** and their State Changes;
- **T5** Perception of **Ledges** and **Save/ Load** Character's Position;
- **T6** Request **Directions** to Current Objective;
- **T7** **Automatic Progression** after Request;
- **T8** Non Playable Character **Perception** and **Interaction**;

Our **findings** resulted from a preliminary qualitative analysis, observations and concluding interviews:

Task completion:
32 out of 56 **completed**,
T1 & T2 had the **worst** rates (10 out of 14).



> **Auditory Landmarks** & **Identification of Areas** give the basis for a confident navigation and mental map creation;

> **Occluded Objects** require detailed feedback of object's shape or condition, with haptic and/ or audio **feedback**;

> Common game specific features require **time to familiarize**, as well as explanation due to possible **unnatural concepts**;

> **Automated Actions** can help with troublesome situations, offering a solution if requested, avoiding forcing them;

> T1 & T2 proved the **difficulty** of navigating **without** dedicated **accessibility** features;

> **Virtual Obstacles** and **Environmental Changes** can be easily conveyed with **haptic** and **audio feedback**;

> Following **directions** is intuitive, although the **camera** perspective and handling needs to be considered when designing these systems;

> **Interaction** & **Perception** of others benefits from properly timed audio feedback and distinct sound effects;

Acknowledgements: This work was supported in part by FCT through the LASIGE Research Unit (UIDB/00408/2020, UIDP/00408/2020), the AccessVR project (2022.08286.PTDC), and the PlayFam project (2022.08895.PTDC).

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

Learning Temporal Patterns for Patient Stratification in ALS

Daniela Amaral[1,2], Diogo F. Soares[1], Pedro Tomás[2], Helena Aidos[1]

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

LASIGE data and systems intelligence

Target Disease: Amyotrophic Lateral Sclerosis (ALS)

- **Amyotrophic Lateral Sclerosis (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.
- This disease has no known cure, and the treatments try to improve the quality of life and the life expectancy of the patients.
- ALS is a **highly heterogeneous disease**, being difficult to determine when a certain treatment should be applied.

⇩

- **Stratification** allows to **find disease-specific subgroups** with similar characteristics and progression.
- **Prognostic** models can be **enhanced** by adding a preliminary **stratification** step to cluster patients into groups and then developing group-specific prognostic models.

Objectives

- Develop a stratification algorithm to capture the temporal dimension of the data, and to group patients.
- Provide clinicians with important insights regarding the disease subgroups, helping to support their decisions.

Data

- Lisbon ALS clinic Dataset, Hospital Santal Maria
- Demographic and temporal data of 919 Patients
- 3 consecutive snapshots

NOVEL TEMPORAL CLUSTERING METHOD SUCCESSFULLY IDENTIFIED GROUPS OF PATIENTS DIFFERING IN DISEASE PROGRESSION

Methodology & Results

Data

Static

Dynamic: 1st appointment, 2nd appointment, 3rd appointment

Patterns' Extraction

Biclustering enables the simultaneous clustering of **two-way** data (patients and features).

Biclustering

Subspace clustering: enables the discovery of discriminative **patterns** from **multidimensional** cohort data.

Virtual Patterns

Virtual Pattern: The Bicluster/Tricluster most representative pattern is the mean object within-cluster and represents its tendency.

The obtained patterns are used to **transform** the **data space** using a **similarity metric**.

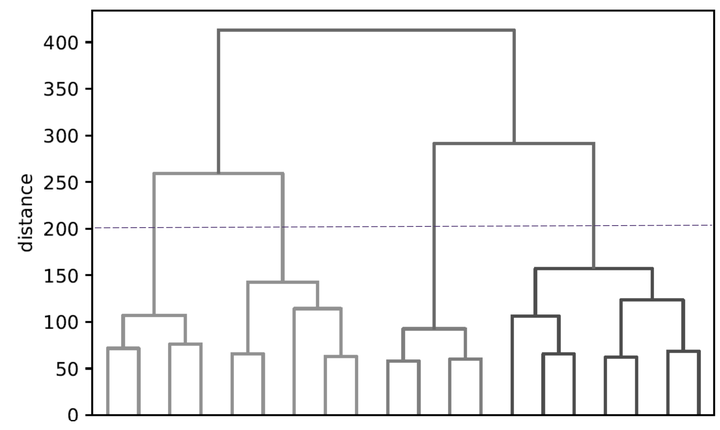
Similarity Matrices: Computed using the Euclidean Distance between virtual patterns and the object patterns (using the same features).

Triclustering enables the simultaneous clustering of **three-way** data (patients and features over time).

Triclustering

Virtual Patterns

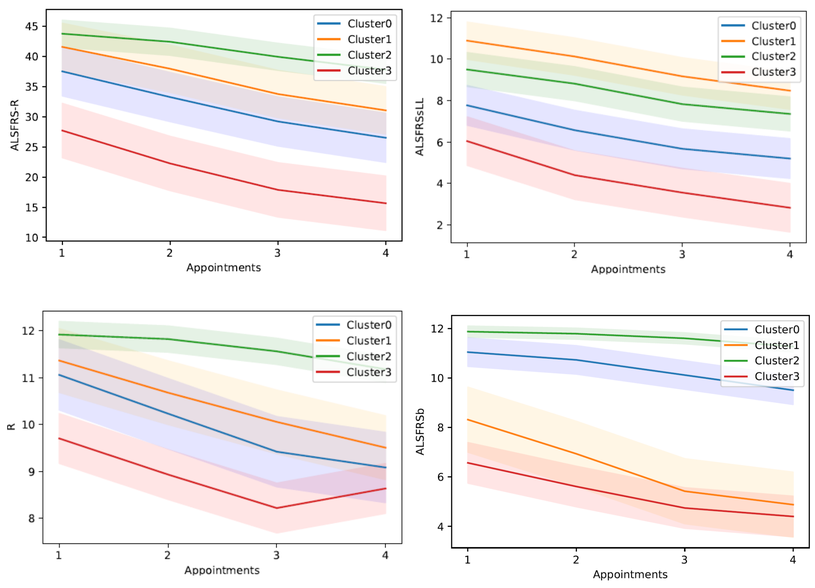
Hierarchical Clustering



Hierarchical clustering generates a **dendrogram** that gives intuition regarding the **most suitable number of clusters**.

Four groups

Groups' Characterization



- Cluster 0 represents the medium progressors with slow bulbar progression;
- Cluster 1 represents the medium progressors with slow limb progression;
- Cluster 2 represents the slow progressors;
- Cluster 3 represents the fast disease progressors.

[1] Soares, D. F., Henriques, R., Gromicho, M., de Carvalho, M., & Madeira, S. C. (2022). Learning prognostic models using a mixture of biclustering and triclustering: Predicting the need for non-invasive ventilation in Amyotrophic Lateral Sclerosis. *Journal of Biomedical Informatics*, *134*, 104172.

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

Download the project!

SCAN ME

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020, the AIpALS project (PTDC/CCI-CIF/4613/2020), and the PhD research scholarship 2020.05100.BD.

Injecting domain knowledge into graph neural networks to predict protein-protein interactions

Laura Balbi, Catia Pesquita

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

LASIGE health and biomedical informatics

Motivation

In recent years, Graph Neural Networks (GNNs) have become the gold standard for node classification and link prediction tasks over graph data. However, in real-world applications such as protein-protein interaction networks, the lack of node features has impaired their performance.

To overcome this challenge, we introduce a novel approach for integrating external biomedical knowledge into Graph Neural Networks through knowledge graph (KG) embeddings. These embeddings map the KG's entities and relationships into a lower-dimensional vector space while preserving their factual information. Our domain knowledge injection approach (DKI) for PPI-based Graph Neural Networks was evaluated in popular benchmark graph learning tasks.

