



Esri International User Conference | San Diego, CA
Technical Workshops | July 11 – 15, 2011

Geocoding - Advanced Techniques

Agatha Wong

Brad Niemand

Sergey Ivanenko

Outline

- **Workflows on improving matching results**
- **Overview of the geocoding engine and matching process**
- **Fine tuning locators**
- **Road ahead for geocoding**

What This Workshop Does Not Cover

- Basic geocoding process
 - Covered in the *Geocoding: An Introduction* Technical Workshop
 - Programming with ArcObjects, Web APIs
 - Address data model and tools
- * Please meet with the geocoding team at the Spatial Analysis Island to discuss about these topics or issues you may have.

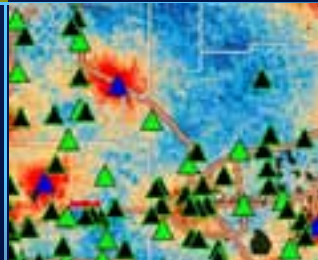
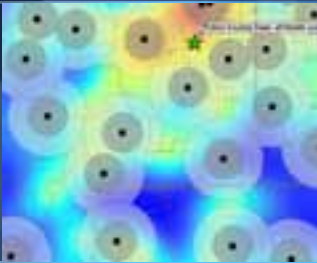
Improving matching results

One of the key questions when working with geocoding:
How to match more addresses?

- Two workflows:
 - Create/manage composite locators
 - Create locators from manually matched addresses

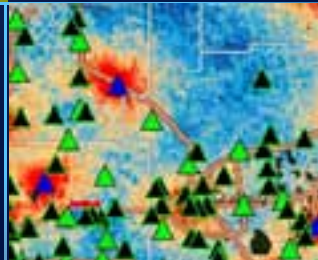
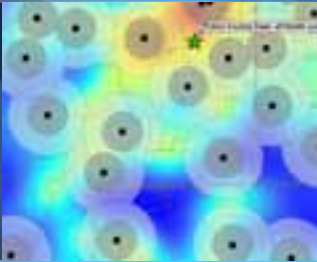
Demo: Composite Locators

- A workflow using Geoprocessing model



Demo: Locators from Manual Matches

- Create point address locators from manual matches



Create locators from manual matches

Assumption: manual-pick addresses contain the true location and address information

Location

Match type

Address

The screenshot displays two windows from the ArcMap application. The top window, titled 'PointAddressesFromGeocodingResults.mxd - ArcMap - ArcInfo', shows a map of Redlands, California, with a red dot indicating a location. The bottom window, titled 'Interactive Rematch - Geocoding_Result *', shows a table of geocoding results. The table has columns: ObjectID, Shape, Status, Score, Match_type, Match_addr, Side, Ref_ID, User_ID, Addr_type, ARC_Street, and ARC_City. The first row is highlighted, showing a match for '209 N EUREKA ST, REDLANDS, CA 92374'. The 'Match_type' column is circled in red. Below the table, the 'Address' field is filled with '209 N EUREKA ST', 'REDLANDS', 'CA', and '92374'. The '56 Candidates' list shows various addresses in Redlands, California, with the top candidate being '209 N Eureka St, Redlands, California, 92374'. The 'Candidate details' section shows the address '209 N Eureka St, Redlands, California, 92374' with a score of 100.00. The 'Geocoding Options...' button is visible at the bottom left.

ObjectID	Shape	Status	Score	Match_type	Match_addr	Side	Ref_ID	User_ID	Addr_type	ARC_Street	ARC_City
10	Point	M	100.00	PP	209 N EUREKA ST		0			209 N EUREKA ST	REDLANDS

Address:

Street or Intersection: 209 N EUREKA ST
City or Place Name: REDLANDS
State: CA
ZIP Code: 92374

56 Candidates

Score	T	Side	Match_addr	FromAddr	ToAddr	PreDir
100.00	L	L	209 N Eureka St, Redlands, California, 92374	201	299	S
96.41	L	L	301 N Eureka St, Redlands, California, 92374	301	377	N
80.33	L	L	379 N Eureka St, Redlands, California, 92374	379	421	N
68.28	L	L	423 N Eureka St, Redlands, California, 92374	423	499	N
68.21	L	L	501 N Eureka St, Redlands, California, 92374	501	699	N
68.02	L	L	701 N Eureka St, Redlands, California, 92374	701	799	N
63.16	R	R	370 N Eureka St, Redlands, California, 92374	370	420	N

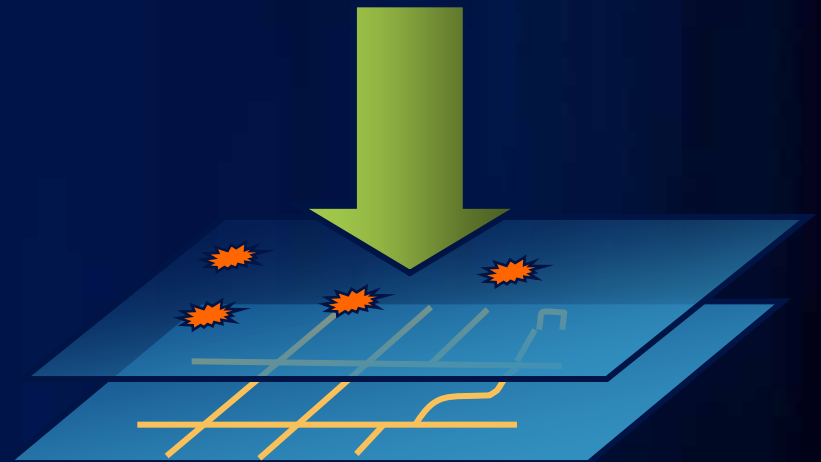
Candidate details:

FromAddr: 201
ToAddr: 299
PreDir: S
PostType: S
StreetName: Eureka
SufType: St
SufDir: S

Create locators from manual matches

Setup:

- **Composite locator consisting of**
 - **Point locator**
 - Built using address point feature class
 - Use **US Address - Single House** style
 - Set **Minimum Match Score** to a high value (95)
 - **Street centerline locator**



Create locators from manual matches

Workflow:

- After each rematch session
 - Select records from the geocoded feature class that are matched using the **Pick Address from Map** tool
 - Append records to the point feature class
 - Rebuild the **point** address locator

