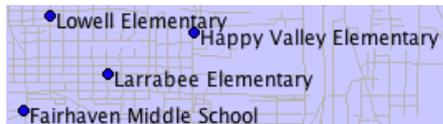


PAGC: Recent Advances and Future Developments

Dan Putler and Walter Sinclair, Anemoi Analytics
Stephen Woodbridge, iMaptools



Postal Address Geo-Coder



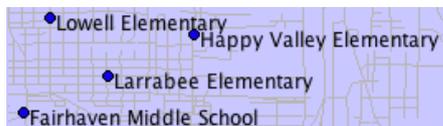
Agenda

- About PAGC and a little about its history
- An overview of what PAGC provides
- A glimpse under the covers of how PAGC works
- The current roadmap for future development
- Examples of how PAGC is currently being used
- Q&A

The PAGC Elevator Pitch 1

- What is PAGC?

The Postal Address Geocoder (or PAGC) is an open source library and web service framework for geocoding locations based on either postal addresses, street intersections, parcels, or landmarks. The system offers a number of features that differentiate it from other open source geocoding software solutions.



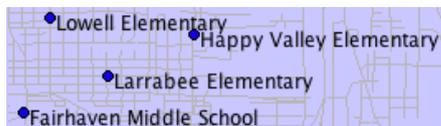
Postal Address Geo-Coder



The PAGC Elevator Pitch 2

- Why should you care?

If you need a tool that takes a textual description of a location, such as an address, landmark, or parcel, PAGC can look that description up in a reference database in order to produce a location (a latitude and longitude), then this is a tool that will provide that service. As a library this can be linked directly into other source code packages to create a seamless integration. As a webservice, you have the flexibility to make requests from a variety of platforms via a simple HTTP request.



Postal Address Geo-Coder



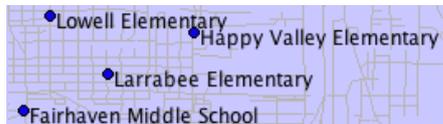
How to Get PAGC

Project Site:

<http://www.pagcgeo.org>

SourceForge Page (the path to SVN):

<http://sourceforge.net/projects/pagc/>

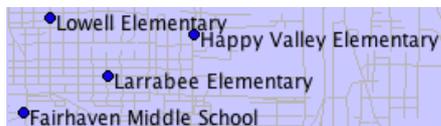


Postal Address Geo-Coder



PAGC – Background

- Development of PAGC began in 2000
- **Objective:** Create a *command line* geocoding program with functionality equivalent to the ArcView 3 geocoder, but make it cross platform
- PAGC becomes an open source project in 2002
- Starting in 2006 PAGC started to transition from a standalone program to a library based suite



Postal Address Geo-Coder



Key Features 1

- PAGC is written in ANSI C, currently runs on all major operating systems, and can be easily ported to nearly all operating systems
- PAGC is pre-configured to use data from several different providers (TIGER/Line, Statistics Canada's Road Network File), and through the creation of a simple configuration file, it can easily work with data from a large number of other providers

Key Features 2

- PAGC makes use of an advanced three-method, probabilistic matching algorithm to provide the greatest possible match rate for street addresses
 - Exact match
 - Soundex
 - Pointer-less trie combined with an edit distance measure
- PAGC can be used with multiple input data sources
 - Shapefiles
 - SQLite/Spatialite database tables
 - PostgreSQL/PostGIS database tables)

Key Features 3

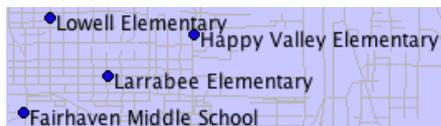
- PAGC works with multiple data stores
 - BerkeleyDB
 - SQLite
 - PostgreSQL (under development)
- PAGC has the ability to provide an ID number in addition to geographic coordinates, which can greatly simplify database operations for location-based applications

Key Features 4

- PAGC allows the user to edit road segment address ranges on the fly, enabling the user to overcome suppressed or altered address ranges frequently encountered in government originated road network layers
- PAGC as a web service allows you to pass an address parsed into fields or as a single text field that will be parsed into fields internally, giving you a lot of control over your requests

What PAGC Does Under the Covers

- **Build:** Create a standardized back-end database of street segments, property parcels, and/or landmarks that can quickly be searched.
- **Match:** Take user input, standardizes it in a way that is compatible with the back-end database, search the database, and return the search results to the front-end software.

