IMET - International Conference on Interactive Media, Smart Systems and Emerging Technologies (2023)
N. Pelechano, F. Liarokapis, D. Rohmer, and A. Asadipour (Editors) *Short Paper*

AnywhereXR - Laying the Foundation for Open Source Embodied Digital Twin Applications

A. Klippel[1], B. Knuiman[1], J. Zhao[1], J. O. Wallgrün[2], and S. Garzon[1]

[1] Wageningen University & Research, The Netherlands
[2] Independent Researcher, Germany

Abstract

Despite recent developments in immersive technologies and data processing through AI, there are still substantial limitations in creating widespread opportunities for embodied digital twins for communication and visualization of physical environments. Open spatial data and particularly 3D spatial data do exist but this does not mean that we can now automatically or semi-automatically create immersive experiences from the existing data, at least not with high fidelity. The goal should be that anyone, not just high-end tech companies, is in the position to create experiential 3D environments for any place on earth. To move a small step closer to this goal, we are discussing the development of a framework that will allow for the creation of embodied digital twins for any place in the Netherlands by only requiring the location's GPS coordinates as input. While the results are not perfect yet, we consider these efforts essential in advancing the democratic use of embodied digital twins and aim to create a widely accessible infrastructure.

CCS Concepts

• ***Human-centered computing*** → *Geographic visualization;* • ***Software and its engineering*** → *Software design engineering;*

1. Introduction

We are living in a data-rich world [KM*16]. The abundance of data we are collecting about our spatial environments is unimaginable. What we are lacking, though, are efficient and human-centric approaches that would allow for creating more intuitive environments for understanding data, making smarter decisions, or communicating environmental or societal challenges efficiently to scientists and non-scientists alike [SPO*19]. While opportunities and knowledge for solutions are available, we lack the means to design engaging forms of information gathering, sharing, and decision-making such as collaborative digital environments and digital games. There are a number of developments that are occurring in parallel that call for the development of a more centralized and open infrastructure for creating new forms of data visualization and better data experiences, sometimes also referred to as data visceralization [LBL*21], that can address and potentially solve these challenges:

1. Immersive technologies (eXtended Reality, XR) have matured. We are witnessing a steadily increasing adoption of immersive technologies by scientists and the general public. In the scientific realm, we see growing numbers of XR-focused or related articles, often outside of technical fields, a testimony to the development and accessibility of the technology and to a more universal embrace of the technology by non-computer scientists.
2. Data, especially spatial data, is available in abundance. From authoritative data sets to global efforts, to private companies and citizen contributions. The collection of data is increasing rapidly, albeit not all data is accessible for proprietary or data structure reasons.
3. Rather than seeing data as something static, recent developments focus on processes. The resulting data structure is referred to as a digital twin [ZWS*22]. Digital twins are combinations of various technologies, including computer-aided design (CAD), Internet of Things (IoT) sensors, data analytics, and artificial intelligence (AI). The purpose of creating a digital twin is to simulate, monitor, and analyze the behavior, performance, and characteristics of the corresponding physical object or system in real-time or over a period of time.

Bringing these three developments together leads to the concept of *embodied digital twins* [KSZ*21] as a new way in which we can communicate both about the world but also with each other. Instead of using symbolic forms of communication, we can now change to also include embodied experiences. Embodied digital twins and their promise to realize human-centric communication and experiences of the overwhelming abundance of data are inspiring companies, educators, and researchers. In the interest of space, we refrain from an extensive review (see, e.g., [FISO19]).

2. Creating the basis of embodied digital twins from open data

To embrace the opportunities that embodied digital twins offer for communication, education, and research, we believe that it

DOI: 10.2312/imet.20231249

is necessary to create open-source, transparent, and reproducible data and software environments (e.g., [Bar22]). This idea is in line with numerous recent efforts aiming at reproducible science (e.g., [Bak16]), FAIR data standards (e.g., [We16]), and empowering people in a data-democratic way. In the following, we will detail our approach, exemplified for the Netherlands. The workflow for creating the virtual 3D environment has been implemented in Unity3D. Figure 1 illustrates the steps involved, showing the four current stages (Land Cover/Use Generation, Incorporating Elevation, Building Generation, and Tree Generation).

