Introduction to Neogeography

Andrew Turner

December 15, 2006

Abstract

Neogeography combines the complex techniques of cartography and GIS and places them within reach of users and developers.

This Short Cut introduces you to the growing number of tools, frameworks, and resources available that make it easy to create maps and share the locations of your interests and history.

Learn what existing and emerging standards such as GeoRSS, KML, and Microformats mean; how to add dynamic maps and locations to your web

site; how to pinpoint the locations of your online visitors; how to create genealogical maps and Google Earth animations of your family's ancestry; or how to geotag and share your travel photographs.

This Short Cut introduces current techniques of online mapping and neogeography. It covers the basic tools for finding out where you are, creating your own maps, and several illustrative projects to get you started and hopefully inspire your own projects.

The code and software examples from this book are available online at **http://mapsomething.com**. This includes a complete listing of the links to sites and resources listed here for easy bookmarking. Feel free to let me know if you have any questions, or want to show off any neat ideas you put together, directly at **shortcut@highearthorbit.com**.

What Is Neogeography?

Cartography enabled and recorded exploration and discovery for ages. It guided sailors across oceans and helped make sense of the wilderness. Like the transformation of paper to hypertext, maps have evolved from arcane lines and arcs on parchment to dynamic dis-

plays of remote geographic information. While geographic information systems (GIS) remained expensive programs, restricted to the use of highly-trained specialists, tools like MapQuest and Yahoo! Maps brought easy-to-use mapping tools to the public. More recently, the release of Google Maps demonstrated to web developers and users the possibilities of navigation and opened a floodgate of interest in online mapping.

Google Maps was not released with an *application programming interface* (API), but developers slowly figured out how to use the maps for their own uses, and eventually Google released a public API. The release of the API allowed developers and users to quickly and easily show geographically based data on shareable maps. Once the primary domain of GIS systems and operators, these new maps, or "mashups," led to a massive increase in their propagation. The very term *mashup* can trace its lineage to the first mapping mashups like Housing Maps and ChicagoCrime. What is it about maps that both fascinates and educates people? Why in the online world of instantaneous global communication and anonymity has "location" become such an important theme?

Neogeography means "new geography" and consists of a set of techniques and tools that fall outside the realm of tradi-

tional GIS, Geographic Information Systems. Where historically a professional cartographer might use ArcGIS, talk of Mercator versus Mollweide projections, and resolve land area disputes, a neogeographer uses a mapping API like Google Maps, talks about GPX versus KML, and geotags his photos to make a map of his summer vacation.

Essentially, Neogeography is about people using and creating their own maps, on their own terms and by combining elements of an existing toolset. Neogeography is about sharing location information with friends and visitors, helping shape context, and conveying understanding through knowledge of place.

Lastly, neogeography is fun. Why else would people create a map of the locations of the television show *24*, or share the location of their first kiss? Never again will you struggle to recall "Where was that photo taken?"

Basic Terminology

Several useful terms are essential to discussing neogeography. Some are familiar, while others have specific contexts within a broader meaning. They are presented here for later referral as you progress through the book and are listed in a logical order so that terms build upon one another:

Coordinates

The geographic coordinates are the absolute position on the Earth (or any body). Typically these are Latitude and Longitude referenced to the WGS84 ellipsoid. Sometimes they will be referred to as "lat" and "lon." Latitude varies North-South, with 0 degrees at the Equator, varies from 90 to −90 degrees towards the poles, and is positive North. Longitude varies East-West, with 0 degrees at the Prime Meridian, varies from 180 to −180 degrees and is positive to East.

Coordinates can be represented in several formats:

- Decimal degrees (DD): 29.975
- Degrees-Minutes-Seconds (DMS): N29° 58' 30''
- Degrees-Minutes (DM): 29° 58.8'

The conversion is straightforward. There are 60 seconds in a minute, and 60 minutes in a degree.

Projection

A projection is required in order to display the three-dimensional Earth, which isn't a sphere but actually an oblate spheroid bulging in the center, onto any other shape. Typically, this is projection onto a two-dimensional map display, where Mercator and Rectangular (i.e., no transformation) projections are the most common. For

example, Google Maps uses the Mercator Projection, which is good for zoomed-in viewing, but causes distortions when zoomed out. It is important to understand the implication of various projections depending on application and also when mixing together mapping providers.