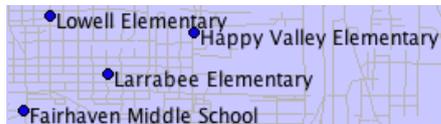


Postal Address Geo-Coder



PAGC's Software Components 1

- **libpagc**: The underlying, API-based workhorse
- **libds**: A data store abstraction layer that allows PAGC to work with both raw and standardized street segments, property parcels, and landmarks in several different underlying formats
 - Shapefile/DBF/BerkeleyDB
 - SpatiaLite/SQLite
 - PostGIS/PostgreSQL (in development)

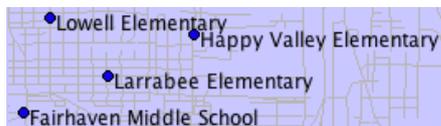


Postal Address Geo-Coder



PAGC's Software Components 2

- Front-end geocoding software to obtain user input and interaction
 - The web geocoder service
 - Command line tool for bulk geocoding, user interaction, and “on the fly” editing of the underlying road segment database
- Front-end build tool for configuring the standardization and loading data into the back-end database



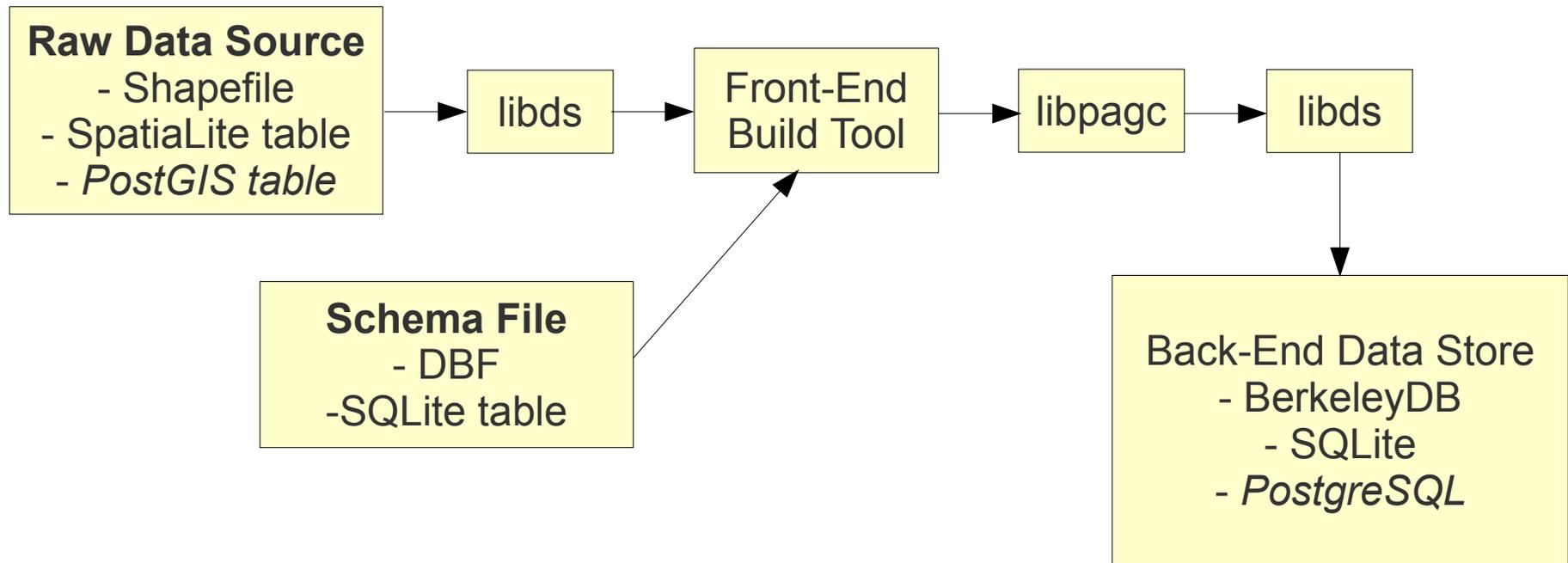
Postal Address Geo-Coder



Other PAGC Components

- Standardization files
 - rules.txt
 - lexicon.csv
 - featwords.csv
 - gazeteer.csv