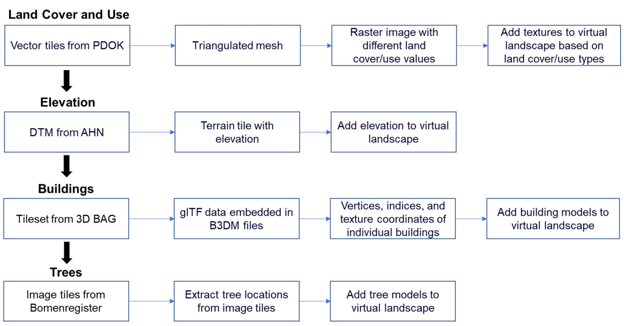

Figure 1: *Our workflow for generating 3D environmental models.*

2.1. Terrain construction from land cover/use data

PDOK (https://app.pdok.nl/vectortile-viewer/) is the official Dutch platform for geo-data sets. In our pilot project, we used vector tiles from PDOK as the main resource for ground layers to realize, for example, different land cover and land use classes. This data consists of concave polygons with holes with their attributes providing information on the natural or man-made features that are present on the ground (e.g., water, grass, and streets; Figure 2, left). As Unity3D cannot handle polygons with holes, our system applies the Earcut algorithm to convert the vector tiles to triangulated polygons, which can be directly used as mesh input in Unity.

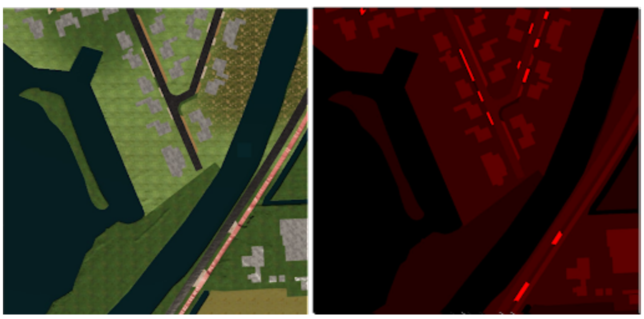

Figure 2: *Left: generation of virtual landscape based on land use and land cover information from PDOK. Right: control image with different red values representing different land cover/use classes.*

Although it is possible to use the triangulated meshes (derived from the vector tiles) directly, this would yield unsatisfying results in areas with high variation of height differences. The vector tiles are originally designed for 2D usage, not taking into account height differences. To overcome this problem, the polygon count can be increased so that enough high resolution is available to integrate the height differences into meshes. However, simply increasing the polygon count would negatively impact the graphical performance and waste computing resources, especially on regions far away from the camera that do not need to be rendered in high detail. So in the current setting, the terrain is rendered as a dynamic mesh that is split up into tiles. Each tile is represented by NxN vertices and can change dynamically to a higher or lower resolution depending on the distance to the camera.

The terrain is textured dynamically via a control image (Figure 2, right) that represents the land use/cover classes for each pixel. In this way, the texturing is independent of the underlying mesh (with a varying number of vertices per tile). The control image is generated by ray tracing a vector tile. For each pixel, a ray is shot and checked against the intersecting triangle, of which the land type is then mapped to a pixel value (e.g., 0 = water, 1 = grass, and 2 = cycling lane). Although control map generation is a fairly time-consuming process (30 seconds per tile in the case of Figure 2, right, which is due to the dynamic sampling in the shader for each tile), the rendering of the terrain mesh is now fast because it can leverage low vertex counts for tiles far away while maintaining high vertex counts near the camera. The final result is a high-quality terrain that performs well (60+ FPS) through the Oculus Quest 2 VR headset connected to a gaming laptop with an NVIDIA 3080 RTX graphics card.