POI (Points of Interest)

Points Of Interest, frequently abbreviated as POI, represents any significant locations such as public buildings, traveler's services, or user-defined waypoints. These may be categorized: restaurant, trail head, friend's house, scenic overlook, or scuba diving site. There is also AOI, Areas of Interest, which may include multiple POI or just a geographic area instead of a single point.

Extents

The bounding box, or farthest latitude and longitude of an AOI. This may also be referred to as just *bbox*. The extents define the Northern and Southern latitude, and the Easter and Western longitude. Extents are the simplest means of specifying the area of interest and are usually used for web service queries.

Tiles

Dynamic, or *slippy* maps, are composed of a set of individual square images. Each image is a tile. Together,

these tiles are placed next to one another, or stitched, to give the impression of a large, sliding map.

Geolocation

The technique of automatically determining the position of something based on measured data. For example, it is possible to locate a computer given its IP address, or a mobile phone based on the observed cell towers. Geolocation is useful for determining where a user or device is without the user having to manually enter this information. GPS is a specific implementation of a geolocation technology.

GPS

GPS (Global Positioning System) really refers to the U.S. military owned and operated satellite network that provides three-dimensional location. GPS is also sometimes used to refer to any means of geolocation that provides geographic coordinates.

Simply put, GPS operates by a network of high-altitude space satellite broadcasting their position and time. Receivers use several of these observed broadcasts to determine its current position and time.

The European operated GPS system, Galileo, is currently expected to be operational by 2010 and will provide

similar, but alternate, functionality to the current GPS.

Geotag

Adding location information to a document, photograph, audio sample, or some other type of data is an example of geotagging. Geotagging formats are not uniform and vary based on the type of document they modify. For example, a photograph can embed location information in the EXIF header of the file itself, or in many web applications the location information can be stored as triple-tags in the user specified tags.

Web Service

A web service is a resource that allows access to data or functionality from a provider—for example, a geocoder that takes an address and returns latitude and longitude is a web service, or a service to ask for all photos within 50 miles of a location. Web services typically use REST (Representational State Transfer) or SOAP (Simple Object Access Protocol) to allow programs and other sites to access the data.

Data Formats

Location can be stored in many formats. With roots in programming, neogeography pushes data to be stored in a clear and readable "plain language" format. Because broadband allows data to be

conveyed quickly, more geographic information can be added to a greater variety of files.

While these formats have adopted a more human-friendly format, they can quickly become large, and complex to write by hand. Therefore, in future sections we'll discuss tools that make it easy to add and read these formats to your projects. For now, this section will introduce you to the overall concepts of the data formats, when it's appropriate to use them, and what their capabilities are.

GPX

Vendors and tools use several standards to describe geographic information. The most common file format standard is GPX, or the GPS Exchange Format. This standard uses an XML file definition to store waypoints and tracks from GPS units. Most GPS receivers internally use their own proprietary file format. Therefore, GPX is a common protocol that allows developers to write tools to convert from and to an application's or device's specific format.

The header of a GPX file stores general information such as the GPX version, the application or device that created the file, and various other XML namespace information:

```
<?xml version="1.0" encoding="UTF-8"?>
<gpx version="1.0"
       creator="GPSBabel - http://www.gpsbabel.org"
       xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
       xmlns="http://www.topografix.com/GPX/1/0"
       xsi:schemaLocation="http://www.topografix.com/GPX/1/0
       http://www.topografix.com/GPX/1/0/gpx.xsd">
```

Below that, there are lists of waypoints, tracks, or routes.

A *waypoint* (`wpt` in the file example below) stores the latitude, longitude and optionally the altitude of the location. The body of the waypoint then includes the name, comment, and symbol:

```
     <wpt lat ="42.277881" lon ="-83.740791">
         <name>01WiFi</name>
         <cmt>ESPRESSO ROYAL</cmt>
         <sym>Tall Tower</sym>
     </wpt>
```

A *track* (`trk` in the file example below) is a successive list of waypoints and usually include a time element. This is useful later for geotagging photos or videos, as well as for creating animations of your tracks by providing a time to synchronize with the camera or recording device:

```
<bounds minlat="42.423656691" minlon ="-83.493977330"
             maxlat="42.502436669" maxlon="-83.144850082" />
<trk>
       <trkseg>
             <trkpt lat="42.500937812" lon="-83.147198063">
```

```
                    <ele>-0.846357</ele>
                    <time>2006-06-26T21:42:38Z</time>
             </trkpt>
             <trkpt lat="42.500988805" lon="-83.147192333">
                    <ele>-1.071942</ele>
                    <time>2006-06-26T21:42:40Z</time>
             </trkpt>
             <trkpt lat="42.500985367" lon="-83.147242180">
                    <ele>-0.691822</ele>
                    <time>2006-06-26T21:42:42Z</time>
             </trkpt>
      </trkseg>
</trk>
```

The Topografix GPX site (**http://www.topografix.com/gpx_resources.asp**) has a list of GPX data as well as additional resources and utilities.

