



Community Experience Distilled

Python Geospatial Development

Second Edition

Learn to build sophisticated mapping applications from scratch using Python tools for geospatial development

Erik Westra

[PACKT] open source*
PUBLISHING community experience distilled

Python Geospatial Development

Second Edition

Learn to build sophisticated mapping applications from scratch using Python tools for geospatial development

Erik Westra

[PACKT] open source 
PUBLISHING community experience distilled

BIRMINGHAM - MUMBAI

Python Geospatial Development

Second Edition

Copyright © 2013 Packt Publishing

All rights reserved. No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without the prior written permission of the publisher, except in the case of brief quotations embedded in critical articles or reviews.

Every effort has been made in the preparation of this book to ensure the accuracy of the information presented. However, the information contained in this book is sold without warranty, either express or implied. Neither the author, nor Packt Publishing, and its dealers and distributors will be held liable for any damages caused or alleged to be caused directly or indirectly by this book.

Packt Publishing has endeavored to provide trademark information about all of the companies and products mentioned in this book by the appropriate use of capitals. However, Packt Publishing cannot guarantee the accuracy of this information.

First published: December 2012

Second Edition: May 2013

Production Reference: 1170513

Published by Packt Publishing Ltd.

Livery Place

35 Livery Street

Birmingham B3 2PB, UK.

ISBN 978-1-78216-152-3

www.packtpub.com

Cover Image by Karl Moore (karl@karlmoore.co.uk)

Credits

Author

Erik Westra

Reviewers

Will Cadell

Richard Marsden

Silas Toms

Acquisition Editor

Kartikey Pandey

Lead Technical Editor

Susmita Panda

Technical Editors

Sharvari Baet

Meenakshi Gupta

Chirag Jani

Project Coordinator

Arshad Sopariwala

Proofreaders

Stephen Silk

Katherine Tarr

Indexers

Hemangini Bari

Rekha Nair

Tejal Daruwale

Graphics

Ronak Dhruv

Abhinash Sahu

Production Coordinator

Manu Joseph

Cover Work

Manu Joseph

About the Author

Erik Westra has been a professional software developer for over 25 years, and has worked almost exclusively in Python for the past decade. Erik's early interest in graphical user-interface design led to the development of one of the most advanced urgent courier dispatch systems used by messenger and courier companies worldwide. In recent years, Erik has been involved in the design and implementation of systems matching seekers and providers of goods and services across a range of geographical areas. This work has included the creation of real-time geocoders and map-based views of constantly changing data. Erik is based in New Zealand, and works for companies worldwide.

I would like to thank Ruth, the love of my life, for all of her support and encouragement.

About the Reviewers

Richard Marsden has over 15 years of professional software development experience. After starting in the field of geophysics and oil exploration, he has spent the last ten years running the Winwaed Software Technology LLC, an independent software vendor. Winwaed specialize in geospatial tools and applications including web applications, and operate the <http://www.mapping-tools.com> website for tools and add-ins for Microsoft's MapPoint product.

Richard also manages the technical aspects of the EcoMapCostaRica.com project for the Biology Department at the University of Dallas. This includes the website, online field maps, field surveys, and the creation and comparison of panoramic photographs.

Richard is also active in the field of natural language processing, especially with Python's NLTK package.

Will Cadell is a principal consultant with Sparkgeo.com. He builds next generation web mapping applications, primarily using Google Maps, geoDjango, and PostGIS. He has worked in academia, government, and natural resources but now mainly consults for the start-up community in Silicon Valley. His passion has always been the implementation of geographic technology and with over a billion smart, mobile devices in the world it's a great time to be working on the geoweb.

Will lives in Prince George, Northern British Columbia, and when he's not writing code or talking about geographic web technology you can find him on a ski hill with his family.

Silas Toms is a GIS programmer with ICF International. His main professional interests are programming in Python and automation of large-scale environmental impact analyses. He has lived in San Francisco for 6 years while finishing a masters in GIS at San Francisco State University on temperature interpolation. He wishes to thank his girlfriend and his family for being supportive of his many interests.

www.PacktPub.com

Support files, eBooks, discount offers and more

You might want to visit www.PacktPub.com for support files and downloads related to your book.

Did you know that Packt offers eBook versions of every book published, with PDF and ePub files available? You can upgrade to the eBook version at www.PacktPub.com and as a print book customer, you are entitled to a discount on the eBook copy. Get in touch with us at service@packtpub.com for more details.