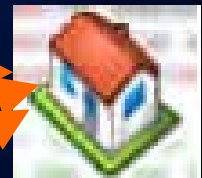
Table

Geocoding_Result_1

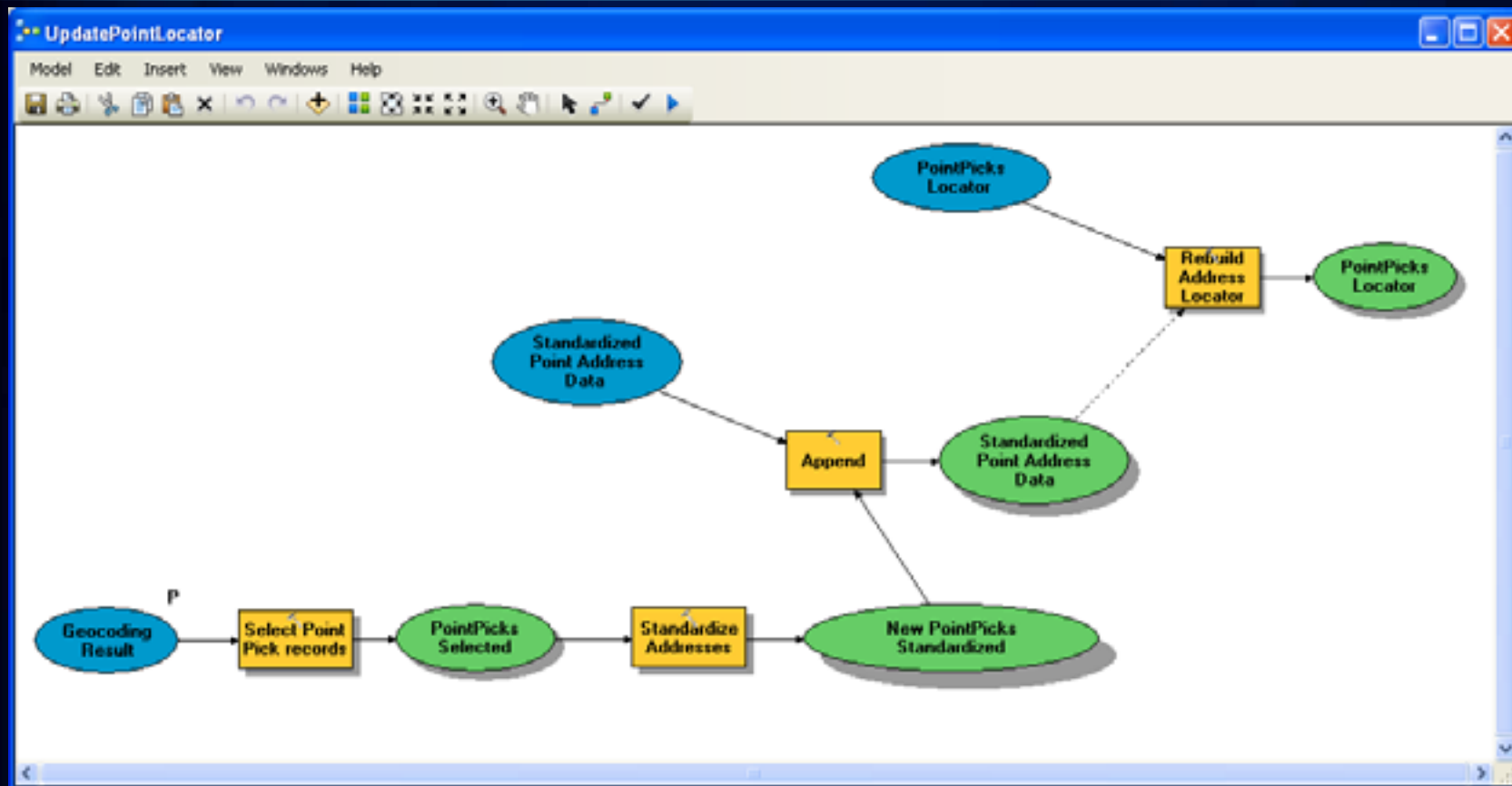
	Status	Score	Match_type	Match_addr	ARC_Street	
▶	M	100	PP		316 GRANT ST	REDL
	M	100	A	319 Grant St, Redlands, California, 92373	319 GRANT ST	REDL
	M	100	PP		5 FIRST ST N	REDL
	M	100	A	110 E State St, Redlands, California, 92373	110 E STATE ST	REDL
	M	100	PP		39 S CENTER ST	REDL

1 (3 out of 24 Selected)

Geocoding_Result_1



Create locators from manual matches



Sample workspace with the Geoprocessing model is available on the UC Proceedings media

<http://proceedings.esri.com/library/userconf/proc11/index.html>

and Esri Geocoding Resource Center

Overview of the geocoding engine at ArcGIS 10

- **Geocoding process**
- **Grammar**
- **Data structure**
- **Aliases**
- **Scoring**

Understanding the basic components of the geocoding engine can help matching addresses more effectively

