Supporting Postal Addresses around the world

Erwin Hom, Mike McKenna

Problem Statement

How does an app handle postal addresses for 200+ countries and 27 languages?

If you start with one country without a global strategy, you will discover lots of challenges.

10/28/15   IUC 39   2

Challenges

Address capture for different use-cases
- customer sign-up, shipping, billing,
- regulatory compliance, etc.

Address display for multiple scenarios
- Single line, multi-line, PO Box, customized

Coarse grained and fine grained address formats

Bad Address Data

😁 Keep app developers from writing address formatting code.

Is this Solution **Good Enough**?

Object Managem

Billing Information

*First/Last Name:

*Street Address:

*City:

*State: US: Please Select State
Non-US:

*Zip-Postal Code:

United States

*Country: Other:

*Day Phone:

Night Phone:

Fax Number:

*Email Address:

☐ Check here if your shipping/billing

Single entry form to collect address data for all countries
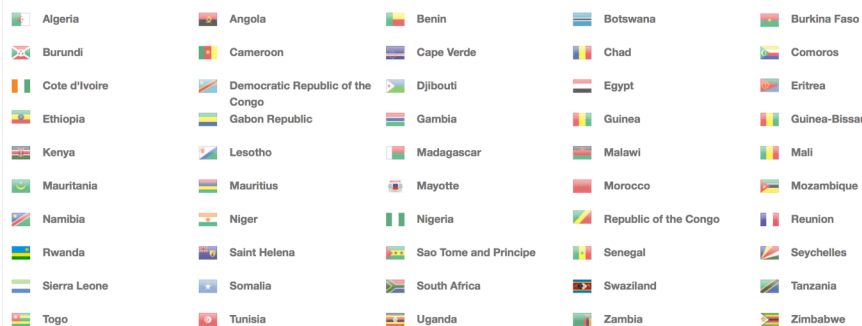Just need to localize the labels.

Where PayPal does business

Africa

Algeria	Angola	Benin	Botswana	Burkina Faso
Burundi	Cameroon	Cape Verde	Chad	Comoros
Cote d'Ivoire	Democratic Republic of the Congo	Djibouti	Egypt	Eritrea
Ethiopia	Gabon Republic	Gambia	Guinea	Guinea-Bissau
Kenya	Lesotho	Madagascar	Malawi	Mali
Mauritania	Mauritius	Mayotte	Morocco	Mozambique
Namibia	Niger	Nigeria	Republic of the Congo	Reunion
Rwanda	Saint Helena	Sao Tome and Principe	Senegal	Seychelles
Sierra Leone	Somalia	South Africa	Swaziland	Tanzania
Togo	Tunisia	Uganda	Zambia	Zimbabwe

Americas

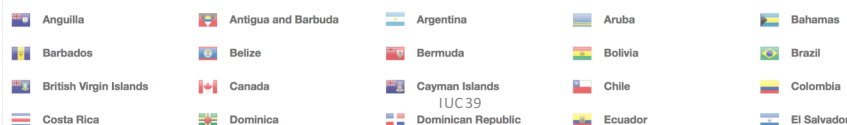|			
Anguilla	Antigua and Barbuda	Argentina	Aruba	Bahamas
Barbados	Belize	Bermuda	Bolivia	Brazil
British Virgin Islands	Canada	Cayman Islands	Chile	Colombia
Costa Rica	Dominica	Dominican Republic	Ecuador	El Salvador

IUC 39   5

What we did

- Get familiar with the US Format
- Take a look at some special things about Address Formats from around world
- Gather Country Address Metadata
- Consider the storage and transmission of address data
- Build Library
- Inform developers

10/28/15   IUC 39   6

US Postal Address Format

Some Postal Address Types (US)