GeoRSS

GeoRSS (**http://georss.org**) is an extension to the common RSS (Really Simple Syndication) used on web sites to notify readers of new articles or updates. GeoRSS adds geographic coordinates and features to RSS and Atom items.

GeoRSS comes in several flavors: W3C Geo, Simple, and GML. Simple was developed to speed the adoption and use of GeoRSS by providing an uncomplicated format that is sufficient for making points, lines, and polygons. W3C Geo exists based on historical use and is limited to just describing single points. As the need for complex geometry become necessary, GML (Geographic Markup

Language) provides the ability to describe complex geographic geometry.

To add GeoRSS Simple to your RSS, you need to first add the GeoRSS namespace reference to your feed definition:

```
<?xml version="1.0" encoding="UTF-8"?>
<rss version="2.0"
       xmlns:content="http://purl.org/rss/1.0/modules/content/"
       xmlns:wfw="http://wellformedweb.org/CommentAPI/"
       xmlns:dc="http://purl.org/dc/elements/1.1/"
       xmlns:georss="http://www.georss.org/georss"
       xmlns:gml="http://www.opengis.net/gml"
>
```

Then, in your RSS entries, you actually add the *georss:point* tag. The most important thing to note is that the latitude and longitude are separated by a single space, and not a comma as is typical in other data formats.

The format of the GeoRSS point tag is:

```
      <georss:point>latitude longitude</georss:point>
```

In an actual RSS entry it would look like:

```
   <entry>
      <title>M 3.2, Mona Passage</title>
      <link href="http://example.org/2005/09/09/atom01"/>
      <id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
      <updated>2005-08-17T07:02:32Z</updated>
      <summary>We just had a big one.</summary>
```

```
        <georss:point>45.256 -71.92</georss:point>
    </entry>
```

As necessary, you can add `radius=""` or `elev=""` attributes to the `georss:` element. Both are assumed to be measured in meters. The radius attribute is useful for giving a scale to the described feature, and elevation is altitude above the WGS84 ellipsoid (typically supplied by a GPS receiver).

In addition to just a geographic definition, GeoRSS uses two optional elements: *featuretypetag* and *relationshiptag* to provide additional metadata to the geographic element.

When noting an element such as a mountain, lake, or residence, the *featuretypetag* is used. The *relationshiptag* relays the geographic coordinates to the actual entity in the RSS item, such as *is-centered-at*:

```
        <georss:point>45.256 -71.92</georss:point>
        <georss:relationship>is-centered-at</georss:relationship>
        <georss:featuretypetag>residential house</<georss:featuretypetag>
```

As mentioned previously, GeoRSS Simple can support more complex geometry such as a line. This may be useful for defining a road or hiking trail:

```
            <georss:line>lat1 lon1 lat2 lon2 lat3 lon3</georss:line>
```

```
        <georss:line>45.256 -110.45 46.46 -109.48 43.84 -109.86</georss:line>
```

A box is useful for defining the extents of an area of interest, or the location of a building:

```
<georss:box>lower_left_lat lower_left_lon upper_right_lat upper_right_lon</georss:box>

            <georss:box>42.943 -71.032 43.039 -69.856</georss:box>
```

or a polygon, which can be any number of sides, and is defined by a set of latitude and longitude points describing the outside boundary of the polygon. The last set of latitude, longitude must equal the first set so that the polygon is closed:

```
<georss:polygon>lat1 lon1 lat2 lon2 lat3 lon3 lat1 lon1</georss:polygon>

<georss:polygon>
       45.256 -110.45 46.46 -109.48 43.84 -109.86 45.256 -110.45
</georss:polygon>
```

As the need for more complex geographic geometry arises, it is worth checking out GML (Geography Markup Language) and how to use it in GeoRSS (**http://www.georss.org/gml.html**.