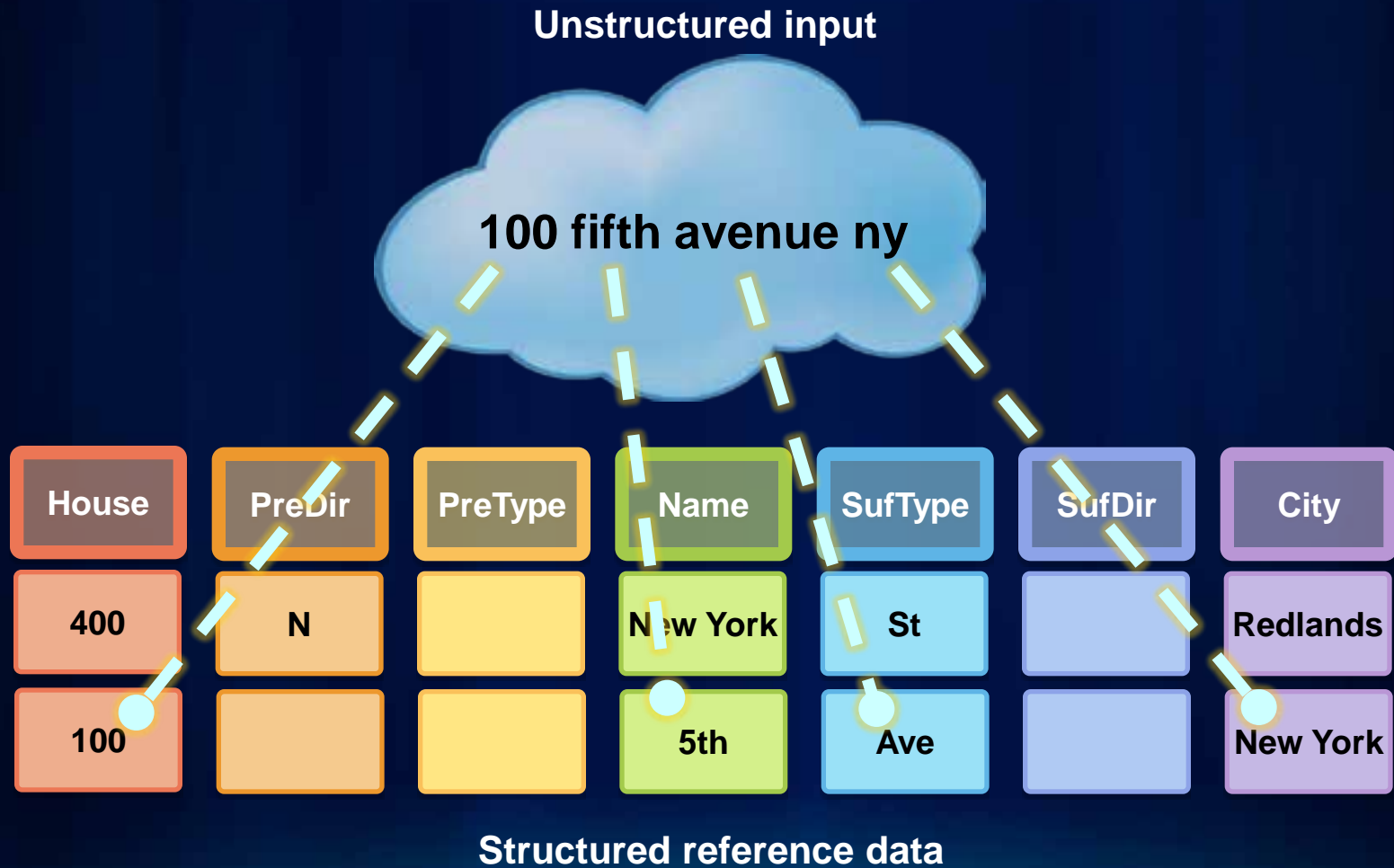
Geocoding process

Unstructured input

100 fifth avenue ny

1600 Pennsylvania Ave NW, Wash. DC
380 New York St, Redlands CA 92101
100 5th Ave, New York NY 10011
1355 N Hrbr Drive, San Diego CA 92101
898 West Mission Bay Drive San Diego California 92109
10 Ave N, Unit 1 Fifth floor

Geocoding process



Geocoding engine : grammar

- All supported forms of addresses explicitly defined in a grammar.

Geocoding engine : grammar (continued)

Grammar example:

```
Address: House StreetName City           // 380 New York St Redlands
        | StreetName "&" StreetName City // Main St & 2nd Ave Springfield
        | SpatialOperator Address ;      // 100 ft SW from 5 Main St San Diego
```

```
StreetName: PreDir PreType Name SufType SufDir ;
```

```
House: number           // 123
      | number "-" number // 17-100
      | number letter ;  // 100A
```

```
PreDir: "N" | "E" | "W" | "S" | "NW" | "SW" | "NE" | "SE" | ;
```

```
PreType: "Ave" | "Hwy" | ;
```

```
SufType: "Ave" | "St" | "Rd" | ;
```

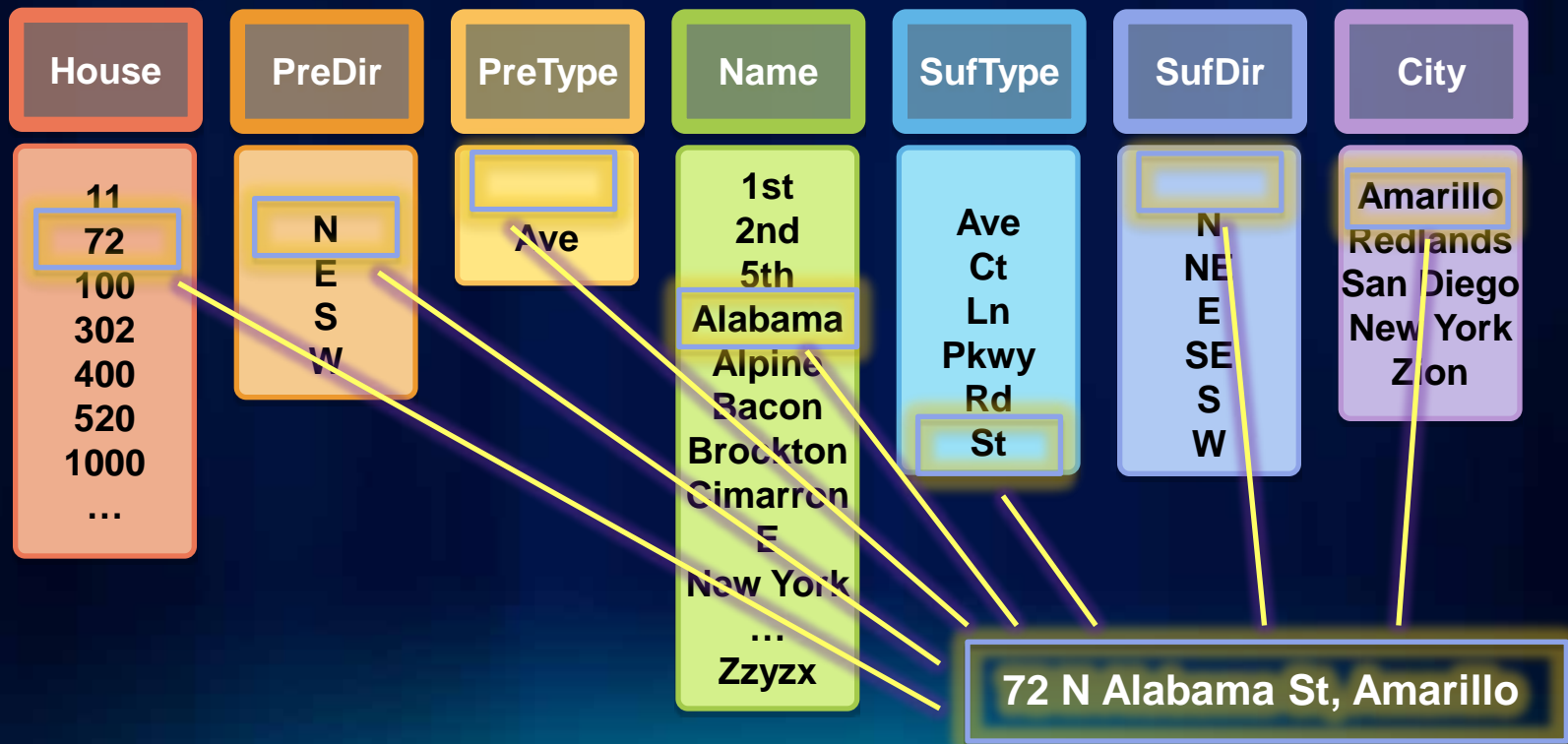
...