At www.PacktPub.com, you can also read a collection of free technical articles, sign up for a range of free newsletters and receive exclusive discounts and offers on Packt books and eBooks.



<http://PacktLib.PacktPub.com>

Do you need instant solutions to your IT questions? PacktLib is Packt's online digital book library. Here, you can access, read and search across Packt's entire library of books.

Why Subscribe?

- Fully searchable across every book published by Packt
- Copy and paste, print and bookmark content
- On demand and accessible via web browser

Free Access for Packt account holders

If you have an account with Packt at www.PacktPub.com, you can use this to access PacktLib today and view nine entirely free books. Simply use your login credentials for immediate access.

Table of Contents

Preface	1
Chapter 1: Geospatial Development Using Python	7
Python	7
Geospatial development	9
Applications of geospatial development	12
Analyzing geospatial data	12
Visualizing geospatial data	14
Creating a geospatial mash-up	16
Recent developments	17
Summary	19
Chapter 2: GIS	21
Core GIS concepts	21
Location	22
Distance	25
Units	27
Projections	29
Cylindrical projections	29
Conic projections	31
Azimuthal projections	32
The nature of map projections	33
Coordinate systems	34
Datums	36
Shapes	37
GIS data formats	39
Working with GIS data manually	41
Summary	49

Chapter 3: Python Libraries for Geospatial Development	51
Reading and writing geospatial data	51
GDAL/OGR	52
GDAL design	52
GDAL example code	55
OGR design	56
OGR example code	57
Documentation	58
Availability	58
Dealing with projections	59
pyproj	59
Design	60
Proj	60
Geod	61
Example code	62
Documentation	63
Availability	64
Analyzing and manipulating geospatial data	66
Shapely	66
Design	66
Example code	68
Documentation	69
Availability	69
Visualizing geospatial data	70
Mapnik	71
Design	72
Example code	74
Documentation	76
Availability	76
Summary	77
Chapter 4: Sources of Geospatial Data	79
Sources of geospatial data in vector format	80
OpenStreetMap	80
Data format	81
Obtaining and using OpenStreetMap data	82
TIGER	84
Data format	86
Obtaining and using TIGER data	87
Natural Earth	88
Data format	89
Obtaining and using Natural Earth vector data	89

Global, self-consistent, hierarchical, high-resolution shoreline database (GSHHS)	90
Data format	91
Obtaining the GSHHS database	92
World Borders Dataset	92
Data format	93
Obtaining World Borders Dataset	93
Sources of geospatial data in raster format	93
Landsat	94
Data format	95
Obtaining Landsat imagery	95
Natural Earth	98
Data format	99
Obtaining and using Natural Earth raster data	99
Global Land One-kilometer Base Elevation (GLOBE)	100
Data format	100
Obtaining and using GLOBE data	101
National Elevation Dataset (NED)	102
Data format	103
Obtaining and using NED data	103
Sources of other types of geospatial data	105
GEOnet Names Server	106
Data format	107
Obtaining and using GEOnet Names Server data	107
Geographic Names Information System (GNIS)	107
Data format	108
Obtaining and using GNIS Data	109
Choosing your geospatial data source	109
Summary	110
Chapter 5: Working with Geospatial Data in Python	111
Pre-requisites	111
Reading and writing geospatial data	112
Task – calculate the bounding box for each country in the world	112
Task – save the country bounding boxes into a shapefile	114
Task – analyze height data using a digital elevation map	119
Changing datums and projections	126
Task – change projections to combine shapefiles using geographic and UTM coordinates	127
Task – change datums to allow older and newer TIGER data to be combined	132
Representing and storing geospatial data	135
Task – define the border between Thailand and Myanmar	136

Task – save geometries into a text file	140
Performing geospatial calculations	141
Task – identify parks in or near urban areas	141
Converting and standardizing units of geometry and distance	146
Task – calculate the length of the Thai-Myanmar border	147
Task – find a point 132.7 kilometers west of Soshone, California	154
Exercises	156
Summary	158
Chapter 6: GIS in the Database	159
Spatially-enabled databases	159
Spatial indexes	160
Open source spatially-enabled databases	163
MySQL	163
PostGIS	168
Installing and configuring PostGIS	168
Using PostGIS	170
Documentation	172
Advanced PostGIS Features	173
Spatialite	174
Installing Spatialite	174
Installing pysqlite	175
Accessing Spatialite from Python	175
Documentation	176
Using Spatialite	177
Spatialite capabilities	178
Commercial Spatially-enabled database	179
Oracle	180
MS SQL Server	180
Recommended best practices	181
Using the database to keep track of spatial references	181
Using the appropriate spatial reference for your data	183
Option 1 – using a database that supports geographies	184
Option 2 – transform features as required	185
Option 3 – transform features from the outset	185
When to use unprojected coordinates	185
Avoiding on-the-fly transformations within a query	186
Don't create geometries within a query	187
Using spatial indexes appropriately	188
Knowing the limits of your database's query optimizer	189
MySQL	189
PostGIS	191
Spatialite	193

