# **Text-to-Tile: Generating 2D RPG Sprites from Text Descriptions**

Nejneriu Mihai + Adrian Nazîru

## **Introduction**

Manually creating tileable game environment sprites (e.g. terrain tiles) is time-consuming and scales poorly​. We propose a text-to-tile generation system that converts natural language descriptions of environmental features (e.g. “mountain”, “decayed forest”, “polluted water”) into 2D RPG-style sprite tiles. Unlike full map generators, our focus is on producing individual tile graphics that seamlessly repeat. This approach leverages advances in natural language processing (NLP) and generative modeling to automate asset creation, eliminating manual tiling effort and enhancing creative flexibility. The outcome will enable game developers and content creators to rapidly obtain custom graphics by simply describing them, greatly accelerating level design and prototyping.

## **Objectives**

* **Natural Language Understanding:** Develop an NLP pipeline to interpret user-provided text descriptions of environments. This will involve parsing the text or extracting semantic features (e.g. identifying terrain type and attributes) to form a structured representation of the desired sprite.
* **Generative Sprite Synthesis:** Fine-tune state-of-the-art generative models (such as conditional GANs or transformer-based diffusion models) to produce small, tileable sprite images matching the input description. Transfer learning will be applied on pre-trained text-to-image models, adapting them to the pixel-art style and constraints of RPG tiles.
* **Tileability and Coherence:** Ensure the generated sprites are seamlessly tileable. This will be achieved by applying tiling constraints during generation (for example, using circular padding in convolutional layers to wrap edges) or by training on tile examples. The resulting image can repeat without visible seams, allowing it to fill maps continuously.
* **Integrated System & Flexibility:** Combine the NLP and image generation components into an end-to-end system that can take any new environment description as input and output a corresponding sprite tile. This flexibility in asset generation will support game development and interactive tools by providing on-demand graphics content.