Towards an Open Source, Systems-Integrating Spatial Decision Support Framework for Urban Public Health Environments

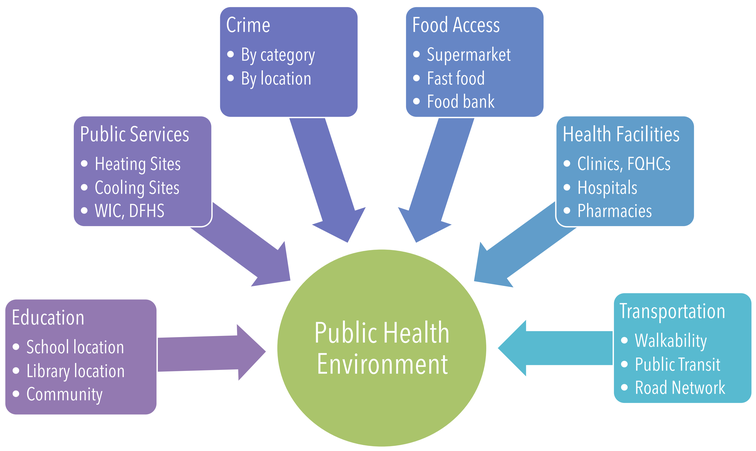
Marynia Kolak | Center for Spatial Data Science | University of Chicago

Introduction

A systems-integrating framework weaving space-time data, spatial analysis and modeling, and decision-support open source platforms was developed to produce a robust, web-based application for use by residents, non-profits, and city government.

The processes of the integration framework, still in development, are:

1) Develop Spatial Data Model
2) Build Spatial Data Warehouse
3) Integrate Space-Time Data & Spatial Analysis APIs
4) Consume in CS Framework with Application Visualization



Spatial Data Model for Public Health

Infrastructure Challenge

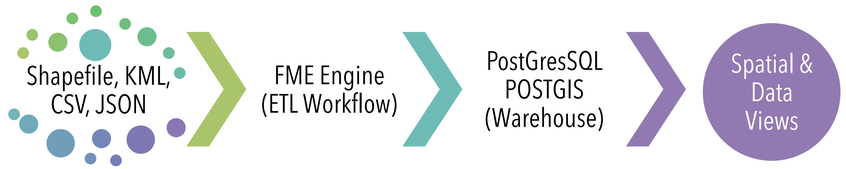
No comprehensive data model to represent the public health environment exists, and yet a need for distributed, interoperable structures to integrate data across health programs and institutions is paramount in a transforming health economy. Common challenges to this infrastructure have included data integration, data standards, interoperability, common terminology, and data confidentiality.*

Data Model Prototype

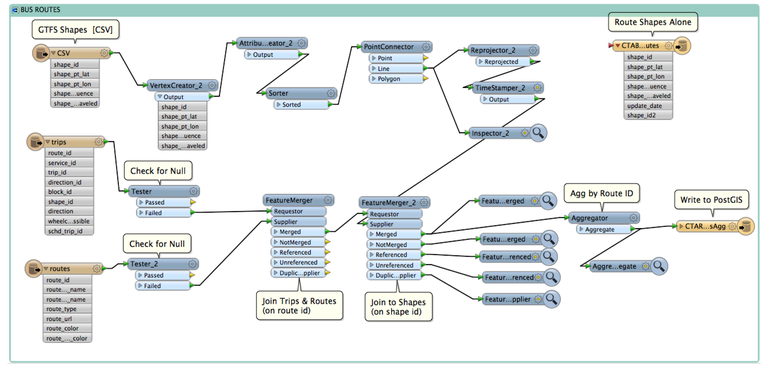
A working spatial data model to represent the public health environment matrix in Chicago was developed from literature review and expert findings. Spatial data and tables include features of the built environment, transportation, city infrastructure, demographic and public health reported data, and facilities relevant to healthy living. Collaboration with public health officials to further develop the complex representation of social determinants of health in the urban environment serves as a core component of the next phase.

