The Changing Landscape of Open Geospatial Data.

“We can harness the power
of existing open data”
says Data Champion, Balázs Dukai.

Software engineer and Data Champion from the Faculty of Architecture and the Built Environment, Balázs Dukai, specialises in the construction of geospatial datasets to generate 3D city models. A ‘3D city model’ is a three-dimensional (3D) digital representation of an urban area that depicts its landscape and architectural infrastructure. The geometry and structure of the terrain, roads, rivers and buildings are described by 3D geospatial data (i.e. coordinates, projections and transformations) for their contextual visualisation using computer modelling software.

Data in all shapes and sizes

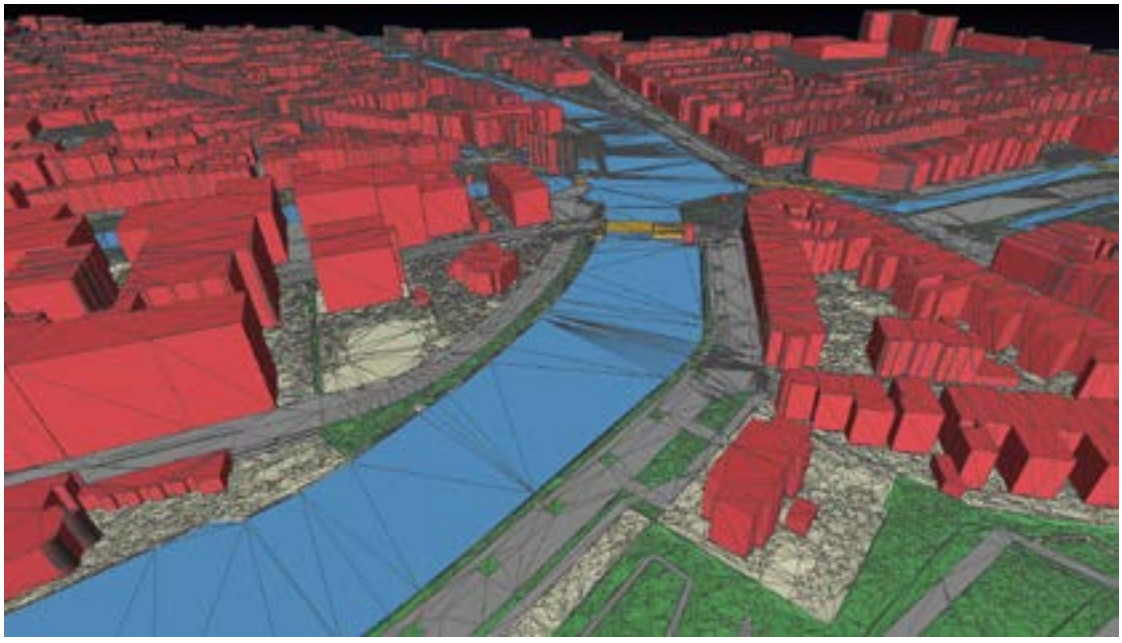
The ability to translate our physical world into a virtual reality has become a valuable asset in the planning, visualisation and management of a wide range of urban applications. By way of example, 3D city models have been used to estimate population size, energy usage, pollutant tracking, solar potential, shadow analysis, wind turbulence and noise impact. With so much to offer, it’s no surprise that there are now more than one thousand virtual city models worldwide.

Rapid population growth and urbanisation is increasing the demand for more complex 3D models that can improve urban development. Dukai discusses some of the current issues surrounding the generation of complex geospatial data, “There are already numerous open geospatial datasets available for use but due to the heterogeneity and complexity in data formats, the exchange between them is difficult at best, but often not possible.” He explains that this problem is confounded by the lack of software support for the utilisation of 3D geospatial data, “3D modelling software is limited to expensive proprietary platforms that benefit many commercial enterprises but few individual researchers.” Dukai expresses his belief that “there is enormous value in existing datasets that are stored away in repositories collecting dust because they are not interoperable – it’s such a waste.”

