Title: Terminology-aware English-Maltese Neural Machine Translation

Supervisor: Prof. Borg Claudia

Main Subject Area: NLP, Machine Translation

Short Description of the proposed FYP: Neural Machine Translation (NMT) systems often struggle with domain-specific and low-frequency terms, especially in under-resourced language pairs like English–Maltese. This project aims to improve the quality of English-to-Maltese NMT by integrating external terminology resources — such as IATE (Interactive Terminology for Europe) database — into the training and inference phases of an MT system.

Moreover, in collaboration with the Department of Translation, Terminology & Interpreting Studies, we will build a website to allow the department's students and staff to upload and process their terminology databases, making the latest contributions in the development of terminology easily accessible. This will allow easier access to well-curated terminology databases for English-Maltese in specific domains.

The student will investigate recent approaches to terminology-aware MT, especially soft and hard constraints in NMT, develop the terminology database and website, and integrate this terminology during fine-tuning or inference, either through constrained decoding, terminology tagging, or vocabulary augmentation strategies.

This project will contribute toward creating more domain-sensitive MT systems for Maltese and provide insights into resource-augmented translation for low-resource languages.

Recommended Resources (including readings): https://aclanthology.org/P19-1294/

https://aclanthology.org/2023.wmt-1.80/

https://aclanthology.org/2023.wmt-1.82/

Prerequisite Knowledge Required: NLP

List any Ethical issues: none, using publicly available datasets

Title: Low-resource Approaches for Text-to-Speech for Maltese

Supervisor: Prof. Borg Claudia

Main Subject Area: NLP, Deep Learning, Speech Synthesis

Short Description of the proposed FYP: Traditionally, Text-to-Speech (TTS) systems require several hours of recorded data by a single person (voice) to be used as training data for a TTS system for a particular language. This is generally challenging to collect for various reasons. In this project, we want to turn our attention to low-resource approaches that can leverage other types of spoken data available and try to apply it to TTS to develop a system that is capable of natural sounding Maltese Speech. In particular, this project investigates how existing Maltese Automatic Speech Recognition (ASR) datasets — typically comprising recordings from multiple speakers — can be reused or adapted for speech synthesis purposes. The main challenge here is that ASR datasets are often not ideal for TTS due to issues like speaker variability, inconsistent audio quality, and limited coverage of prosodic variation. However, recent developments in multi-speaker TTS, speaker embedding techniques, and transfer learning open up promising avenues for repurposing such data. The student will first conduct a review of recent low-resource and multi-speaker TTS techniques, including models like VITS, Tacotron 2 with speaker embeddings, and YourTTS. Using one of these toolkits (e.g., ESPnet, Coqui TTS, or NVIDIA NeMo), the student will prepare and preprocess an existing Maltese ASR dataset — performing tasks such as forced alignment between text and audio, speaker normalization, and embedding extraction. The goal will be to train and evaluate a TTS model that can produce intelligible and natural-sounding Maltese speech, even if it is trained from imperfect and heterogeneous data. The evaluation will focus both on naturalness (using metrics like Mean Opinion Score or MOSNet) and intelligibility (e.g., through re-synthesis WER or listener tests). This project will contribute toward developing speech technologies for under-resourced languages like Maltese, and will provide a valuable methodology for leveraging speech data that was not originally intended for synthesis.

Recommended Resources (including readings):

https://aclanthology.org/2022.aacl-main.56/
https://aclanthology.org/2023.paclic-1.27/
https://aclanthology.org/2022.acl-long.472/

Prerequisite Knowledge Required: NLP

List any Ethical issues: All datasets that will be used are publicly available for research purposes.

Title: Sentiment and Emotion Analysis for Maltese

Supervisor: Prof. Borg Claudia

Main Subject Area: NLP, Sentiment Analysis

Short Description of the proposed FYP: Sentiment Analysis stops short at identifying a comment as being Positive, Negative and Neutral. Emotion Analysis goes a step further and tries to classify comments into further subcategories (Joy, Trust, Surprise, Anticipation, Anger, Disgust, Fear, Sadness).

In this project we will analyse the current state-of-the-art in Sentiment Analysis and begin to move towards Emotion Analysis for Maltese.

Currently, we only have labelled data for Sentiment Analysis in Maltese. Thus in the first phase of the project, we will investigate how this data can be extended with emotion annotations. To do this, we will explore emotion taxonomies (e.g., Plutchik's Wheel of Emotions) and emotion annotation schemes used in English and other languages. We will then examine how these can be adapted for Maltese and applied to existing Maltese texts such as news comments, social media posts, or opinion forums.

The project will involve (i) transferring emotion classification models trained in high-resource languages (using multilingual models like XLM-R or mBERT), (ii) training a custom classifier using a manually or semi-automatically annotated dataset in Maltese. Emphasis will be placed on the challenges of fine-grained emotion classification, handling code-switching, and low-resource data augmentation (e.g., backtranslation, prompt-based generation).

The student will evaluate the emotion classifier using standard metrics (e.g., precision, recall, F1-score) and compare results against simpler sentiment classification baselines. The end goal is to provide a proof-of-concept system capable of automatically tagging user-generated Maltese content with emotion labels — a step forward for affective computing in low-resource languages.

Recommended Resources (including readings):
https://aclanthology.org/2024.eacl-long.18/
https://aclanthology.org/2023.wassa-1.26/
https://aclanthology.org/2024.lrec-main.1233/

Prerequisite Knowledge Required: NLP

List any Ethical issues: Publicly available datasets will be used.

Title: Automatic Transcription of Maltese Historic Documents

Supervisor: Prof. Borg Claudia

Main Subject Area: NLP

Short Description of the proposed FYP: Handwriting recognition can be seen as one of the first image to text tasks - a research area that began in the early 80s. However, whilst clean handwriting is easy to automatically transcribe and digitalise, older handwritten documents, especially those written with inconsistent orthography, stylized scripts, or degraded ink, introduce a level of complexity not easily handled by standard OCR or handwriting recognition systems.

This project will specifically look at old manuscripts written in Maltese and will continue to build upon an existing collaboration with the Library at the University of Malta. The primary aim is that of building a model that is able to assist librarians in the digitalisation process of manuscripts in its archive.

In order to achieve the overall goal, research will seek to improve the preprocessing pipelines (e.g., image binarization, segmentation, and line extraction), train or fine-tune handwriting recognition models and then use post-processing techniques to fix any OCR inconsistencies. Experiments will also look at the differences in model performance between line-based approaches to full-page approaches and determine what is the best approach to take in a low-resource scenario such as Maltese.

The final objective is to build a more accurate and robust transcription tool that supports University of Malta Library in its effort to digitise and preserve of Malta's cultural and linguistic heritage.

Recommended Resources (including readings):
https://aclanthology.org/2022.cltw-1.9/

https://aclanthology.org/2023.emnlp-main.953/

https://aclanthology.org/2023.alp-1.12/

Prerequisite Knowledge Required: NLP

List any Ethical issues: none

Title: Talk2MeLatin

Supervisor: Prof. Borg Claudia

Main Subject Area: NLP, Chatbot

Short Description of the proposed FYP: Learning Latin — a classical language no longer spoken natively — poses unique challenges for modern students, particularly in acquiring vocabulary, understanding syntax, and mastering inflectional morphology. While digital tools for Latin grammar and vocabulary exist, interactive conversational systems are rare.

This project proposes the development of a Latin-learning chatbot designed to help students practice reading, writing, and basic conversational Latin in an engaging and interactive way. This project will be carried out in collaboration with the Department of Classics & Archaeology.

The chatbot will be built using existing Latin NLP resources (e.g., the Classical Language Toolkit (CLTK), Latin treebanks, or UD Latin corpora). We will explore different techniques aimed at building a chatbot, potentially integrating LLM components and fine-tuning a model capable of understanding and generating a simple set of sentences in Latin as a proof-of-concept (e.g., greetings, translations, grammatical explanations). A lightweight chat interface (e.g., web or Telegram bot) will be developed and tested with beginner Latin students to gather feedback on usability and learning effectiveness.

The chatbot will be evaluated both for linguistic correctness (through expert or rule-based validation) and pedagogical value (via student feedback). The outcome will be a prototype tool to support Latin learners in a way that's interactive, accessible, and linguistically sound.

