Wellness indicator MVP Test

DIALOG has published the Community Wellbeing framework with an intent to apply design thinking to improve the wellbeing of communities through building design. Having a definition for metrics and indicators though are only a start. The significant challenge of how to support the desired indicators with meaningful data still remains. DIALOG may be able to use external third party contractors to assist with this. However, while this will work, it fails to provide DIALOG with a sustainable competitve advantage. Data and analytics and in particular spatial analysis needs to become a core competancy at DIALOG if DIALOG wants to be a player in this space.

There are a broad spectrum of metrics and indicators outlined in the framework. A detailed assessment of the feasibility and potential approach to computing these metrics is beyond the scope of this short study. At a high level, the metrics can be broken down into the following:

- Soft metrics that are more empircal in nature and may require end-user surveys or studies
- Metrics that can be derived from open or government data sources
- Metrics that will require input from AEC firm based internal data sources

DIALOG has only limited data available in a form that would make it possible to calculate metrics. Soft metrics would require a great deal further research and may also require specialized development of apps or other forms of surveys to collect data from stakeholders.

The goal of this short-cycle study therefore is to go through the excercise of using the Wellbeing Framework to create a Minimum Viable Product example of a visualization dervied from internal DIALOG data and Open Data Sets that rely on immediately available open data sets and internally available DIALOG sources.

Wellbeing Metrics MVP Goals

In order to gain a baseline understanding of the nature of the challenge facign DIALOG's Wellness indicator project, a short-cycle Minimum Viable Product project was understaken with a limited scope and time.

The goal of this project were as follows:

- Select a single urban region to focus on where DIALOG has buidlings
    - Calgary is a good candidate as it is a reletively small contained urban center to test
    - DIALOG has several large projects in the Calgary region to draw data from
- Select one of the 18 domains, and a limited set of a few specific metrics to see if it were possible to assemble relevant data from combining open sources and DIALOG sources
    - There are many 'soft' metrics that may require surveys or other forms of direct user data contribution. These will be very challenging to aquire data for.
    - The MVP therefore will focus on selecting a domain and associated metric that uses redily available 'hard' data
- Assess and Map out a workflow associated with the cleanup, assembly and integration of data
    - There is an internal workflow process assocaited with assembling and provisioning data DIALOG may have about a building. For this, basic data will be pulled for a single project
    - The MVP will focus on pulling together data surrounding the project from open sources for things such as amenities etc.
- Determine potential methodology for contributing data back to the community from DIALOG sources

For the MVP therefore, the following base metrics are possible:

- 1D 1: Public/Common area and primary circulation routes around project: public art, retail
- 3B 1: Walking distance to cycle path
- 2A: Project is walking distance to health support services
- 2B: Project is walking distance from healthy food options
- 2C: 2: Project is close proximity to outdoor spaces for quiet repose

Of these, Walk Score food options was chosen for simplicity to build out the base process.

Walk Score is typically computed using a proprietary algorithm. DIALOG may be able to use Walk Score as a partial metric for understanding communities. Ideally the baseline walk score would be augmented and blended into the Wellness framework. The goal is to see if it is possible to construct a methodology for computing specific walk scores for communities immediately surrouding buidlings. While these data are available on the Walkscore website, the intent of this experiement is to validate a baseline methdology against a known metric. This then can provide a foundation for processing and analyzing other forms of metrics for wellbeing and/or blending or aumenting baseline walkscore beyond its simple foundation.

Walk Scores

Walk Score® is scaled linearly, ranging from:

- 0 to 24 “car-dependent” (car required for almost all errands),
- 25–49 “car-dependent” (car required for most errands),
- 50–69 “somewhat walkable” (car required for some errands),
- 70–89 “very walkable” (car not required for most errands), to
- 90–100 “walker's paradise” (car not required for errands) (Walk Score, 2012).

Walk Score® is calculated by determining a raw score out of fifteen, normalizing that score from zero to one hundred, and deducting two penalties for low intersection density (ID) and high average block length (ABL) (Walk Score, 2012).

Walk Score® = Raw Score/15 x 6.67 - (ID - ABL)

The raw score is composed of nine amenity categories of walking destinations (grocery, restaurants, shopping, coffee shops, bank services, schools, entertainment, bookstores, and parks) each weighted from one to three points based on low, medium, or high importance for walking in six research articles referenced by Walk Score® (Cerin et al., 2007; El-Geneidy and Levinson, 2011; Iacono et al., 2010; Lee and Moudon, 2006; Moudon et al., 2006; Piekarski, 2009; Walk Score, 2012). Based on this literature, we selected indicators from the CHBE project for each amenity category (Table 1). Our indicator selection was straightforward for grocery, restaurants, coffee shops, bank services,schools, and bookstores. We included specialty food stores under shopping (Cerin et al., 2007; Lee and Moudon, 2006), all green spaces under parks (Cerin et al., 2007), and both cultural and sporting activities under entertainment (El-Geneidy and Levinson, 2011; Iacono et al., 2010).