Handles
ambiguous addresses

No "standardization"

Geocoding engine : data structure

- Locator stores a snapshot of reference data
- List of unique values for each reference data field
- Many-to-many relationship between fields



Geocoding engine : data structure

Name

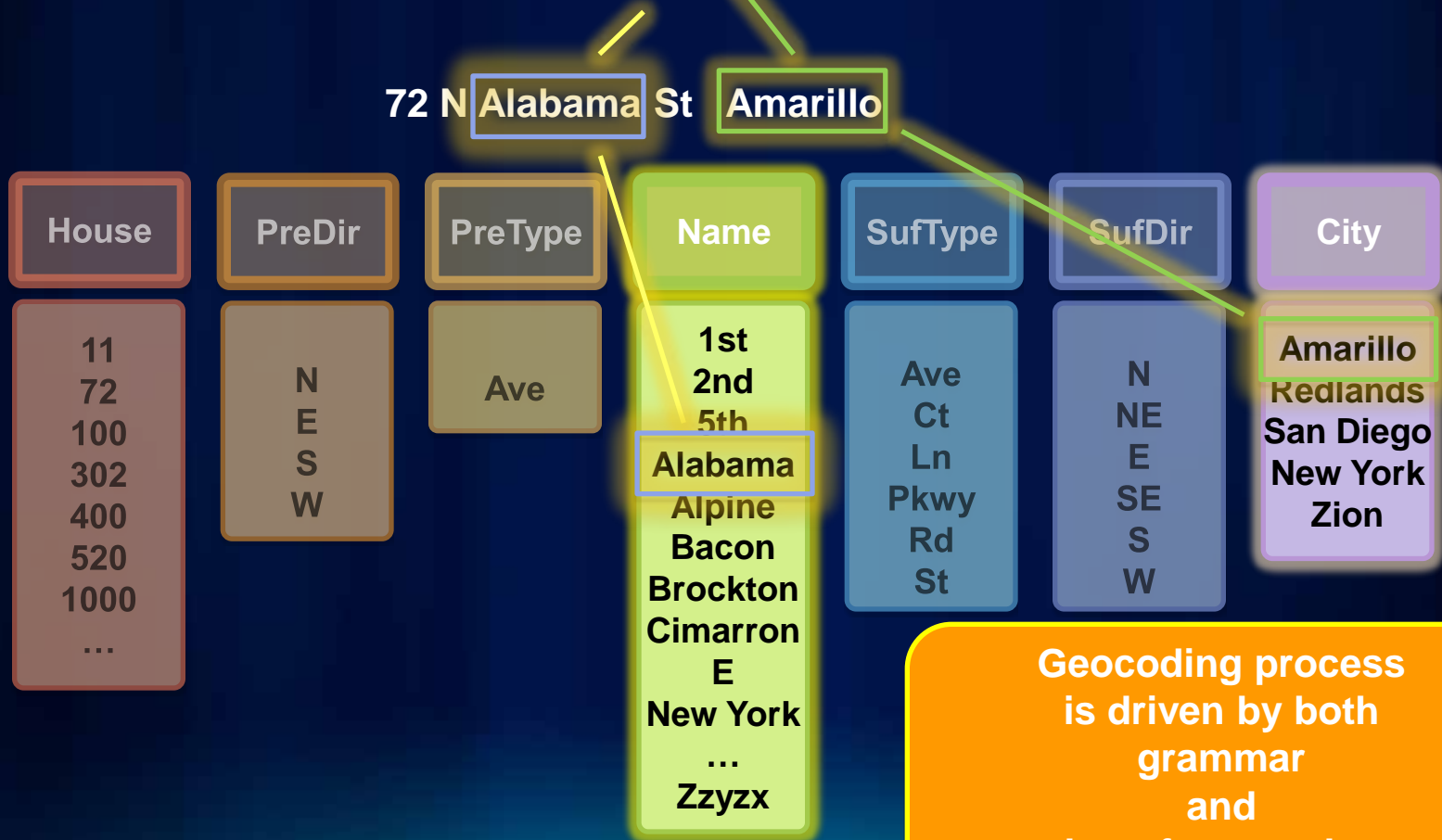
City

Geocoding engine : data structure

Address: House StreetName City

...

StreetName: PreDir PreType Name SufType SufDir ;



Geocoding engine : data structure (continued)

Reference Data is an extension of the grammar

SufType: "Ave" | "St" | "Rd" | "Ct" | "Ln" | "Pkwy" | ;

PreDir	PreType	Name	SufType	SufDir	City
N E S W	Ave	1st 2nd 5th Alabama Alpine Bacon Brockton Cimarron E New York ... Zzyzx	Ave Ct Ln Pkwy Rd St	N NE E SE S W	Amarillo Redlands San Diego New York Zion

Reference data help dealing
with some
local addressing oddities
without the need for locator customization

Geocoding engine : aliases

- **Common abbreviated forms (aliases)**
 - “Mt”, “Mtn”, “Mount”, “Mountain” mean the same in the context of a City name
 - Same for “View” vs. “Vw”
 - There are many ways to spell “Mountain View”
 - **Mountain View**
 - **Mt View**
 - **Mtn View**
 - **Mount View**
 - **Mountain Vw**
 - **Mt Vw**
 - **Mtn Vw**
 - **Mount. Vw**