Recommended Resources (including readings):
https://aclanthology.org/2023.clinicalnlp-1.9/
https://aclanthology.org/2022.findings-emnlp.219/
https://aclanthology.org/W15-4618/

Prerequisite Knowledge Required: NLP

List any Ethical issues: none envisaged

Title: VIVA-AI: Verifying Integrity in Verbal Assessment using AI

Supervisor: Prof. Montebello Matthew

Main Subject Area: GenAI, Assessment Integrity & Conversational AI

Short Description of the proposed FYP: This project will investigate and research AI techniques to design and implement a GenAI-powered system to conduct and evaluate oral exams in real-time. The system should assess the student's responses using criteria such as coherence, subject knowledge, and critical thinking. It should also include integrity verification mechanisms — such as conversational consistency checks, real-time behavioural analysis (e.g., hesitation patterns, eye contact if using video), and contextual follow-ups — to detect potential third-party assistance or AI-generated help. Integration of voice-based LLMs or transcription models (like Whisper) is encouraged. Evaluate system performance against live human examiner assessments.

Recommended Resources (including readings): none

Prerequisite Knowledge Required: none

List any Ethical issues: Testing involving participants will require prior permission.

Title: GenAI Tutor: A Personalized Learning Assistant

Supervisor: Prof. Montebello Matthew

Main Subject Area: AI & Adaptive Learning

Short Description of the proposed FYP: Develop a GenAI-powered chatbot that adapts explanations and practice questions based on a student's learning style, pace, and performance in a given subject (e.g., maths or programming). Integrate basic NLP profiling to personalise interactions.

Recommended Resources (including readings): none

Prerequisite Knowledge Required: none

List any Ethical issues: none

Title: MultiLinguaLearn: AI-Powered Language Localisation

Supervisor: Prof. Montebello Matthew

Main Subject Area: Multilingual Education & NLP

Short Description of the proposed FYP: This FYP employs AI techniques to build a system that takes educational content (e.g., a short lesson transcript) and translates it into multiple languages using GenAI, with localized cultural and idiomatic tuning while comparing output with standard translation tools.

Recommended Resources (including readings): none

Prerequisite Knowledge Required: none

List any Ethical issues: none

Title: EduGen: Course Content Generator with Guardrails

Supervisor: Prof. Montebello Matthew

Main Subject Area: AI Ethics & Content Generation

Short Description of the proposed FYP: The project involves research and investigation into AI techniques to create a mini platform that uses Generative AI to draft lecture outlines, slides, and activity sheets for a specific subject. Include a human-in-the-loop "guardrail" mechanism to vet and refine content for accuracy, bias, and appropriateness.

Recommended Resources (including readings): none

Prerequisite Knowledge Required: none

List any Ethical issues: none

Title: AutoQuiz: Generative AI for Intelligent Assessment

Supervisor: Prof. Montebello Matthew

Main Subject Area: GenAI, Assessment & Learning Analytics

Short Description of the proposed FYP: The proposed FYP involves the use of AI techniques to design a tool that generates quizzes from course material (slides, lecture notes, or textbooks) using LLMs, with automatic difficulty scaling, levels of employing GenAI as a student assistant, as well as answer explanations. The project will ensure ethical use and academic integrity of the educational content while evaluating its accuracy and usefulness in real learning environments.

Recommended Resources (including readings): none

Prerequisite Knowledge Required: none

List any Ethical issues: none

Title: Quantifying uncertainty in affect modelling

Supervisor: Dr Makantasis Konstantinos

Main Subject Area: machine learning, affect modelling

Short Description of the proposed FYP: Recognising human emotional states is a challenging yet fundamental task in affective computing. The primary computational challenge arises from the subjective nature of emotions, which impacts the performance of emotion recognition models and necessitates methodologies for assessing prediction uncertainty. This project investigates how advanced machine learning algorithms can quantify uncertainty in emotion recognition models. The research will apply the conformal prediction framework to develop models that not only recognise emotions but also report confidence levels for their predictions. The implemented conformal prediction algorithms will be evaluated against conventional machine learning approaches, comparing both computational efficiency and recognition accuracy. This evaluation will utilise two publicly available datasets (AGAIN and either RECOLA or SEWA). The project will deliver two key outputs: 1) A comprehensive evaluation of the conformal prediction framework for emotion recognition and 2) An open-access code repository containing all implemented algorithms.

Recommended Resources (including readings): [1] Shafer, Glenn, and Vladimir Vovk. "A tutorial on conformal prediction." Journal of Machine Learning Research 9.3 (2008).

[2] Makantasis, Konstantinos, Antonios Liapis, and Georgios N. Yannakakis. "The pixels and sounds of emotion: General-purpose representations of arousal in games." IEEE Transactions on Affective Computing 14.1 (2021): 680-693

Prerequisite Knowledge Required: Python

List any Ethical issues: none

Title: Learning invariant representations in affect modelling

Supervisor: Dr Makantasis Konstantinos

Main Subject Area: machine learning, affect modelling

Short Description of the proposed FYP: Affect modeling seeks to uncover the complex relationship between affect elicitation, manifestation, and annotation. Currently, the ground truth in affect studies relies heavily on subjective affect labels, which inherently contain biases due to the subjective nature of emotions. Each unique combination of user, task, and annotator creates a distinct data generation distribution, introducing systematic bias into the collected data. While collecting larger datasets is a common approach to address these limitations, this is often impractical. This study proposes an alternative framework by exploring how invariant data representations can reduce or potentially eliminate bias in affect modeling. The research will specifically investigate slow features analysis as a method for deriving data representations that remain consistent across different user-task-annotator combinations. These learned invariant representations will be integrated with various machine learning models and evaluated for their effectiveness in affect modeling accuracy, both within and across different user-task-annotator combinations. The project will deliver two key outputs: 1) A comprehensive evaluation of slow features analysis for learning invariant representations in emotion recognition and 2) An open-access code repository containing all implemented algorithms.

Recommended Resources (including readings): [1] Wiskott, Laurenz, and Terrence J. Sejnowski. "Slow feature analysis: Unsupervised learning of invariances." Neural computation 14.4 (2002): 715-770.

[2] Makantasis, Konstantinos, Antonios Liapis, and Georgios N. Yannakakis. "The pixels and sounds of emotion: General-purpose representations of arousal in games." IEEE Transactions on Affective Computing 14.1 (2021): 680-693

Prerequisite Knowledge Required: Python

List any Ethical issues: none

Title: Integrating privileged information in different affect modelling paradigms

Supervisor: Dr Makantasis Konstantinos

Main Subject Area: machine learning, affect modelling

Short Description of the proposed FYP: Despite recent advances in Affective Computing powered by sophisticated deep learning algorithms, a significant challenge persists: reliably transferring affect models trained on laboratory data (in-vitro) to real-world settings (in-vivo). Laboratory environments yield high-quality, precisely measured multimodal data, while real-world affect measurements are compromised by various environmental and experimental factors. This information disparity between in-vitro and in-vivo conditions severely limits the practical application of affect models in real-world scenarios.

This study aims to address this gap using the Learning Using Privileged Information (LUPI) framework, which is specifically designed for scenarios where different amounts of information are available during model training versus testing phases. The research will systematically integrate LUPI principles into three distinct affect modeling paradigms: 1) Affect modeling as a classification problem, 2) Affect modeling as a regression problem, and 3) Affect modeling as a ranking problem. The LUPI-based algorithms will be rigorously evaluated against conventional machine learning approaches that do not utilise privileged information. This comparative analysis will assess both computational efficiency and modelling accuracy using two publicly available datasets (AGAIN and either RECOLA or SEWA). The project will deliver two key outputs: 1) A comprehensive evaluation of the LUPI framework for emotion recognition, and 2) An open-access code repository containing all implemented algorithms.

Recommended Resources (including readings): [1] Makantasis, Konstantinos, et al. "From the lab to the wild: Affect modeling via privileged information." IEEE Transactions on Affective Computing (2023).

[2] Makantasis, Konstantinos, Antonios Liapis, and Georgios N. Yannakakis. "The pixels and sounds of emotion: General-purpose representations of arousal in games." IEEE Transactions on Affective Computing 14.1 (2021): 680-69
[3] Vapnik, Vladimir, and Akshay Vashist. "A new learning paradigm: Learning using privileged information." Neural networks 22.5-6 (2009): 544-557

Prerequisite Knowledge Required: Python

List any Ethical issues: none

Title: Shape-to-Path Conversion and Drawing with a Mobile Robot

Supervisor: Dr Galea Ingrid

Main Subject Area: Robotics

Short Description of the proposed FYP: This project explores how a low-cost mobile robot can transform shapes into physical drawings using motion control. The system, built using the Elegoo Smart Robot Car Kit V4.0, will be equipped with a pen and servo lifting mechanism, allowing the robot to draw on large sheets of paper as it moves.

