IV WPSC
IV WORLD PLANNING SCHOOLS CONGRESS JULY 3-8TH 2016 RIO, BRAZIL

An intuitive "Scientific Workflow System" for spatial planning

Tiberiu Florescu, Ph.D;[1] Cătălin N. Sârbu, Ph.D;[2] Andrei Mitrea, Ph.D;[3] Corina T. Chirilă;[4] Alin D. Chiș, Ph.Dc;[5] Vlad Cocheci, Ph.Dc;[6] Alexandru V. Costin, Ph.D[7]

[1, 2, 3, 4]*School of Urban Planning, 'Ion Mincu' University of Architecture and Urban Planning, Bucharest, Romania.*

tiberiuflorescu [at] gmail.com; sarbu52 [at] yahoo.com; andrei.mitrea [at] uauim.ro; corina.chirila [at] yahoo.com;

[5]*School of Geography, University of Bucharest, Bucharest, Romania,*

alin.d.chis [at] gmail.com;

[6]*School of Geography, Babeș Bolyai University, Cluj-Napoca, Romania,*

vladcocheci [at] gmail.com;

[7]*Neoton Design, Cluj-Napoca, Romania,*

neotondesign [at] gmail.com.

1. The Problem

Recent years have witnessed a sharp increase in complex spatial planning problems. Consequently, almost every research exercise has to face the rise of vast new bodies of information, which become increasingly difficult to handle. At the same time, big data expands rapidly through recurrent additions of new datasets.

In principle, such readily available data should pave the way for greater transparency in planning, research and decision-making. Unfortunately, the increased availability of data comes with a huge drawback: data is usually available in many flavours of varying quality. Therefore, one needs to clean, normalize and filter such data, prior to connecting it with other sets of already processed information. This is most certainly a necessary exercise. However, it subjects the researcher to the arduous task of cleaning and sorting data, which is normally a time consuming, repetitive and often boring task. Moreover, it represents just a fraction of the entire problem he or she aims to address.

In addition, these new kinds of research require an inter-disciplinary approach: statistics, mathematics, informatics, geography, economics and sociology, urban and

regional planning, as well as law and politics, to name only a few. Any researcher involved in such activities therefore needs to work with different tools, each coming with its own files, formats and language. It quickly becomes very difficult to keep track of all the data within one's workspace.

2. Objectives

We recognize that the problem described above is twofold: Firstly, we need a reliable analytical tool for tackling spatial problems, and secondly, we need reliably processed information, in order to shorten research time. In addition, we strongly believe that this tool should harbour the possibility to cater for the needs of other researchers.

Hence, our proposed solution is as follows: To begin with, we wish to create a single open source work environment with basic functionalities. This work environment should come with the possibility of easily constructing one's own custom functionalities.

At the same time, such a work environment should easily connect to a database structure for storing and retrieving data. Hence, a subsequent step would then be to host a database to store 'clean', normalized data, available online, ready to be used via the application. We believe there is no need for everyone to process the same data over and over again. We therefore wish to focus on sharing already published data, thereby aiding future projects.

Hence, the application will constantly be connected to this database, thereby offering the option for saving the workflow, as well as for import – export operations. General functionality thus includes the following: data collection and processing; data storage; data retrieval and filtering; data analysis; graphical representation; text processing; the possibility of exporting data at each stage of the research project; data sharing; workflow loading and saving; and workflow templates.

With respect to the application itself, we wish to avoid creating a novel application from scratch, as we do not have the necessary time and resources. Instead, we would like to create a common framework for selected existing open source projects, on which they can work and interact.

We are also aware that we need to provide an intuitive homogenous graphical user interface, in order to increase the speed and quality of research.

Thus, the application fits into the 'Scientific Workflow System' category. There are already other relatively similar tools available, but they were started from other fields of research and do not comprehend all of the functionalities needed in spatial planning and research.

3. Methodology

Our language of choice in solving the problem sketched above is the Julia Programming Language. We have come to appreciate Julia for its easiness of use, its inherent speed and for the possibility to integrate other programming languages.

At the time of writing, we are developing both the application and its associated database in an incremental manner, as we encounter difficulties within our own research projects.

Specifically, we shall make a list of required tools, and subsequently check every source to test whether it is compatible with our programming language of choice. We shall then proceed to create the connections between these existing tools and Julia. If the language of the source code does not prove desirable, we shall try to find alternatives. If they are non-existent, we shall create them anew.

We are working on templates to store different kinds of data. The prime database of choice is the RDBMS PostgreSQL, but we shall also need noSql DB, for storing graphs. Another database we would like to use is SciDB, a database created for working with arrays and matrices.

With respect to the graphical user interface, it will need to be created from scratch.

4. Main Results and Contributions

As a final result, we expect to have a unified open source work environment for students and scientists alike, as well as a readily available interconnected database. Both of them should provide great improvements in the speed and the quality of research. In addition, the resulting research work will gain full transparency, since both data and workflow are available for others to check.

Both tools should, in principle, have a high degree of applicability, mainly in the field of planning and research. Thus, the application is built with robustness in mind, in order to tackle even the most demanding tasks. At the same time it supports extensions and customisations, in order to suit other researchers' needs.