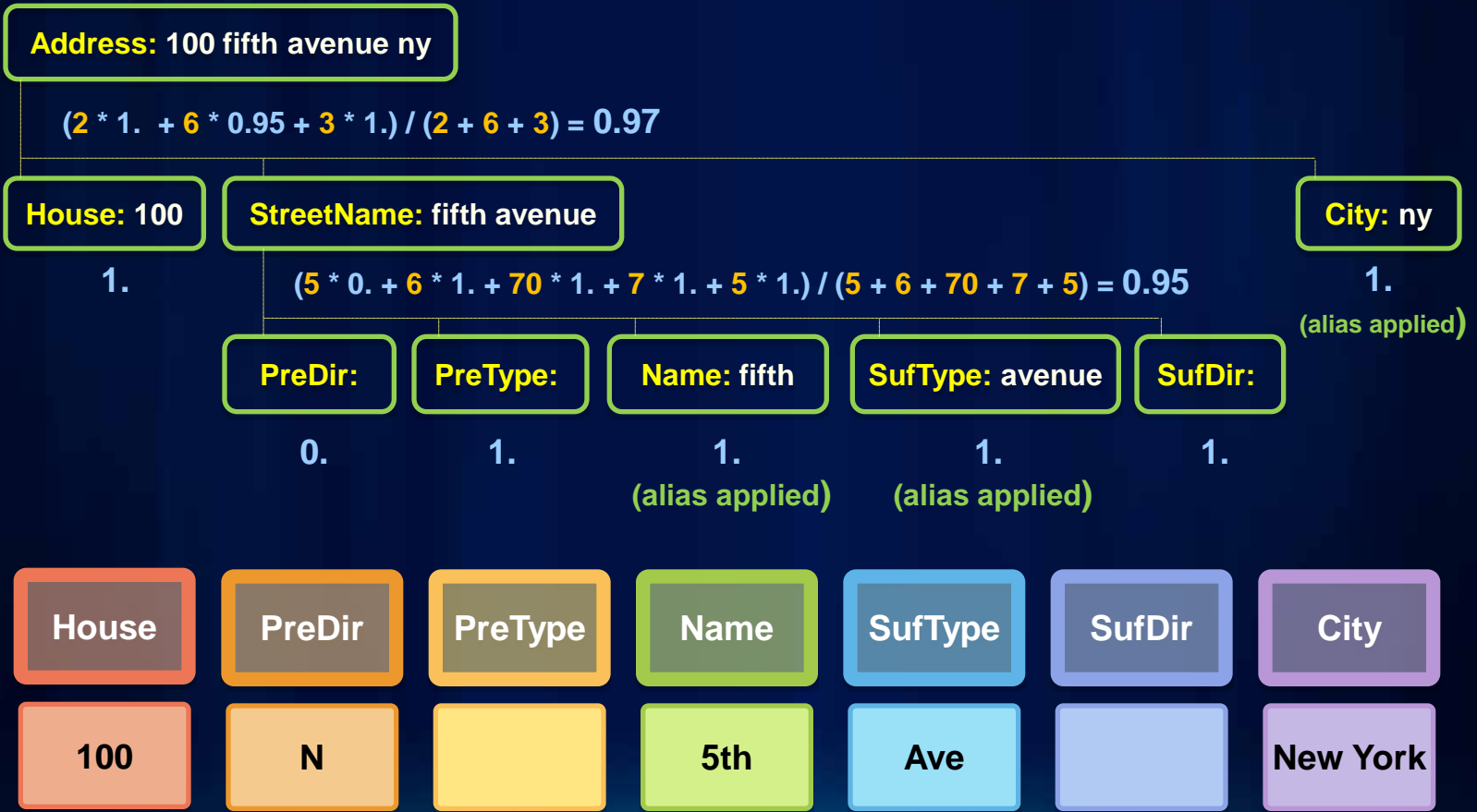
Geocoding engine : scoring

- Grammar has information about relative importance of the components

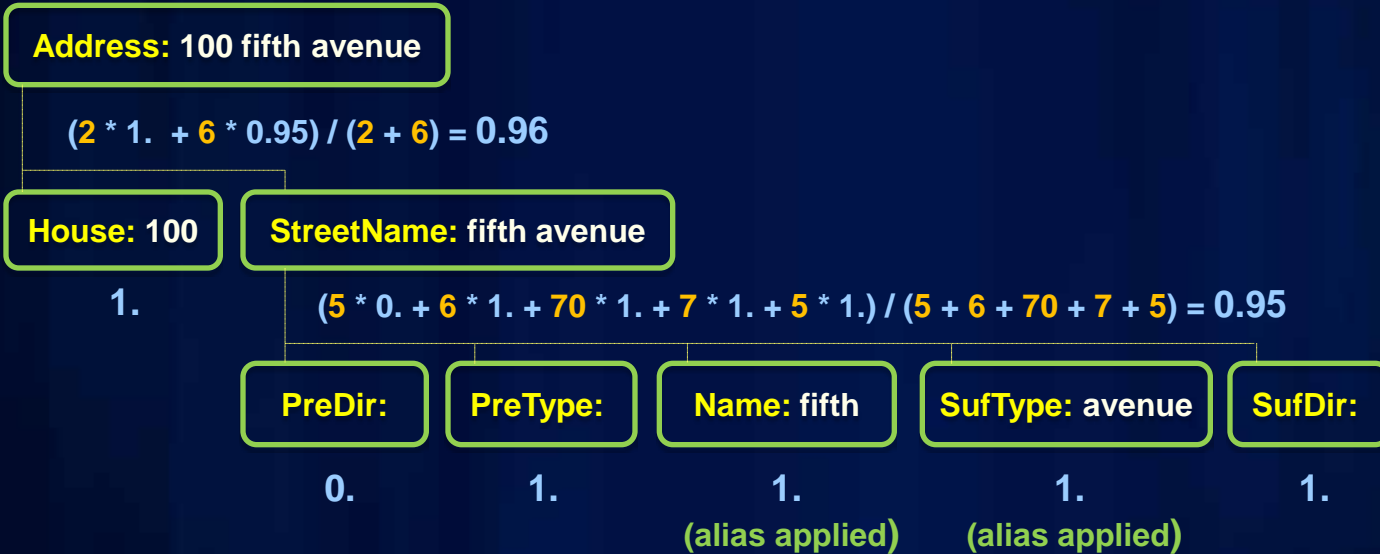
```
Address: House2 StreetName6 City3  
        | House2 StreetName6  
        ...  
StreetName: PreDir5 PreType6 Name70 Suftype7 SufDir5 ;
```

There is no penalty
if some of the components
(i.e. City, ZIP)
are missing from the input address

Geocoding engine : scoring (continued)



Geocoding engine : scoring (continued)

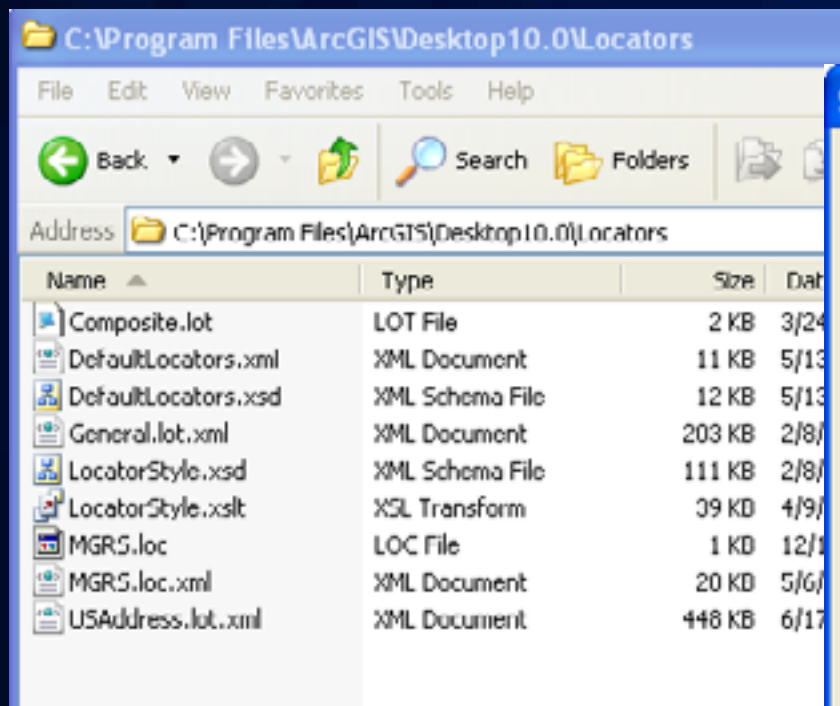


House	PreDir	PreType	Name	SufType	SufDir	City
100	N		5th	Ave		New York

ArcGIS 10 Geocoding engine summary

- Handling of ambiguous addresses
 - Example: **10 West Ridge Dr**
- No penalty for missing zone information
- Can parse any input field, not just Street Name
- Returns matching address as it is written in the reference data
- Single line input
- Highly configurable
 - Single XML file