The student will implement:

a) A path translation algorithm to convert shapes into a sequence of robot-friendly motion commands (e.g., turn, move forward, lift/lower pen)

b) A pen-lifting mechanism controlled by a servo motor for handling pen contact

c) A motion execution system that interprets the path and drives the robot to draw

Recommended Resources (including readings): none

Prerequisite Knowledge Required: Arduino/C++ programming (motor control, servo control, digital logic)

Basic geometry and coordinate transformations (e.g., turning angles, distances) Python scripting (for vector-to-path translation and shape processing)

Recommended: Student has completed ARI2202.

List any Ethical issues: none

Title: MapBot: A Low-Cost Robot for Indoor Mapping and Navigation

Supervisor: Dr Galea Ingrid

Main Subject Area: Robotics

Short Description of the proposed FYP: This project focuses on developing a low-cost, autonomous robot that can explore and map indoor environments using basic sensor fusion and intelligent navigation strategies. Built on the Elegoo Smart Robot Car Kit V4.0, the robot combines ultrasonic sensing, IMU-based orientation estimation (MPU6050), and time-based motion tracking to approximate its position in the environment (dead reckoning). As it navigates, it detects obstacles and constructs a 2D occupancy grid map by transmitting sensor data to a computer, where mapping is handled in real time using Python.

While this system does not implement full SLAM, it adopts a SLAM-inspired approach to mapping and navigation within the constraints of affordable hardware. The robot makes basic AI-driven decisions, such as obstacle avoidance and space coverage, to autonomously explore unknown environments. The project includes system design, real-time sensor integration, AI decision-making, and evaluation of the mapping results against known layouts.

Recommended Resources (including readings):

https://github.com/AtsushiSakai/PythonRobotics
Elegoo Smart Car Kit V4.0 Documentation

Prerequisite Knowledge Required: Arduino/C++ programming (motor control, sensor integration, serial communication)

Basic geometry and coordinate transformations (e.g., dead reckoning, turning angles, position updates)

Python scripting (for sensor data processing, real-time map plotting)

Recommended: Student has completed ARI2202.

List any Ethical issues: none

Title: RoboTag: Chase and Evade with Two Mobile Robots

Supervisor: Dr Galea Ingrid

Main Subject Area: Robotics

Short Description of the proposed FYP: This project explores a physical implementation of the classic game of tag using two mobile robots built with the Elegoo Smart Robot Car Kit V4.0. One robot acts as the chaser, the other as the evader. The focus is on developing and evaluating multi-agent behaviors, where each robot uses onboard sensors (e.g., ultrasonic, IR) and decision logic to either pursue or evade in real time.

The student will implement:

A real-time chasing strategy (e.g., greedy pursuit with obstacle avoidance)

An evasion strategy (e.g., fleeing, random dodging)

Distance-based interaction using ultrasonic sensors to detect proximity Behaviour switching when a "tag" occurs (e.g., role reversal if distance < threshold) The game will take place in a simple, enclosed arena marked with tape or soft boundaries.

Recommended Resources (including readings):
https://dl.acm.org/doi/10.1145/3640824.3640843
https://dl.acm.org/doi/10.1145/3636534.3694724
Elegoo Smart Car Kit V4.0 Documentation

Prerequisite Knowledge Required: Arduino/C++ programming (motor control, sensor integration, state management, and serial communication) Sensor-based decision-making (ultrasonic distance sensing, behavior switching, collision detection)

Recommended: Student has completed ARI2202.

List any Ethical issues: none

Title: AgentVerse: A Multi-Domain Generative AI Framework

Supervisor: Prof. Dingli Alexiei

Main Subject Area: Artificial Intelligence (AI) – Multi-Agent Systems, Generative AI, Human-Computer Interaction

Short Description of the proposed FYP: The proposed final year project focuses on developing and exploring AgentVerse, an advanced framework designed to facilitate the creation of domain-specific generative AI agents. AgentVerse implements multi-agent systems with adaptive interaction models, enabling the seamless adaptation of AI agents across various contexts, including education, healthcare, business automation, and the creative industries. Drawing inspiration from ecosystems such as Stanford's Smallville and OpenAI's Swarm, AgentVerse embodies a modular and extensible environment where each agent possesses specialised expertise. These agents interact dynamically to solve complex problems, automate workflows, and provide tailored insights. The framework emphasises user experience, ethical AI practices, and scalability, empowering users to harness the potential of generative AI across diverse applications. The project will involve the design and implementation of a prototype multi-agent framework. This system will demonstrate the collaborative capabilities of specialised agents in a selected domain, showcasing their ability to interact, adapt, and provide solutions in real-time. The project will also address the ethical considerations of deploying such systems, ensuring that the developed agents adhere to responsible AI practices.

Recommended Resources (including readings): Park, J., et al. (2023). Generative Agents: Interactive Simulacra of Human Behavior.
Chen, W., Su, Y., Zuo, J., et al. (2023). AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors. https://github.com/openai/swarm
https://rikiphukon.medium.com/stanford-smallville-is-officially-open-source-9882e3fbc981

Prerequisite Knowledge Required: Proficiency in Python programming. Understanding of artificial intelligence concepts, particularly in multi-agent systems and generative models. Familiarity with machine learning frameworks and tools. Basic understanding of ethical considerations in AI development.

List any Ethical issues: Bias and Fairness: Ensuring that the AI agents do not perpetuate or amplify existing biases present in training data. Transparency: Providing clear explanations of agent decision-making processes to users. Privacy: Safeguarding user data and ensuring compliance with data protection regulations. Autonomy and Control: Maintaining appropriate levels of human oversight over autonomous agent actions. Accountability: Establishing mechanisms to hold systems accountable for their actions and decisions.

Title: HoloTutor: AI-Guided Hand Movement Training in Mixed Reality

Supervisor: Prof. Dingli Alexiei

Main Subject Area: Human-Computer Interaction (HCI), Virtual Reality (VR), Artificial Intelligence (AI)

Short Description of the proposed FYP: The proposed project, HoloTutor, aims to develop an immersive training system that captures expert hand movements using the Meta Quest 3's advanced hand-tracking capabilities. The system will create a comprehensive dataset of expert actions by recording these movements within a specific context, such as operating a home appliance. This data will then be utilised to train an AI model that guides novices through the same tasks in a virtual environment and an augmented reality setting. The system will provide real-time feedback to users, correct mistakes, and reinforce proper techniques. This approach leverages the natural interaction afforded by hand tracking, eliminating the need for physical controllers and enhancing the learning experience. HoloTutor seeks to improve the efficiency and effectiveness of skill acquisition in various practical domains by focusing on the replication and correction of hand movements.

Recommended Resources (including readings):
https://developers.meta.com/horizon/documentation/unity/unity-handtracking-overview/
https://www.frontiersin.org/journals/virtual-reality/articles/10.3389/frvir.2021.728461/full

Prerequisite Knowledge Required: Proficiency in virtual reality development platforms, especially Unity.

Understanding of machine learning principles, particularly supervised learning and imitation learning.

List any Ethical issues: Data Privacy: Ensuring that recorded hand movement data is anonymised and securely stored to protect user privacy. Informed Consent: Obtaining explicit consent from participants whose movements are recorded for training purposes.

Bias and Inclusivity: Ensuring that the AI model is trained on diverse datasets to prevent bias and ensure applicability across different user groups.

Dependence on Technology: Addressing the risk of over-reliance on AI guidance, which may impact the development of independent problem-solving skills.

Title: AgentCall: Development of an AI-Powered Conversational Agent for Automated Virtual Meeting Participation and Summarisation

Supervisor: Prof. Dingli Alexiei

Main Subject Area: Artificial Intelligence, Natural Language Processing, Human-Computer Interaction

Short Description of the proposed FYP: The proposed project, AgentCall, aims to design and implement an AI-driven conversational agent capable of autonomously participating in virtual meetings on behalf of a user. Leveraging Large Language Models (LLMs), the agent will be orchestrated through predefined scripts to engage in dialogues during meetings conducted via platforms such as Google Meet. Post-meeting, the agent will transcribe the conversation and generate a concise summary, which will be delivered to the user. The development process will involve integrating speech recognition for real-time transcription, employing LLMs for natural language understanding, response generation, and utilising summarisation techniques to distil key information from meetings. The project will also address challenges related to real-time communication, context management, and ethical considerations in AI-mediated interactions.