2.2. Elevation

AHN (https://www.ahn.nl/) is a database containing the National Height Model of the Netherlands. It is an openly available data set of the entire Netherlands and acquired through airborne laser scanning (LiDAR). To add elevation to the virtual terrain, the AHN 3 dataset is employed for this project to create Digital Terrain Models (DTMs) of specified locations in the Netherlands. Compared to the AHN 1 and 2 datasets, AHN 3 has a higher sampling point density (or higher raster resolution), which can give the terrain representation a precision of approximately 0.5m^2 at ground level.

2.3. Buildings

3D BAG (https://3dbag.nl/en/viewer) is the 3D extension of the BAG, the register of buildings and addresses in the Netherlands. 3D BAG provides structural information about almost all buildings in the Netherlands as point clouds. The desired area can be defined interactively (depending on the viewpoint) and downloaded from the 3D BAG web viewer as a tileset.json file. The tileset itself is a quadtree holding references to .b3dm files in its leaf nodes. The b3dm stands for Batched 3D models and is an established format for storing and streaming large-scale city models. The .b3dm files embed glTF data that contain batch tables displaying the properties of multiple objects using vertex attributes. From these tables, the vertices along with their indices of each building model can be obtained to create meshes in Unity.

The texture coordinates of individual buildings are not included in the .b3dm files. Textures can be directly mapped on flat roofs without orientation transformation, but for walls and angled roofs that are not parallel to the ground, a texture must be rotated and rescaled to make it aligned with the specified plane (Figure 3, left). In order to determine where to place textures, a series of vector operations are performed to map the positions (horizontal and vertical) and extents (width and height) of textures onto given building facades. (i.e., UV mapping). Furthermore, the orientations of different triangle surfaces of a building representation are compared to judge whether neighboring triangles belong to the same surface. If so, the same texture would be applied. We reduce the precision of normal vectors to one decimal point to leave some tolerance to the subtle orientation differences in neighboring triangles for surface determination. Figure 3 (right) shows an example result.

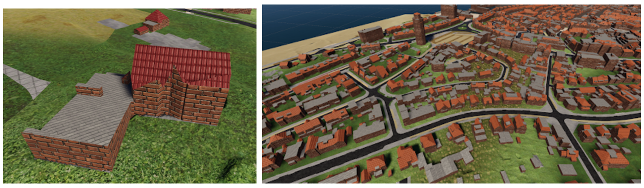

Figure 3: *Left: Texture mapping for building facades. Right: Building model implementation for Zandvoort, The Netherlands.*

2.4. Trees

Trees are crucial for numerous environmental applications and societal challenges. The Bomenregister (https://bomenregister.nl) portal provides information about tree distributions in the Netherlands. The downloaded dataset is a collection of image tiles in which individual trees are present as distinct islands of green colored pixels, and the color intensity indicates the tree size (e.g., the lightest green pixels as shown in the top left of Figure 4 represent bushes). To obtain tree positions from the image retrieved from Bomenregister, the flood-fill algorithm is employed to inspect each pixel's position and its color. Only those neighboring pixels with the same color intensity are identified to belong to a single tree, and their centroid is taken as the tree's location.

As the flood-fill algorithm takes only one image into consideration at a time, it is not effective in handling trees located along the edge of adjacent tiles (Figure 4, bottom left). Therefore, pixel islands that span across two or more image tiles are reconnected in a post-processing step so that the outcome will only be a single tree. Figure 4 (right) shows the implementation results that were obtained for the Grebbeberg, Rhenen, The Netherlands. Figure 5 shows the final outcome from a higher-elevation viewpoint.