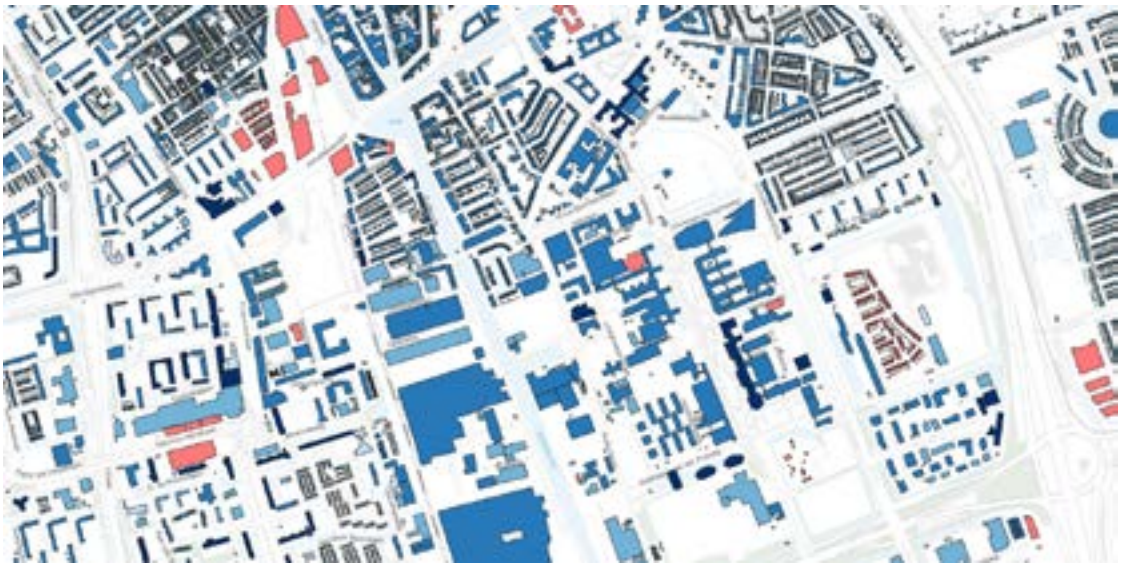
Dusting off data

Harnessing the power of existing geospatial data to produce 3D city models has been at the forefront of Dukai's research agenda. At the 3D Geoinformation research group at TU Delft, he is working on 3dfier, an open software that transforms 2D topographical datasets to 3D by lifting every polygon using height information from point clouds. The software combines a digital map of the Netherlands (The Basic Registration of Large-Scale Topography, BGT) with the digital elevation model for the Netherlands (The Actueel Hoogtebestand Nederland, AHN) to create a volumetric 3D representation of the urban environment. So far, 3dfier is used by the Dutch national mapping agency (Kadaster) to reconstruct the 3D model of the Netherlands — Check out the animation to discover how 3dfier was used to create Delft's 'digital twin'!

3dfier has also been used to create the 3D Registration of Buildings and Addresses (3D BAG) using the original Dutch BAG register and the digital elevation model (AHN). This map provides details of a building's address, it's construction and registration date, and it's intended use. By clicking on a building footprint, you can learn more about a building's attributes, such as ground surface or roof height. The 3D BAG dataset is automatically updated each month to provide users with an up-to-date record of the latest building stock and elevation information as buildings are built, registered and demolished. 3D BAG data is free and available to download and can be accessed via our 4TU.Centre for Research Data repository.



Caption: An example of a 3D city model created by TU Delft software



Caption: 3D BAG is the most detailed, openly available dataset on buildings and addresses in the Netherlands.

In collaboration with Rijkswaterstraat, RIVM and Kadaster, Dukai's research group are reusing a collection of datasets (i.e. BGT, AHN and 3D BAG) to investigate how 3D geospatial data on noise sources can be automatically generated for the whole of the Netherlands. This project is a great example of a relevant urban application that leverages existing geospatial data. "The development aims to improve efficiency by reusing datasets thereby eliminating the need to acquire new data," says Dukai. "Of equal importance is the opportunity to standardise input data for noise studies. The collection of 3D geospatial data by different research groups using various methods has led to undesirable differences in results which cannot be compared unambiguously."

Data harmonisation on the horizon

Dukai is an advocate for standardising open geospatial data and supporting software, and is determined to improve data quality by ensuring that it is findable, accessible, interoperable and reusable (FAIR). To this end, he is involved in the ongoing development of 'CityJSON', a JSON-based exchange format for the CityGML data model. CityGML, or Geography Markup Language, is an international standard produced by the Open Geospatial Consortium (OGC) that is used to store and exchange 3D city models. However, it's XML-based exchange format has drawbacks; it's complex and lacks software support. Dukai makes the case for switching to a simpler alternative, "CityJSON has been designed with programmers in mind. It's compact and easy-to-use whether you're reading or creating datasets. With one click you can convert CityGML to JSON files which are supported by a growing number of software platforms."

CityJSON's compatible software component, cijo (CityJSON input/output) is a Python command-line interface (CLI) programme designed to process, manipulate and validate CityJSON files. Since

both CityJSON and cijo are open to the vibrant research community via GitHub, iterative workflows facilitate their continual improvement through the collaborative efforts of those who wish to contribute to the common goal of harmonising geospatial data. (To learn more about #CityJSON, (see Dukai's GeoPython2019 conference slides).

Becoming a Geospatial Data Champion

Our conversation ended with Dukai's final thoughts on the changing landscape of open geospatial data and a positive reflection on his impact as a Data Champion. "As a software engineer, I understand the pain of having to work with proprietary data. Whilst making data 'open' doesn't necessarily guarantee quality, it does guarantee transparency and allows people to learn from one another and work together to accelerate research." He adds, "As an individual researcher, becoming a Data Champion has allowed me to acquire valuable skills and knowledge that I wouldn't have otherwise encountered. As part of a wider research community, becoming a Data Champion has allowed me to connect my group with a diverse network with opportunity for collaboration."

Mapping the future landscape…

Dukai's research group have recently been collaborating with TU Delft Associate Professor and Research Director, Frank van der Hoeven; Research coordinator, Susan Ng-A-Tham, and Data Stewards; Yan Wang (Faculty of Architecture and the Build Environment), Kees den Heijer (Faculty of Civil Engineering and Geosciences) and Jeff Love (Faculty of Industrial Design and Engineering), to develop a Geoportal that will facilitate the visualisation of geospatial data via maps…Indeed, this could include geospatial data archived in our 4TU. Data repository!

Citation

This showcase is a selection from the following publication:

Clare, Connie. (2019). The Real World of Research Data [Book]. Zenodo. http://doi.org/10.5281/zenodo.3584373.