MR JOHN DOE 123 MAGNOLIA ST HEMPSTEAD NY 11550-1234 UNITED STATES OF AMERICA	MISS JANICE SMITH **PO BOX 34** DULUTH MN 55803-0034 UNITED STATES OF AMERICA
JOHN DOE **ABC COMPANY** 1401 MAIN ST FALLS CHURCH VA 22042-1441 UNITED STATES OF AMERICA	SMITHSONIAN INFORMATION **PO BOX 37012** SI BUILDING, ROOM 153, MRC 010 WASHINGTON, D.C. 20013-7012 UNITED STATES OF AMERICA

Elements of the US Address Format

Addr Line 1	BLDG, APT, STE, or UNIT	PO Box Number
Addr Line 2	[HOUSE NUMBER][ ][STREET NAME (+ TYPE)]	
	CITY, STATE_ABBR ZIP+4 Code	
	USA	

The Country Name is included in the International Format.

10/28/15   IUC39   9

Where do these conventions come from?

These conventions are documented in

"Publication 28, Postal Address Standards"

  http://pe.usps.gov/text/pub28/28c2_002.htm

Note that letters are uppercased for OCR readability.

Standards allow your letters, postcards, junk-mail to be **processed** (going through the system) **quickly**

  http://pe.usps.gov/text/pub28/28apa_003.htm

10/28/15   IUC39   10

ADDRESS DISPLAY FORMATS
Wrap Up or Wrap Down

MS MILDRED DOE
12 E BUSINESS LN STE 209
KRYTON TN 38188-0002

MS MILDRED DOE
BUSINESS LN STE 209
KRYTON TN 38188-0002

Wrap Up Street No

Wrap Down Zip Code

MS MILDRED DOE
STE 209
12 E BUSINESS LN
KRYTON TN 38188-0002

MS MILDRED DOE
BUSINESS LN STE 209
KRYTON TN
38188-0002

Guidelines on when to wrap up or wrap down a US address can be found here

10/28/15 — IUC39 — 11

PERSONAL INFORMATION | ACCOUNT NO.

SOCIAL SECURITY NO. | MARITAL STATUS: CHECK ONE If you reside in or are relying on property in a community property state or if you are applying for a secured credit or joint account.

☐ UNMARRIED | ☐ MARRIED | ☐ SEPARATED | ☐ UNMARRIED

AST NAME (JR./SR.) | FIRST NAME | INITIAL | LAST NAME (JR./SR.)

YEARS AT THIS ADDRESS | CURRENT STREET ADDRESS | APT. NO. | YEARS AT THIS ADDRESS

ER'S LICENSE NO. | CITY | STATE | ZIP | DRIVER'S LICENSE NO.

DENTS (NOT INCLUDING | DATE OF BIRTH | HOME PHONE ( ) | NO. OF DEPENDENTS (NOT INCLUDING YOURSELF)

EMPLOYMENT INCOME

Paystub with this Application)

CAPTURING ADDRESS DATA

MONTHLY | MONTHLY SALARY

PHONE ( ) | PHONE ( )

DATE | POSITION / TYPE OF WORK | START DATE

10/28/15 — IUC39 — 12

Address Entry Forms

Street Address Line 1

Street Address Line 2

City   State   ZIP Code

Number* Street*

Building Name   Additional Guidance

City   State   ZIP Code

`\d{5}([\-]\d{4})?`

AL – Alabama
AK – Alaska
AZ – Arizona
CA – California
CO – Colorado
CT – Connecticut
DE – Delaware
...

Fine-grain Capture

10/28/15   IUC 39   13

ADDRESS ENTRY FORMS
Parse key address fields

Street Address Line 1 ⟹ Apt   Floor   Building

Street Address Line 2 ⟹ Street Number   Street Name

2798 West Main Street #12
Minneapolis MN 23876 ⟹

Apt	12
Street Number	2798
Street Name	West Main Street
City	Minneapolis
State	MN
Postal Code	23876