Scores within each category were attenuated by a close approximation of the Walk Score® distance decay function awarding:

- 100% of the possible maximum points to amenities located within a network walkshed distance of 0.25 miles (400 m or 5 min walk)
- 75% within 0.5 miles (800 m or 10 min),
- 40% within 0.75 miles (1200 m or 15 min),
- 12.5% within 1.0 mile (1600m or 20min) of each location (Walk Score, 2012)

The weighting of three categories (restaurants, shopping, and coffee shops) reflects the number of destinations available (or "depth of choice") (Walk Score, 2012).

Finally, the Walk Score® intersection density (ID) function was used to deduct a maximum 5% penalty for 60 intersections per square mile and the Walk Score® average block length (ABL) function was used to deduct the Same maximum of 5% for N195 m length per block.

General Walkscore Algorithm Ranks:

- Grocery 3 Grocery stores; ethnic food stores
- Restaurants 3 Fast food counters/restaurants; full service/hotel/ethnic restaurants; banquet halls; outdoor dining; bars/nightclubs; other food outlets
- Shopping 2 Big box shops; shoppingmalls; stripmalls; bakeries; butcher shops; delicatessens; farmers' markets
- Coffee shops 2 Coffee shops
- Bank services 1 Commercial banks; financial services
- Schools 1 Elementary/junior high schools; high schools; universities; other schools
- Entertainment 1 Auditoriums/concert halls; theatres; museums; movie theatres; games rooms; gyms/fitness centres; indoor/outdoor hockey arenas; indoor/outdoor pools; wading pools; tennis courts; basketball nets; community gardens;
- other recreational spaces/public places
- Bookstores 1 Bookstores; libraries
- Parks 1 Playgrounds; spray decks; playing fields; open green spaces; golf courses; lakes/ponds; fountains/reflecting ponds; campgrounds; streams/rivers/creeks/canals;

Intersection Density and Average Block Length

The Walk Score® intersection density (ID) function is used to deduct a maximum 5% penalty for < 60 intersections per square mile.

- 1600 m radius = 8042477.193 m^2 = 3.1052 mi^2
- ID = #Intersections/3.1502

Intersection density (intersections per square mile): over 200: no penalty

- 150-200: 1% penalty
- 120-150: 2% penalty
- 90-120: 3% penalty 60-90: 4% penalty under 60: 5% penalty

The Walk Score® average block length (ABL) function is used to deduct the same maximum of 5% for > 195 m length per block. Average block length (in meters):

- under 120 m: no penalty 120-150 m: 1% penalty

- 150-165 m: 2% penalty

- 165-180 m: 3% penalty 180-195m: 4% penalty over 195m: 5% penalty

- Field Validation Of Walk Score Study

- Validation of Walk Score for estimating access to walkable amenities

Limitations of Walk Score, and The Potential For Graph Database Augmentation Approach to Wellbeing

Walk Score is limited and excludes aspects that limits its applicability to the desired metrics outlined in DIALOG's Community Wellbeing Framework. The walkscore algorithmis an approximation, and doesn't include such factors as topography, street design, available public transportation or bodies of water in its rankings. It calculates using linear distance, which means it doesn't factor in barriers such as lakes, rivers, open spaces, closed spaces or routing.

The Walk Score algorithm also doesn't rate a neighborhood on visual aspect: a factor that is a clear apsect of DIALOG's Wellbeing framework, but rather how easy it is to live car-free. If you live near a nature preserve or hiking trails, that greenery may improve your quality of life, but it won't improve your walk score.

Walk Score was created in a time when data was more limited and less available, and the toolsets more primitive than are now available.

The ready availability of Open Street Data, Government data and even with the limitied and admittedly poor state of DIALOG's internal data, it still would be conceptually possible to construct a far more powerful algorithmic approach to computing the outlined indicators and metrics.

For example: the Open Street Map data is in fact a graph, it becomes possible to increase the accuracy and power of walk score type calculations by using graph analtyics methods. The A* algorithm for example makes it possible to use route finding between buidling and amenities instead of relying on abstractions for assessing walkability of a neighborhood. In effect - the computation theoretically becomes more accurate as the algorithm is in effect sending an 'agent' down the actual path between the building and all potential amenities.

This could provide a powerful alternatives for making assessments of community well being by actually having agents walk the graph and computing a set of not single but combinatorial metrics along the route.

The agent pattern could be run continuously over time as new data sources became available. This would provide a powerful set of dashboards for continuously assessing the wellbeing of communities.

Open Data Sources

Mapping Base Data: Open Street Map

Open Street Map is an open source globally assembled spatial database. It is well supported, has a broad set of tools both open and closed source and is extensible by individuals and organizations. It has established definitions and tagging structures as well, and is broadly used by government bodies as a foudnation for adjacent open data sets. Open Street map sets are easily blended with other data sets external and internal. All mapping systems including ESRI, Geoserver, neo4j etc. can be tied into Open Street Maps data sets as well.

