



# HOUSE

## About the project "Historical Ontology of Urban Space" (HOUSE)

The Tadeusz Manteuffel Institute of History of the Polish Academy of Sciences [hereinafter referred to as IHPAN] implements the project entitled "Historical Ontology of Urban Space" (HOUSE). This project is financed by the National Agency for Academic Exchange [hereinafter referred to as NAWA] under the International Academic Partnerships program (project number PPI / APM / 2019/1/00053 / U / 00001). The deadline for implementing the project under the contract signed with NAWA is 30 September 2021, but we are currently planning to postpone this deadline until the end of 2021. This is a procedure that will allow us to complete three of the five project tasks.

These three tasks cover a whole series of project activities, among which the most important are two **conferences**, four **workshop meetings** and a **dozen internships** and research stays, which hopefully shall take place without any obstacles in 2021. We have planned to make it possible that such internships at IHPAN Warsaw be attended by representatives of our partners and our associates in partner institutions. Thanks to these events, we will build a strong international cooperation network for researchers dealing with and interested in the history of the city. We are also planning several publications about which we shall write more soon. At the moment we are preparing a conference about the current state of research concerning the space and history of the city. We are changing its formula as in the current situation we can exchange knowledge and experience only remotely. Details of the conference "Historical Ontology of Urban Space: State of the Art" can be found on the project website.

We wish to complete the other two remaining tasks by the end of September 2021, which is why we plan to spend funds allocated for this purpose **by the end of September 2021**.

To put it simply, the HOUSe project consists of **two main parts**: 1) **the scientific one (and this is the case in this document**; the current situation will not affect the tasks since we have planned from the beginning that they can be partially implemented remotely) 2) the one involving international cooperation (carried out by means of conferences, workshops and scientific internships; the implementation of this part of the project in the current situation will be affected - we will postpone the deadline for scientific internships, and we will organize at least one conference remotely). At this point it is worth emphasizing that we plan to **register speeches at the conference**, the venue of which has been moved to our place. This will enable our colleagues to get acquainted with **the current state of research** in the field that interests us. If we manage to handle this matter as planned, the conference will be held at the turn of April and May 2020. Unfortunately, due to the current situation, we must remain patient in this case. We implement the project with foreign partners from Germany, France, Great Britain and Italy, which has two main effects: 1) we must be able to harmonize our data with data from non-Polish data sets, 2) the entire project is conducted in English. **The list of our partners** is provided in the draft version of **the project website**, which will be developed and about which we write below at: <https://atlas.ihpan.edu.pl/gis/urbanonto/>. Project activities will also be visible (recently started) on the network: we denote them #urbanonto.

The aforementioned **scientific part of the project concerns the history of the city**, and more specifically: the changes that occurred in the city space from the late Middle Ages (from the 16<sup>th</sup> century) to the beginning of the 20<sup>th</sup> century along with permanent elements that we think can constitute the city. Let us add: a European city. **In the attachment we provide the list of materials**: historical sources (including cartographic materials) and studies that are the basis of our research. They mainly concern Warsaw, for we will focus on this city in the project. The data collected in the project will refer to **the historical space of Warsaw**, but we will use the administrative boundaries of today's city to delineate this space, as it will provide us with a very wide range of research. In other words: we are interested in all elements / objects that were located in the past in the area enclosed by the borders of Warsaw in the 21<sup>st</sup> century. It should be emphasized that what is understood by data - in simplified terms - are those elements of the city space that can be represented on the city plan by means of cartographic symbolization. We want the results of our work to be applied when working on a series of historical atlases of cities conducted under the auspices of the International Commission on Urban History. The main goal of the series is to create a foundation for a comparative study of urban space in the past, as well as to disseminate knowledge on this topic: <https://www.historiaurbium.org/activities/historic-towns-atlases/>. This has an impact on **the model of the data** we collect in the project, as explained below. We want to build **a system** that will allow different teams dealing with city atlases to work in a uniform pattern, to enable future exchange of data sets and to improve comparative research.

