



UNIVERSITI PUTRA MALAYSIA

***ENHANCING GEOGRAPHIC COORDINATES REPRESENTATION
STANDARD FOR REVERSE GEOCODING WEB SERVICES***

MOHAMMED MUSTAFA ABDULRAHMAN AL-HABSHI

FK 2020 89



**ENHANCING GEOGRAPHIC COORDINATES REPRESENTATION
STANDARD FOR REVERSE GEOCODING WEB SERVICES**

By

MOHAMMED MUSTAFA ABDULRAHMAN AL-HABSHI

**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia,
in Fulfilment of the Requirements for the Degree of Doctor of Philosophy**

September 2018

COPYRIGHT

All material contained within the thesis, including without limitation text, logos, icons, photographs, and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

EENHANCING GEOGRAPHIC COORDINATES REPRESENTATION STANDARD FOR REVERSE GEOCODING WEB SERVICES

By

MOHAMMED MUSTAFA AL-HABSHI

September 2018

Chairman : Associate Professor Helmi Zulhaidi Mohd. Shafri, PhD
Faculty : Engineering

Reverse-geocoding is Location Base Services (LBS) that supports mobile and web applications. Using a given geographical coordinates, it provides human-readable addresses and other related information. These days, regardless of Geographic Information System (GIS) and Spatial Data Infrastructure (SDI), reverse-geocoding web services provide the backbone of many of web, mobile, and Internet-of-Things (IoT) applications. These developments in Information and Communication Technologies (ICT) founded standardization diversity, where differences between web services providers and organizations committees, such International Standardization Organization (ISO) and Open Geospatial Consortium (OGC). This research interest in ISO6709 “standard representation of geographic point”. The standard provides guidelines in term of geographic point location representation. While its main concern in having a constancy in data exchange, it lacks providing the basis of decimal degrees digits precision quality measurement. The digit precision has been mentioned in the ISO6709, but the basis of geospatial precision was not mentioned. In addition, the standard doesn’t provide guidelines related impact of precision quality on data storage and internet bandwidth, where the majority of coordinates’ format is text. Therefore, further effective standardization is in need. The research approach to tackle the precision quality of decimal degrees by utilizing European Petroleum Survey Group (EPSG) dataset of Coordinate Reference System and use it to identify the effect of rounding decimal digits in a metric measurement. The research developed a standardized Intuitive Geospatial Relational Query Model (IGRQM) that can be used to determine latest unique Coordinate Reference System (CRS) for a single geographic point. The IGRQM contains of 3341 spatial features with 6526 CRS and was used within a geospatial query process and tested against 100 generated and distributed geographic points with addresses around the globe, where the decimal digits are between 10 and 14. The development process was exploration research to discover patterns of non-unique CRS through qualitative analysis and embed them as rules of CRS exclusions in the IGRQM. The results were each of the 100 generated points have a single latest metric unique CRS, where the distinctive count of CRS is 93. The 100 points with their CRS were used to measure the

impact of decimal digits rounding on metric distance. These measurements summary was validated with World Geodetic System (WGS) 1984 spheroid results. Although the difference doesn't show high variation, the research findings provided an evidence and a method that respect and consider counties developed CRS. In addition, a simple table as guidelines was developed to show the bytes size of the decimal digit in scaling. Benchmarking the guidelines found the data is maintained within about $68\% \pm 1\%$ of the original size without any algorithm compression. This study provides some basis for consideration when addressing geographic standards rather arbitrary.



Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia
sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

MENINGKATKAN PIAWAIAN PERSEMBAHAN KOORDINAT GEOGRAFI UNTUK PERKHIDMATAN WEB PENGEOKODAN TERBALIK

Oleh

MOHAMMED MUSTAFA AL-HABSHI

September 2018

Pengerusi : Profesor Madya Helmi Zulhaidi Mohd. Shafri, PhD
Fakulti : Kejuruteraan