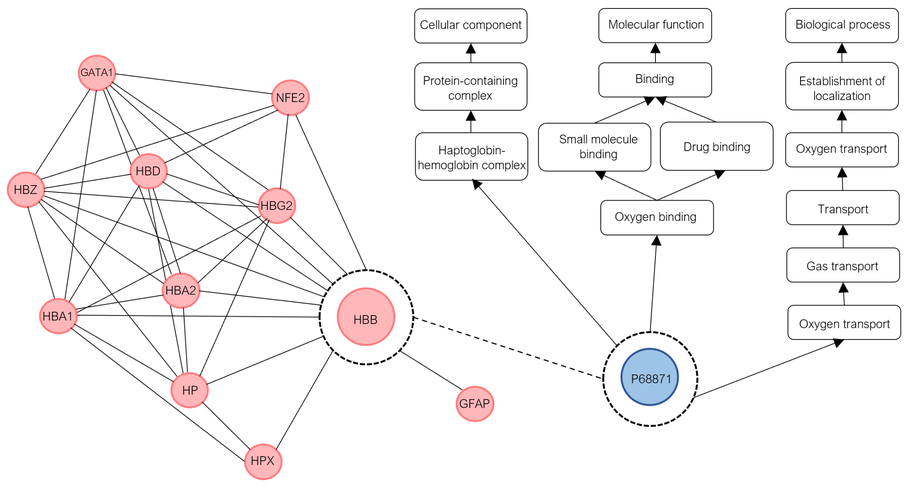


Figure 1. Subgraph of STRING showing the interactions for the human protein HBB and subgraph of the Gene Ontology KG showing classes that describe HBB. Note that in each graph, HBB is not identified with the same identifier and this requires a mapping to be established.

GNNs are a general formalism for deep neural networks that take graph data as input. Its defining feature is that it employs a type of neural message-passing in which vector messages are exchanged between nodes and updated on each iteration.

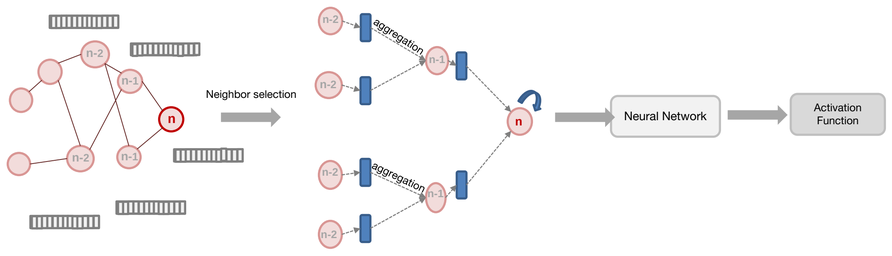


Figure 2. Basic GNN architecture and node representation processing procedures.

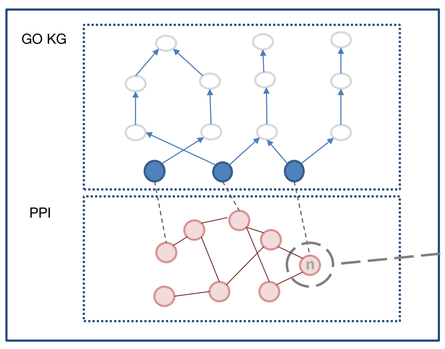
Methodology

Two graph mining tasks were employed over two PPI network benchmark datasets: link prediction and node classification.

Task	Graph	# nodes	# edges
node classification	PPI Graph	132,534	39,561,252
	KG CC	176,800	499,961
	KG All	176,800	1,363,187
	KG R2	176,800	798,086
link prediction	PPI Graph	576,289	30,326,273
	KG All	620,555	5,989,238

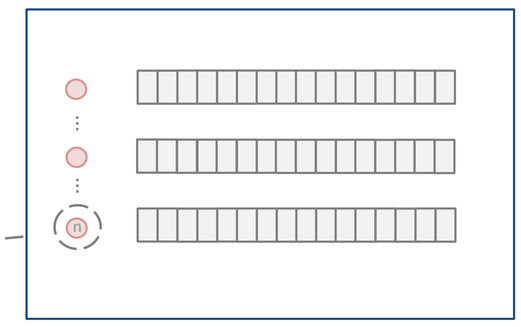
Table 1. Summary of graph composition for Node Classification and Link Prediction tasks.

Knowledge Graph construction



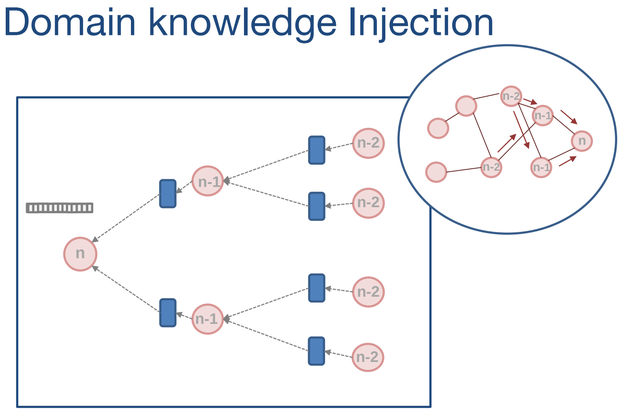
Several KGs were built through different mapping strategies of the PPI networks' proteins with the GO classes that describe them.

KG Representation Learning



Different embedding methods (e.g. RDF2Vec) were used to generate highly informative node representations based on node connectivity along their local neighborhoods in the biomedical KGs.

Domain knowledge Injection



DKI feeds GNNs with the embeddings as numerical features for each node. GNNs can now access the semantic information encoded in the embeddings and explore it as node representations during message-passing along the PPI network.

Results

Model	ROC-AUC	Accuracy
GraphSage	0.772	0.874
GraphSage +DKI	**0.980**	0.990
NGNN+GCN	0.402	0.937
NGNN+GCN+DKI	**0.440**	0.886
NGNN+GraphSage	0.649	0.923
NGNN+ GraphSage+ DKI	**0.991**	0.991
DeeperGCN	0.812	0.921
DeeperGCN+DKI	**0.994**	0.983

Table 2. Results pertaining to 4 of the neural approaches evaluated for the node classification task. "+DKI" refers to our approach to the including RDF2Vec KG embeddings as features. The best results for each GNN-based approach are in bold.

Model	Hits@100	Accuracy
GCN	0.175	0.899
GCN+DKI	**0.240**	0.982
GraphSage	0.179	0.979
GraphSage+DKI	**0.280**	0.980
NGNN+GCN	0.383	0.910
NGNN+GCN+DKI	**0.437**	0.917
NGNN+GraphSage	0.412	0.894
NGNN+GraphSage+DKI	**0.439**	0.902

Table 3. Performance results for 4 of the GNN models evaluated for link prediction task. "+DKI" refers to our approach to the including RDF2Vec KG embeddings as features. The best results for each neural approach are in bold.

Experiments were conducted with a total of 10 state of the art GNN approaches and 5 representative knowledge graph embedding methods.

Results show that the domain knowledge injection approach substantially improves performance on both node classification and link prediction tasks. They also surpass state of the art methods on both tasks proving the effectiveness of generating powerful node features from external sources of knowledge to augment GNN approaches.

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

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020, by the European Commission through the ProjectName project (REFREFREF), and FCT grant SFRH/BD/REFREF/YYYY

Holistic Biomedical Knowledge Graph Integration

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

Motivation

Biomedical research is increasingly data-centric as multi-domain high-throughput data is becoming commonplace. Ontologies provide context to data and are critical to enable interoperability, but there are significant challenges.