For a list of GeoRSS sources, check out:

- Mapufacture: **http://mapufacture.com**
- PlaceDB: **http://placedb.org**

KML

Another XML format is KML (Keyhole Markup Language) (**http://earth.google.com/kml**), which was developed by Keyhole Technologies and is now owned by Google. Keyhole created the KML file format, Google Earth's predecessor. Since acquiring Keyhole, Google greatly extended the use and capability of KML, which can now define three-dimensional geometry for creating geospecific buildings.

Similar to GeoRSS, KML's major difference is coordinate order, which is longitude, latitude:

```
<Point>
  <coordinates>-90.8694,48.2545</coordinates>
</Point>
```

Like GeoRSS, KML can also define complex geometries:

```
<Polygon id="ID">
  <!-- specific to Polygon -->
  <extrude>0</extrude>                    <!-- boolean -->
  <tessellate>0</tessellate>              <!-- boolean -->
  <altitudeMode>clampToGround</altitudeMode>
    <!-- kml:altitudeModeEnum: clampToGround, relativeToGround, or absolute -->
  <outerBoundaryIs>
    <LinearRing>
      <coordinates>...</coordinates>      <!-- lon,lat[,alt] -->
    </LinearRing>
  </outerBoundaryIs>
  <innerBoundaryIs>
    <LinearRing>
```

```
        <coordinates>...</coordinates>          <!-- lon,lat[,alt] -->
      </LinearRing>
    </innerBoundary>
```

While you would probably never write KML yourself, many tools exist to generate KML from other file formats, or via a drawing program.

One unique feature to KML is the ability to embed 3-D visualization models into the KML file. This is how users are insert buildings, objects, and geographic annotations (Sticky Notes for maps) within their files.

The Keyhole BBS (Bulletin Board System) (**http://bbs.keyhole.com**), also known as the Google Earth Community, is an active group of users and developers sharing KML files, how-tos, and ideas, and serves as a clearinghouse for KML information.

Microformats

Microformats are used in web pages to identify common data such as people, places, or events. When browsers show a web page, they are translating the web page's Hypertext Markup Language (HTML) for graphical rendering. The actual words or information presented are immaterial to the browser. Microformats add meaning to the HTML by providing a standardized schema applied to the class and ID of HTML attributes, permitting

manipulation of this information by other programs.

For example, the hCard Microformat defines how to mark up contact information for a person. Once the information is marked up in a common, meaningful way, other tools can then recognize that information and add the contact to your address book, find her social networking accounts (such as LinkedIn, Plaxo, or even MySpace), send har an email, or instant message her.

For Neogeography there are currently two interesting Microformats: *adr* and *geo*. *adr* is the definition of an address:

```
        <div class="adr">
         <div class="street-address">23 Main St.</div>
         <div class="extended-address">Suite 104</div>
         <span class="locality">Northville</span>,
         <span class="region">MI</span>
         <span class="postal-code">48167</span>
         <div class="country-name">U.S.A.</div>
        </div>
```

geo defines a geographic coordinate in latitude and longitude:

```
<span class="geo">
        <span class="latitude">42.4266</span>,
        <span class="longitude">-83.4931</span>
       </span>
```

Either of these formats can also be embedded in another Microformat in order to give that broader Microformat geographic context. For example, putting an *adr* in an hCal event would tell you where that event was taking place:

```
<span class="vevent">
 <a class="url" href="http://conferences.oreillynet.com/where/">
  <span class="summary">Where 2.0 Conference</span>:
  <abbr class="dtstart" title="2007-5-29">May 29</abbr>-
  <abbr class="dtend" title="2007-10-31">30</abbr>,
 at the <span class="location adr">Fairmont Hotel,
       <span class="locality">San Jose</span>,<span class="region">CA</span>
       </span>
 </a>
</span>
```

Technorati (**http://technorati.com**) is the primary supporter of Microformats and provides numerous tools and aggregators for them.

For more information on various Microformats, how to use them, and how to help develop new ones, check out Brian Suda's O'Reilly Short Cut "Using Microformats."

Advanced formats

So far, we've just discussed emerging and entry-level technologies and formats. There are many more full-featured, but also more complex, standards such as WMS, WKT, and GML that provide very high-fidelity data representations and can handle complex geometry.