Pengekoda belakang adalah perkhidmatan berasaskan lokasi (LBS) yang menyokong mudah alih dan web aplikasi. Dengan menggunakan koordinat geografi yang telah diberi, ia menyediakan alamat-alamat yang boleh dibaca manusia dan maklumat lain yang berkaitan. Pada masa sekarang, tanpa mengira Sistem Maklumat Geografi (GIS) dan Infrastruktur Data Spasial (SDI), perkhidmatan web pengekode belakang menyediakan tulang belakang banyak aplikasi web, mudah alih dan IoT. Perkembangan-perkembangan dalam teknologi maklumat dan komunikasi (ICT) ini berasaskan kepelbagaian piawaian, di mana perbezaan antara pembekal-pembekal perkhidmatan web dan organisasi Jawatankuasa seperti Organisasi Piawaian Antarabangsa (ISO) dan Konsortium Geospasial Terbuka (OGC). Penarikan penyelidikan ini dalam ISO6709 "perlambatan piawai titik geografi". Piawaian ini menyediakan panduan dari segi perlambatan lokasi titik geografi. Manakala kebimbangan utama perlambatan lokasi titik geografi dalam memiliki penukaran data yang berterusan, ia terdapat kekurangan dalam menyediakan asas ketepatan pengukuran kualiti dalam digit darjah perpuluhan. Ketepatan digit telah disebutkan di dalam ISO6709, tetapi ketepatan bagi dasar geospasial tidak disebut. Tambahan pula, piawaian tersebut tidak menyediakan garis panduan yang berkaitan dengan kesan kualiti ketepatan pada data simpanan dan internet jalur lebar, di mana majoriti format koordinat adalah dalam bentuk teks. Oleh itu, standardisasi yang lebih berkesan adalah amat diperlukan. Pendekatan penyelidikan ini adalah untuk menangani kualiti ketepatan bagi darjah perpuluhan dengan menggunakan dataset Kumpulan Tinjauan Petroleum Eropah (EPSG) dari Menyelaras Sistem Rujukan (CRS) dan menggunakannya untuk mengenal pasti kesan bundaran digit perpuluhan dalam ukuran metrik. Kajian ini telah menghasilkan intuitif Geospasial hubungan pertanyaan Model (IGRQM) yang boleh digunakan untuk menentukan CRS unik yang terkini untuk satu titik geografi. IGRQM mengandungi 3341 ciri-ciri spasial dengan 6526 CRS dan telah digunakan dalam proses pertanyaan geospasial dan diuji dengan menggunakan 100 titik geografi yang telah dijana dan titik-titik geografi bersama alamat yang telah diedarkan di seluruh dunia, di mana digit perpuluhannya adalah di antara 10 dan 14. Proses penghasilan adalah penerokaan penyelidikan untuk mengetahui pola-

pola CRS yang tidak unik melalui analisis kualitatif dan menerapkannya sebagai peraturan pengecualian CRS dalam IGRQM ini. Keputusan di mana setiap 100 titik yang dijana mengandungi satu metrik CRS unik, di mana kiraan CRS yang tersendiri adalah 93. 100 titik dengan CRS mereka telah digunakan untuk mengukur kesan ke atas bundaran angka perpuluhan pada jarak metrik. Ringkasan daripada pengukuran-pengukuran ini telah disahkan dengan menggunakan keputusan spheroid Sistem Geodetik Dunia 1984 (WGS84). Walaupun perbezaannya tidak menunjukkan perubahan yang tinggi, hasil penyelidikan ini telah menyediakan bukti dan kaedah yang menghormati dan mempertimbangkan daerah-daerah yang telah membangunkan CRS. Di samping itu, satu jadual mudah sebagai panduan telah dihasilkan untuk menunjukkan saiz bait angka perpuluhan dalam bentuk skala. Penanda aras bagi panduan telah menunjukkan data tersebut adalah sama dalam kira-kira $68\% \pm 1\%$ daripada saiz asal tanpa sebarang pemampatan algoritma. Kajian ini menyediakan beberapa asas untuk dipertimbangkan apabila mengamalkan piawaian geografi dan tidak melakukannya secara sewenang-wenangnya.

ACKNOWLEDGEMENTS

First, I would like to express my appreciation to all those who contributed to the conversation that has helped this thesis come into being. Numerous people from the Geospatial Information Science Research Centre, Department of Computer and Communication Engineering, Department of Biological and Agricultural Engineering and Department of Civil Engineering, in particular Geomatics Engineering Unit. I would like to thank Affendi Suhaili, Goma Ahmad, Bala Jahun Gambo, Moahammed Idris, Tajuddeen Abiodun Adelodun, Habib Ahmad and Nisreen Sofo for their cooperation and suggestions.

I would like to thank my former supervisor – Assoc.Prof. Dr. Ahmad Rodzi and my current supervisor Assoc.Prof. Dr. Helmi Zulhaidi Mohd. Shafri and other co-supervisors Prof. Dr. Noordin Ahmad and Prof. Dr. Abdul Rashid b. Mohamed Shariff – and to other members of the Geomatic unit – Prof. Dato Dr. Shattri Mansor- for continued guidance, specific comments provided for various drafts. I am also grateful to those who volunteered to assist me in collect and share the data used in the thesis, notably.