Fine tuning locators



Displaying Locator Contents on Internet Browser

US Address

Locator style for US addresses

Format version: 10

[-] Grammar

[-] Aliases

StreetNameAliases

rl	"rl" "real"
lincoln	"lincoln" "abraham lincoln" "abe lincoln" "a lincoln"
alt	"alt" "alternate"
armor	"armor" "armrd" "armd" "armored"
_ave	_ave
batallion	"batallion" "bn"
_bch	_bch
ramirez	"ramirez" "benito a ramirez" "b a ramirez" "b ramirez"
_bnd	_bnd
beltway	"beltway" "belt"
benito	"benito" "bnito"
bonita	"bonita" "bonnita"
bus	"bus" "business highway" "bs hwy" "busi" "business" "bus"
_cyn	_cyn
cemetery	"cemetery" "cmtery"
_ctr	_ctr
central	"central" "cntrl"
complex	"complex" "cmplex"
col	"col" "colonel"
cottonwood	"cottonwood" "cottonwd"
_crk	_crk
crownview	"crownview" "crown view"
depot	"depot" "dep" "dept"
div	"div" "division" "divide"
executive	"executive" "exec"
garcia	"garcia" "felix garcia" "f garcia"
felipe	"felipe" "felpe"
flagstone	"flagstone" "flagstn"

[-] Suffix types

_acrd	"Acrd"
_aly	"Aly"
	"Allee"
	"Alley"
	"Ally"
	"Al"
_anx	"Anx"
	"Annex"
	"Annx"
	"Anex"
_arc	"Arc"
	"Arcade"
_avct	"Avct"
	"Avect"
	"Avenue" _ct
	"Ave" _ct
	"Av" _ct
_avdr	"Avdr"
	"Avedr"
_ave	"Ave"
	"Ave"
	"Av"
	"Aveflr"
	"Aven"
	"Avenue"
	"Avenue"
	"Avn"
	"Avnue"
	"Avenue of the"
	"Ave of the"
	"Av of the"

Done

My Computer

Fine tuning locators (with XML Editor)

- Aliases

```
<section desc="Aliases">
  <alias_list name="StreetNameAliases">
    ...
    <alias_def>
      <alias_def>
        <alt>Hts</alt>
        <alt>Height</alt>
        <alt>Heights</alt>
        <alt>Hgts</alt>
        <alt>Hghts</alt>
        <alt>Hgt</alt>
        <alt>Ht</alt>
      </alias_def>
    </alias_def>
  </alias_list>
</section>
```

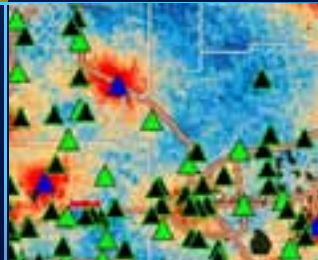
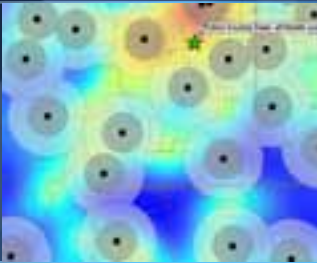
- Character mapping (spelling variations)

```
<spelling>
  <character_equivalency_table name="default">
    <noise_chars list=".,'`&#180;&#39;&#8216;&quot;; AOEIUYaoeiuY&#xC0;&#xC1;
    <allowed_penalties list="0,1,1,1,2,2,2,3,3,3,3,4,4,4,4,5,5,5,5"/>
    <double_chars list="bcdfighijklmnpqrstvwxyz" cost="0.1"/>
    <entry from="-'" to="-'" spell="90" cost="0.1" type="group"/>
    <entry from="sch" to="sh" spell="90" cost="0.1" type="simple"/>
    <entry from="sh" to="sch" spell="90" cost="0.1" type="simple"/>
    <entry from="mi" to="my" spell="90" cost="0.1" type="simple"/>
    <entry from="my" to="mi" spell="90" cost="0.1" type="simple"/>
    <entry from="yn" to="in" spell="90" cost="0.1" type="simple"/>
  </character_equivalency_table>
</spelling>
```

Fulton County Dept. of Health and Wellness/District 3, Unit 2, 04

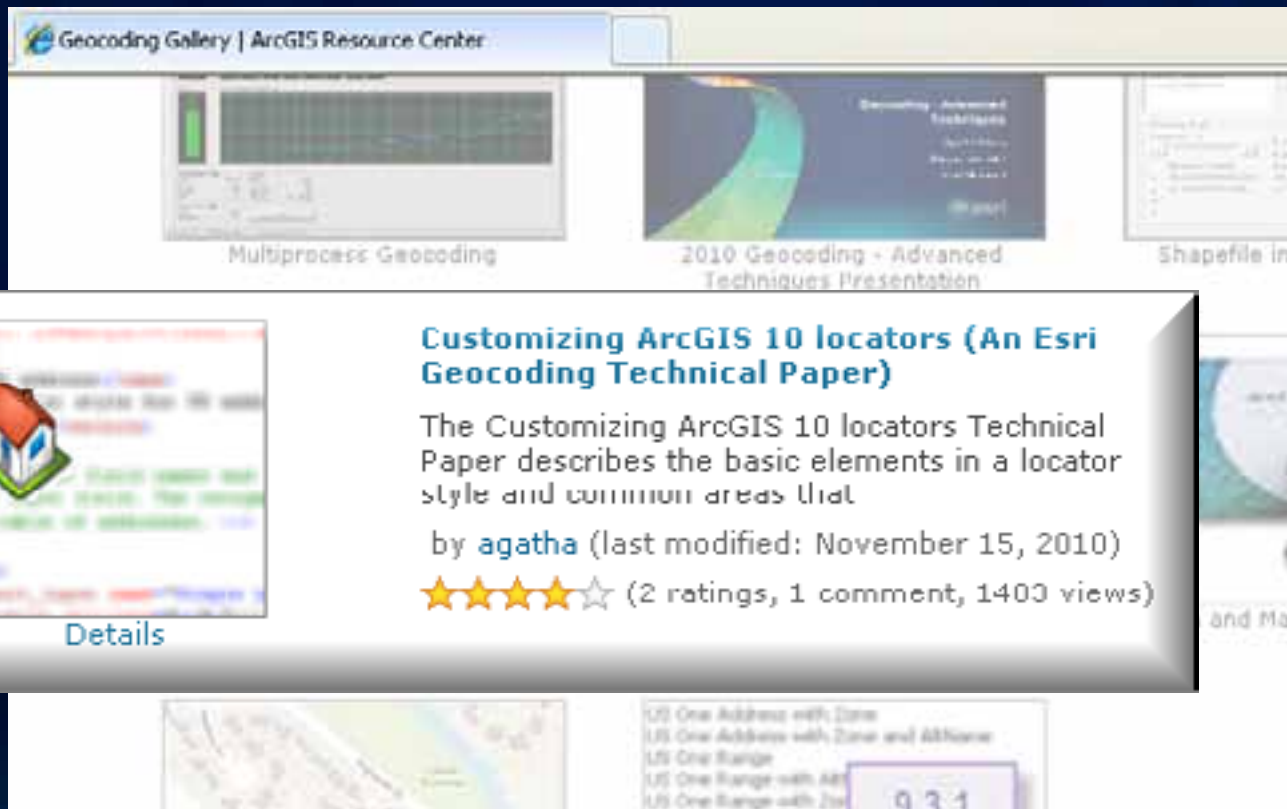
Demo: Find locations

- Spatial offset
- Dealing with challenging addresses



Customizing ArcGIS 10 Locators

- An Esri Geocoding Technical Paper
- <http://resources.arcgis.com/gallery/file/geocoding>



The screenshot displays the 'Geocoding Gallery | ArcGIS Resource Center' interface. It features a grid of resource thumbnails. The first row includes 'Multiprocess Geocoding', '2010 Geocoding - Advanced Techniques Presentation', and 'Shapefile Int'. The second row features a 'Details' pop-up for 'Customizing ArcGIS 10 locators (An Esri Geocoding Technical Paper)', which includes a house icon, a description, author information, and a star rating. Below the pop-up, there are thumbnails for 'US One Address with Zone', 'US One Address with Zone and Address', 'US One Range', 'US One Range with Address', and 'US One Range with Zone'.