10/28/15   IUC 39   14

ADDRESS ENTRY FORMS
Address Normalization

Address Correction and Standardization

2211 North First Street S Jose CA 11201	→	2211 N 1st St San Jose CA 95131-2021

ADDRESS ENTRY FORMS
Auto fill

Legal business name		Business phone
Street address (no PO Box)		
San Jose	CA	95112|

Summary - Supporting US Postal Address

- more than one address type
- the address entry form and the address display format are similar but the entry form can be different for mobile apps.
- use drop-down list (states) to ease input
- mandatory/optional fields
- parse key address fields

10/28/15   IUC 39   17

Address Entry Forms

Address Validation

- use regex to validate postal code. Error message to guide them.
- Does the address exist?
- Does the postal code match with city and state?
- Is the house number valid?

10/28/15   IUC 39   18

SUPPORTING ADDRESS FORMATS FROM AROUND THE WORLD

10/28/15   IUC 39   19

What we started with

Universal Postal Union
http://www.upu.int/en/activities/addressing/postal-addressing-systems-in-member-countries.html

Global Sourcebook for International Data
http://www.grcdi.nl/gsb/global%20sourcebook.html

Address Doctor
Lookup Address Format
https://www.addressdoctor.com/en/countries-data/address-formats.html
Verify Address
http://lookup.addressdoctor.com/lookup/default.aspx?lookup=1&lang=en

10/28/15   IUC 39   20

Postal Address standards

International Standards:
- Universal Postal Union's S42 Standard
- ISO 19 19160 Address_Standards
- OASIS

Country Standards:
- US FDGC Address
- BS 7666
- Australia
- Canada
- Singapore
- Swiss Post
- …

10/28/15   IUC 39   21

International Address Formats?

"**There is no internationally applicable address format**. As in almost every country, the address format differs.

Even if differences between the respective formats seem small, not only the correct elements of an address, but also their hierarchical alignment play a big role.

Only with the right format, you can be sure that a mailing gets to its addressee."

www.**addressdoctor**.com/en/countries-data/address-formats.html

10/28/15   IUC 39   22

Japan – Address Elements

Addressee	Miki Murakami
Organization	Google Japan Inc.
Address Line 1	26-1 Sakuragaoka-cho
Address Line 2	Cerulean Tower 6F
City	Shibuya-ku
Prefecture	Tokyo
Postal code	150-8512

10/28/15   IUC39   23

Japan – Address Entry Form (**Native**)

[*Postal code*]
Prefecture
[*City / Ward / Town / Village*]
[*Street name, street number*]
[*Building name, floor, room number* (optional)]

Your postal code must
contain 7 digits in the
following format: 112-0001

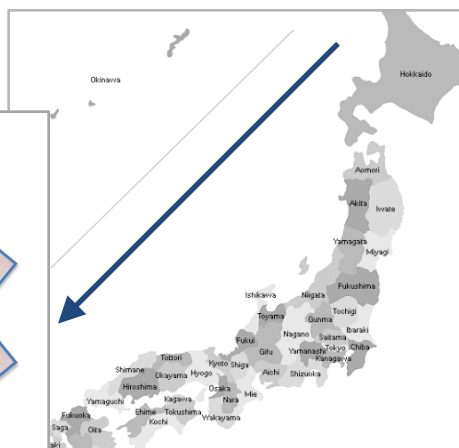10/28/15   IUC39   24

Listing of Prefectures

geographic order
~~dictionary order~~

Hokkaido	Aichi
Aomori	Akita
Iwate	Aomori
Miyagi	Chiba
Akita	Ehime
Yamagata	Fukui
Fukushima	Fukuoka
Ibaraki	Fukushima
Tochigi	Gifu
Gunma	Gunma
Chiba	Hiroshima
Tokyo	Hyogo
...	...