The Build Phase



Note: *Italics* indicate that use of this format is under development

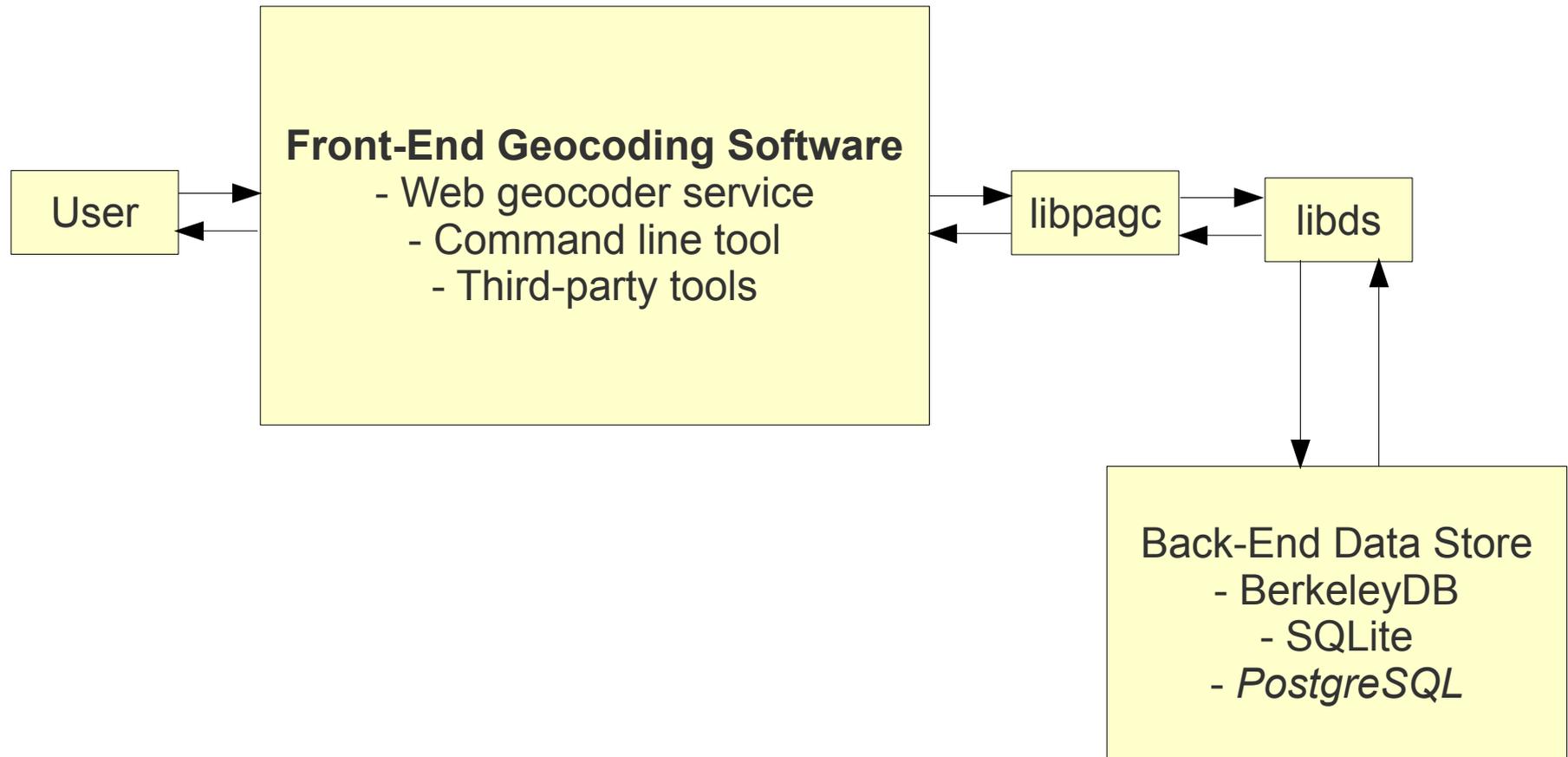
The Schema (Configuration) File

	A	B	C	D	E	F
1	ATTRIB,C,13	COMPARE,C,26	NAME1,C,10	NAME2,C,7	NAME3,C,9	NAME4,C,7
2	HOUSE	NUMBER_INTERVAL_LEFT_RIGHT	LFROMADDR	LTOADDR	RFROMADDR	RTOADDR
3	PREDIR	CHAR_SINGLE	PREDIRABRV			
4	PRETYP	CHAR_SINGLE	PRETYPABRV			
5	STREET	CHAR_SINGLE	NAME			
6	SUFDIR	CHAR_SINGLE	SUFDIRABRV			
7	SUFTYP	CHAR_SINGLE	SUFTYPABRV			
8	QUALIF	CHAR_SINGLE	QUALIFIER			
9	CITY	ALT_CHAR_LEFT_RIGHT	PPLACEL	PPLACER	CPLACEL	CPLACER
10	PROV	CHAR_LEFT_RIGHT	STATEL	STATER		
11	POSTAL	POSTAL_LEFT_RIGHT	ZIPL	ZIPR		
12	SOURCEID	NO_COMPARISON	TLID			
13	GEOMETRY	SPATIALITE_WKB	the_geom			

Building PAGC Standardized Data

- Analyze your data and decide on a schema
- Create a schema (configuration) file
- Join the data if needed
 - tiger_street_join utility
 - SELECT INTO statements for SpatiaLite tables
- Run the pagc_build_schema utility to standardize and load the records

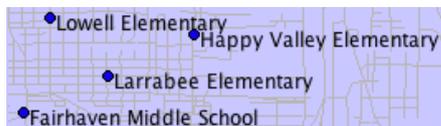
The Match Phase



Note: *Italics* indicate that use of this format is under development

The PAGC Roadmap 1

- Determine ways to improve performance with **large** data sets
- Revitalize the command line tools for bulk geocoding
- Simplify the build and setup process
- Better support for languages other than English
 - UTF8 support if/as needed
 - Support for country specific gazetteers, abbreviations, etc

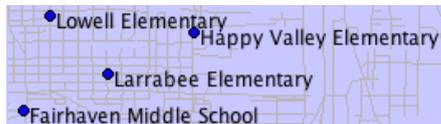


Postal Address Geo-Coder



The PAGC Roadmap 2

- Improve user and administrator documentation