At IHPAN we have hitherto implemented one project involving the construction of domain ontology. The "Ontological basis for the construction of historical geographic information systems" (OntoHGIS) project was completed in 2019. The project website along

with part of the documentation can be found at: <https://ontohgis.pl/przyklad-strona/ontohgis/> We have also gathered the results of our work in this completed project in a comprehensive monograph, the preprint of which will be available on request of those who contact the project manager of OntoHGIS, prof. Bogumił Szady ([szady@kul.lublin.pl](mailto:szady@kul.lublin.pl)) directly or through us ([urbanonto@ihpan.edu.pl](mailto:urbanonto@ihpan.edu.pl)). In the implementation of this project, we cooperated with dr hab. Paweł Garbacz and dr hab. Agnieszka Ławrynowicz. Their publications, which may be useful in the work on the UrbanOnto system, are available in the IHPAN library.

In the HOUSe project, we want to use the experience from the OntoHGIS project and build a system similar to the city space described above. We want the starting point in this work to be objects, classes of objects and categories of classes of objects included in the Topographic Object Database (BDOT10k) referring to the area of Warsaw within its contemporary administrative borders. However, using this starting point, we want to work retrogressively. In the HOUSe project, we use historical data, which are often not in the modern database or they denote objects in different ways from their contemporary successors / counterparts. The most important part of the ontology being built will therefore be **to organize and integrate contemporary / historical concepts / objects**, both in semantic terms (understood here as concept mapping, e.g. a gas station from the beginning of the 20<sup>th</sup> century and a modern one), and logical terms (understood here as the assignment of individual objects to relevant classes and categories). We expect that in this way a thesaurus will be created, in which individual entries will not only be classified and hierarchized, but also connected with more complex mereological and mereotopological relations. On the basis of the preliminary research, we expect an effect in the form of **a four-degree polyhierarchical structure**, but this is not a final assumption due to the ongoing process of collecting data to supply the planned database. Historians and cartographers from the HOUSe project team at IHPAN in close cooperation with the ontologist / ontologists whose work will be organized by an external contractor will be responsible for coming up with definitions of individual concepts. The basis for creating the definition of historical objects not included in BDOT10k and existing in the past in Warsaw will be historical sources and studies, including normative sources allowing to analyze legal and administrative structures regulating the city's functioning in the past (compare the part on the data model below). For example, we want to isolate and define concepts such as "building" or "plot" in such a way that these definitions relate to a specific period of time. An ideal situation occurs when it is possible to define and place in the conceptual structure the concept applicable to the entire period under analysis (from the 16<sup>th</sup> to the 20<sup>th</sup> cent.). Owing to our previous experience (resulting from the implementation of the OntoHGIS project), however, we know that these will be rare cases. We expect individual concepts to change over time, because some or all of their attributes will have changed - for example, the object "allotment garden" from BDOT10k will most likely not map to the object "garden" depicted in the 19<sup>th</sup>-century city plan or the object "garden" being part of the city plot in the 16<sup>th</sup> century. In such a situation, it will be necessary to introduce either an additional class (the class "permanent cultivation" will not match the garden of the 16<sup>th</sup> century), or a higher-order concept - e.g. "garden" for all three historical objects. Details of the system in which BDOT10k was created can be found in the publication:

<http://www.gugik.gov.pl/projekty/gbdot>. We expect that as a result of the ongoing works, domain ontology will be created. It can be applied to any European city from its beginnings to the present in the thematic scope included in the series of historical atlases of cities published under the auspices of the International Commission on Urban History. Therefore, it must be in two languages - Polish and English. The translation of the individual terms making up the ontology is provided by the contracting authority in consultation with the ontologist / ontologists whose work is organized by the contractor.

An important element of the created system will be such a preparation of data along with metadata that **the FAIR principles** (Findable, Accessible, Interoperable, Reusable) are met. You can read more about it here: <https://www.go-fair.org/fair-principles/> The HOUSE project that we implement is financed from public money, which is why we want all results to be accessible and open, as much as possible. We hope to achieve this by adopting the appropriate data model.

