TRAIL Summer Workshop' 25
Project Proposal

Full Name of Team Leader	Kevin El Haddad (Wal4XR Senior Researcher) Daniele Bonato (Wal4XR Senior Researcher) Maxime Cauze (Wal4XR Senior Researcher) François Rosoux (Wal4XR Senior Researcher)
Project Title	Building an Efficient Immersive Custom Environment using AI
Profile of the Team Leader(s)	The leadership team has experience in automated 3D modeling, Multiview generation, physical-digital work interaction, applied machine learning for industrial use-case, human-agent interaction technologies which cover all the requirements for this project
Abstract	This project aims at developing and evaluating an AI-based toolkit that facilitates the creation of a custom 3D environment. To have value, these environments require 3D assets adapted to the desired outcome as well as realistic interaction between users and objects. The latter relies on a lot of different parameters like the sense of embodiment, an accurate localization/mapping of objects and users, an adapted representation of objects and scenes, etc. For this work, we chose to focalize on the localization and mapping of users and objects due to its crucial functionality across applications and because it would allow us to explore several different features (object detection, depth estimation, localization and mapping algorithms, etc.) and aspects like the real-time and Head Mounted Device (HMD) embedded needs. So, the development of the toolkit will focus on two main features. The first one is a 3D asset generation set of models that allow for the creation of 3D assets that will populate the environment. The main approaches considered here will be the generation of 3D assets from prompts and from images of real-life assets. The assets here could be either object, faces, bodies or scenes. Two main technologies will be considered here: Gaussian splatting approach and the Neural Radiance Field (NeRF) approaches. One of the main limitations of such approaches is that they are unable to segment semantically the scene/object that has been modelized. Our current work show that Gaussian Splatting is the most promising one [1-3]. The second one is the localization and mapping of the users and objects in the scene which will leverage on several state-of-the-art systems for Simultaneous Localization and Mapping (SLAM) algorithms, 3D and 6D object localization and depth estimation to build and compare several systems to localize and map a user in a scene. While this is not the case for the first features, the second one needs to be able to run ideally in real-time and on the HMD. There is a need for benchmarks for localization algorithms specifically running on HMD in the literature. So the comparison of the different approaches will make what is to the best of our knowledge the first benchmark on which future work can build.
Project Objectives	• Develop a 3D asset generation toolkit:

	○ Compare prompt-based and image-based 3D asset generation models for objects and scenes on several criteria, namely: the fidelity of the end result to the original asset, the resources required by the model to run and the ability to semantically segment the assets. ○ Develop a semantical segmentation method • Develop and compare several user localization and mapping approaches ○ Develop at least two approaches (the amount will depend on the number of participants): based on Visual SLAM, SLAM, object detection and depth estimation ○ Benchmark all the developed approaches by comparing them according to their capability on running on HMD (several will be considered) and their accuracy Several models of the 3D generation features will already have been implemented before the workshop as well as some work on the semantic segmentation. The team work during the workshop will focus on normalizing the output format of the models and continue the work on the segmentation. Concerning the localization and mapping, the methods to be implemented will be defined and pre-tested. The preparation of the benchmark datasets will start several weeks before the beginning of the workshop depending on the selection of the approaches (which itself will depend on the number of Wal4XR participants funded for the TSW). The work during the workshop will focus on adaptation of the solutions in the HMD setup and their benchmarking. The end result will be a documented open-source toolkit facilitating the creation of immersive custom environments.
Project Dataset	For the object semantic segmentation: We will use the ULB Toystable dataset (https://zenodo.org/records/5055543) which features Lambertian scenes with good visibility conditions and challenging complex geometries including small amounts of fur. For preliminary experiments on non-Lambertian scenarios, we will use the dataset featuring a transparent torus (https://zenodo.org/records/5048275), which presents significant challenges for segmentation algorithms due to transparency effects. For the localization and mapping on HDM feature: Several datasets and benchmark for SLAM algorithms exist like SLAMBench2 [4]which is a benchmarking framework designed to evaluate SLAM algorithms and integrate state-of-the-art systems or ARKitScenes [5] which is a dataset of indoor scenes and objects. These will be adapted for testing the systems considering the HMD hardware limitations.
Background Information	One of the objectives of the Wal4XR project is to ease and accelerate the adoption of eXtended Reality technologies in Wallonia. For this the boundaries of the state-of-the-art in different domains. An open-source toolkit facilitating the creation of virtual environment is a high value tool for several industrial societal use-cases (technician or expert training, prototyping, digital twins, remote maintenance, co-presence experiences, etc.). AI-based technology exists to build that can be leveraged for this but are not quite adapted for XR applications (considering the computation and storage

	requirements, considering the specific use-cases, the behavior of users in such environments, etc.). Our project aims at contributing to tackling this issue by attempting to improve the existing tools or rather adapting them for XR use-cases.
Bibliographic References	[1] Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Transactions on Graphics, 42(4). (https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/) [2] Shen, Q., Yang, X., & Wang, X. (2025). FlashSplat: 2D to 3D Gaussian Splatting Segmentation Solved Optimally. In A. Leonardis, E. Ricci, S. Roth, O. Russakovsky, T. Sattler, & G. Varol (Eds.), Computer Vision – ECCV 2024 (Lecture Notes in Computer Science, Vol. 15080). Springer, Cham. (https://doi.org/10.1007/978-3-031-72670-5_26) [3] Ye, M., Danelljan, M., Yu, F., & Ke, L. (2024). Gaussian Grouping: Segment and Edit Anything in 3D Scenes. In European Conference on Computer Vision (ECCV). [4] Bodin, B., Wagstaff, H., Saecdi, S., Nardi, L., Vespa, E., Mawer, J., ... & O'Boyle, M. F. (2018, May). Slambench2: Multi-objective head-to-head benchmarking for visual slam. In 2018 IEEE International Conference on Robotics and Automation (ICRA) (pp. 3637-3644). IEEE. [5] Baruch, G., Chen, Z., Dehghan, A., Dimry, T., Feigin, Y., Fu, P., ... & Shulman, E. (2021). Arkitscenes: A diverse real-world dataset for 3d indoor scene understanding using mobile rgb-d data. arXiv preprint arXiv:2111.08897.
Detailed Work Plan	Although team leaders are mentioned, the work here will be implemented by several researchers from the Wal4XR project depending on the funding obtained. The planning that follows considers 5 full-time researchers on the project: Week 1: 1- Implementation of pre-selected image to 3D asset generation models 2- Development of semantic segmentation approach (most probably only on the Guassian splatter approaches) 3- Implementation of the pre-defined localization and mapping systems 4- Adaptation of the systems in point three for HMD environment Week 2: 5- Evaluation and comparison of the 3D asset generation as well as the localization and mapping systems based on the abovementioned criteria 6- Integration of all 3D asset generation in a single GUI (if there is enough time left) 7- Development of a benchmark for future evaluation and comparison of newly developed localization and mapping systems for HMD 8- Code cleaning and documentation of the results The team will operate in parallel tracks with daily sync meetings. Specialists will onboard according to project phase and need, maximizing productivity within the short timeframe.
Other Remarks	

Optional: Add any relevant figure for the project