Geocoding Gallery | ArcGIS Resource Center

Multiprocess Geocoding

2010 Geocoding - Advanced Techniques Presentation

Shapefile Int

Customizing ArcGIS 10 locators (An Esri Geocoding Technical Paper)

The Customizing ArcGIS 10 locators Technical Paper describes the basic elements in a locator style and common areas that

by [agatha](#) (last modified: November 15, 2010)

★★★★☆ (2 ratings, 1 comment, 1403 views)

Details

US One Address with Zone

US One Address with Zone and Address

US One Range

US One Range with Address

US One Range with Zone

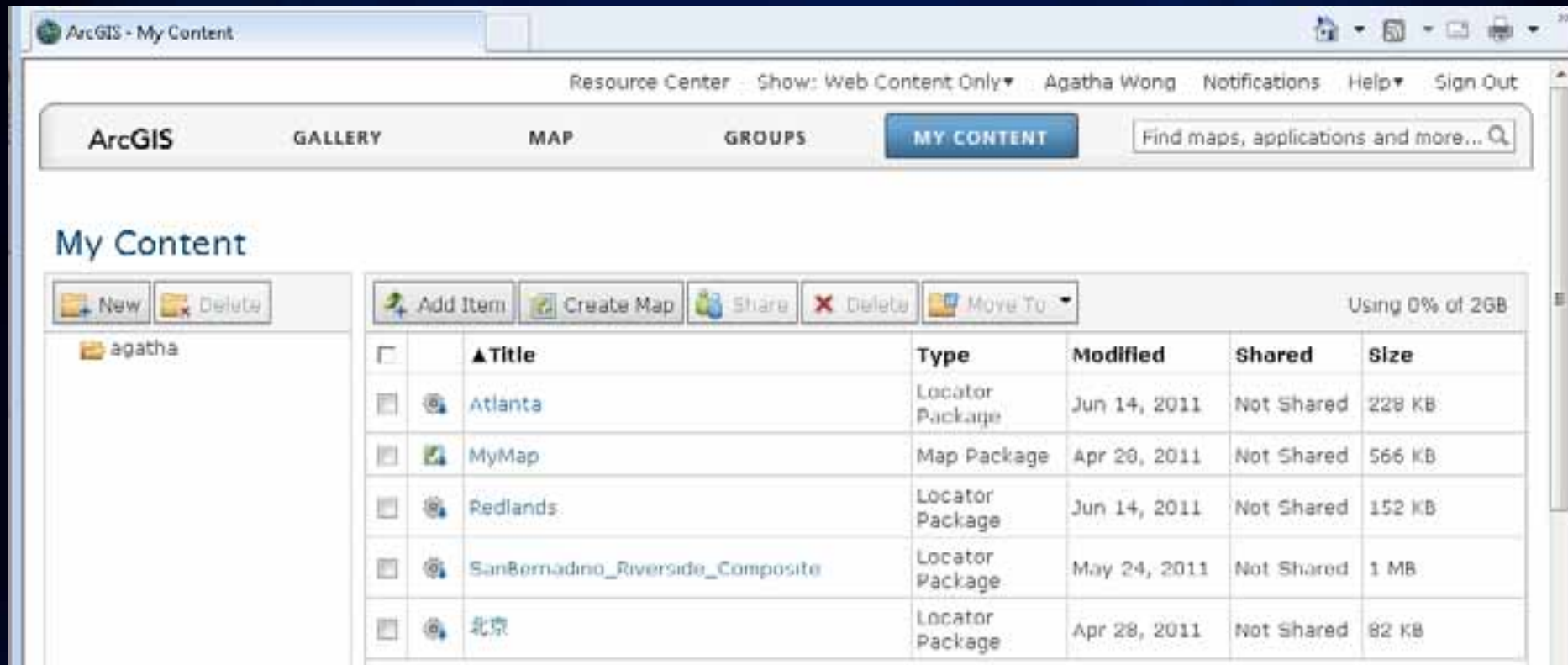
931

Road ahead for Geocoding at ArcGIS 10.1x

- **Better geocoding performance**
 - Batch matching
 - REST API

Road ahead for Geocoding at ArcGIS 10.1x

- Locator Package / Sharing via ArcGIS.com



The screenshot displays the ArcGIS My Content web interface. At the top, there's a navigation bar with 'ArcGIS', 'GALLERY', 'MAP', 'GROUPS', and 'MY CONTENT' (highlighted). A search bar on the right says 'Find maps, applications and more...'. Below the navigation bar, the 'My Content' section is visible. It includes a left sidebar with a folder named 'agatha'. The main content area has a toolbar with 'Add Item', 'Create Map', 'Share', 'Delete', and 'Move To' buttons. To the right of the toolbar, it says 'Using 0% of 2GB'. Below the toolbar is a table listing content items.

<input type="checkbox"/>	▲ Title	Type	Modified	Shared	Size
<input type="checkbox"/>	Atlanta	Locator Package	Jun 14, 2011	Not Shared	228 KB
<input type="checkbox"/>	MyMap	Map Package	Apr 20, 2011	Not Shared	566 KB
<input type="checkbox"/>	Redlands	Locator Package	Jun 14, 2011	Not Shared	152 KB
<input type="checkbox"/>	SanBernadino_Riverside_Composite	Locator Package	May 24, 2011	Not Shared	1 MB
<input type="checkbox"/>	北京	Locator Package	Apr 28, 2011	Not Shared	82 KB

Road ahead for Geocoding at ArcGIS 10.1x

- Geocode Services in the Cloud



Find Address Candidates: (Japan_JPNnumber) - Windows Internet Explorer

http://ec2-50-19-166-229.compute-1.amazonaws.com:6080/arcgis/rest/services/Japan_JPNnumber/GeocodeServer/findAd

Favorites Find Address Candidates: (Japan_JPNnumber)

ArcGIS REST Services Directory

[Home](#) > [services](#) > [Japan_JPNnumber \(GeocodeServer\)](#) > [findAddressCandidates](#)

Find Address Candidates: (Japan_JPNnumber)

Prefecture:

Gun:

City:

BlockAddress:

Single Line Input:

Out Fields:

Output Spatial Reference:

Format:

Address Candidates: (# address candidates : 200)