- Provide language binding for libpagc
 - Python
 - Java
- Library changes to support embedding PAGC into a databases systems such as PostgreSQL



Postal Address Geo-Coder



Metropolitan Mosquito Control District

mmcd
METROPOLITAN MOSQUITO CONTROL DISTRICT

Zoom to Address Select Site Go to Site# Print Help

Jump To City:

Map Layers Information

Zoom to Address
House number and street name are required. City and Zip are optional. Input address will be geocoded to nearest location.

Address:

City:

Zip Code:

Or search by Intersection. Both street names are required for an intersection search.

Street 1:

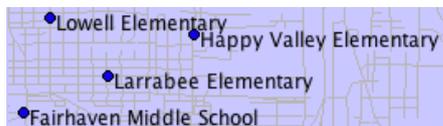
Street 2:

Or search by Point of Interest / Landmark name, such as a park, school, or lake.

Place Name:

200 m 1000 ft

GeoMOOSE 2.2 X,Y: 459441.09962, 4948647.73299 Lat, Lon: 44.690, -93.512 USNG: 15TVK5944148647 Scale: 1:4951



Postal Address Geo-Coder



Baldwin County Public Schools

Baldwin County Public Schools
Building Excellence

Jump To:

Map Layers Results FAQ-Help

Find An Address

Enter Street Number:

Enter Street Name:

Cancel Go!

100000 ft
20 mi

Powered by GeoMOOSE 2.2 X,Y: 2198657.69840, 419026.25735 Scale: 1:788368

Anemoi Analytics Prototype

Search **Go** or **Locate**

Info | **Map**

- ▶ Help
- ▶ Neighborhood Summary
- ▶ Schools and Education
- ▶ Housing
- ▼ Demographics

Sex

Female: 45% and Male: 55%

Age

Age Distribution

This Area	Age Group	Total US
4	80 and Over	4
4	70 to 79	5
6	60 to 69	8
9	50 to 59	13
14	40 to 49	15
31	30 to 39	13
27	20 to 29	14
6	10 to 19	14
4	Under 10	14