Working with geospatial databases using Python	195
Prerequisites	195
Working with MySQL	195
Working with PostGIS	199
Working with SpatiaLite	201
Comparing the databases	204
Summary	205
Chapter 7: Working with Spatial Data	207
About DISTAL	207
Designing and building the database	211
Downloading the data	216
World Borders Dataset	216
GSHHS	216
GNIS	216
GEOnet Names Server	217
Importing the data	217
World Borders Dataset	218
GSHHS	220
US place name data	222
Worldwide places' name data	227
Implementing the DISTAL application	229
The shared "database" module	232
The "select country" script	233
The "select area" script	236
Calculating the bounding box	241
Calculating the map's dimensions	242
Setting up the data source	245
Rendering the map image	246
The "show results" script	248
Identifying the clicked-on point	249
Identifying features by distance	250
Displaying the results	259
Application review and improvements	262
Usability	262
Quality	264
Place name issues	264
Lat/Long coordinate problems	264
Performance	266
Finding the problem	266
Improving performance	269
Calculating the tiled shorelines	271

Using tiled shorelines	280
Analyzing performance improvement	282
Summary	282
Chapter 8: Using Python and Mapnik to Generate Maps	285
Introducing Mapnik	286
Creating an example map	292
Mapnik in depth	296
Data sources	296
Shapefile	297
PostGIS	298
Gdal	300
Ogr	301
SQLite	302
OSM	303
MemoryDatasource	304
Rules, filters, and styles	304
Filters	305
Scale denominators	307
"Else" rules	309
"Also" rules	309
Symbolizers	310
Drawing lines	310
Drawing polygons	316
Drawing labels	319
Drawing points	329
Drawing raster images	333
Using colors	335
Maps and layers	336
Map attributes and methods	337
Layer attributes and methods	338
Map rendering	339
MapGenerator revisited	341
The MapGenerator interface	342
Creating the main map layer	343
Displaying points on the map	344
Rendering the map	345
What the map generator teaches us	346
Map definition files	346
Summary	350
Chapter 9: Putting It All Together – a Complete Mapping System	353
About ShapeEditor	353
Designing ShapeEditor	357
Importing a shapefile	358

Selecting a feature	360
Editing a feature	362
Exporting a shapefile	362
Prerequisites	362
The structure of a Django application	363
Models	365
Views	366
URL dispatching	366
Templates	369
Setting up the database	371
Setting up the ShapeEditor project	373
Defining ShapeEditor's applications	375
Creating a shared application	375
Defining data models	377
Shapefile	378
Attribute	378
Feature	379
AttributeValue	379
The models.py file	380
Playing with the admin system	384
Summary	391
Chapter 10: ShapeEditor – Implementing List View, Import, and Export	393
<hr/>	
Implementing the "list shapefiles" view	393
Importing shapefiles	398
The "import shapefile" view function	399
Extracting the uploaded shapefile	403
Importing the shapefile's contents	405
Open the shapefile	405
Add the Shapefile object to the database	406
Define the shapefile's attributes	407
Store the shapefile's features	408
Store the shapefile's attributes	410
Cleaning up	413
Exporting shapefiles	414
Defining the OGR shapefile	415
Saving the features into the shapefile	416
Saving the attributes into the shapefile	418
Compressing the shapefile	420
Deleting temporary files	420

Returning the ZIP archive to the user	420
Summary	422
Chapter 11: ShapeEditor – Selecting and Editing Features	423
Selecting a feature to edit	424
Implementing Tile Map Server	425
Setting up the base map	435
Tile rendering	438
Using OpenLayers to display the map	444
Intercepting mouse clicks	449
Implementing the find feature view	454
Editing features	460
Adding features	468
Deleting features	471
Deleting shapefiles	472
Using ShapeEditor	474
Further improvements and enhancements	474
Summary	475
Index	477

Preface

With the explosion of map-based websites and spatially-aware devices and applications, geospatial development is becoming increasingly important. The geospatial market is growing rapidly, and as a Python developer you can't afford to be left behind. In today's location-aware world, all commercial Python developers can benefit from an understanding of geospatial concepts and development techniques.

