L-Università ta' Malta
Faculty of Information &
Communication Technology | Department of Artificial Intelligence

ARI3333: Practical Project

"Developing a Content Generation System for Creative Storytelling"

Project Overview:

You are asked to develop a complete AI-powered content generation system capable of creating coherent, multi-modal creative stories by combining text, image, and audio generation models. The project involves building a system that can generate a fictional story, complete with narrative elements (text), illustrations (images), and background music or sound effects (audio).

This project will challenge you to integrate various generative AI models (e.g., GPT for text, GANs or Diffusion Models for images, and WaveNet or Jukebox for sound). You must also think critically about the interaction between modalities to ensure a cohesive and meaningful output.

You have the possibility to choose to work on the project as a single individual, or in a group of up to 3 members. **This Google form** will allow you to choose your preferred modality.

Key Objectives:

o   Multi-Modal Generative Models Integration:
  *   Implement and integrate generative models for text, image, and audio generation.
  *   Ensure synchronization and consistency between different content types.

o   Creative and Critical Problem-Solving:
  *   Develop a cohesive narrative using an AI-based text generation model.
  *   Generate images (e.g., DALL-E, StyleGAN) and audio (e.g., WaveNet, Jukebox) that align with the generated story.
  *   Critically evaluate and refine the generated content for creative coherence.

o   User Interaction and Feedback Loop:
  *   Build an interface that allows users to interact with the system and provide input (e.g., specifying a story theme or character).
  *   Implement a feedback loop that allows users to refine or edit generated content, encouraging iterative improvement.

o   Ethical Considerations:
  *   Consider ethical issues related to the generation of content, such as biases in storytelling or content appropriation.
  *   Include features to filter inappropriate content (e.g., violent or offensive themes) generated by the model.

Project Components:

o   Text Generation Module – ***for those you opt on individual submission OR 1ˢᵗ member in a group***
- Use a language model like GPT-3 or a fine-tuned GPT-2 to generate a fictional story based on user input (e.g., theme, character, or setting).
- Ensure the text is coherent, with a structured narrative (beginning, middle, end).
- Add flexibility for users to provide parameters such as genre (sci-fi, fantasy, mystery), tone (serious, humorous), or specific plot points.

o   Image Generation Module – ***for individual OR 2ⁿᵈ member in a group***
- Use an image generation model like DALL-E, StyleGAN, or a diffusion model to generate illustrations based on the story's content.
- Generate character designs, scenes, or cover art in alignment with the narrative.
- Employ fine-tuning or pre-trained models for specific styles (e.g., realistic, cartoonish).

o   Audio Generation Module – ***for individual OR 3ʳᵈ member in a group***
- Integrate an audio generation model like OpenAI's Jukebox or WaveNet to create soundtracks or ambient sound effects that complement the story.
- Design different audio tracks based on story elements such as mood, setting (e.g., background music for an eerie or exciting scene), or character interactions (e.g., sound effects for specific events).

o   User Interface (UI) – ***in all cases***
- Create an easy-to-use UI that allows users to:
    i.   Provide story inputs (themes, keywords, characters).
    ii. Review the AI-generated content (text, images, and audio).
    iii. Offer feedback and refine the content based on user preferences.
    iv. Ensure a smooth, interactive experience by allowing users to regenerate parts of the story, illustrations, or audio.

o   Evaluation and Improvement Module – ***opportunity for anyone to gain extra credit***
- Incorporate a feedback system that allows users to rate and refine the generated content.
- Implement fine-tuning mechanisms that improve subsequent generations based on user interactions.
- Integrate video generation

Additional Information:

- o System Design and Integration:
  - You will need to design a workflow where different generative models (text, image, and audio) interact harmoniously.
  - Describe your thought process to address the challenges of integrating various models, ensuring consistency between the different modalities.

- o Data and Model Evaluation:
  - You need to evaluate the outputs of each model (e.g., whether the generated image matches the narrative description or whether the soundtrack enhances the mood).
  - You will also need to critically assess model limitations, biases in generation, and strategies to improve the model outputs through iteration.

- o Ethical and Social Considerations:
  - Reflect on the ethical implications of generated content. For example, ensure the avoidance of biased or harmful narratives and imagery.
  - Show how you developed a filtering mechanism to handle inappropriate or offensive content generated by the models, promoting responsible AI use.

- o Creativity in Storytelling:
  - Creatively think about the artistic and narrative cohesion of the generated content.
  - Employ unique narratives, experimenting with different genres, and utilizing AI as a creative tool rather than a mere utility.

Possible project choices:

	Text Gen		**Image Gen**		**Audio Gen**
Individual	Student 1	OR	Student 1	OR	Student 1
Team of 2	Student 1	AND	Student 2		
Team of 3	Student 1	AND	Student 2	AND	Student 3

Deliverables by each individual student (even those in groups):

    i.   Final System Prototype:
- a. A working content generation system that users can interact with
- b. The system should demonstrate the ability to integrate a modality (text, image, audio, video seamlessly)

    ii.   Project Report:
- a. A detailed technical report documenting:
    - Design choices for each generative model.
    - The overall system architecture and how different models were integrated.
    - Challenges encountered (e.g., coherence, bias) and how they were addressed.
    - Ethical considerations and mitigation strategies.

    iii. Presentation & Demo:
- a. A live presentation or recorded demo showcasing the system's capabilities.
- b. Personal evaluative comments on the creative and technical decisions, including lessons learned and potential improvements.

Assessment Criteria:

- Creativity and Cohesion (30%) – How well the final prototype works well to create a cohesive and compelling story.

- Technical Implementation (40%) – The complexity, efficiency, and functionality of the system, including model selection, training, and optimization.

- Ethical Consideration (10%) – Evaluation of how well the project addresses the ethical implications of generative AI.

- User Interaction and Interface (10%) – Quality and usability of the user interface and user feedback mechanisms if implemented.

- Project Report & Presentation (10%) – Clarity, depth, and insight demonstrated in the report and presentation.