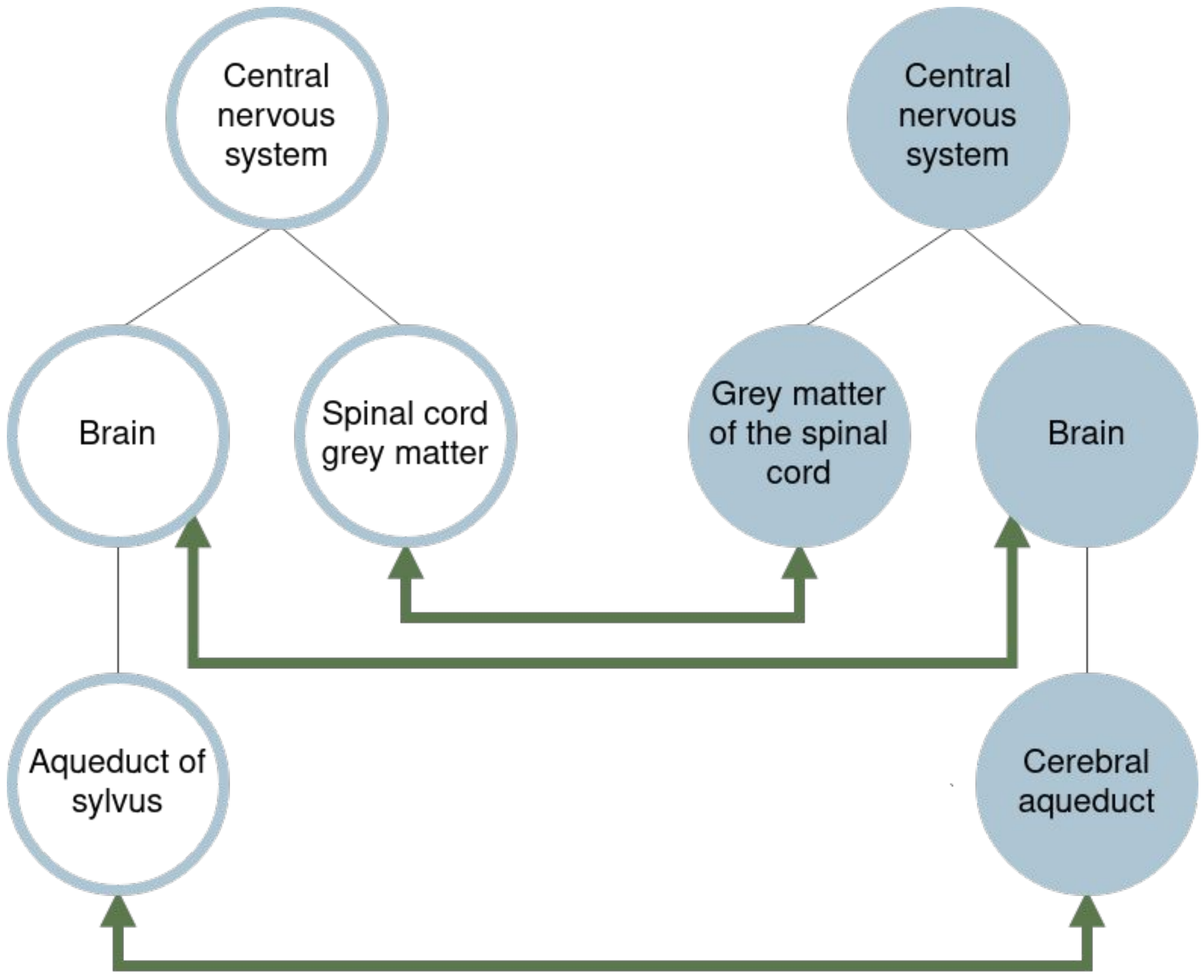
Challenges

→ **Scalability:** hundreds of biomedical ontologies with 10k+ classes
→ **Completeness:** many ontologies have overlapping domains and follow different models
→ **Correctness:** overlap found automatically must be validated

Objective

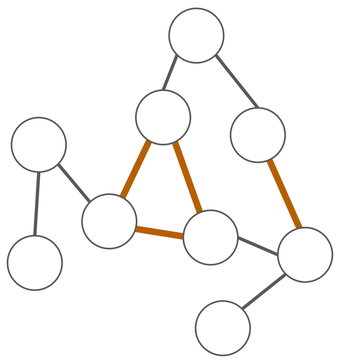
Matching multiple biomedical ontologies, both efficiently and accurately, to produce a comprehensive and coherent Knowledge Graph.

Ontology Matching (OM)



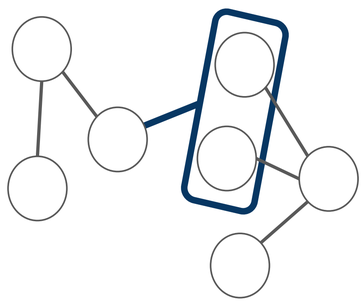
Holistic OM

Extension of **ontology matching for multiple ontologies resulting in an alignment**. Usually done by combining pairwise alignments, which is not scalable.



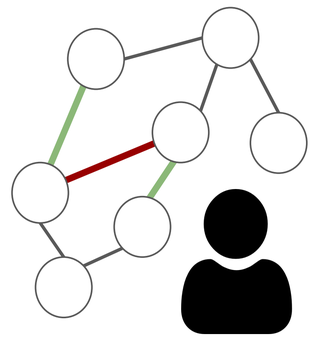
Complex OM

Extension of ontology matching where instead of ontology entities, **mappings can include logical expressions analogous to class and property expressions.**

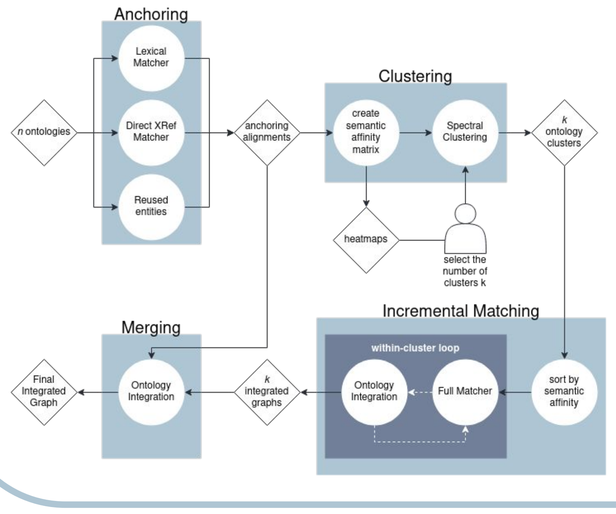


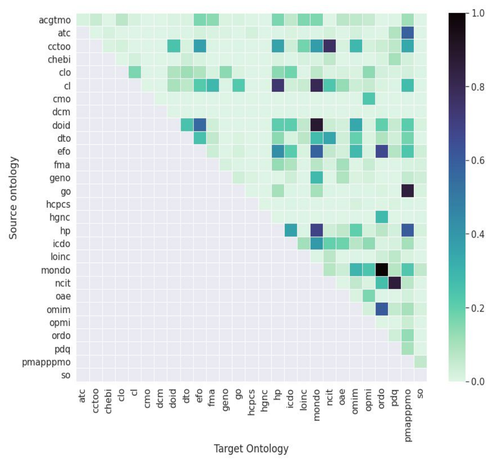
Interactive OM

User expertise can be used to validate the alignment. Users can detect and remove/correct mappings, as well as provide additional mappings.



Combination of two holistic ontology matching approaches **achieves a reduction of almost 50% in total runtime** when matching 28 biomedical ontologies into a single alignment.





Strategy	Load	Match	Total	Mappings	Tasks
GPA	11:47	19:51	31:37	554547	378
CIA	12:52	03:04	15:56	417131	378* + 24

*Runtime (hh:mm). GPA: global pairwise alignment. CIA: within cluster incremental alignment. * from the Anchoring step*

Acknowledgements: This work was supported by FCT through the LASIGE Research Unit (UIDB/00408/2020 and UIDP/00408/2020).

The KATY project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017453.

Can ChatGPT explain protein interactions better than Knowledge Graphs?

Rita T. Sousa, Sara Silva, Catia Pesquita
LASIGE, Faculty of Sciences, University of Lisbon, Portugal

LASIGE health and biomedical informatics

Is there an interaction between protransforming growth factor α (TGF-α) and disks large homolog 2 (DGL2)?

Literature

Although there is no known interaction between these proteins, there is evidence of an interaction between highly similar proteins [1]: TGF-β is regulated by Dlg5 and both proteins activate the MAPK cascade.

[1] Chiara Andolfi, Claudio Tiribelli, Devis Pascut. Recent Hints on the Dual Role of Discs Large MAGUK Scaffold Protein 5 in Cancers and in Hepatocellular Carcinoma. Front. Biosci. (Landmark Ed) 2022, 27(5), 164.