Recommended Resources (including readings): https://dxiaochuan.medium.com/summarising-your-meeting-with-chatgpt-and-langchain-8eb646cfcdd1
https://arxiv.org/abs/2410.18624?utm_source=chatgpt.com

Prerequisite Knowledge Required: Proficiency in Python, with experience in libraries such as TensorFlow or PyTorch for implementing machine learning models. Familiarity with NLP concepts, including language modelling, text summarisation, and dialogue systems. Ability to integrate and utilise APIs, particularly those related to speech recognition and virtual meeting platforms.

List any Ethical issues: Privacy and Consent: Recording and transcribing meetings involve handling sensitive information. Ensuring that all participants are informed and have consented to the recording is paramount. Data Security: Safeguarding the stored transcripts and summaries against unauthorised access or breaches is critical to maintain confidentiality. Transparency: The AI agent's presence and role in meetings should be transparent to all participants to avoid deception or misunderstandings. Bias and Fairness: The LLMs employed may inadvertently perpetuate biases present in their training data. Measures should be taken to identify and mitigate such biases to ensure fair and unbiased interactions. Reliability: Over-reliance on AI-generated summaries without human oversight could lead to misinterpretations. Establishing mechanisms for human review is advisable.

Title: A Coverage Path Planning System for Aerial Vehicles

Supervisor: Dr Bajada Josef

Main Subject Area: Planning and Reinforcement Learning

Short Description of the proposed FYP: Coverage path planning is the problem of finding an efficient path that maximizes the coverage of an area under some constraints. These constraints could be requirements such as covering a percentage of the area, a minimum overlap between paths, time available to execute the path, and battery life. This problem has various applications, such as search and rescue, exploration and discovery, mapping, remote sensing and photogrammetry.

In this project we will explore the current state-of-the art techniques for coverage path planning and adapt them to Unmanned Aerial Vehicles such as drones. These include algorithms based on classical approaches, together with reinforcement learning techniques. We will initially look into 2D spaces, with the plan to move to 3D environments. A multi-agent approach can also be investigated. We will review existent benchmarks and analyse the advantages and trade-offs of algorithms, determining which is the most effective solution for this problem.

Recommended Resources (including readings): Tan, C. S., Mohd-Mokhtar, R., & Arshad, M. R. (2021). A comprehensive review of coverage path planning in robotics using classical and heuristic algorithms. IEEE Access, 9, 119310-119342.

Kumar, K., & Kumar, N. (2023). Region coverage-aware path planning for unmanned aerial vehicles: A systematic review. Physical Communication, 59, 102073.

Aggarwal, S., & Kumar, N. (2020). Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges. Computer communications, 149, 270-299.

Cabreira, T. M., Brisolara, L. B., & Paulo R, F. J. (2019). Survey on coverage path planning with unmanned aerial vehicles. Drones, 3(1), 4.

Prerequisite Knowledge Required: Search and Planning Algorithms. Reinforcement Learning.

List any Ethical issues: none

Title: An Autonomous Underwater Navigation System

Supervisor: Dr Bajada Josef

Main Subject Area: Planning and Reinforcement Learning

Short Description of the proposed FYP: Autonomous Underwater Vehicles have several applications, such as seabed mapping and geological surveys, wreck exploration and archaeological discovery, underwater infrastructure inspection, and ecological monitoring. They can be used to extend underwater missions, which would otherwise be restricted in duration and depth if performed by humans or require significantly higher costs and risks.

In this project we will investigate the problem of autonomous underwater navigation, and the challenges it carries with regards to visibility, localization and drift. We will explore the sensors that are available to infer the surrounding environment, such as Sonar, depth sensor and cameras. We will analyze different approaches, including classical search-based techniques and also reinforcement learning algorithms. The project will make use of simulation frameworks to facilitate the development, testing and evaluation of the techniques employed.

Recommended Resources (including readings): Zhang, B., Ji, D., Liu, S., Zhu, X., & Xu, W. (2023). Autonomous underwater vehicle navigation: A review. Ocean Engineering, 273, 113861.

Jalal, F., & Nasir, F. (2021, January). Underwater navigation, localization and path planning for autonomous vehicles: A review. In 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST) (pp. 817-828). IEEE.

Christensen, L., de Gea Fernández, J., Hildebrandt, M., Koch, C. E. S., & Wehbe, B. (2022). Recent advances in ai for navigation and control of underwater robots. Current Robotics Reports, 3(4), 165-175.

Guo, Y., Liu, H., Fan, X., & Lyu, W. (2021). Research progress of path planning methods for autonomous underwater vehicle. Mathematical Problems in Engineering, 2021(1), 8847863.

Prerequisite Knowledge Required: Search, Planning and Reinforcement Learning

List any Ethical issues: none

Title: Next-Basket Prediction System

Supervisor: Dr Bajada Josef

Main Subject Area: Machine Learning, Recommender Systems

Short Description of the proposed FYP: Next-basket Prediction involves predicting the items in a user's next set of purchases, referred to as a basket. This is especially useful in recurrent retail settings such as groceries and supermarkets. Predicting the items a user is likely to buy next can provide various opportunities, such as personalized offers and automatic basket recommendations.
In this project we will explore various techniques to predict the items a user is likely to buy next from the previous history of the user and other users with similar buying patterns.
We will investigate techniques such as Recurrent Neural Networks (RNNs), Collaborative Filtering and Time-Series Analysis. We will have a special focus on grocery shopping, since this has a high element of repeat item purchasing. We will analyse one of more datasets such as the Instacart dataset or the Acquire Valued Shoppers Challenge dataset, and determine which Machine Learning techniques are most effective for this problem.

Recommended Resources (including readings): Mozhdeh Ariannezhad, Sami Jullien, Ming Li, Min Fang, Sebastian Schelter, and Maarten de Rijke. 2022. "ReCANet: A Repeat Consumption-Aware Neural Network for Next Basket Recommendation in Grocery Shopping." In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '22). Association for Computing Machinery, New York, NY, USA, 1240–1250. https://doi.org/10.1145/3477495.3531708
Ming Li, Sami Jullien, Mozhdeh Ariannezhad, and Maarten de Rijke. 2023. "A Next Basket Recommendation Reality Check." ACM Trans. Inf. Syst. 41, 4, Article 116 (October 2023).
Hu, Haoji, et al. "Modeling personalized item frequency information for next-basket recommendation." Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval. 2020.
Faggioli, G., Polato, M., & Aiolli, F. (2020, July). Recency aware collaborative filtering for next basket recommendation. In Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (pp. 80-87).

Prerequisite Knowledge Required: Supervised and Unsupervised Machine Learning

List any Ethical issues: none

Title: True News: An AI-Powered Mobile Application for Balanced News Aggregation and Personalised Consumption

Supervisor: Prof. Dingli Alexiei

Main Subject Area: Artificial Intelligence, Natural Language Processing, Mobile Application Development, Human-Computer Interaction

Short Description of the proposed FYP: The need for reliable and balanced news consumption has become important in an era characterised by information overload and the proliferation of misinformation. "True News" aims to address this challenge by developing a mobile application that aggregates news from pre-selected, reputable sources and employs advanced Large Language Models (LLMs) to generate concise, balanced, and well-researched summaries. The application will present these summaries and provide users with the original sources, enabling transparency and fostering trust.

Key features of the application include:

Customisable News Consumption: Users can tailor their news intake preferences, choosing between text, voice, alerts, or a combination thereof, ensuring accessibility and convenience.

Sentiment Analysis and Statistics: Each news item will be accompanied by sentiment analysis and relevant statistics, offering users more profound insights into the news content.

Historical Contextualisation: The application will link current news items to previous related articles, providing users with a comprehensive understanding of ongoing narratives.

The project will involve designing and implementing the mobile application, integrating news aggregation mechanisms, developing LLM-based summarisation pipelines, and incorporating user interface features that support customisation and interactivity. Emphasis will be placed on ensuring the ethical use of AI, maintaining data privacy, and delivering a user-centric experience.