Working with geospatial data can get complicated because you are dealing with mathematical models of the Earth's surface. Since Python is a powerful programming language with high-level toolkits, it is well-suited to geospatial development. This book will familiarize you with the Python tools required for geospatial development. It introduces basic geospatial concepts with a clear, detailed walkthrough of the key concepts such as location, distance, units, projections, datums, and geospatial data formats. We then examine a number of Python libraries and use these with freely-available geospatial data to accomplish a variety of tasks. The book provides an in-depth look at the concept of storing spatial data in a database and how you can use spatial databases as tools to solve a variety of geospatial problems.

It goes into the details of generating maps using the Mapnik map-rendering toolkit, and helps you to build a sophisticated web-based geospatial map editing application using GeoDjango, Mapnik, and PostGIS. By the end of the book, you will be able to integrate spatial features into your applications and build a complete mapping application from scratch.

This book is a hands-on tutorial. It teaches you how to access, manipulate, and display geospatial data efficiently using a range of Python tools for GIS development.

What this book covers

Chapter 1, Geospatial Development Using Python, gives an overview of the Python programming language and the concepts behind geospatial development. Major use-cases of geospatial development and recent and upcoming developments in the field are also covered.

Chapter 2, GIS, introduces the core concepts of location, distance, units, projections, shapes, datums, and geospatial data formats, before discussing the process of working with geospatial data manually.

Chapter 3, Python Libraries for Geospatial Development, explores the major Python libraries available for geospatial development, including the available features, how the library is organized, and how to install and use it.

Chapter 4, Sources of Geospatial Data, investigates the major sources of freely-available geospatial data, what information is available, the data format used, and how to import the downloaded data.

Chapter 5, Working with Geospatial Data in Python, uses the libraries introduced earlier to perform various tasks using geospatial data, including changing projections, importing and exporting data, converting and standardizing units of geometry and distance, and performing geospatial calculations.

Chapter 6, GIS in the Database, introduces the spatial capabilities of PostGIS, MySQL, and SQLite. It discusses best practices for storing different types of spatial data, and looks at how to access these databases from Python.

Chapter 7, Working with Spatial Data, works through the design and implementation of a complete geospatial application called DISTAL, using freely-available geospatial data stored in a spatial database. It investigates the performance of this application and then works to optimize it using best-practice techniques.

Chapter 8, Using Python and Mapnik to Produce Maps, gives an in-depth look at the Mapnik map-generation toolkit, and how to use it to produce a variety of maps.

Chapter 9, Putting it all Together: a Complete Mapping Application, introduces the "ShapeEditor", a complete and sophisticated web application built using PostGIS, Mapnik and GeoDjango. We start by designing the overall application, and then build the ShapeEditor's database models.

Chapter 10, ShapeEditor: Implementing List View, Import, and Export, continues the implementation of the ShapeEditor system, concentrating on displaying a list of imported shapefiles, along with the logic for importing and exporting shapefiles via a web browser.

Chapter 11, ShapeEditor: Selecting and Editing Features, concludes the implementation of the ShapeEditor, adding logic to let the user select and edit features within an imported shapefile. This involves the creation of a custom Tile Map Server, and the use of the OpenLayers JavaScript library to display and interact with geospatial data.

Bonus chapter, Web Frameworks for Geospatial Development, examines the concepts of web application frameworks, web services, JavaScript UI libraries, and slippy maps. It introduces a number of standard web protocols used by geospatial applications, and finishes with a survey of the tools and frameworks available for building geospatial applications that run via a web interface.

You can download this chapter from: http://www.packtpub.com/sites/default/files/downloads/15230S_Bonuschapter.pdf

What you need for this book

To follow through this book, you will need to have Python Version 2.5 to 2.7. You will also need to download and install the following tools and libraries; full instructions are given in the relevant sections of this book:

- GDAL/OGR
- GEOS
- Shapely
- Proj
- pyproj
- MySQL
- MySQLdb
- SpatiaLite
- pysqlite
- PostgreSQL
- PostGIS
- pycopg2
- Mapnik
- Django