Rich Pinder
Los Angeles Cancer Surveillance Program
Links and Slides – google ‘Rich Pinder USC’

rpinder@usc.edu

NAACCR Annual Meeting
Incorporating realtime geocoding systems into data entry routines

Presented at the NAACCR Annual Meeting
June 2007

Detroit, June 2007
Data entry – main focus here

- Increase accuracy
- Increase ease/speed
- Zip codes & Area Codes & Postal info
- Standardize CityName (bulk mailings)
An example from the past:
MARY BO PEEP
123 FARMYARD LANE
VAN NUYS, CA 91411-

AddValDate 6/4/2007 03

Detroit, June 2007
- Zip code based data
  - Postal code info; Area codes; lat/long (of the Post Office);
    Population, by race; Household count; Income; House value

- Subscription – updated monthly
- $150/yr
- US file ~ 9mb
http based Geocoding

- Various Accuracy/Completeness options
  - Census Tiger → NavTeq & TeleAtlas, Enhanced
  → Assessor Parcel data
- Popular interfaces to mapping resources
  - Yahoo API / Google API
I like to begin all software programs with a common set of predefined modules, units, components and development tools……
I like to begin all software programs with a common set of predefined modules, units, components and development tools……
Ajax (also known as AJAX), shorthand for "Asynchronous JavaScript and XML", is a development technique used for creating interactive web applications. The intent is to make web pages feel more responsive by exchanging small amounts of data with the server behind the scenes, so that the entire web page does not have to be reloaded each time the user requests a change. This is intended to increase the web page's interactivity, speed, functionality, and usability. Ajax is also usable on many operating systems and architectures as it's based on JavaScript and XML.
Ajax (also known as AJAX), shorthand for "Asynchronous JavaScript and XML", is a development technique used for creating interactive web applications. The intent is to make web pages feel more responsive by exchanging small amounts of data with the server behind the scenes, so that the entire web page does not have to be reloaded each time the user requests a change. This is intended to increase the web page's interactivity, speed, functionality, and usability. Ajax is also usable on many operating systems and architectures as it's based on JavaScript and XML.

XMLHttpRequest (XHR) is an API that can be used by JavaScript, and other web browser scripting languages to transfer XML and other text data to and from a web server using HTTP, by establishing an independent communication channel between a web page's Client-Side and Server-Side.
Leaving Ajax now……
Web services examples....
$php
function request_cache($url, $dest_file, $timeout=43200) {
    if(!file_exists($dest_file) || filemtime($dest_file) < (time()-$timeout)) {
        $stream = fopen($url,'r');
        $tmpf = tmpnam('/tmp','YWS');
        file_put_contents($tmpf, $stream);
        fclose($stream);
        rename($tmpf, $dest_file);
    }
}

function yahoo_geo($location) {
    $q = 'http://api.local.yahoo.com/MapService/V1/geocode';
    $q .= '?appid=plerdorfo&location=1'.rawurlencode($location);
    $tmp = '/tmp/yws_geo_'.md5($q);
    request_cache($q, $tmp, 43200);
    libxml_use_internal_errors(true);
    $xml = simplexml_load_file($tmp);
    $ret['precision'] = (string)$xml->Result['precision'];
    foreach($xml->Result->children() as $key=>$val) {
        if(strlen($val)) $ret[(string)$key] = (string)$val;
    }
    return $ret;
}
?>
Server PHP Script

```php
<html>
<head>
<title>GeoCoding API Example</title>
</head>
<body>
<form action="/php/ymap/geo1.php" method="GET">
<input type="text" size="80" name="location" />
</form>

<?php
include '../geo.inc';
if(!empty($_REQUEST['location']) ) {
    $a = yahoo_geo($_REQUEST['location']);
    echo "<pre>
        print_r($a);
    </pre>
";
}
?>
</body></html>
```
Beverly Hills, CA

Array
{
    [precision] => city
    [Latitude] => 34.074169
    [Longitude] => -118.399292
    [City] => BEVERLY HILLS
    [State] => CA
    [Country] => US
}
Array
{
    [precision] => street
    [Latitude] => 34.06684
    [Longitude] => -118.38274
    [Address] => 8750 WILSHIRE BLVD
    [City] => BEVERLY HILLS
    [State] => CA
    [Zip] => 90211
    [Country] => US
}
NAACCR GeoCoding API Example - Mozilla Firefox

Array
{
    [precision] => address
    [Latitude] => 34.066963
    [Longitude] => -118.387864
    [Address] => 9000 WILSHIRE BLVD
    [City] => BEVERLY HILLS
    [State] => CA
    [Zip] => 90211-1809
    [Country] => US
}
The White House

Array
{
    [precision] => address
    [Latitude] => 38.897610
    [Longitude] => -77.036377
    [Address] => The White House
    [City] => Washington
    [State] => DC
    [Zip] => 20006
    [Country] => US
}
Detroit, June 2007
Future Options:

• Use validated info for data entry only
• Same, but capture Lat/Long data
• Same, but build (acquire?) a second webservice, to use Lat/Long to store Census Tract
• Same, but customize to include additional census data to enhance record.
links

Google API & Geocoding Tutorial
Google & Bulk Geocoding
Zipcode Database - monthly updated subscription
Yahoo API & Ajax – mapping
Great GoogleEarth / GIS site !
Census Tract boundary files - from US Census
Incorporating realtime geocoding systems into data entry routines
Regional registries routinely use GIS batch processing software to append geocoded location information to the cancer record. With the availability of accurate reference data files, and the ubiquitous use of high speed internet connections, could we consider automating the geocoding step up front? Perhaps as routine data entry steps occur, a tool could geocode (and validate) the information being added in real time.
The presenter will describe existing sources for address based reference file lookup and the steps involved to interface to them. At the time of writing, the two we will examine are the Google Maps API and the Geocoder.US web based interface. Then the presenter will discuss feasibility of developing a standalone hardware and software solution to provide this functionality directly.

Detroit, June 2007