10/28/15 IUC 39 25

ADDRESS DISPLAY FORMATS

Zero Width Space

Construction of Single-line Japanese address with ZWSP

<Postal Code><ZWSP><City><ZWSP><Ward><ZWSP>
<District><ZWSP><Block><ZWSP><Room>

What the end user sees

〒100-8994<ZWSP>東京都<ZWSP>中央区<ZWSP>八重洲<ZWSP>一丁目<ZWSP>5番3号

Single line format wrapped using ZWSP

〒100-8994
東京都中央区
八重洲
一丁目5番3号

Tōkyō-to	Chūō-ku	Yaesu	1 Chōme	5-ban 3-gō

10/28/15 IUC 39 26

United Kingdom – Address Elements

Organization	Starbucks Coffee Co
Address Line 1	Building 4, Chiswick Park
Address Line 2	566 Chiswick High Road
Post town	LONDON
County	
Postcode	W4 5YE
	United Kingdom

10/28/15   IUC39   27

UK -- Address Entry Form

[*Address line 1*]
[*Address line 2* (optional)]
[*Town / City*]

County (optional)

[*Postcode*]

Please enter a valid UK
postcode using uppercase
letters — e.g. TW9 1EH or
BPFO 772 (for British
Armed Forces).

10/28/15   IUC39   28

Customer Challenge

How do you get this type of address into a coarse grain (line1-line2) address entry form?

Addressee identification	
Building	
Dependent thoroughfare	
Thoroughfare	
Double dependent locality	
Dependent locality	
POST TOWN	
UNITED KINGDOM	

10/28/15   IUC 39   29

Country Address Metadata

... is a **description** of a Country's Postal Address Format.

Address Metadata is **essential** for Address Formatting.

http://unicode.org/reports/tr35/tr35-info.html**#Postal_Code_Validation**

Deprecated in v27.

Please see other services that are kept up to date, such as:

- http://i18napis.appspot.com/address/data/US
- http://i18napis.appspot.com/address/data/CH
- ...

10/28/15   IUC 39   30

Country Address Metadata

I18napis Address Data Service is a **CLDR** for Postal Address Metadata.

http://i18napis.appspot.com/address -- landing page

http://i18napis.appspot.com/address/data returns list of countries supported (2-Letter Country Codes)

http://i18napis.appspot.com/address/data + <2-Letter Country Code> returns address metadata in JSON format.http://i18napis.appspot.com/address/data/ZZ (Default) should be included in each country's address metadata.

10/28/15 IUC 39 31

Japan Address Metadata

Country Name	"name": "JAPAN"
Country code	"key": "JP"
Native Format	"fmt": "〒%Z%n%S%C%n%A%n%O%n%N"
Latin Format	"lfmt": "%N%n%O%n%A%n%C, %S%n%Z"
Address Elements Required	"require": "ACSZ"
Address Elements to be displayed as UPPER-cased.	"upper": "S"
ZIP Code Validation Pattern	"zip": "\\d{3}-?\\d{4}
ZIP Code Examples	
Prefecture Names in Latin Script	"sub_lnames": "Hokkaido~Aomori~Iwate~Miyagi~Akita~Yamagata~…"
Prefecture Keys	"sub_keys": "北海道~青森県~岩手県~宮城県~秋田県~山形県~…"

10/28/15 IUC 39 32

%n

"fmt": "%N**%n**%O**%n**%A**%n**%C, %S %Z"

- Multi-line = <line break>
- Single-line = Space / No Space
- Multi-line = Stylized output (HTML, CSS)

griffin.COARSE_ADDRESS_FORMAT_DISPLAY_MULTI.**semanticOutput**());

10/28/15   IUC 39   33

Griffin Address API – Format Single

Current Language: ja-JP

Current Country: JP

Result:

〒150-8512 東京都渋谷区桜丘町26-1セルリアンタワー6階

function (locality) {

var griffin = locality.griffin,
    address = {
      'line1':'桜丘町26-1',
      'line2':'セルリアンタワー6階',
      'state':'東京都',
      'city':'渋谷区',
      'postcode':'150-8512',
};
return griffin.formatCoarseAddress(address, griffin.COARSE_ADDRESS_FORMAT_DISPLAY_SINGLE);