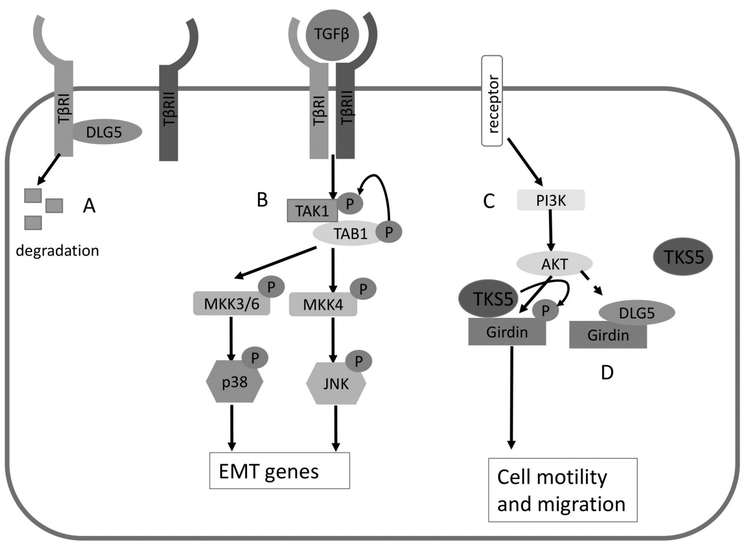
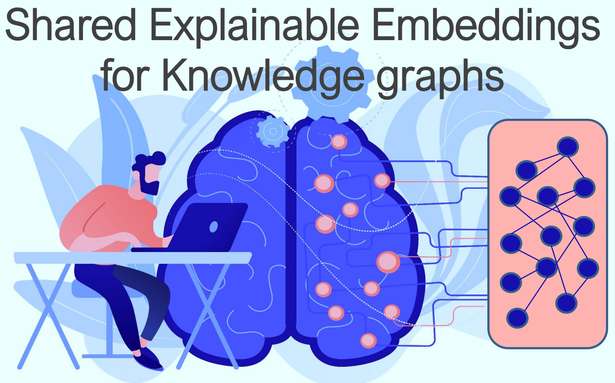


Figure 1: Role of Dlg5 and TGF-β in cancer prevention (extracted from [1]).

SEEK

Shared Explainable Embeddings for Knowledge graphs



The SEEK explanations highlight their co-location in the basolateral plasma membrane and their participation in the MAPK cascade.

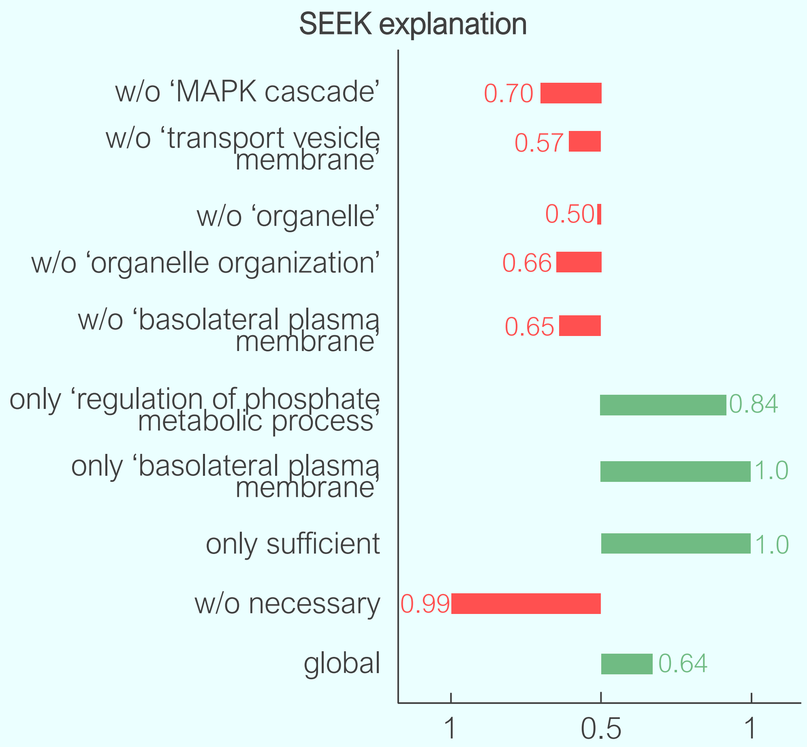
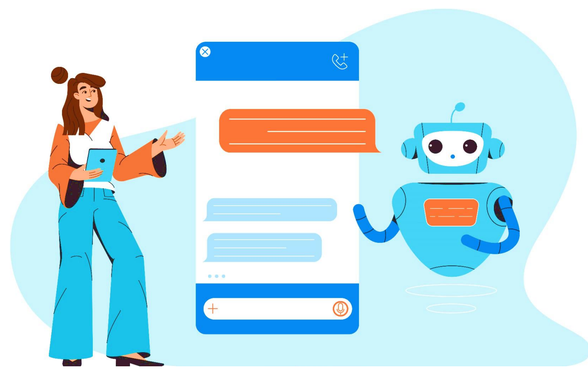


Figure 2: Bar chart using different sets of disjoint common ancestors to represent the pair. Each bar represents the likelihood returned by the ML model of the predicted class being correct (red for class 0 and green for class 1).

ChatGPT



R Do TGF-α and DLG2 interact?

There is currently no evidence to suggest a direct interaction between TGF-α and DLG2. While TGF-α and DLG2 may be involved in similar cellular processes, there is no known physical interaction between them.

R Can you explain why TGF-α and DLG2 interact?

TGF-α and DLG2 are two proteins that **have been found to interact in various cellular contexts**. Additionally, studies have suggested that the interaction between TGF-α and Dlg2 may play a role in cancer development and progression.

R Can you explain why TGF-α and DLG2 do not interact?

As an AI language model, I don't have access to the latest research developments and **cannot give a definitive answer** to this specific question without more information on the context of the interaction you are referring to.

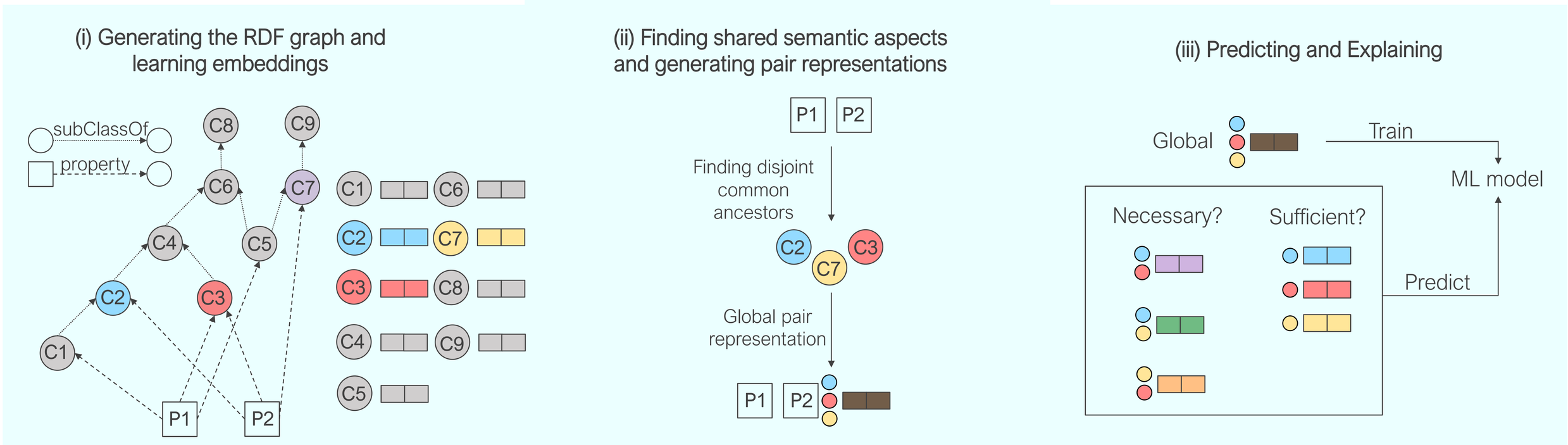


Figure 3: Overview of SEEK.