In addition to deriving data to act as a foundation for Wellbeing, DIALOG could become a contributor to the Canada Open Street Map Building inventory project. By doing this, DIALOG would be able to contribute data regarding all the projects it has worked on to build out this Canada-wide data set. This would not only build reputation, but make available a foundational set of data that could be used to feed back into DIALOG's own designs.

A standard process is well established that enables data to be added to the master Open Street Map datasets.

Open Street Map has a utility called JOSM that allows for remote control of Open Street Map. With this, it becomes possible to create a data value stream workflow that would enable DIALOG to take its internal building data and load it directly into open street map.

Open Street Map Resources

- Open Street Map Site:: Main open street map site - account can be setup to add, download or edit open street map data
- Editing and Adding Open Street Data:: Blog post on editing open street map data
- Open Street Map Editing Walkthrough:: Walkthrough of open street map data editing
- JOSM Java Open Street Map editor Editor for working on open street map data
- You tube video for import process for OSM Video tutorial of import process. Contributors can sign up for an import account and use this to bring data into open street map
- Open Street Map Tag information Open Street Map Tag definitions
- Open Street Map Biking Export For Ottawa Ottawa open street map dataset for cycling

Open Data Building Data Sources

Open street map has been adopted by the federal government as a mechanism for capturing and storing data associated with buildings, communities, cities etc. There is an initiative undertaken by Ottawa and other communities to crowdsource the input of these data through the Canada Open Street Map Building inventory project.

- Federal facility green house gas data: datasets for federal buildings and fleets

for green house gas emissions
- Canada Open street Map Building Inventory Project for capturing building data for all of canada
- Canada buiding 2020 data sets
- Open Database Of Buildings
- Calgary open street map buliding task Task associated with mapping Calgary buildings

DIALOG Data

While DIALOG has a considerable amount of data avialable in its various internal systems, DAILOG faces significant challenges assocaited with being able to effeictvely use any of these data.

dataLink - A Platform for DIALOG Data

In nearly all instances of the indicators outlined in the Community Wellbeing Framework propose using drawings as a source for key underlying data. The objective of dataLink was to be a foundation for exactly purpose: automating the harvesting and integration from key DIALOG sources and integrating them so that

The goal of the dataLink project was not to be a search tool for DIALOG. The intent was for dataLink be DIALOGs data platform - a system for unifying all of DIALOGs project and buliding data and providing a single API to access this rich dataset. The search utility was an additional benefit.

The dataLink would have acted as a means of pulling together all of DIALOGs data, enabling stakeholders to clean the data sets up, contribute additional key data about projects and associated buildings including locations, footprints, key building aspects etc.

These unified data would have acted as a source to pipe data into open street map, green database or other DIALOG sponsored data related project.

Without the dataLink, any Community Wellbeing framework related initiative will suffer from requiring costly manual effort to support it.

Wellness Workflow Using DataLink:

The workflow would have been as follows:

- dataLink's databots continuously integrate data from key DIALOG data sources
- Currently building geometry and footprints are not immediately available, but could have been derived from using Reports and ML processing to extract baseline building outlines
- Building outlines could also be contibuted by DIALOG users either via upload of polygon, or through digitizing in the dataLink app
- This would have provided a combination of the baseline building characteristcs and the key geometry in addition to all other relevant project data
- These data could be extracted to an open street map style layer and imported into JOSM tool
- A DIALOG user would then push these into open street map for contribution to key data sets for AIA 2030 and overall building statistics
- These data would then be accessible immediately for use in wellness indicator reports and calculations along with all the additional data.

Visualization Tools

- Uber Deck.gl: Uber's powerful spatial visualizaiton javascript framework. This tool is desgined for website developers to create powerful interactive spatial visualizaitons
- Uber Kepler: Uber's simple user friendly quick visualizaiton tool. Data can be added using simple CSV spatial files as layers to quickly build visualizations
- Quantum GIS: Open source powerful general purpose GIS tool. Just as good as ESRI for most spatial analytics and spatial data manipulation and translation

Data Tools

- Neo4j Open Street Map importer: java based importer that pulls open street map data into neo4j. Open Street Map is actually a graph, so this importer brings data directly into neo4j using the underlying spatial graph represenation. A lot of potential here for combining spatial open street map data with other sorts of non-spatial data for wellness analytics spatial to non-spatial. If DIALOG had continued with dataLink, it would have been possible for example to blend

DIALOG's internal project/building data with Open Street map data.

- Spacetime Reviews spatial query/graph example: Example of using spatial queries against neo4j to plot graphs of data associated with spatial and non-spatial data
- Blog post on neo4j spatial and time data: Example of using neo4j spatial and javascript to make an interactive spatial website