**Data model:** as noted above, the data collected in the HOUSE project primarily concern Warsaw from the 16<sup>th</sup> century to the beginning of the 20<sup>th</sup> century. The data collected in the project will describe **the space of historic Warsaw**. As the city's boundaries changed in different time sections, we shall use retrogression: we will use the administrative boundaries of today's city and look at those objects that from the 16<sup>th</sup> century to the beginning of the 20<sup>th</sup> century appeared, disappeared, or were constantly present in the space enclosed by the borders of Warsaw of the 21<sup>st</sup> century. It should be emphasized again that what we understand by the data interesting to us in the HOUSE project - in simplified terms - are those elements / objects of the city that can be presented on the city plan by means of cartographic symbolization. The specific nature of this data may be their variability in time. Due to the fact that we work with foreign partners dealing with similar issues in relation to cities from other parts of Europe, it may also turn out that individual elements (more precisely: their functions or semantic scope) may also differ spatially. For example: analyzing old city books (the 16<sup>th</sup> century), lists of residents (the 18<sup>th</sup> century), cadastral plans (the 19<sup>th</sup> century), address books (the early 20<sup>th</sup> century), we collect references regarding streets / roads / paths / routes / alleys and other communication routes with their attributes (for example: 1) building material: paved street, beaten road, 2) construction: dead-end street, built-up avenue, 3) location: street / suburban road or a representative prospectus running in the city centre, 4) size: small, narrow, long) in order to organize the structure of this type of information later- thanks to ontology. The starting point for working on the data model is the 10k Topographic Object Database (BDOT10k). On the one hand, we want to prepare data in a model that is successfully used throughout the country, which will provide us with knowledge about the possibilities of integrating historical and contemporary data. On the other hand - exactly the opposite - we want to test the scope of use of this model for historical data. We know that this model is not sufficient for the data collected in the HOUSE project for several reasons:

- 1) It is a model referring to Poland and therefore it cannot be used as a data model for a city space in Europe - in this case, when developing a data model for the HOUSE project, we should use the experience of our partners who will provide information on analogous data models from their countries.

- 2) It is a model referring to contemporary objects, classes and categories of objects and as such it does not take into account historical objects which no longer exist - in this case, when developing the data model for the HOUSE project, one should use the experience of Tomasz Panecki, PhD, who developed the foundations of the historical data model based on historical topographic maps, but with the limitation that this model does not refer in sufficient detail to the city space. The dissertation "The concept of the structure of the database of historical topographic objects" by Tomasz Panecki, Phd, can be found here: <https://depotuw.ceon.pl/handle/item/2643>
- 3) It is a static model, which does not take into account changes in time related to both the functions of individual objects and their attributes - in this case, when developing the data model for the HOUSE project, one should take advantage of the experience developed within the OntoHGIS project, especially in terms of the attribute of "time" determined by the beginning and ending moment (the occurrence of the phenomenon, the appearance of information in the source, etc.) expressed by a date with different accuracy (for example: the 16th century, 1567, 17 July 1567, September 1567).
- 4) This model uses **identifiers** that are understandable to a Polish-speaking user, but it is difficult to maintain this scheme of creating identifiers for a database that is to be useful for international teams working on individual historical atlases of cities in different languages. In this case, an alternative scheme for the creation of identifiers should be developed, perhaps based on TERYT identifiers or other. It is possible to design local identifiers, which in the future may set a standard for data collected, developed and published by teams working on subsequent issues of the historical atlas of cities. The contractor is responsible for this part of the work in close cooperation with the contracting authority.

Historians and cartographers from IHPAN who provide data to the database deal with the analysis of historical sources and the acquisition of data from these sources. These will be data constituting the main population of the database, collected initially in the form of .xls or similar files, .mdb or similar files, or .shp files, and then entered directly into the application intended for it. Such an application will become part of the UrbanOnto IT system we are working on. We anticipate that historians involved in the project will have collected data by the end of May 2021, with some of the data being provided to the company that guarantees IT support for the project as soon as the contract is signed, at the stage of developing the data model and creating the database structure. These will be data from a selected map of Warsaw from the 19<sup>th</sup> century and a selected written source. The data will be submitted in a format agreed directly with the task's contractor (e.g. .shp files for data obtained from the city plan and an .xls or .mdb file with data from a written source ordered according to the data model from BDOT10k). We plan all the data to be used to develop project publications and then, as we hope, the data will be used for subsequent projects.

Wieslawa Duzy, PhD