Shape:

Point:

X: 139.74064867245562

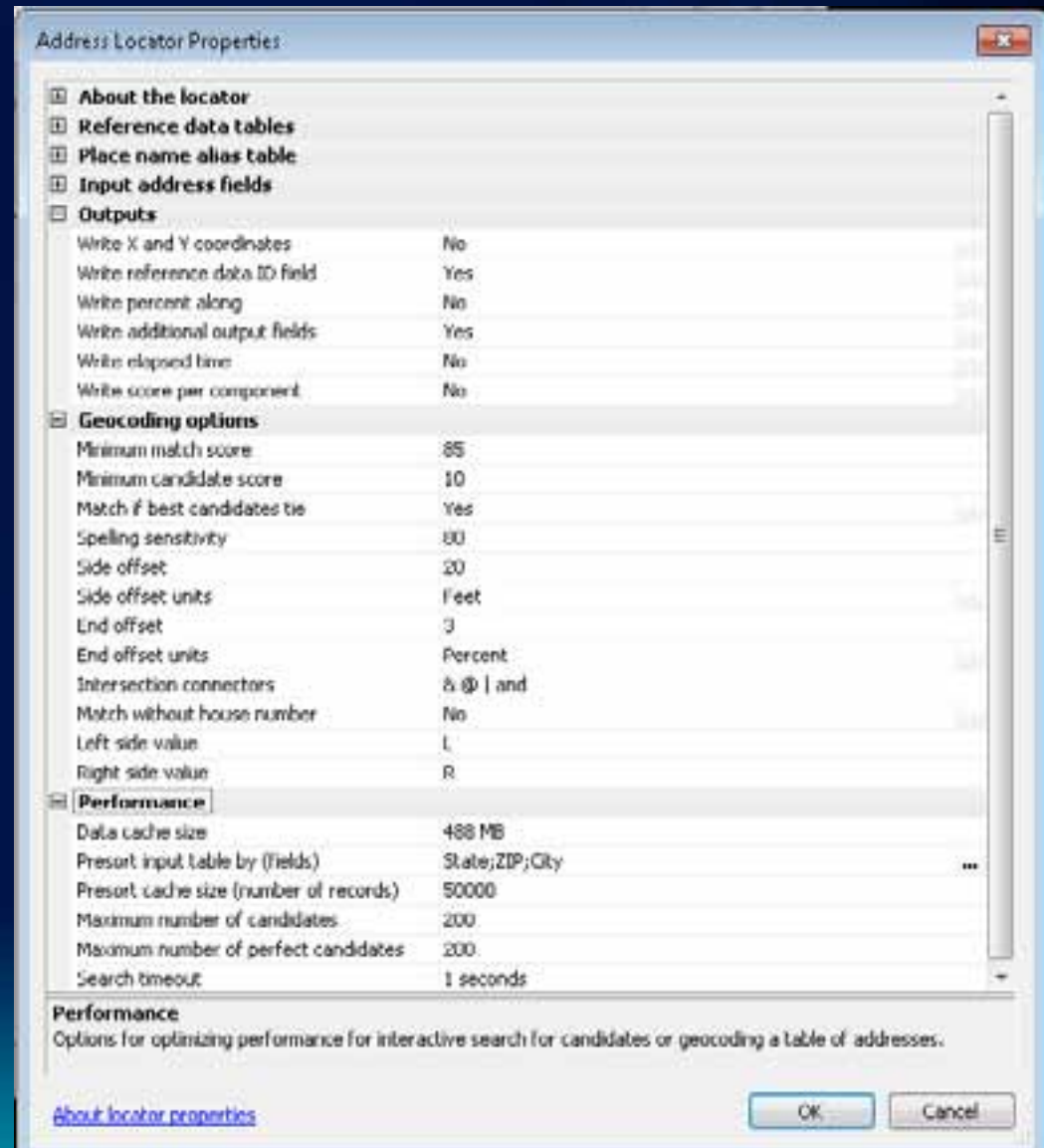
Y: 35.67921413682062

Score: 100.0

Address: 東京都千代田区平河町 2-7-1

Road ahead for Geocoding at ArcGIS 10.1x

- New Locator Properties user interface



Road ahead for Geocoding at ArcGIS 10.1x

- **Enhanced searching**
 - Search by extent
 - Within the current map extent
 - Return candidates within a bounding box
- **Return candidates based on rank (simple rank, or population)**
 - Example:
 - New York, NY
 - New York, TX

Road ahead for Geocoding at ArcGIS 10.1x

- **Global geocoding initiatives – international locator styles at Esri Resource Centers, and geocode services on ArcGIS.com.**

Resources and References

- ESRI Resource Centers

<http://resources.arcgis.com>



The screenshot shows the ArcGIS Resource Center website interface. At the top, there is a navigation bar with links for 'Content management', 'User management', and 'Help'. The main header features the 'ArcGIS Resource Center' logo, 'Help', 'Blogs', and 'Forums' links, along with a search bar. Below the header, a green banner indicates the current page is 'Home » Geocoding' and shows the 'Version: 10.0'. A 'toggle tools' link is also present. On the left side, a sidebar menu lists 'Geocoding' with sub-links for 'Locators', 'Finding Locations', 'Geoprocessing', 'Help', 'Forums', 'Gallery', 'Videos', and 'Knowledge Base'. A 'Get Support' button is located at the bottom of the sidebar. The main content area is titled 'What is Geocoding?' and contains a paragraph explaining the process: 'Geocoding is the process of transforming a description of a location—such as an address or a place name—to a location on a map. You can geocode by entering one location description at a time or by providing many at once in a table. The resulting locations are output as geographic features with attributes, which can be used for mapping and spatial analysis.' Below this text is a link to 'See What's new in geocoding in ArcGIS 10.' To the right of the text is a small map showing a residential area with red buildings and green trees.

Content management User management Help 230 / 11 Log out agarha

ArcGIS Resource Center Help Blogs Forums

Home » Geocoding Version: 10.0

View current Revisions toggle tools

Geocoding

- Locators
- Finding Locations
- Geoprocessing

Help Forums Gallery Videos Knowledge Base

Get Support

What is Geocoding?

Geocoding is the process of transforming a description of a location—such as an address or a place name—to a location on a map. You can geocode by entering one location description at a time or by providing many at once in a table. The resulting locations are output as geographic features with attributes, which can be used for mapping and spatial analysis.

See [What's new in geocoding in ArcGIS 10.](#)

Additional Geocoding Sessions

- ***Geocoding – An Introduction***
 - ***Wednesday 1:30 PM, (Room 14B) – Offering II***
- ***ESRI Showcase Software Island Demo Theater***
 - ***Road Ahead – Geocoding (Tuesday 5:00 PM)***
 - ***From a table of addresses to locations on the map (Thursday 11:30 AM)***
- ***Visit the Spatial Analysis Island in the Exhibit Hall***

Questions ?

&

Thank you for attending

Please fill out and submit your evaluation form

www.esri.com/sessionevals



esri