3. Evaluation & limitations

We are working on a comprehensive evaluation framework as a combination of objective assessments of (spatial) data quality, such as qualitative errors and the available resolution of spatial information, and subjective experiences. The latter could be characterized as a location-based *experiential Turing test* where participants familiar with a location are exposed to the real and virtual location

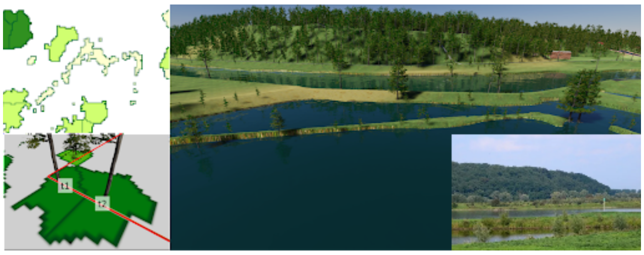

Figure 4: *Top left: Image tile from Bomenregister using pixel islands to indicate tree distributions. Bottom left: The pixel islands, t1 and t2, are split across two image tiles (separated by red lines). In this condition, the naive implementation of the flood-fill algorithm would create a total of four trees, which are supposed to be two trees only (one per island). Right: Landscape representation with trees in Unity (and the bottom right is a photo of the actual landscape).*

Figure 5: *Birds-eye view of the final 3D model.*

consecutively and assess the level of fidelity. According to the original ideas of Turing, the ultimate goal would be an environment that becomes indistinguishable from a real environment. As indicated, this could be done for actual environments but also for potential future environments. While there are options that deliver higher levels of fidelity than our approach, it is important to stress that the main focus of our work is on the ability to create embodied digital twins (semi) automatically for any location.

We started evaluating our work using a stratified random sampling approach in the province of Gelderland. We selected 12 locations in either urban, agricultural, or "natural" (as much as this is possible in the Netherlands) areas. The example of the hill in Figure 4 (right) is one such example. We are in the process of defining a comprehensive matrix that assesses the objective quality of the data as well as the visual and experiential fidelity.

In addition to the stratified sampling approach, we are well aware of some of the shortcomings of our approach, for example:

- **Integration of iconic landmarks such as bridges.** While the visualization approach works for larger and more generic areas, iconic landmarks are important to include in more detail as they are essential in understanding places and for orientation purposes. Possible solutions are the use of sources of publicly available 3D models such as Sketchfab or Turbosquid.

- **Facades of houses.** Information from 3D BAG allows for the generic modeling of facades but not a high-fidelity representation of the visual appearance of buildings. There are numerous opportunities to extract this information from publicly or semi-publicly available sources.

4. Discussion and Outlook

While far from perfect or reaching Turing-test levels of fidelity, our approach to automatically and at large scale create immersive environments is advancing the basis for open-source embodied digital twins. Previous approaches to using open-access data for immersive experiences (e.g., [EHK*18, KESD21]) started to discuss the potential but fell short of reaching higher levels of fidelity and were quite limited in their scope of data used (e.g., only OpenStreetMap data).

It is still a long way to democratize the creation of anywhere XR experiences. Even with the simple examples that we have discussed and in a data-rich environment like the Netherlands, there are still challenges. Some concern the data itself, some are organizational, and some are rooted in technology. While built environments have benefited from a planned infrastructure, inventories of natural environments only slowly emerge. These efforts are spurred by advances in sensing technologies such as satellite remote sensing and the need to catalog changes in the environment at a global scale. However, satellite data allows only for identifying tree locations but not tree species. We are confident that with the combination of advanced sensing, AI, and alternative data sources, such as citizen-science efforts, the collection of data for natural environments will dramatically advance in the future.

The ultimate goal of our approach has to be to extend the framework to more areas. While still challenging, the rapid developments in spatial data infrastructure allow for a very concrete vision for this effort. At the European level, we find, for example, developments such as INSPIRE (https://inspire.ec.europa.eu/), an initiative of the European Commission. While still in development, countries who are participating agreed on a joint standard for geo-data to create a spatial data infrastructure to support environmental policies and assessments. This is important as spatial data are traditionally organized by countries independently. This open data infrastructure is already in place and the data we used for our realization, such as PDOK, is almost completely compatible.

Open source data and code play an important role in creating transparency and security allowing everyone to access, evaluate, and alter code. We consider it crucial to create open source platforms for environmental data as a basis for solving societal problems as it allows for avoiding vendor lock-in, enables the sharing of knowledge, creates grounds for rapid innovation, and facilitates community and collaboration.