U LISBOA Ciências ULisboa FCT Fundação para a Ciência e a Tecnologia MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA LASIGE driven by excellence

Acknowledgements: Catia Pesquita, Sara Silva, Rita T. Sousa are funded by the FCT through LASIGE Research Unit (ref. UIDB/00408/2020 and ref. UIDP/00408/2020), and the FCT PhD grant (ref. SFRH/BD/145377/2019).It was also partially supported by the KATY project which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017453.

Links vs. Nodes: do richer Knowledge Graphs improve gene-disease association predictions?

Catarina Canastra, Susana Nunes, Cátia Pesquita

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE health and biomedical informatics

Gene-disease association prediction

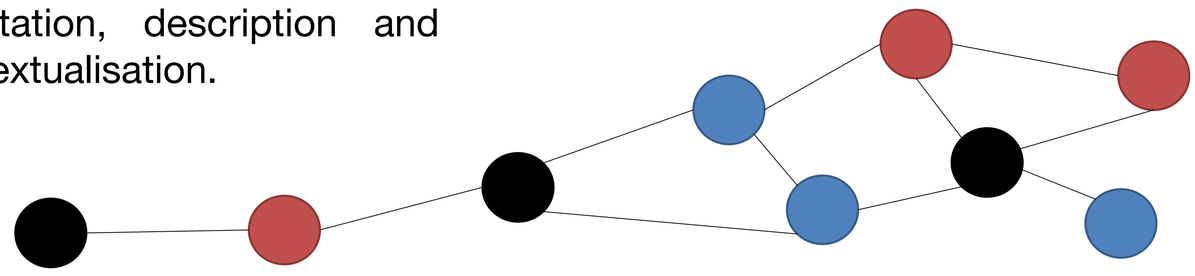
- Identifying disease-associated genes is the basis for a comprehensive study of the disease mechanisms.
- Studies on gene-disease associations have provided diverse data, which can facilitate the gene-disease association prediction.
- Integrating diverse data is critical to developing high-accuracy prediction models.

Ontologies

Computational structures that provide a way to model domains through a set of concepts (entities) and relationships. The use of ontologies has been important to provide data annotation, description and contextualisation.

Knowledge Graphs

Directed labeled graphs with labels that capture the meaning of relations between entities (nodes). It flexibly structure different domains providing data integration.



Methodology

Knowledge Graph Embedding Algorithms
- RDF2Vec
- OPA2Vec

Vector Operators
- Cosine Similarity
- Hadamard
- Concatenation
- Average
- Weighted-L1
- Weighted-L2

Machine Learning Classifier
- Random Forest
- 10-fold cross-validation
- **Evaluation:** median and IQR of [WAF, Precision, Recall]

Hypothesis to test

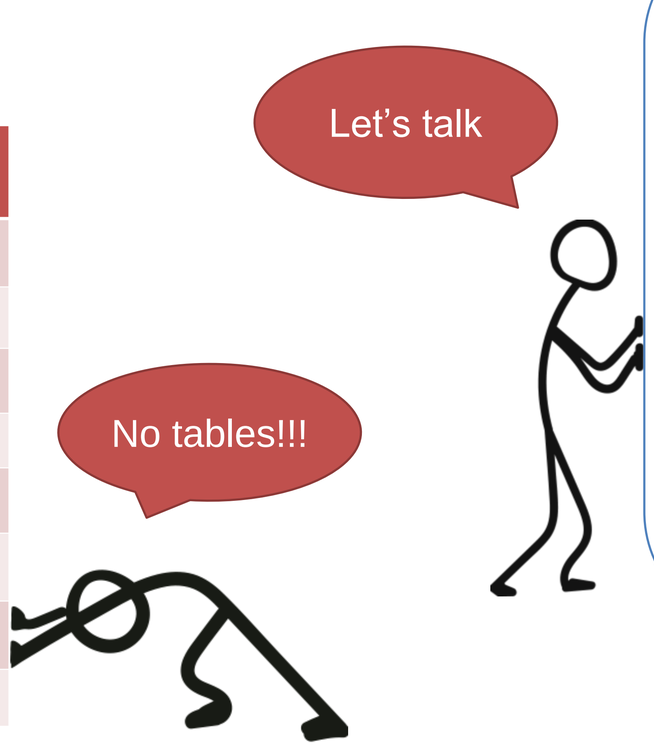
- Logical definitions
- Mappings
- Disease Ontology (DO)

Results

Example results median performance RDF2Vec:

KGs / Operators	Cosine Similarity	Hadamard
GO + HP	[0.687, 0.767, 0.555]	…
GO + HP*	[0.683, 0.765, 0.557]	…
GO + HP**	[**0.690**, 0.772, 0.561]	…
GO + HP***	[0.684, 0.748, **0.570**]	…
GO + HP + DO	[0.684, 0.755, 0.559]	…
GO + HP* + DO	[0.683, 0.765, 0.550]	…
GO + HP** + DO	[0.681, 0.768, 0.543]	…
GO + HP*** + DO	[0.680, **0.788**, 0.526]	…

GO - Gene Ontology
HP - Human Phenotype ontology
* - Logical definitions
** - Mappings
*** - Logical definitions and mappings



- Hadamard operator outperforms other operators.
- RDF2Vec achieves better results.
- The simplest version **without** logical definitions reach superior performance than the version with more links of these.
- However, Knowledge Graphs with mappings and logical definitions provides richer information to predict gene-disease associations.
- Adding DO does not significantly improve results.

Conclusions

- RDF2Vec with the gene and disease vectors combined by Hadamard operator achieves better performance when the Knowledge Graph include logical definitions and mappings.
- DO does not provide such relevant information beyond data on genes and disease phenotypes.

What comes next?

Thus, we are going to

- Gene-disease association prediction as a **link prediction** task.
- Integrate curated **gene-disease associations** from DisGeNET.
- Add a **protein-protein interaction** network in some Knowledge Graphs.
- Analyse how different Knowledge Graph characteristics impact on the performance of **Knowledge Graph embedding** algorithms and **Graph Neural Network** Models.

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

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020.

Cooperative Autonomous Driving in Simulation

Gonçalo Costa, José Cecílio, António Casimiro

LASIGE, Faculdade de Ciências da Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

Motivation

- Autonomous vehicles are being designed to provide high degree of safety.
- They will allow to reduce the number of accidents.
- Actual vehicles do not communicate with each other to get more information.
- Each vehicle is wasting resources computing the same information.

Context

- We will use SVL, an ultra-realistic simulator, to create real-world scenarios, implement and test a protocol for cooperative driving.
- An autonomous driving (AD) stack (Apollo) is used to emulate the control system of autonomous vehicles.
- Car to car communication is emulated using a central publisher/subscriber broker supporting message exchange.