}

Run Code

10/28/15   IUC 39   34

Griffin Address API - Format

```
Current Language:  ja-JP

Current Country:   JP
```

```
Result:
〒150-8512
東京都渋谷区
桜丘町26-1
セルリアンタワー6階
```

Run Code

```
var griffin = locality.griffin,
    address = {
        'line1':'桜丘町26-1',
        'line2':'セルリアンタワー6階',
        'state':'東京都',
        'city':'渋谷区',
        'postcode':'150-8512',
};
return griffin.formatCoarseAddress(address, griffin.COARSE_ADDRESS_FORMAT_DISPLAY_MULTI);
```

10/28/15   IUC 39   35
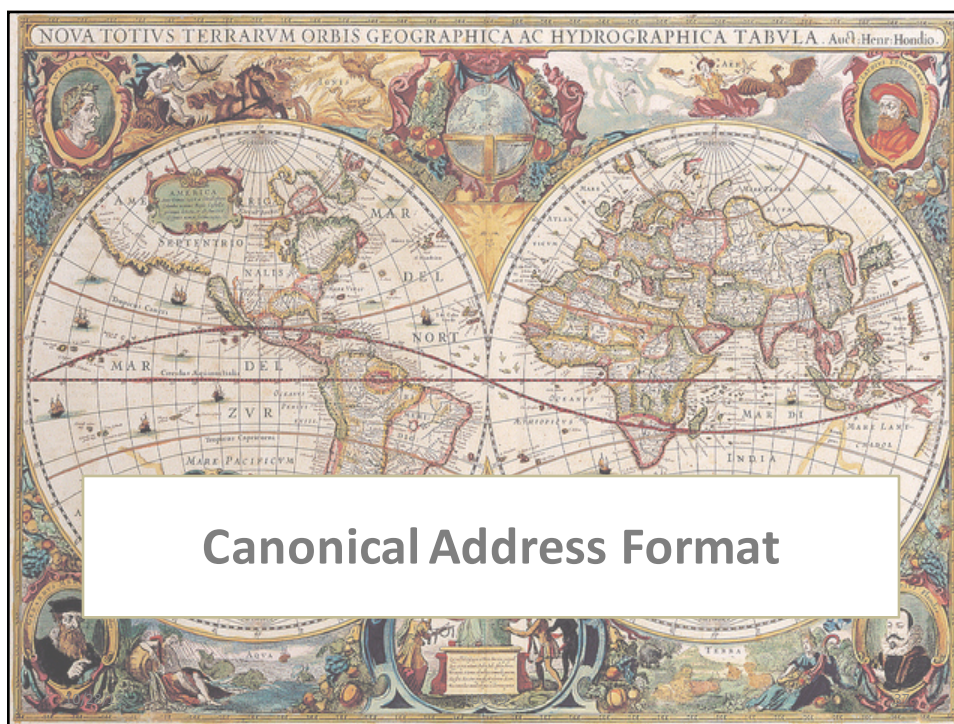
Lessons Learned

i18napis Address Data Service:

• Doesn't provide Default Field labels

• Postcode Error Message
• Sometimes Address Format might not match your Customer
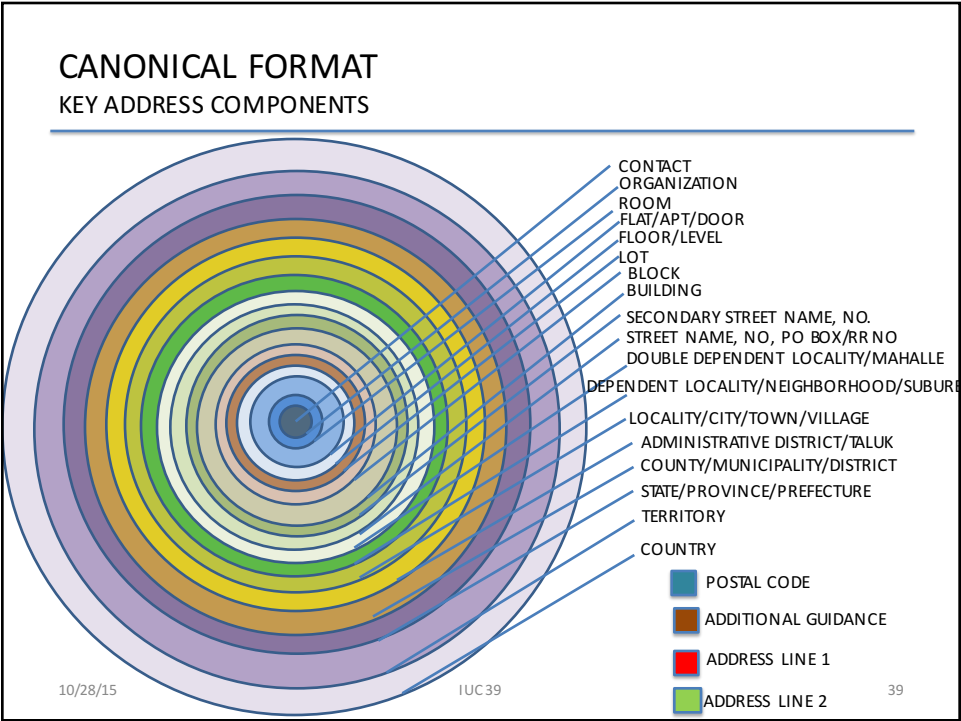 Requirements. (override the reference)

10/28/15   IUC 39   36

Canonical Address Format

Canonical Address Format?

- a single common schema to be used for interchange (JSON), processing and storage (database schema)
- locale-agnostic
- all locale-specific fields should map to one or more fields of the canonical format
- composed of finite number of fields

CANONICAL FORMAT
KEY ADDRESS COMPONENTS

CONTACT
ORGANIZATION
ROOM
FLAT/APT/DOOR
FLOOR/LEVEL
LOT
BLOCK
BUILDING
SECONDARY STREET NAME, NO.
STREET NAME, NO, PO BOX/RR NO
DOUBLE DEPENDENT LOCALITY/MAHALLE
DEPENDENT LOCALITY/NEIGHBORHOOD/SUBURB
LOCALITY/CITY/TOWN/VILLAGE
ADMINISTRATIVE DISTRICT/TALUK
COUNTY/MUNICIPALITY/DISTRICT
STATE/PROVINCE/PREFECTURE
TERRITORY
COUNTRY

- POSTAL CODE
- ADDITIONAL GUIDANCE
- ADDRESS LINE 1
- ADDRESS LINE 2

10/28/15 · IUC 39 · 39

Canonical Address Format

Canonical	BR	JP	US	GB
Organization				
Address Line 1				
Address Line 2				
Dependent Locality	Neighborhood			
City				Post town
State		Prefecture	State	County
Postal code	CEP	Postal Code	ZIP Code	Post code
Country	Brazil	Japan	USA	United Kingdom
Script	Latin	Japanese, Latin	Latin	Latin

10/28/15 · IUC 39 · 40

What we learned

- Use **country standards** to define the locale specific formats. (Refer to Google Address Metadata)
- Use of centralized canonical schema for Address data storage, processing and transmission.

- Use (build) an API to access country address metadata to do the formatting.
- Adding support for a new country means adding metadata for that country.

