International Cartographic Association
Association Cartographique Internationale

Exploring the potential of NLP technology to generate textual descriptions of space for indoor navigation applications

Krzysztof Lipka [a,*], Dariusz Gotlib [a], Kamil Choromański [a]

[a] *Warsaw University of Technology, Faculty of Geodesy and Cartography, Department of Cartography*

* *author for correspondence, krzysztof.lipka@pw.edu.pl*

Keywords: cartography of building interiors, artificial intelligence, space descriptions, NLP, navigation

Abstract:

For many years, the number of applications and research related to assisted navigation in buildings has steadily increased. Some of them use textual descriptions of space, called geo-descriptions in short, in navigation. Geo-descriptions are specially constructed multi-level textual descriptions associated with a space (a point, a line, an area), which can be displayed on the screen or read, for example, by a speech synthesiser at the appropriate moment in the localisation or navigation process, as well as during spatial analysis of an object. They complement the traditional cartographic transmission of information in graphic form. They can be beneficial in tourism applications, in assisting special services and in assisting people with disabilities (including blind people).

An example of a geo-description is: *"You are in the Main Building of the University Campus on the second floor in the northern part of the building. You are in the corridor leading to the Main Library. In the opposite direction you can walk to the Main Auditorium and to the elevator adapted for people with disabilities. This part of the building also houses the teaching laboratories of the Department of Geodesy and Cartography and the Center for International Cooperation. The main exit is about 80 meters from where you are."*

Currently, geo-descriptions are created manually. However, this is a time-consuming and complex process. The research conducted has shown that there are not yet solutions that lead to the results assumed by the authors of this article. It is not enough to use existing Natural Language Processing (NLP) techniques by issuing a command like "Generate geo-description." The paper shows that using NLP techniques, the process of generating geo-descriptions can be partially automated. Therefore, research has been undertaken to automate this process to the extent possible. Automation can involve at least three stages:

1. Selection of places that should be described.
2. Extraction of features of the area.
3. Creation of a linguistically correct and natural description.

For this, it was decided to use both classical methods of spatial analysis known from GIS systems and methods of artificial intelligence. The first stage of the research focused on the use of NLP techniques (FLAN-T5 XXL language model, ChatGPT 3.5) and the analysis of maps of building interiors. In the field of natural language processing, the so-called large language models play an important role. This is a type of artificial intelligence algorithm that uses deep learning techniques and large data sets to: understand, summarise, generate and predict new content.

In the first step of the proposed method, spatial parameters such as, for example, coordinates and place names, mutual topological relations between objects, distances and azimuths from landmarks, technical information about objects, spatial statistics, information about restrictions on people's movement, etc., are extracted using GIS algorithms. In the second step, from the acquired set of features, a command is constructed which is then to go to the NLP for execution. The way this command is formulated affects the result achieved, so it is iteratively determined what the schema of this command should be and what words should be used when forming it. In the third step, a response in the form of an expected description of the space is generated using NLP techniques.

In order to achieve the desired goal, a number of experiments and applications were made to support the research.
A prototype application was made in the form of a plug-in to the QGIS application for collecting information about the

features of the space around a designated place or route. Then, according to the adopted schemes, commands were generated that went to the FLAN-T5 XXL model and to ChatGPT.

The results obtained were then evaluated by comparing them with a benchmark, that is, a geo-description generated by an expert and considered correct. This was done both at the expert evaluation level and through automatic semantic similarity testing. In order to validate the methodology and technology adopted, a GeoDescTest mobile application was created to test the generated geo-descriptions by users. The application was developed in the Android Studio environment in the Kotlin language, and the building map used in it was generated using ArcGIS SDK technology. The building models were developed at Warsaw University of Technology as part of the "Building Accessibility Maps of Warsaw University of Technology Project".

The idea behind the application is to select from the places marked on the building map the one that matches the geo-description given by the application (Figure 1). At the same time, there are many more marked places on the map than there are geo-descriptions. The authors of this publication assumed that the accuracy of the location of points on the map largely reflects the accuracy of geo-descriptions. The better the geo-description is constructed, the easier it is to indicate the corresponding location on the map the first time. However, the quality of the map and the ability of those taking part in the survey to read it should be taken into account when analysing the conclusions.

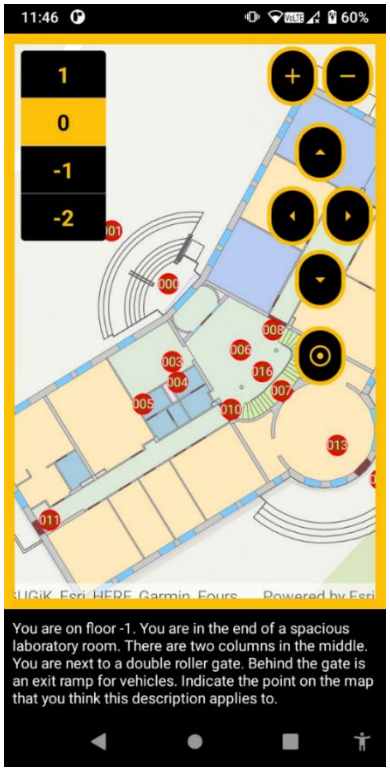

Figure 1. GeoDescTest application for geo-description testing (own elaboration)

The article will present the detailed results of tests conducted on the first test group, which were performed to verify the validity of the adopted methodology and technology.

More than a few dozen students participated in the study. All those who took part in the application testing completed the task, and the purpose of the study and the way the application was used were not in doubt. Although the survey was not a target task, the first data was acquired, the analysis of which will allow adequate preparation of the target survey for a much larger audience in different buildings. It is worth noting that the methodology adopted in the study is not tied to a specific language model. The proposed solution can be applied to others, including emerging newer language models.

https://doi.org/10.5194/ica-abs-7-90-2024