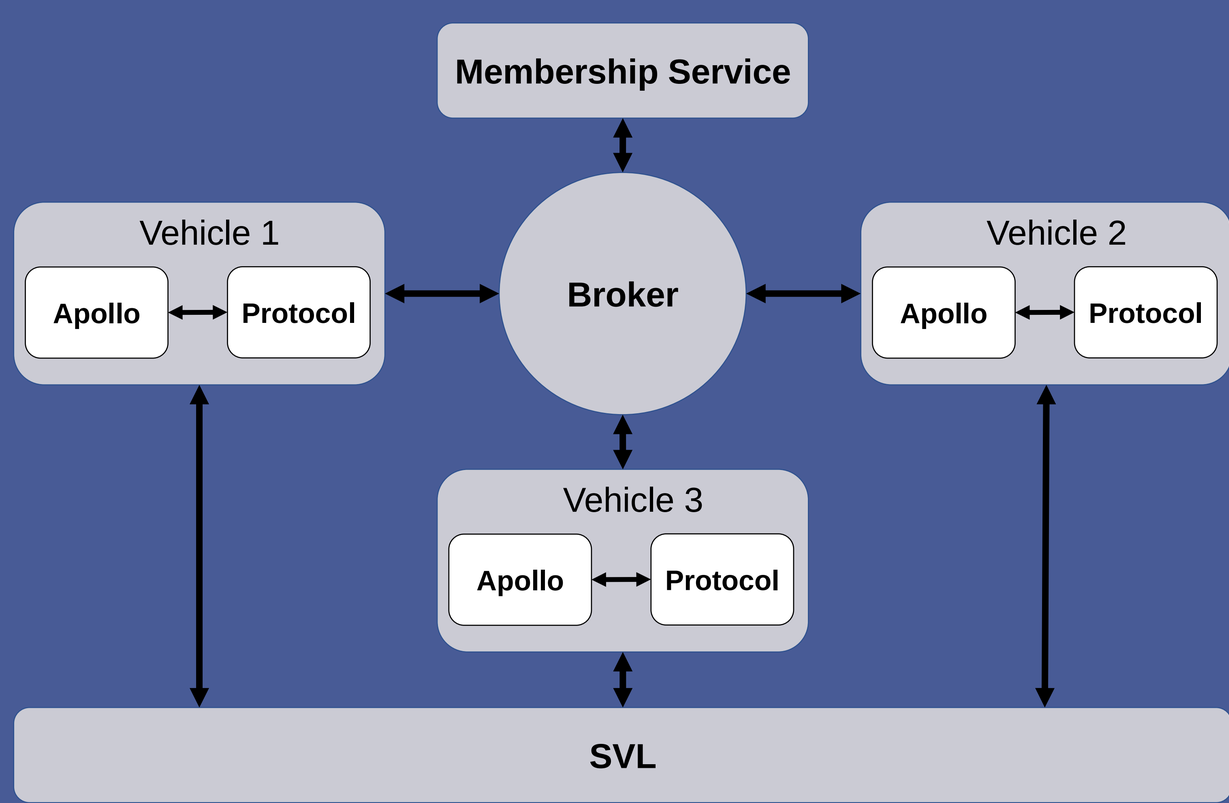
Objectives

- Define, implement, and test an architecture for a simulation environment where cooperative autonomous driving protocols can be tested.
- Implement and test a protocol for vehicle cooperation, applied to a road crossing scenario.
- Perform a reliability and resilience evaluation of the cooperative AD protocol through fault injection.

Protocol

- Vehicles have different priorities, depending on the manoeuvre they intend to perform.
- Vehicles communicate with each other to grant or deny a request for crossing an intersection.
- They also communicate with a Membership service to obtain their membership.
- This membership includes all other vehicles with which it is necessary to cooperate to cross the intersection safely.

Architecture



Challenges

- Installation and configuration of the Apollo AD stack.
- Implementation and integration of the protocol with the Apollo stack.
- Definition and realization of tests to evaluate the implementation correctness and performance.
- Definition and execution of purposeful communication failure scenarions to evaluate the protocol's reliability.

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

Acknowledgments: This work was supported by the LASIGE Research Unit (ref.UIDB/00408/2020 and ref. UIDP/00408/2020), and by the European Union's Horizon 2020 research and innovation programme under grant agreement No 871259 (ADMORPH project).

RA-Troia, Augmented Reality in Public Spaces

Mário Ferreira, Ana Paula Cláudio, M. Beatriz Carmo

LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

LASIGE cyber-physical systems

Introduction

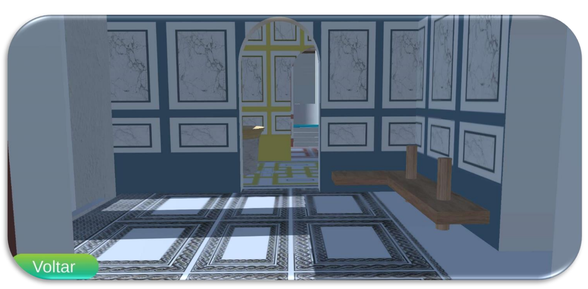
Augmented and virtual reality are technologies that are becoming more popular, and can help not only protect, but also promote cultural heritage. By using these technologies, we can enhance the visit to an archaeological site. We used materials from different sources to create the virtual models of 2 buildings from the Roman ruins of Troia and developed an Android app with 3 different types of experiences to show what the buildings looked like without damaging the site.

Augmented Reality

The first experience uses Augmented Reality, and from a viewpoint, we can see the virtual model of the building in its original place, size and shape.

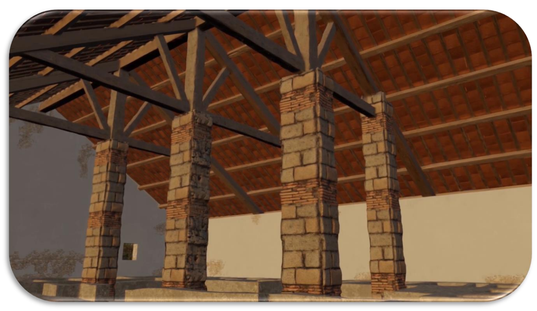
Virtual Navigation

The second experience uses Virtual Reality to allow the visitor to explore the virtual building from inside. Two variations of this experience were created: one where the visitor navigates freely, other where he is anchored at the centre of the building.



Virtual Visit

The third experience shows a video with a navigation around and inside the virtual building.
The scenario includes virtual humans in daily routines.



Evaluation

Two user tests were conducted:
1) At the archaeological site, with experts, to evaluate AR and VR experiences, regarding their usability and suitability to support site visits;
2) In an open space, to evaluate 2 different navigation modes in the virtual buildings: free navigation of the user in the building vs fixing a position for the user and allowing him to look around.

Conclusion and Future Work

The RA-Troia app allows to explore the virtual reconstruction of the thermae building and a salt factory of the Roman ruins of Troia.
In future work, the app can be extended including more buildings and the evolution over time of their construction phases.
In the evaluation tests, the users gave a very positive feedback and revealed their satisfaction in exploring the virtual buildings.

Acknowledgements: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UIDB/00408/2020 and UIDP/00408/2020

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

Adding target structure for improved *in silico* drug discovery in GPCRs

Pedro Oliveira[1], Rita Guedes[2] Andre O. Falcao[1]

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

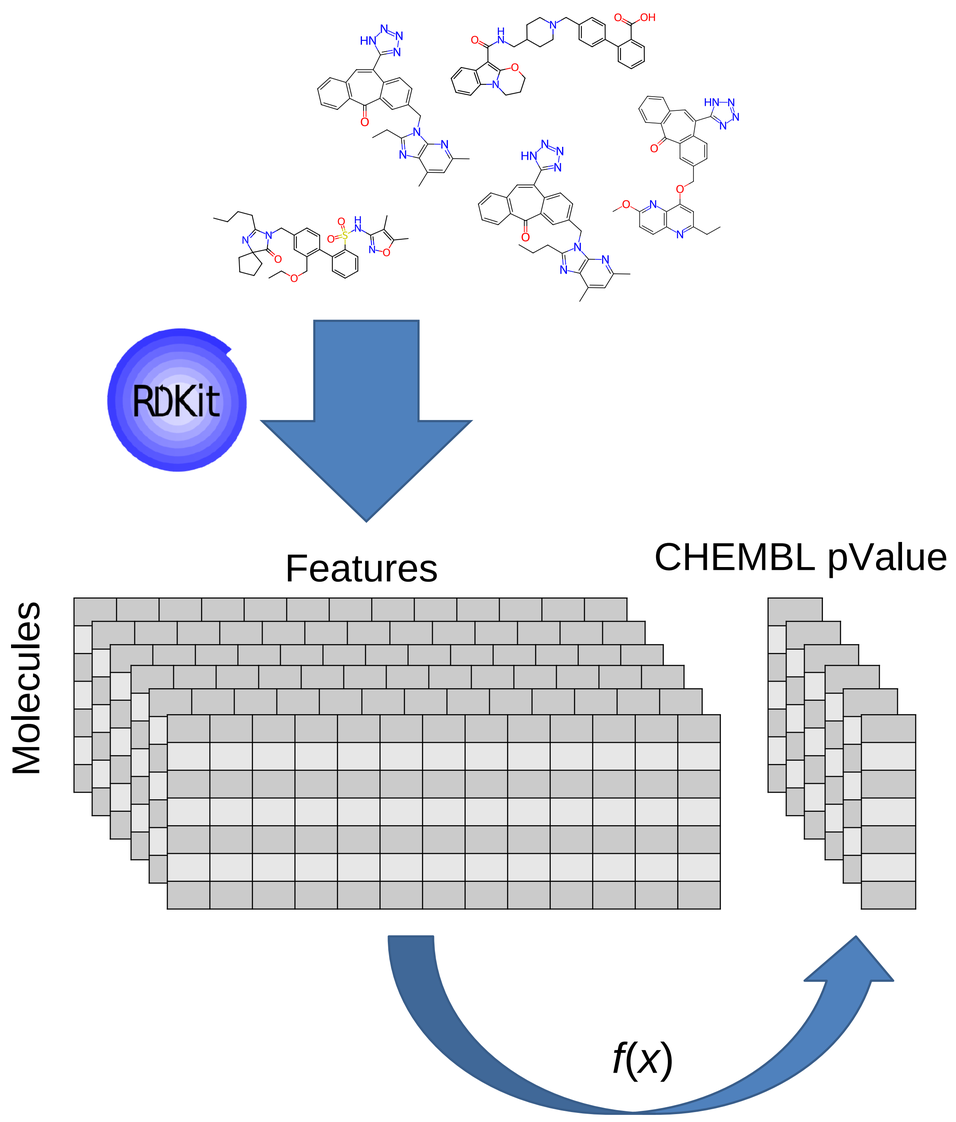
LASIGE health and biomedical informatics