10/28/15   IUC39   41

What we also learned

Considerations for Address Display:
- Order address elements based on country and layout
- Spacing , punctuation, casing and line breaks
- Wrap up or down

Considerations for Address Entry:
- Different layouts for different scenarios
- Localized field labels
- Address Validation
- Address Normalization
- Script Restriction

10/28/15   IUC39   42

Getting Started

- Get Familiar with Countries you are targeting
    - Look at Country Metadata
- Make sure you have a Canonical Address Format that works for you.
- Will you support more that one language for that country?
    - Address Labels and localization
    - Script restriction
- Capture Locale Address: Locale Address-> Canonical Address
- Display Locale Address: Canonical Address -> Locale Address

10/28/15   IUC 39   43

Appendix

10/28/15   IUC 39   44

Links to I18napi Address Data Service

- http://i18napis.appspot.com/address
- http://i18napis.appspot.com/address/data
- http://i18napis.appspot.com/address/data/ZZ
- http://i18napis.appspot.com/address/data/   + {US,JP,BR,UK,…}

Open Source Address Library

https://github.com/googlei18n/libaddressinput
https://github.com/googlei18n/libaddressinput/blob/master/README.md

10/28/15   IUC 39   45

About **Country** Address Formats

USPS – Publication 28

http://pe.usps.gov/text/pub28/28c2_002.htm
http://www.royalmail.com/find-a-postcode

10/28/15   IUC 39   46

Australia – Address Entry Form

[*Unit number or Building name*]
[*Residential address* (optional)]
[*City / Suburb*] [State / Territory ▲▼] [*Postcode*]

```
Your postcode must be
4 digits.
```

10/28/15   IUC39   47

Brazil – Address Elements

Addressee	Marina Costa e Silva
Street Name, Street Number	**Rua** Afonso Canargo, **805**
Neighborhood	Santana
City - State	Guarapuava - **PR**
CEP (Brazilian Postal code)	85070-200

10/28/15   IUC39   48

BR - Address Entry and Display Formats

[*Postal code*]
[*Address*]
[*Neighborhood* (optional)]
[*City*]

State

Your postal code must
be 8 digits.

line1
line2
city - state
POSTCODE

line1, line2, city - state, POSTCODE

10/28/15   IUC39   49

China – Address Display Formats

- line1
- line2
- city
- **Province**
- POSTCODE

Local Script:

- **Province city**
- line1
- line2
- POSTCODE

中国200021上海市**上海市**淮海中路300

10/28/15   IUC39   50

China – Address Entry Form (**Native**)

Province / Municipality ▲▼

[City / County]
[Address line 1]
[Address line 2 (optional)]
[Postal code]

```
Your postal code must
be 6 digits.
```

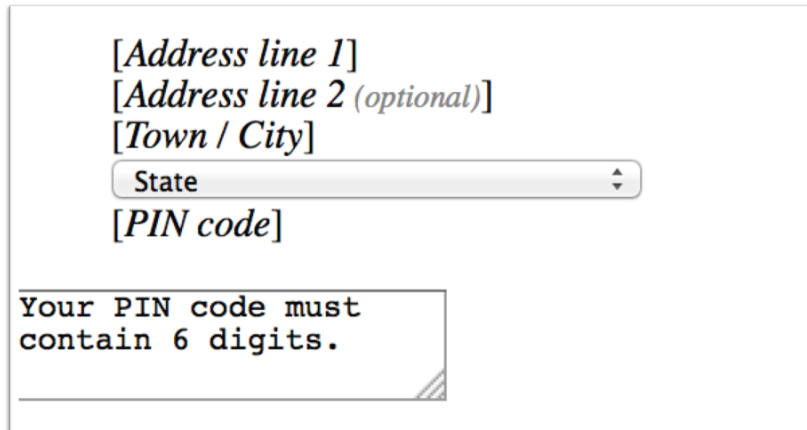## India – Address Elements

Addressee	Raj Patel
Organization	Popular Electronics
Address Line 1	Main Road
Additional Guidance (Address Line 1)	Astro High Complex, Near Old Taxi Stand Main Road
Town/City	Ranchi
State	Jharkhand
PIN code	834001
	INDIA

India -- Address Entry Form

[*Address line 1*]
[*Address line 2* (*optional*)]
[*Town / City*]

State ▲▼

[*PIN code*]

Your PIN code must
contain 6 digits.

End