Urban Spatial Data Warehouse Build

A spatial data warehouse was established as a spatial data infrastructure to integrate and standardize complex data in a common interface that comprises a public health environment matrix of the City of Chicago.



Over a hundred datasets were extracted, transformed, and loaded into a PostGresSQL/POSTGIS warehouse using the Feature Manipulation Engine (FME), pulling from both open and proprietary data formats and transformed into standardized, commonly projected features. Spatial and data views were then developed to drive front-end visualization for a decision support prototype. PgModeler was used to visualize the data model of the warehouse, and views loaded into OpenJump GIS for testing.



ETL workflow transforms CSV bus stop data to bus route spatial file stored in PostGRES/POSTGIS.

This spatial infrastructure was developed as a first phase to the process. ETL workflows will be translated to python scripts in a Complex Systems Framework to allow for open source flows and operations.

Integrate Data & Analysis Platforms

An effective decision support framework for complex, dynamic environments like cities require abilities for on-the-fly analysis and incorporation of dynamic data sources. Open source platforms that support data extraction, data analysis in a spatial context, workflow, & visualization were investigated and connected in a systems-integrating framework.

Plenario API

Plenario (at plenar.io) is a spatio-temporal open data platform in its alpha stage. It serves as an automated ETL builder that aggregates data in a single spatial and temporal index.

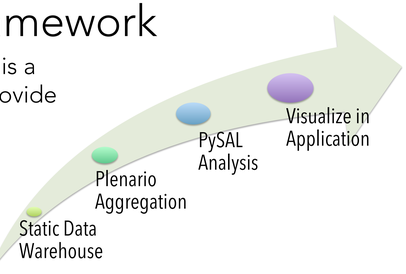
PySAL API

PySAL (at pysal.org) is an open source library of spatial analysis functions written in Python intended to support the development of high level applications.

Complex Systems Framework

The Complex System Framework (CSF) is a flexible software system designed to provide decision-makers with live data, intuitive controls, and cost-benefit analysis (at complexsystemsframework.com).

CSF OPEN SOURCE DECISION TECHNOLOGY

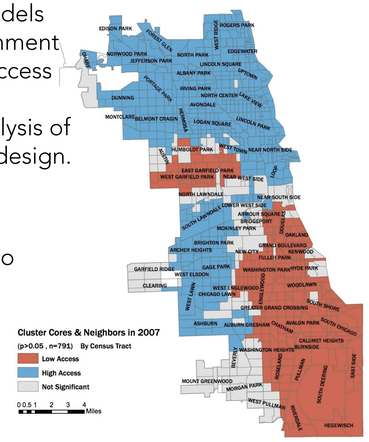


Food Access in Chicago

For our case study, a support system models components of the public health environment in Chicago, visualizes food access and access clusters using different transportation methods, and would offer on-the-fly analysis of that environment with human-centered design.

Challenges & Considerations

- Most effective incorporation of Plenario aggregation with PySAL analysis?
- Need for long-term engagement and varied ownership
- Security & quality varies across data
- Inherent loss of quality by representation: how does this affect decisions for policy-making?



Clusters of Food Access in 2007

The Team

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

Core team members are Marynia Kolak, Julia Koschinsky, and Luc Anselin of the University of Chicago's Center for Spatial Data Science and Robert Pahle of ASU, and Raed Mansour at the Chicago Department of Public Health. The warehouse was built with Alexander Stepanov and Geri Miller from John Hopkins University, and Satyender Goel of CHITREC at Northwestern University. This project wouldn't be possible without the additional support of Derek Eder and the Plenar.io team.

*Richardson, D. B., Volkow, N. D., Kwan, M. P., Kaplan, R. M., Goodchild, M. F., & Croyle, R. T. (2013). Spatial Turn in Health Research. *Science*, 339(6126), 1390–1392.

marynia.kolak@gmail.com