Background

Virtual screening techniques have been used for some time and are a popular way to reduce the costs and time of the drug discovery process. One of these techniques is the quantitative structure-activity relationship (QSAR) which has the advantages of not being computationally expensive and being able to screen the increasingly available chemical libraries with good results.

Classical QSAR

- 1 Target -> 1 Model
- Multiple Targets - >Multiple models



Preliminary results

Results with normal QSAR:

	RMSE	R^2
Gradient Boosting	0.1945	0.5686
Support Vector Machines	0.1692	0.6710
Random Forests	0.1717	0.6609

Results for Random Forest models with the solution proposed (2 sets of different protein features):

	RMSE	R^2
Features 1	0.1709	0.6621
Features 2	0.1676	0.6734
Features 1+2	0.1674	0.6731

Motivation

Although good results have been obtained with QSAR models, there are some limitations:

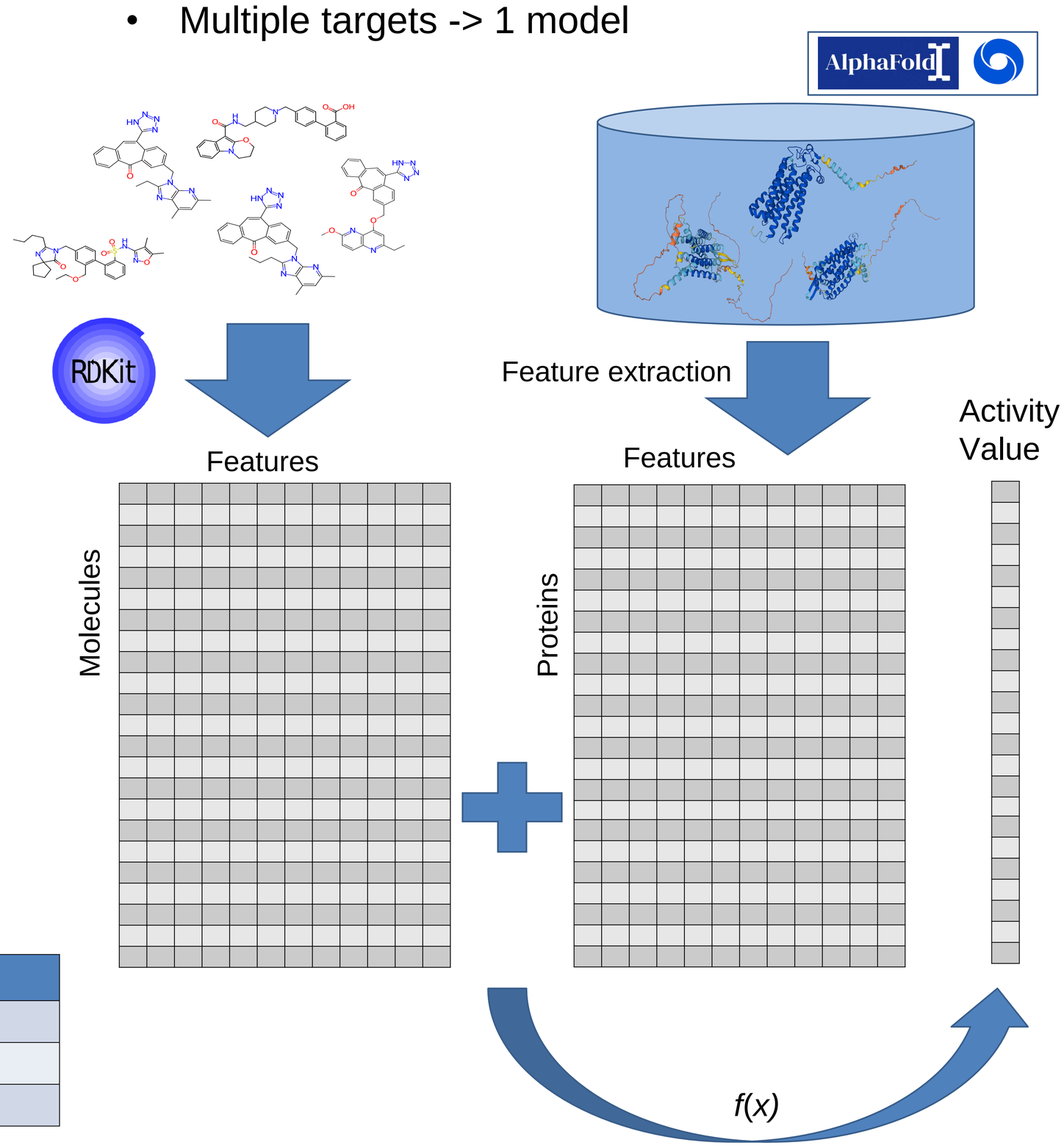
- These models only take into account the structural information about the molecules
- They require that previous assays had been conducted on the desired target to have bioactivities of each molecule available

This means that they can't be utilized in targets for which no or very few assays have been conducted.

With the development of AlphaFold, there is now a way to obtain proteins' structural information with good accuracy for virtually every protein.

Solution Proposed

- Multiple targets -> 1 model



Advantages

- Eliminates the requirement for assay results to be available
- Can help to understand the what are the protein features that most impact bioactivity

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

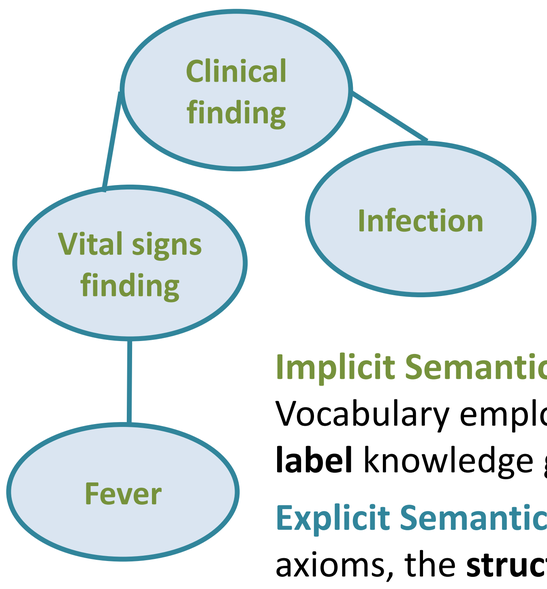
Knowledge Graph Alignment with Implicit and Explicit Semantics

Natércia Soeiro[1], Cátia Pesquita[1], Daniel Faria[2]

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

LASIGE health and biomedical informatics

Implicit and Explicit Semantics



Implicit Semantics: Vocabulary employed to **label** knowledge graph entities.

Explicit Semantics: The KG's rules and axioms, the **structure** of the graph.