Recommended Resources (including readings): https://docs.langchain.com/

https://docs.streamlit.io/

Prerequisite Knowledge Required: To successfully undertake this project, the student should possess:

Programming Skills: Proficiency in Python, with experience in libraries such as Pandas, NumPy, and frameworks like Flask or Django.

Machine Learning and NLP: Understanding of machine learning concepts, particularly in natural language processing, including familiarity with LLMs and tools like Hugging Face Transformers.

Mobile Application Development: Experience in developing mobile applications using platforms such as React Native or Flutter.

API Integration: Ability to work with RESTful APIs for data retrieval and integration.

Data Analysis: Skills in data preprocessing, analysis, and visualisation.

List any Ethical issues: The development and deployment of "True News" must consider the following ethical aspects:

Bias in AI Models: LLMs can inadvertently perpetuate biases present in training data. It is crucial to implement mechanisms to detect and mitigate such biases to ensure balanced news summaries.

Misinformation and Source Credibility: Aggregated news sources must be vetted for credibility to prevent the dissemination of false information. The application should provide transparency regarding source selection criteria.

User Privacy: Collecting user preferences and interaction data necessitates strict adherence to data privacy regulations, including GDPR compliance. User data should be anonymised and securely stored.

Transparency and Accountability: Users should be informed about the AI's role in content summarisation and have access to original sources to verify information.

Title: Bike Fit and Posture Estimation using Machine Learning

Supervisor: Dr Guillaumier Kristian

Main Subject Area: Machine learning, machine vision

Short Description of the proposed FYP: Bicycle sizing, geometry, and rider's physical characteristics are crucial elements in obtaining a comfortable fit and improving enjoyment, safety, and performance.

In this project, we aim to build a machine learning and vision based system that is able to calculate bicycle geometry and rider posture from images and/or video. This has important real-world applications including:

(1) Many people purchase bicycles online. While bicycle geometry databases exist, it is often difficult or impossible to determine one's comfort level or posture from this alone. Also bicycle sizing (e.g., small, medium, large) varies widely by manufacturer and bicycle type (e.g., road, gravel, mountain) -- most manufacturers make size recommendations based on rider's height alone. Furthermore, these recommendations do not take into account other parameters such as wheel size, crank, or stem length. This system will allow users to estimate their fit on a variety of bikes before committing themselves to a purchase.

(2) A good and comfortable bike fit can be achieved on existing bikes with simple and relatively inexpensive modifications such as seatpost height, saddle angle, crank arm length, stem length, and handlebar width. Such a tool can examine a rider's posture on a bike and make recommendations (e.g., to lower the saddle) to obtain an ideal fit.

Amongst the deliverables aimed for, a tool that can 'virtually' place a rider on a selected bike will be developed and allow for the tuning of parameters to observe how they affect fit and comfort.

Recommended Resources (including readings): Material will be provided during the start of the FYP

Prerequisite Knowledge Required: Preferably machine learning and vision study units

List any Ethical issues: None

Title: A DFA Learning Visualisation Toolkit

Supervisor: Dr Guillaumier Kristian

Main Subject Area: Machine Learning, Grammatical Inference, Search Algorithms, Heuristics, Algorithm Design

Short Description of the proposed FYP: Grammatical inference is the task of learning a formal grammar from strings that belong to the language and strings that do not. In this project, we will focus on the identification of regular languages (as deterministic finite-state automata, DFAs) from training sets consisting of both positive and negative examples. This inference task has many real-world applications including robotics, data mining, structural pattern recognition, speech recognition, and bioinformatics.

DFA learning algorithms typically involve instantiating and interacting with complex data structures. Furthermore, the operations involved on these data structures can be conceptually hard to understand. This is especially true as problem instances become harder (e.g., target DFAs become large and/or the density of training sets becomes sparser). The primary aim of this project is to extend an existing DFA learning toolkit (DfaGo) with a comprehensive set of visualization tools. This will help academics as well as people who are new to the area better understand the behaviour of common DFA learning algorithms, the underlying data structures, operations on them, how heuristics work, and help design new algorithms.

Recommended Resources (including readings): Material will be provided at the start of the FYP

Prerequisite Knowledge Required: Ideally, machine learning and formal languages study units

List any Ethical issues: None

Title: A Multimodal AI Approach to Analysing Nationality Representation in Maltese Online News

Supervisor: Dr Seychell Dylan

Main Subject Area: Vision and Language Technology

Short Description of the proposed FYP: This project explores how artificial intelligence can be used to analyse the representation of different nationalities in Maltese online news. Using a combination of computer vision and natural language processing techniques, the project examines both visual and textual portrayals of people from various national backgrounds.
Computer vision models will identify people in news images and classify potential nationality indicators based on visual cues, analysing factors such as frequency of appearance, context, and framing to reveal potential biases. Meanwhile, natural language processing techniques will be employed to quantify the frequency of mentions of different nationalities in news articles, the contexts in which they appear, and to assess the associated sentiment and framing.
The comparison of the findings from both analyses aims to provide a comprehensive understanding of how nationality is represented in Maltese online news. The insights generated will be valuable for media organisations, policymakers, and the public, contributing to a more informed and equitable media landscape in Malta, particularly regarding the portrayal of both Maltese nationals and foreign nationals.

Recommended Resources (including readings): The supervisor will provide datasets, access to experts and journalists, and any relevant literature.

Prerequisite Knowledge Required: The student needs to follow the following courses: ARI3129 - Advanced Computer Vision for AI and ARI3900 - Ethics and Artificial Intelligence

List any Ethical issues: The student should be mindful of ethical considerations such as data privacy, potential biases in AI models, and the risk of reinforcing national stereotypes. Particular attention must be paid to avoiding problematic proxies for nationality identification that could reinforce racial profiling. The project should acknowledge the limitations of visual identification of nationality and focus on contextual indicators rather than relying solely on physical appearance. Transparency regarding the methodology and responsible interpretation of findings is crucial to prevent misrepresentation and potential harm to individuals or national groups. The deliverables must also be careful not to contribute to divisive discourse around immigration and national identity.

Title: Interactive News Recommendation Algorithm Explorer

Supervisor: Dr Seychell Dylan

Main Subject Area: Recommendation Engines, Web Technology

Short Description of the proposed FYP: This project develops a web-based educational tool that illustrates how recommendation algorithms influence the news content users see online. The student will create an interactive web dashboard that allows users to manipulate various parameters of a simulated news recommendation algorithm and observe in real-time how these changes affect content selection and presentation. The system will incorporate content-based and collaborative filtering approaches, allowing users to explore how different features, such as user demographics, engagement patterns, and content characteristics, influence recommendations. By adjusting sliders, toggles, and other interface elements representing algorithmic parameters, users can visualise how seemingly minor modifications to the algorithm can significantly alter the news articles presented to them. The dashboard can also include visualisations showing content diversity, political leaning, topic distribution, and other metrics that change dynamically based on user adjustments. Additionally, it will provide educational explanations about the mechanisms behind recommendation systems and their broader societal implications. This project aims to demystify the "black box" nature of recommendation algorithms for the general public, promoting digital literacy and critical thinking about information consumption in an age of algorithmic curation. It will serve as a valuable resource for educators, policymakers, and concerned citizens seeking to understand the technology that shapes news consumption.

Recommended Resources (including readings): The supervisor will provide datasets, access to experts and journalists, and any relevant literature.

Prerequisite Knowledge Required: The student needs to follow the following courses: ARI3129 - Advanced Computer Vision for AI and ARI3900 - Ethics and Artificial Intelligence

List any Ethical issues: The student should consider ethical implications, such as potential biases in the training data used for the recommendation engine simulation, which could inadvertently perpetuate stereotypes or misinformation. Care must be taken to provide balanced representations of different news sources and political perspectives. While the goal is educational, the project should avoid oversimplification of complex algorithmic processes that might lead to misunderstandings about how actual recommendation systems function. Additionally, the dashboard should clearly communicate that it is a simulation and may not perfectly replicate the proprietary algorithms used by major platforms.

Title: IoT-Powered Multimodal AI Assistant Platform

Supervisor: Dr Seychell Dylan

Main Subject Area: Multimodal AI, Internet of Things

Short Description of the proposed FYP: This project creates a proof-of-concept system that demonstrates how multimodal AI models can interact with sensor data from connected devices. The student will build a desktop-based system that collects data from basic sensors (webcam, microphone, and simple environmental sensors) and processes this information using an open-source multimodal AI model, such as Google's Gemma.