Finally, I owe a huge debt of gratitude to all those who have given me the support and encouragement to prioritise this research above all other competing demands, and who were then duly understanding when I was absorbed in my work. To my family in Yemen, My Father who pass away (2014), My Mother and My Brother and his family, I'd like to say thanks for the consistent and unconditional support that has helped me get to this point many, many years to come.

This thesis was submitted to the Senate of the Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Helmi Zulhaidi bin Mohd Shafri, PhD

Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Noordin Ahmad, PhD

Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Abdul Rashid bin Mohamed Shariff, PhD

Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

ZALILAH MOHD SHARIFF, PhD

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 13 August 2020

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	iii
ACKNOWLEDGEMENTS	v
APPROVAL	vi
DECLARATION	viii
LIST OF TABLES	xii
LIST OF FIGURES	xiii
LIST OF APPENDICES	xv
LIST OF ABBREVIATIONS	xvi
 CHAPTER	
 1 INTRODUCTION	 1
1.1 Background	1
1.2 Justification	2
1.3 Problem Statement	2
1.4 Research Aim	3
1.5 Research Questions	3
1.6 Objectives	4
1.7 The Knowledge Gap and Research Contribution	4
1.8 Research Scope and Limitation	4
1.9 Thesis Layout	5
 2 LITRATURE REVIEW	 6
2.1 Introduction	6
2.2 LBS and the Role of Reverse Geocoding Web Service	6
2.3 Aspects of Reverse Geocoding service	7
2.3.1 Input Parameters in the Request	7
2.3.2 Returned Results in Response	8
2.3.3 Software Implementation Nature	8
2.4 Online and Public Reverse Geocoding web services	9
2.4.1 Google Geocoding API	12
2.4.2 Yahoo PlaceFinder	12
2.4.3 Microsoft Bing Location API	13
2.4.4 ESRI ArcGIS API	14
2.4.5 Nominatim Open Street Map (OSM)	17
2.4.6 Geonames	17
2.4.7 Here API	18
2.5 Standardization Issues with Reverse Geocoding	20
2.5.1 Current Related Standards to Reverse-Geocoding and Their Limitations	21
2.5.2 ISO 6709:2008 “Standard Representation of Geographic Location by Coordinates”	23
2.6 Other’s related works	27
2.7 Summary	28

3	METHODOLOGY	29
3.1	Introduction	29
3.2	Developing a Geospatial Query Model for Determining Latest Unique Metric CRS	31
3.2.1	The EPSG dataset with the geospatial features and the establishment of its database	31
3.2.2	Establishing Initiative Geospatial Relational Query Model (IGRQM)	35
3.2.3	Validating the IGRQM	35
3.3	Resolving Divergences Integrations of EPSG Dataset Attributes with its Geospatial Feature	38
3.4	Generating random bounded points with Diversity of location addresses	39
3.5	Discovering Rules of Uniqueness and integrating them within Geospatial IGRQM	43
3.6	Measuring Digits Efficiency on Distance Differences between the Ellipsoid CRS and Latest Unique Metric CRS	43
3.7	Benchmarking the Efficiency of decimal degrees digits data size	45
3.8	Summary	47
4	RESULTS AND DISCUSSION	48
4.1	Introduction	48
4.2	The Initiative Geospatial Relational Query Model (IGRQM)	48
4.3	IGRQM qualitative validation through similarities and differences between it and EPSG-registry results	50
4.4	The optimized IGRQM model test with 100 random bounded points	57
4.5	CRS Classification based Geographic points' distribution	62
4.6	Understanding the digits nature of the generated random bounded points	64
4.7	Measurements for Standardization Digits Requirements for Reverse Geocoding Service	65
4.7.1	Pairing Data Size and digits precision	66
4.7.2	Benchmarking the guidelines	68
4.8	Summary	69
5	CONCLUSIONS	70
5.1	Introduction	70
5.2	Findings and Achievements	70
5.3	Contributions and Novelty	71
5.4	Future Work	72
5.5	Recommendation	73
	REFERENCES	74
	APPENDICES	86
	BIODATA OF STUDENT	101
	LIST OF PUBLICATIONS	102