They are key components in the underlying geospatial technology, but are beyond this scope of this book. I suggest following up with *Web Mapping Illustrated* by Tyler Mitchell (O'Reilly) for more in-depth information on these advanced concepts.

Right formats for the right job

Each of the formats above have different strengths and weaknesses. They can be used together or separately depending on your application and needs.

GPX is primarily a file format for working with GPS units. It lacks features for better annotation and describing actual geometries. This should be used for archiving and sharing data with devices.

GeoRSS is best used for sharing geographic updates and individual items since it is limited to describing a single RSS item such as a blog post, sensor reading, story, photograph, or the tracking of a moving object. GeoRSS cannot currently handle groups or sets of points or geometries.

KML is very well suited for storing and displaying sets of data such as favorite locations, plot areas, and geospecific buildings.

Microformats should be limited to displayed data that will be read by a human. It doesn't serve as well as a ma-

chine-only format, and without an hCard or hEvent, there isn't currently a way to associate an *adr* or *geo* with its contextual information.

Software Tools

Now that you know what types of data are useful, it is important to identify the various tools you'll need to work with the data. Using a text editor is always an option, but only rarely the best solution. Instead, find a good program that supports the various data formats you will be using. If something is missing, try adding it yourself and sharing it back to the neogeography community.

Arguably the most useful tool you'll need in neogeography is GPSBabel (**http://www.gpsbabel.org**). It is a cross-platform application that can convert to and from a large number of geographic data formats. GPSBabel can also upload or download data directly with your GPS unit. This might include support tracks and routes in addition to simple waypoints.

Because GPSBabel is a command-line utility, some users may prefer a graphical user interface (GUI). Users of most platforms can find simple GUIs widely available. In addition to converting between formats, GPSBabel can also apply filters to the data as it is converted. Filters are useful for splitting long tracks

into shorter ones where a long time or distance gap exists between points, such as when you have multi-day trips, or take a long break and don't want to store the many GPS points at a single location. Also, you can apply filters to optionally include only waypoints with a specific radius, to smooth tracks, or to remove high-error points from the track log. These filters are useful for finding POI within a specific area, or removing GPS points that have high error due to poor satellite signals.

For easy GPX file management, sorting, and organizing waypoints, Windows users can use Easy GPS (**http://www.easygps.com**), and Mac users should check out MacSimpleGPS (**http://www.macsimplegps.com**).

Currently, most RSS feeds do not contain the GeoRSS extensions, but they may contain useful geographic information. Geonames' RSS-to-GeoRSS converter (**http://www.geonames.org/rss-to-georss-converter.html**) is a great tool for parsing and converting these RSS feeds to GeoRSS ones that can be used in your applications.

Google Earth (**http://earth.google.com**) is a 3-D geographic visualization tool that uses KML for locations, geometry, overlays, and histories. Google Earth also understands network links to remote KML sources, which is very useful for data

that is continuously updating, such as weather tracking.

WorldWind (**http://worldwind.arc.nasa.gov/**) is another 3-D geographic tool developed by NASA for scientific visualization and uses the beautiful satellite imagery for their tiles.

These are a few of the tools that will ensure a productive and fun time, with new tools being developed every day. Subscribe to bookmarking feeds such as **http://del.icio.us/tag/geo** or **http://del.icio.us/inbox/starhill_blend**, or join one of the communities mentioned in the resources section at the end of the book to find new sites and tools as the show up.

Hardware Tools

Neogeography does not have specific "tools of the trade" so much as repurposing devices you already use in a geo-centric way. For example, you may want to use a camera to take photographs of places that you visit. An audio recorder can capture the noises or music of an area. Laptops, handheld computers, and mobile phones all assist in locating, tagging, annotating and viewing your creations. But none of these tools were necessarily designed with "geography" as their purpose. Be creative and find new uses for common devices.

GPS Units

GPS units have become prolific and relatively inexpensive. You can purchase a simple USB GPS unit to attach to your laptop for as little as $30 USD.

There are a couple of major factors to consider when choosing a GPS device:

Display
Do you want to use the unit to display maps, tracks, and information?

Logging
Some GPS units have built-in logging, so that when you turn the unit on it starts storing your location as tracks.

Bluetooth
Wireless connection is useful if you want to use the unit with a mobile phone, handheld computer, or just not deal with cables.