The system will demonstrate how different types of data (visual, audio, numeric) can be combined to provide the AI model with a more complete context, allowing for more informed output. The project will focus on creating a functional pipeline rather than complex IoT architecture, with the desktop computer handling the processing and serving as the hub for connected sensors.

As an extension, students may explore adapting the concept to Android devices, leveraging built-in sensors (camera, microphone, accelerometer) to create a mobile version of the system. This would demonstrate how readily available consumer devices can serve as platforms for multimodal AI applications.

The final deliverable will be a demonstration showing how sensor inputs enhance AI understanding and reasoning in a selected domain, such as health. This will include a simple interface for configuring input sources and viewing the AI's interpretations and responses.

Recommended Resources (including readings): The supervisor will provide datasets, access to hardware, access to experts, and any relevant literature.

Prerequisite Knowledge Required: The student needs to follow the following courses: ARI3129 - Advanced Computer Vision for AI
ARI3900 - Ethics and Artificial Intelligence

List any Ethical issues: The student should address ethical considerations including data privacy concerns when collecting sensor data, being transparent about what information is being captured and processed, and implementing appropriate data handling practices. The project should focus on creating a system that processes data locally where possible to minimise privacy concerns, and should clearly communicate to users when sensors are active. As this is a proof-of-concept, careful consideration should be given to the limitations of the system and potential biases in the AI models used, which should be openly acknowledged in any demonstrations or documentation.

Title: AI-Assisted Newspaper Digitisation

Supervisor: Dr Seychell Dylan

Main Subject Area: Computer Vision

Short Description of the proposed FYP: This project creates a streamlined system to convert a limited collection of scanned newspaper issues into an accessible digital web archive. The student will work with a manageable dataset (approximately one week to one month of newspaper issues) to develop and test an effective digitisation pipeline.

The system will use both traditional OCR techniques and modern AI vision models (such as Gemini Flash) to extract text from scanned newspaper pages. The student will develop algorithms to identify and separate individual news articles within each page, recognising headlines, body text, and image captions. For each identified article, the system will generate a standalone web page with properly formatted content.

The final deliverable will be a simple web archive that presents each digitised article with its original text, images where available/possible, and metadata (date, page number, section). The archive will include basic search functionality and chronological browsing options. This proof-of-concept system will demonstrate how AI vision technology can make historical print media accessible in the digital age.

Recommended Resources (including readings): The supervisor will provide datasets, access to hardware, access to experts and journalists, and any relevant literature.

Prerequisite Knowledge Required: The student needs to follow the following courses:
ARI3129 - Advanced Computer Vision for AI
ARI3900 - Ethics and Artificial Intelligence

List any Ethical issues: The student should address ethical considerations including ensuring accurate representation of the original text when converting to digital format, and acknowledging any limitations or errors in the OCR process. The system should clearly indicate when text extraction might be uncertain or incomplete. Since this is a limited proof-of-concept, the project should document recommendations for scaling the approach responsibly, including where human oversight would be beneficial for quality control if the system were to be expanded to larger archives.

Title: Art & AI

Supervisor: Dr Vanessa Camilleri

Main Subject Area: Multimodality, Explainable AI

Short Description of the proposed FYP: To develop an AI-powered tool that analyzes visual artworks and predicts whether they were created by a human or generated by an AI system, using both visual and textual features. The project will also explore the explainability of the model's decisions and its potential to generate basic art critiques, contributing to discussions around creativity, authorship, and authenticity in AI-generated art.

Recommended Resources (including readings): none

Prerequisite Knowledge Required: none

List any Ethical issues: none

Title: Enhancing medicX with GenAI

Supervisor: Dr Abela Charlie

Main Subject Area: Machine Learning, Knowledge Graphs, Generative AI

Short Description of the proposed FYP: Demand for clinical decision support systems in medicine and self-diagnostic symptom checkers has substantially increased in recent years. Existing platforms rely on knowledge bases manually compiled through a labor-intensive process or automatically derived using simple pairwise statistics. The Malta Medicine's Authority addresses Malta's regulatory landscape and involves alignment with the European Medicines Agency but at the same time needs is still partially dependent on the UK supply.

In this dissertation you will investigate how to automatically or semi-automatically, learn a knowledge graph for Maltese healthcare by leveraging on multiple sources such as medical journals and medical imagery, through the use of NLP and Generative AI. This work will build over existing work done on the medicX system and related knowledge graph medicX-KG.

A possible use case for this knowledge graph is to serve as a tool for pharmacy students to be able to learn about a medicines by navigating through comprehensive information about the condition.

The system can be accessed both through a web interface as well as a mobile interface.

Recommended Resources (including readings): 1. Farrugia, L. and Abela, C., 2020, January. Mining drug-drug interactions for healthcare professionals. In Proceedings of the 3rd International Conference on Applications of Intelligent Systems (pp. 1-6).

2. Farrugia, L., Azzopardi, L.M., Debattista, J. and Abela, C., 2023. Predicting Drug-Drug Interactions Using Knowledge Graphs. arXiv preprint arXiv:2308.04172.

3. Farrugia, L., Azzopardi, L.M., Debattista, J. and Abela, C., 2024. medicX-KG: A Knowledge Graph for Pharmacists' Drug Information Needs.

Prerequisite Knowledge Required: none

List any Ethical issues: none

Title: Building a Continuous Integration Pipeline for Knowledge Graph Maintenance in medicX-KG

Supervisor: Dr Abela Charlie

Main Subject Area: Data Integration; Knowledge Graph; Continuous Integration

Short Description of the proposed FYP: medicX-KG is a biomedical knowledge graph (KG) developed to support pharmacists by integrating data from multiple trusted sources, such as the Malta Medicines Authority (MMA), DrugBank, and the British National Formulary (BNF).

However, biomedical knowledge is constantly evolving: new drugs are approved, indications change, and adverse effect information is updated. Without regular updates, the KG risks semantic drift, where old knowledge becomes inconsistent, outdated, or incorrect relative to current clinical realities.

To address this, we propose building a Continuous Integration (CI) pipeline for medicX-KG that automatically ingests updates from external sources and validates them before integration. This ensures the knowledge graph remains accurate, consistent, and clinically relevant over time.

The aim is therefore to design and implement an automated CI pipeline that continuously updates medicX-KG with minimal manual intervention, ensuring the knowledge graph remains reliable and up-to-date.

Recommended Resources (including readings): 1. Rossanez, A., Dos Reis, J.C., Torres, R.D.S. and de Ribaupierre, H., 2020. KGen: a knowledge graph generator from biomedical scientific literature. BMC medical informatics and decision making, 20, pp.1-24.

2. Xu, J., Kim, S., Song, M., Jeong, M., Kim, D., Kang, J., Rousseau, J.F., Li, X., Xu, W., Torvik, V.I. and Bu, Y., 2020. Building a PubMed knowledge graph. Scientific data, 7(1), p.205.

3. Noy, N., Gao, Y., Jain, A., Narayanan, A., Patterson, A. and Taylor, J., 2019. Industry-scale Knowledge Graphs: Lessons and Challenges: Five diverse technology companies show how it's done. Queue, 17(2), pp.48-75.

Prerequisite Knowledge Required: none

List any Ethical issues: none

Title: Retrieval-Augmented Generation for Enriching the medicX Knowledge Graph with Explainable Pharmacological Knowledge

Supervisor: Dr Abela Charlie

Main Subject Area: Machine Learning, Knowledge Graphs, Generative AI

Short Description of the proposed FYP: The effectiveness of knowledge graphs in pharmacology depends on the quality, breadth, and traceability of the information they represent.
This dissertation explores how Retrieval-Augmented Generation (RAG) can be used to semi-automatically enhance the medicX Knowledge Graph by extracting and integrating structured pharmacological knowledge from trusted sources such as SmPCs, DrugBank, and PubMed. The aim is to improve both the coverage and explainability of the graph for downstream use in clinical and educational settings.
The student will investigate the development of a RAG-based pipeline that retrieves relevant evidence and generates subject–predicate–object triples aligned with the medicX-KG schema. This includes experimentation with document retrieval, prompt engineering, and triple accuracy. A second area of focus will involve the design and testing of a provenance-aware integration module that links each new triple to its originating document and textual context, enhancing trust and transparency. The system's enrichment impact will be evaluated by measuring the factual correctness of generated triples (via expert annotation), their novelty relative to existing graph content, and their usability in exploratory search tasks.
This project will demonstrate how generative AI, when guided by retrieval and grounded in domain-specific resources, can safely extend knowledge graphs in pharmacology.