US Census Bureau, 2009 ACS Figure © Anemoi Analytics

Race and Ethnicity

Hispanic or Latino: 9% Middle

Navigate | **Measure** | **Identify** | |

-122.453 37.773

Ashbury Street / Haight Street, San Francisco, California
Address is approximate

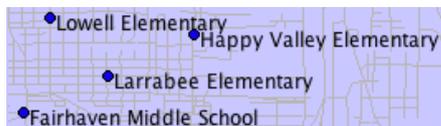
Haight Street

Ashbury Street

Haight Street

Ashbury Street

Google © 2011 Google - Terms of Use



Postal Address Geo-Coder



MetroGIS Geocoder for ArcGIS



MetroGIS Geocoder Tool

Geocoder Tool for ArcMap

The MetroGIS Geocoder Tool for ArcMap was created so users could easily use the MetroGIS Geocoder Web Service within ESRI's ArcMap 9.3 Software. The tool allows users to enter an address, geocode it using the MetroGIS Geocoder Web Service, display the results and zoom ArcMap to each selected result. This tool is currently only compatible with ArcMap 9.3.

The screenshot shows a software window titled "MetroGIS Geocoder Service". It has a text input field containing "600 Country Trail East, Jordan, 55352" and a "Locate Address" button. Below the input field is an "Address example: 600 Country Trail East, Jordan, 55352" and a "Click for Help" link. A table displays the geocoding results with columns for Zoom To, Address, Place, Zip, Lat, and Long.

	Zoom To	Address	Place	Zip	Lat	Long
▶	Zoom To	600 COUNTRY TRAIL	SPRING LAKE T...	55352	44.688341	-93.5077
	Zoom To	1125 COUNTRY TRAIL	SAND CREEK T...	55352	44.679264	-93.5454
	Zoom To	475 COUNTRY TRAIL	SAND CREEK T...	55352	44.677686	-93.5288
	Zoom To	1225 COUNTRY TRAIL	SAND CREEK T...	55352	44.680518	-93.5460
	Zoom To	2250 COUNTRY TRAIL	SAND CREEK T...	55352	44.68325	-93.5694

Thanks to Our Sponsors



MetroGIS



Sharing Information Across Boundaries.

Serving the Minneapolis/St. Paul Metropolitan Area [Home](#) | [Search](#) | [Contact Us](#)

About MetroGIS

- [What is MetroGIS?](#)
- [Deliverables & Outcomes](#)
- [Business Planning](#)
- [What's New](#)
- [Major Accomplishments](#)
- [Annual Reports](#)
- [Affiliations](#)
- [Awards](#)
- [Grants](#)
- [History](#)

Benefits

- [Overview](#)
- [Short Quotes](#)
- [Testimonials](#)
- [Studies](#)
- [Performance Measurement](#)

GIS Data and Applications

- [Overview](#)

Applications > Geocoder

- [Address Look-up Made Easy - The MetroGIS Geocoder Project](#)
- [Using the Geocoder Web Service in your Application](#)
- [Build Your Own Geocoder Service](#)
- [Help Improve the Geocoder Service](#)
- [FAQ](#)
- [Sample](#)

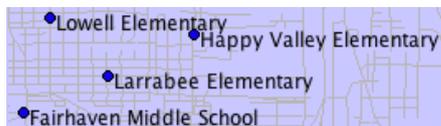
Address Look-up Made Easy – The MetroGIS Geocoder Project [top of page](#)

Do you need address spatial look-up in your application?

- MetroGIS has sponsored development of a Web Service that takes Address or Intersection requests and returns closest matches with their latitude-longitude.
- Use this service to add geocoding to a web map or any application that needs spatial info for addresses, without having to store your own local copy of the underlying data.
Look-up by:
 - street address
 - intersection (2 street names)
 - landmark name

- MetroGIS and the Metropolitan Mosquito Control District

- Social Sciences and Humanities Council of Canada



Postal Address Geo-Coder



Questions?

The PAGC Project

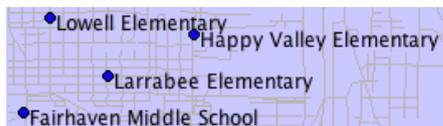
<http://www.pagcgeo.org>

Anemoi Analytics

<http://www.anemoianalytics.com>

iMaptools

<http://www.imaptools.com>



Postal Address Geo-Coder