Complementary to open source software and data are data standards. We need data standards to enable interoperability, increase data integrity and quality, facilitate data exchange and integration, analyze data, and for creating scalable solutions.

Finally, our approach largely focuses on the use of Unity as a platform for creating embodied digital twins. There are other options that could further enhance the idea of an open platform for embodied digital twins such as Unreal and/or R. There are several R packages available that offer efficient tools for the generation of landscapes such as Rayshader [MW23] or TerrainR [MBA22] that can be imported into game engines. We have started exploring these options, too, as we are interested in creating Free and Open Source Software (FOSS), but at the same time see that there is not one way yet that allows for the most efficient generation of embodied digital twins.

References

[Bak16] BAKER M.: 1,500 scientists lift the lid on reproducibility. *Nature News 533*, 7604 (May 2016), 452. doi:10.1038/533452a. 2

[Bar22] BARBA L. A.: Defining the role of open source software in research reproducibility. *Computer 55*, 8 (aug 2022), 40–48. doi:10.1109/MC.2022.3177133. 2

[EHK*18] EDLER D., HUSAR A., KEIL J., VETTER M., DICKMANN F.: Virtual reality (vr) and open source software: A workflow for constructing an interactive cartographic vr environment to explore urban landscapes. *Kartographische Nachrichten 68* (02 2018). doi:10.1007/BF03545339. 4

[FISO19] FLAVIÁN C., IBÁÑEZ-SÁNCHEZ S., ORÚS C.: The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research 100* (2019), 547–560. doi:10.1016/j.jbusres.2018.10.050. 1

[KESD21] KEIL J., EDLER D., SCHMITT T., DICKMANN F.: Creating immersive virtual environments based on open geospatial data and game engines. *KN - Journal of Cartography and Geographic Information 71*, 1 (2021), 53–65. doi:10.1007/s42489-020-00069-6. 4

[KM*16] KUROSE J., MARZULLO K., ET AL.: The federal big data research and development strategic plan. *USA: Big Data Senior Steering Group (SSG, Executive Office of the President, National Science and Technology Council. pp. 1-46* (2016). 1

[KSZ*21] KLIPPEL A., SAJJADI P., ZHAO J., WALLGRÜN J. O., HUANG J., BAGHER M. M.: Embodied digital twins for environmental applications. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences V-4-2021* (2021), 193–200. doi:10.5194/isprs-annals-V-4-2021-193-2021. 1

[LBL*21] LEE B., BROWN D., LEE B., HURTER C., DRUCKER S., DWYER T.: Data visceralization: enabling deeper understanding of data using virtual reality. *IEEE Transactions on Visualization and Computer Graphics 27*, 2 (Feb. 2021), 1095–1105. doi:10.1109/TVCG.2020.3030435. 1

[MBA22] MAHONEY M. J., BEIER C. M., ACKERMAN A. C.: terrainr: An r package for creating immersive virtual environments. *Journal of Open Source Software 7*, 69 (2022), 4060. doi:10.21105/joss.04060. 4

[MW23] MORGAN-WALL T.: *rayshader: Create Maps and Visualize Data in 2D and 3D*, 2023. https://www.rayshader.com, https://github.com/tylermorganwall/rayshader. 4

[SPO*19] SKARBEZ R., POLYS N. F., OGLE J. T., NORTH C., BOWMAN D. A.: Immersive analytics: Theory and research agenda. *Frontiers in Robotics and AI 6* (2019). doi:10.3389/frobt.2019.00082.

[We16] WILKINSON M., ET AL.: The fair guiding principles for scientific data management and stewardship: Comment. *Scientific Data 3* (2016). doi:10.1038/sdata.2016.18. 2

[ZWS*22] ZHANG Z., WEN F., SUN Z., GUO X., HE T., LEE C.: Artificial intelligence–enabled sensing technologies in the 5g/internet of things era: From virtual reality/augmented reality to the digital twin. *Advanced Intelligent Systems 4*, 7 (2022). doi:10.1002/aisy.202100228. 1