Recommended Resources (including readings): 1. Rezaei, M.R., Fard, R.S., Parker, J., Krishnan, R.G. and Lankarany, M., 2025. Adaptive Knowledge Graphs Enhance Medical Question Answering: Bridging the Gap Between LLMs and Evolving Medical Knowledge. arXiv preprint arXiv:2502.13010.
2. Matsumoto, N., Moran, J., Choi, H., Hernandez, M.E., Venkatesan, M., Wang, P. and Moore, J.H., 2024. KRAGEN: a knowledge graph-enhanced RAG framework for biomedical problem solving using large language models. Bioinformatics, 40(6), p.btae353.
3. Farrugia, L., Azzopardi, L.M., Debattista, J. and Abela, C., 2024. medicX-KG: A Knowledge Graph for Pharmacists' Drug Information Needs.

Prerequisite Knowledge Required: Machine Learning, Knowledge Graphs, NLP

List any Ethical issues: none

Title: Leveraging LLMs and Knowledge Graphs for Generating Multilingual, Compliant Medicinal Information Leaflets

Supervisor: Dr Abela Charlie

Main Subject Area: Machine Learning, Medical Image Segmentation, VLMs

Short Description of the proposed FYP: Medicinal product leaflets must communicate complex drug information in a clear, accurate, and patient-accessible manner across national languages. In Malta, this requirement is uniquely shaped by dual-language obligations (English and Maltese) and a hybrid regulatory landscape that incorporates both EMA and UK-derived products.
This dissertation investigates how Large Language Models (LLMs), when grounded in structured knowledge such as medicX-KG and regulatory SmPCs, can be used to automatically generate bilingual medicinal leaflets that conform to regulatory standards.
The work will focus on structuring pharmacological facts—such as indications, contraindications, and side effects—into semantically faithful and linguistically accessible content, while incorporating controlled medical terminologies (e.g., ATC, MeSH, and Maltese glossaries) to promote terminological consistency.
The student will experiment with different generation strategies and terminology integration approaches, and evaluate how these affect factual accuracy, linguistic fidelity, and regulatory alignment. Human expert feedback will be used to assess and iteratively refine the quality and compliance of the generated outputs, contributing to the broader aim of safe, inclusive, and explainable AI in pharmacology.

Recommended Resources (including readings): 1. Farrugia, L., Azzopardi, L.M., Debattista, J. and Abela, C., 2025. medicX-KG: A Knowledge Graph for Pharmacists' Drug Information Needs.
2. Jeblick, K., Schachtner, B., Dexl, J., Mittermeier, A., Stüber, A.T., Topalis, J., Weber, T., Wesp, P., Sabel, B.O., Ricke, J. and Ingrisch, M., 2024. ChatGPT makes medicine easy to swallow: an exploratory case study on simplified radiology reports. European radiology, 34(5), pp.2817-2825.
3. Moslem, Y., Romani, G., Molaei, M., Haque, R., Kelleher, J. D., & Way, A. (2023). Domain Terminology Integration into Machine Translation: Leveraging Large Language Models. Proceedings of the Eighth Conference on Machine Translation (WMT), pages 902–911.

Prerequisite Knowledge Required: Machine Learning, NLP

List any Ethical issues: none

Title: Designing a Continuous Integration Pipeline for Reliable and Scalable Maintenance of the medicX Knowledge Graph

Supervisor: Dr Abela Charlie

Main Subject Area: Data Integration; Knowledge Graph; Continuous Integration

Short Description of the proposed FYP: Biomedical knowledge graphs such as medicX-KG play a crucial role in supporting pharmacists with structured, semantically rich information from trusted sources like the Malta Medicines Authority (MMA), DrugBank, and the BNF. However, pharmacological knowledge evolves rapidly—new drugs are approved, indications shift, and adverse effect data is continuously updated—making static graphs prone to semantic drift.
This dissertation investigates the design and development of a Continuous Integration (CI) pipeline that automates the ingestion and validation of new data into medicX-KG.
The student will need to explore methods for retrieving and transforming updates from pharmacological sources into structured RDF triples aligned with the medicX-KG ontology. Further investigation involves developing validation and reconciliation mechanisms to detect and resolve conflicts, duplicates, or schema inconsistencies prior to integration. Finally, evaluate the enrichment pipeline by measuring triple correctness, novelty, and usability through snapshot comparisons and expert review.
This project contributes to maintaining a reliable, up-to-date, and explainable biomedical knowledge graph, advancing methods for graph-based AI in pharmacology.

Recommended Resources (including readings): 1. Stear, B.J., Mohseni Ahooyi, T., Simmons, J.A., Kollar, C., Hartman, L., Beigel, K., Lahiri, A., Vasisht, S., Callahan, T.J., Nemarich, C.M. and Silverstein, J.C., 2024. Petagraph: A large-scale unifying knowledge graph framework for integrating biomolecular and biomedical data. Scientific Data, 11(1), p.1338.
2. Ilyas, I.F., Rekatsinas, T., Konda, V., Pound, J., Qi, X. and Soliman, M., 2022, June. Saga: A platform for continuous construction and serving of knowledge at scale. In Proceedings of the 2022 international conference on management of data (pp. 2259-2272).
3. Kumar, P., Metzger, V. T., Purushotham, S. T., Kedia, P., Bologa, C. G., Lambert, C. G., & Yang, J. J. 2025. KG2ML: Integrating Knowledge Graphs and Positive Unlabeled Learning for Identifying Disease-Associated Genes. medRxiv : the preprint server for health sciences, https://doi.org/10.1101/2025.03.17.25323906

Prerequisite Knowledge Required: Knowledge Graphs, Machine Learning

Title: TrafficBot: Simulating Smart Traffic Navigation with Reinforcement Learning

Supervisor: Dr Galea Ingrid

Main Subject Area: Robotics

Short Description of the proposed FYP: This project aims to develop a small-scale mobile robot that simulates smart traffic navigation using Q-learning. Using the Elegoo Smart Robot Car Kit V4.0, the robot will operate within a simplified road layout featuring intersections and lanes. The Q-learning model will be trained in simulation on a PC, and the resulting Q-table will be deployed on the robot for real-world navigation. The robot will learn to choose optimal actions (e.g. turn, stop, go straight) based on its current state (e.g. obstacle ahead, at intersection), simulating intelligent vehicle behaviour in a traffic system.

Recommended Resources (including readings): 1. Sutton, R. S., Barto, A. G. (2018). Reinforcement Learning: An Introduction. The MIT Press
2. Simple Q-Learning tutorials in Python e.g. https://towardsdatascience.com/q-learning-for-beginners-2837b777741/
3. Elegoo Smart Car Kit V4.0 Documentation

Prerequisite Knowledge Required: Python programming
Fundamentals of reinforcement learning (Q-learning algorithm)
Basic robotics concepts
Arduino programming

Recommended: Student has completed ARI2202.

List any Ethical issues: none

Title: CleanSweep: The Learning Cleaner Bot

Supervisor: Dr Galea Ingrid

Main Subject Area: Robotics

Short Description of the proposed FYP: CleanSweep is a proposed smart cleaning robot that will be built using the Elegoo Smart Robot Car Kit V4.0. It is intended to simulate autonomous floor-cleaning behaviour within a grid-based environment. The robot will begin by using a simple random walk combined with obstacle avoidance to explore the area and identify "dirt" tiles, which will be represented by printed dark markers on the grid. When a dirty tile is detected using the IR sensor, the robot will simulate cleaning by flashing an LED and recording the tile as cleaned. A Q-learning model will be trained in simulation to enable the robot to optimise its movement strategy over time, reducing overlap and improving overall coverage. The project will compare the effectiveness of the learned behaviour against the baseline reactive approach.

Recommended Resources (including readings): 1. Sutton, R. S., Barto, A. G. (2018). Reinforcement Learning: An Introduction. The MIT Press
2. Simple Q-Learning tutorials in Python e.g. https://towardsdatascience.com/q-learning-for-beginners-2837b777741/
3. Elegoo Smart Car Kit V4.0 Documentation

Prerequisite Knowledge Required: Python programming
Fundamentals of reinforcement learning (Q-learning algorithm)
Basic robotics concepts
Arduino programming

Recommended: Student has completed ARI2202.