Problem

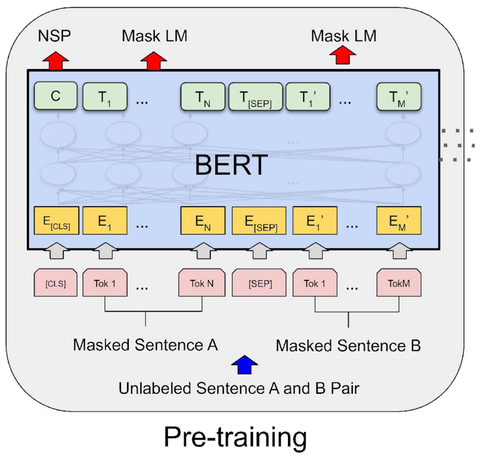
Knowledge graph (KG) alignment, or Ontology Matching, is necessary to combine different KGs, which allows for integrating knowledge held by different parties. However, current state-of-the-art techniques are not sufficient to find all existing matches between two KGs, as they depend as they depend fully on the information present in ontologies.

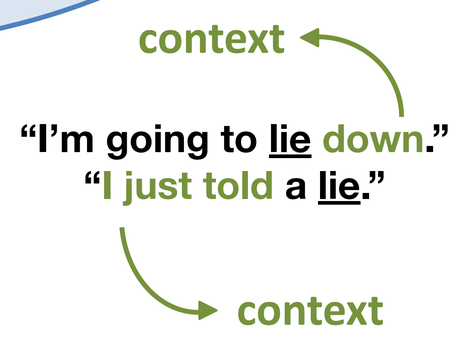
Solution

Integrate **Large Language Models (LLMs)** as an external source of information into the knowledge graph alignment process to complement state-of-the-art approaches.

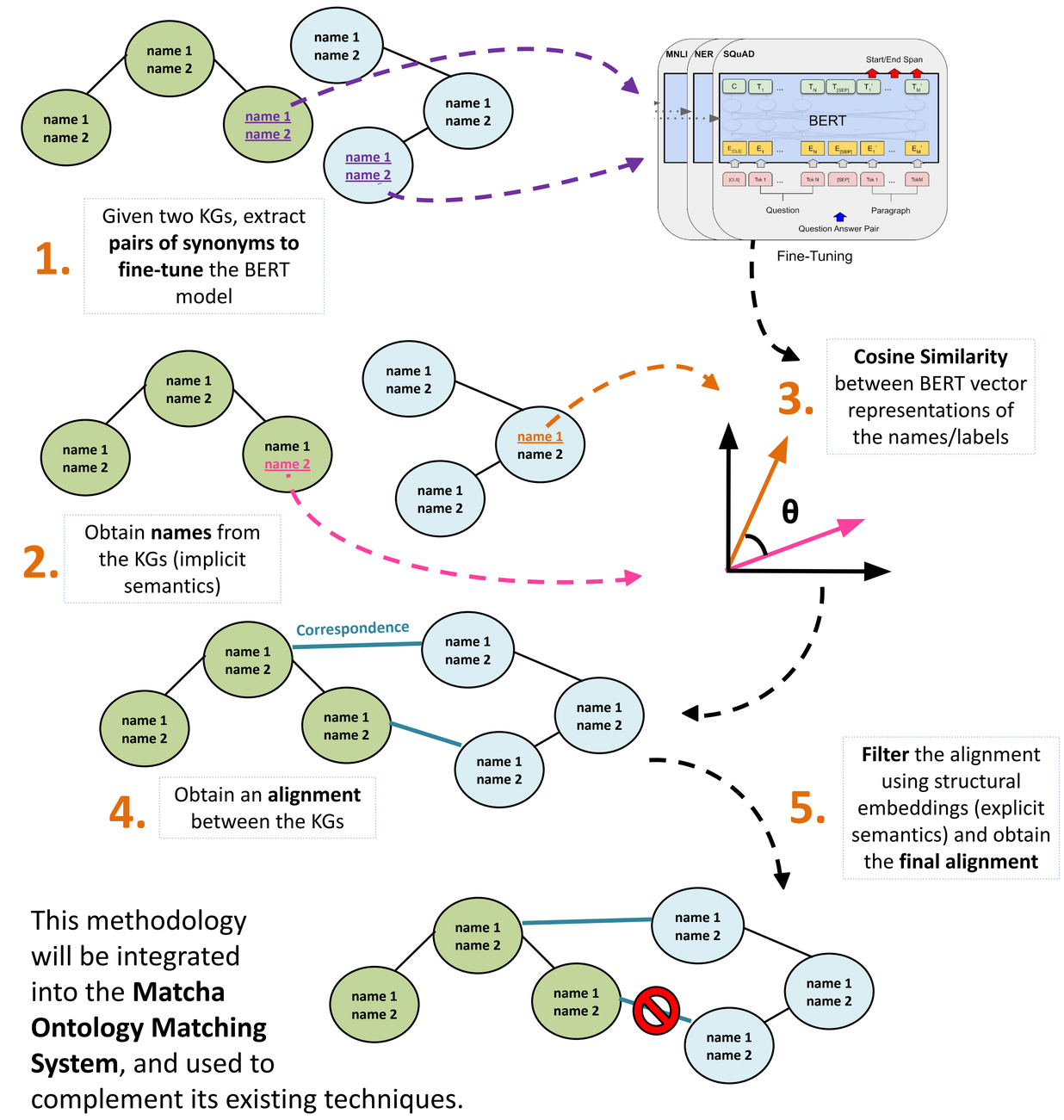
BERT

BERT, Bidirectional Encoder Representations from Transformers, is a NLP pre-trained algorithm, that **can generate context-aware word/sentence embeddings**, which are a form of vocabulary representation. BERT can also be **fine-tuned** for specific tasks.



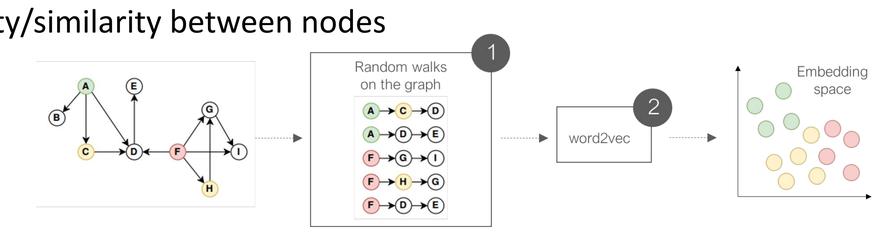


Methodology



This methodology will be integrated into the **Matcha Ontology Matching System**, and used to complement its existing techniques.

The filter used in step 5 uses the **RDF2Vec approach** to generate structural embeddings, and uses the same cosine similarity method to determine structural proximity/similarity between nodes



Preliminary Results

The table below showcases results that compare **Matcha's found correspondences to a reference set**, when testing with the Anatomy ontologies provided by the OAEI.

BERT Type	Precision	Recall	F-score	Found	Correct	New
Without	95.10%	93.00%	94.10%	1482	1410	-
Global	53.20%	93.40%	67.80%	2662	1416	6
Local	64.7%	93.70%	76.50%	2194	1420	10

- BERT Type "Without": Matcha's results using only it's **default matchers**;
- BERT Type **"Local":** Matcha with a BERT matcher that looks for correspondences in the **neighborhood** of other matched classes;
- BERT Type **"Global":** a BERT matcher that looks for them **naively**.
- Neither of the BERT Matchers has been fine-tuned;
- Minimum similarity score used as threshold for the BERT Matchers: 0.9.

The BERT Matchers are showing promising results so far, as they **can find new matches** between these ontologies, but the number of incorrect mappings **justifies the need** to implement the fine-tuning and filtering steps.

Forthcoming Work

The next steps consist of **developing the RDF2Vec Filter**, and **implementing the BERT fine-tuning** into Matcha, alongside the already implemented BERT Matcher. In the future, this method will also be tested with larger ontologies, with the goal of improving the state-of-the-art in KG Alignment.

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

Acknowledgements: This work was supported by the European Commission through the *ProjectName* project (REFREFREF) and by the Fundação para a Ciência e a Tecnologia (FCT) under LASIGE Strategic Project UID/CEC/00408/2019, the *ProjectName* project (REFREFREF), and FCT grant SFRH/BD/REFREF/YYYY.