Based on these factors, there are four major types of GPS units: Handheld, Vehicle, Logger, and Puck or GPS Mouse. A Handheld GPS unit is a full-featured device with map display, routing, tracking, and a user interface. It can be used independently of any other devices and is the most useful when out in the field. Good Handheld units include the Garmin Geko or Garmin 60CSx

Vehicle GPS units are very similar to handheld units, except that in-vehicle units typically are larger, have car mounts, and assume your position is along roadways. Many cars now come with navigation systems built-in. The benefit of a built-in navigation system is that there are typically additional sensors, such as wheel speed sensors, magnetometers, and accelerometers that aid in navigation when pure GPS location is unavailable, such as in tunnels. However, these built-in systems are also not upgradeable and do not allow for as much modification and uploading of new data. Examples of vehicle GPS units include the Garmin Nüvi and the Magellan RoadMate.

Figure 1. GPS receivers come in different shapes in sizes: GPS logger (top left), handheld GPS (right), vehicle nav system (bottom left)

A Logger GPS is a small, simple device with no connections or mapping display. Typically, you just turn the device on, and it starts storing locations to a track. Therefore, you can toss it into your bag and forget about it. When you have com-

pleted your trekking, you can then connect the GPS unit to a computer or other device and download the track. Additionally, many GPS Loggers have Bluetooth, so if you want to display your current location, you can connect a mobile phone via Bluetooth to the GPS and display maps, current position, satellites, and so on. Good loggers include the Wintec WBT-200, NaviGPS, or the Delorme Bluelogger.

A Puck, or GPS Mouse is a simple GPS receiver with a cable for connecting to a computer to actually read and store the positioning data. These are the simplest devices and also don't cost very much. They have also gotten very small and easy to interface with embedded devices, such as microcontrollers, or radio controlled vehicles.

Check out the OpenStreetMap GPS reviews page for good discussion on GPS units: **http://wiki.openstreetmap.org/index.php/GPS_Reviews**.

Cameras

There are very few digital cameras that feature actual GPS or geolocation hardware. The Ricoh Pro G3 was the first widely available camera to have an onboard GPS sensor and automatic geotagging of your photographs. The Eye-Fi Eye-Film SD memory card should be available in 2007 and features WiFi connectivity for automatically uploading

and geotagging your photos when you are near a wireless network. By embedding the geolocation technology in the common memory card, any camera can gain geolocation ability of photographs.

In general, any camera can be used for Neogeography projects. Later on we will discuss how to automatically and manually add geotags to your photos to locate and map them. You can also use print film that is scanned or any kind of picture.

Other hardware

Neogeographers experiment with a lot of other devices in their projects.

Mobile phones can be used to geolocate people by recording the observed cell tower data. The Boost Mobile and Nextel mobile phones provide an interface for developers to use the geographic location in their applications. Symbian OS mobiles also provide a limited interface to the cell location information.

Handheld computers and tablets such as the Nokia 770 (**http://nokia.com/770**) are being used by neogeographers to display maps using Maemo Mapper (**https://garage.maemo.org/projects/maemo-mapper/**) and GPS. Also, handheld devices can record audio, video, and text for geotagging and uploading or displaying.

Apple and Nike released a Sport Kit (**http://www.apple.com/ipod/nike/**), which is an electronic pedometer that controls your iPod music player based on the distance you have run. Neogeographers are experimenting with using this device for track distance measurement and location augmentation.

Björn Hartmann's "GPS + Google Maps Mash-up in 42 Lines of Code" (**http://regexp.bjoern.org/archives/000186.html**) is a great example of a quick but slick hack for building your own GPS device and using the data.

Finally, geolocation components and technology are continuing to drop in price, size, and power requirements. Therefore, future generations of devices will begin to have GPS built-in as a standard features.

GEOPARSING ✎

Batch geocoding usually relies on us having a fairly well organized list of locations. However, it is more likely we have documents or web sites we want to get the locations from. Perhaps it is a document analyzing various real estate options in the area or a news web site.

GeoParsing is the process of converting a natural language document (free of specific markup) and extracting geographic locations and their contexts. For an excellent example, check out GutenKarte (**http://gutenkarte.org/**), which parses classic literature and generates maps. You can see an automatically generated map of Verne's *Around the World in Eighty Days*, or Tolstoy's *War and Peace*.

MetaCarta Labs provides a GeoParser API. Using this API, you can GeoParse your own documents or web sites and generate maps, images, or feeds.

See:

http://labs.metacarta.com/GeoParser/documentation.html