List any Ethical issues: none

Title: Dynamic Bus Routing for On-Demand Public Transportation

Supervisor: Dr Bajada Josef

Main Subject Area: Discrete Optimization, Reinforcement Learning

Short Description of the proposed FYP: On-Demand Public Transportation promises to improve the service levels of traditional public transportation by enabling passengers to pre-book their journey and have the transportation vehicle routes adapt dynamically to passenger demand. This approach could ensure that passenger load and route lengths are distributed across vehicles, and pickup waiting times and journey times are minimized. While this solution has been available for small vehicles such as cabs, scaling it to larger capacity vehicles has the potential to make public transportation more attractive for users, reduce traffic and pollution.

In this project, the student will develop a system which, given a number of pickup and drop-off waypoints from different customers, will assign the right sequence of waypoints for each public transportation vehicle. It will take advantage of passengers making use of the same waypoints to optimize the routes of multiple buses and minimize pickup waiting time and journey time.

A number of approaches will be investigated, including traditional discrete optimization techniques such as Tabu search and Genetic Algorithms, together with more recent experimental approaches, such as using Reinforcement Learning. The results from this thesis will be a comprehensive analysis of the performance of different algorithms as scenarios become more complex with more customer trips and more vehicles to schedule accordingly.

Recommended Resources (including readings): Raza, Syed Mohib, Mohammad Sajid, and Jagendra Singh. "Vehicle Routing Problem using Reinforcement Learning: Recent Advancements." Advanced Machine Intelligence and Signal Processing. Singapore: Springer Nature Singapore, 2022. 269-280.

Prerequisite Knowledge Required: none

List any Ethical issues: none

Title: Reinforcement Learning for Real-time Adversarial Games.

Supervisor: Dr Bajada Josef

Main Subject Area: Reinforcement Learning

Short Description of the proposed FYP: In this project we will explore the use of RL algorithms to create intelligent agents that play adversarial real-time games. In such environments, all the players can take actions simultaneously, rather than waiting for their turn. Such games present various challenges, such as state representation, action encoding, action masking and selecting the right reward signals to guide algorithms to learn an effective policy.

For this project, one game will be selected, and the student will be expected to use or build a simulator for the game and build a framework through which a policy to play the game well will be learnt. The student will be expected to review existent literature on similar games, and take inspiration from their ideas how to model the right state information, encode the possible actions to take in each state, craft the rewards for each action, and integrate the environment with reinforcement learning algorithms. The performance of these algorithms will be analysed and compared, as part of the evaluation of the project.

Recommended Resources (including readings): Sutton, R. S. and Barto, A. G., "Reinforcement Learning: An Introduction", 2nd Edition, The MIT Press, 2018.

Prerequisite Knowledge Required: ARI2204 - Reinforcement Learning
ARI3212 - Advanced Reinforcement Learning (recommended 3rd year study unit)

List any Ethical issues: none

Title: Pairwise Transformation Methods for Preference-Based Models of Affect

Supervisor: Dr Makantasis Konstantinos

Main Subject Area: machine learning, affect modelling

Short Description of the proposed FYP: Recognising human emotional states is a fundamental yet challenging task in affective computing. The primary computational difficulty stems from the inherently subjective nature of emotions, which introduces subjectivity bias and adversely affects the performance of emotion recognition models. A prominent strategy for mitigating this bias involves treating emotion labels as relative, thereby reformulating emotion recognition as a preference learning task. This approach necessitates pairwise data transformation, which substantially increases the dataset size. This project aims to explore and propose various methods for performing pairwise transformations to develop models that achieve an optimal balance between computational efficiency and emotion recognition accuracy. In particular, the research will focus on methodologies based on probabilistic sampling and preference orderings for data transformation. The datasets generated through these different pairwise transformation techniques will be used to train affect models, whose performance will be evaluated in terms of both computational efficiency and recognition accuracy. Publicly available datasets, such as AGAIN and RECOLA, will be utilised for this evaluation. The project will deliver two key outputs: 1) A comprehensive evaluation of different pairwise transformation techniques and 2) An open-access code repository containing all implemented algorithms.

Recommended Resources (including readings): [1] Pinitas, Kosmas, et al. "RankNEAT: outperforming stochastic gradient search in preference learning tasks." Proceedings of the Genetic and Evolutionary Computation Conference. 2022.
[2] Makantasis, Konstantinos, Antonios Liapis, and Georgios N. Yannakakis. "The pixels and sounds of emotion: General-purpose representations of arousal in games." IEEE Transactions on Affective Computing 14.1 (2021): 680-693

Prerequisite Knowledge Required: Python

List any Ethical issues: none

Title: From Manifestations to Representations: Contrastive Learning Approaches for Affect Modelling

Supervisor: Dr Makantasis Konstantinos

Main Subject Area: machine learning, affect modelling

Short Description of the proposed FYP: Affect modeling has traditionally been understood as the process of mapping measurable affective manifestations from multiple modalities of user input to corresponding affect labels. This mapping is typically achieved through end-to-end (manifestation-to-affect) machine learning approaches. However, an alternative perspective is to first train general-purpose data representations and then utilise these representations for affect modeling. This project seeks to investigate this approach. Specifically, the project will explore the application of various contrastive learning techniques, both self-supervised and supervised, for learning data representations relevant to affect modeling tasks. The research will implement different contrastive learning methods to derive informative representations both within individual users and across different users. These learned representations will subsequently be used to train affective models, whose performance will be assessed in terms of affect modelling accuracy. Publicly available datasets, such as AGAIN and RECOLA, will be employed for this evaluation. The project will deliver two key outputs: 1) A comprehensive evaluation of different contrastive learning techniques within the problem of affect modelling and 2) An open-access code repository containing all implemented algorithms.

Recommended Resources (including readings): [1] Pinitas, Kosmas, et al. "Supervised contrastive learning for affect modelling." Proceedings of the 2022 International Conference on Multimodal Interaction. 2022.
[2] Khosla, Prannay, et al. "Supervised contrastive learning." Advances in neural information processing systems 33 (2020): 18661-18673.
[3] Jaiswal, Ashish, et al. "A survey on contrastive self-supervised learning." Technologies 9.1 (2020): 2.

Prerequisite Knowledge Required: Python

List any Ethical issues: none

Title: Towards Automatic Ontology Learning from LLM-Derived Conceptual Structures in Historical Legal Texts

Supervisor: Dr Abela Charlie

Main Subject Area: Knowledge Graphs, LLMs

Short Description of the proposed FYP: The manual construction of ontologies for historical legal archives is resource-intensive, requiring both deep domain expertise and technical proficiency in semantic modelling.
This dissertation explores how Large Language Models (LLMs) can assist in the semi-automatic generation of ontology scaffolds from historical legal texts, particularly those associated with Malta's notarial manuscripts.
The aim is not to replicate the full ontology engineering workflow, but to investigate whether LLMs can extract, group, and structure domain-relevant concepts into class hierarchies and property assertions suitable for expert refinement.
The project will begin by prompting LLMs with excerpts from transcribed notarial deeds and evaluating their ability to return structured triples or conceptual clusters. These will then be processed and visualised using ontology-compatible tools and visualisations. The student will need to evaluate these outputs against expert-created or CIDOC-CRM-aligned ontologies using ontology quality metrics including class coverage, depth, and precision. The broader objective is to assess the viability of LLMs as early-stage collaborators in ontology bootstrapping for digital humanities projects.
This research complements the goals of the Notarypedia Knowledge Management Platform, project.

Recommended Resources (including readings): 1. Ellul, C., Abela, C., & Azzopardi, J. (2019). NotaryPedia: A Knowledge Graph of Historical Notarial Manuscripts. In OTM 2019 Conferences, Springer.
2. Xu, H., Chen, Y., Liu, Z., Wen, Y. and Yuan, X., 2022. TaxoPrompt: A Prompt-based Generation Method with Taxonomic Context for Self-Supervised Taxonomy Expansion. In IJCAI (Vol. 22, pp. 4432-4438).
3. Babaei Giglou, H., D'Souza, J. and Auer, S., 2023, October. LLMs4OL: Large language models for ontology learning. In International Semantic Web Conference (pp. 408-427). Cham: Springer Nature Switzerland.
4. Chen, Boqi, Fandi Yi, and Dániel Varró. "Prompting or fine-tuning? a comparative study of large language models for taxonomy construction." In 2023 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C), pp. 588-596. IEEE, 2023.

Prerequisite Knowledge Required: Knowledge Graphs, Machine Learning