LIST OF TABLES

Table	Page
2.1 Summary of advantages and disadvantages of using online reverse geocoding services	9
2.2 Geocoded results for Yahoo, Google and ESRI geocoding Services	10
2.3 Alignment of all accuracy-levels, leading to normalized accuracy-ranking	11
2.4 Subset of the responses metadata related to location address	13
2.5 Yahoo Street & XStreet metadata in reverse geocoding responses	13
2.6 Microsoft Bing Location Address components	14
2.7 ArcGIS Online World Geocoding service's Location Address components	15
2.8 Corrected confrontation among geocoders' positioning results	18
3.1 EPSG dataset (non-geospatial) with records count	33
3.2 The Unique CRS with some of relational data for GISRC location	39
4.1 The one point with three CRS located in Canada	59
4.2 First point of two that have two CRS located in Sri Lanka	60
4.3 Second point of two that have two CRS located in Turkey	61
4.4 Points count <i>per</i> CRS	63
4.5 Summary of distance measurement between points <i>per</i> CRS	64
4.6 Decimal degrees presentation precision measurements up 10 digits based on latest unique metric CRS	65
4.7 Decimal degrees presentation precision measurements up 8 digits Based on WGS84 spheroid	66
4.8 The proposed guideline table for decimal degree digits precisions and data size for geographic coordinate's textual presentation	67
4.9 Benchmark of the guideline with accuracy of one centimeter (7 digits)	68

LIST OF FIGURES

Figure		Page
2.1	Role of reverse geocoding in the OpenLS information Model	7
2.2	Conventional Reverse Geocoding vs Online Web Services	10
2.3	Screenshot of ArcGIS Online Geocoding Service Coverage	15
2.4	Overview of Response Schema	19
2.5	Overview of Location Type Response schema	19
2.6	Standards Framework, courtesy of OGC (2008)	21
2.7	UML Model for coordinates representation, (ISO/TC 211, 2009)	25
3.1	Research Methodology	30
3.2	Approach for developing a Query Model to determine latest CRS	32
3.3	ESPG extracted Entity Relationship between tables	34
3.4	ESPG shapefile schema and PostGIS SRS schema	34
3.5	Method to count the differences and similarities in each Table and domain accuracy to point	36
3.6	ESPG Registry Scraping process using XPath and Regex	37
3.7	Algorithm flow diagram for Generating Geographic random bounded point with a valid address	42
3.8	Illustration of the approach to measure digits efficiency	44
3.9	Approach for measuring decimal degree digits efficiency based on the distributed latest unique CRS	45
3.10	The approach to benchmark the guidelines and the algorithm development for applying the guidelines	46
3.11	Diagram of a regular expression of decimal degrees	47
4.1	Initiative Geospatial Relational Query Model (IGRQM)	49
4.2	Similarities and differences between EPSG-registry and IGRQM in term of kind	51

4.3	Classification of differences based on coverage domain reclassification	52
4.4	Sunburst graph summarizing IGRQM columns, their distinctive attributes count and whether to drop them in term of evolving the model	53
4.5	The SQL code of IGRQM first pre-processing phase	55
4.6	Sunburst graph summarizing IGRQM fields, their distinctive attributes count after the first preprocessing stage	56
4.7	The updated IGRQM distinctive attributes count with results summary of Step two and three	56
4.8	Results summary pie charts for samples points CRS query using IGRQM data model for the first trial (described as step one) and the second trial (known as step two)	58
4.9	The last three CRS names case that are assigned to a single point	59
4.10	The Third step to determine Latest Unique CRS for the last three points	61
4.11	Random Geographic points distribution based on Mersenne Twister Algorithms	62
4.12	CRS classification based the random geographic points and EPSG dataset	63
4.13	Measurement of Decimal digits variation among the 100 points latitude and longitude	65
4.14	Illustration of digit precision that is beyond the need of cadastral mapping, using CRS:2801 (NAD83(HARN) / Louisiana South), USA, LA. Where (a) showing all variations of digits location with a residential property, (b), (c), (d) are zooming respectively showing that extra digits in such coordinates are meaningless	67
4.15	Comparison of geospatial text files' size before and after applying the guidelines	68
4.16	The efficiency of applying the guidelines	69

LIST OF APPENDICES

Appendix	Page
A Hardware And Software Used for Development	86
A 1 Hardware	86
A 2 Software	86
B SQL Codes for Creating the Model, Analysis	87
B 1 The IGRQM with fixed criteria	87
B 2 The SQL query model with geospatial Criteria	87
B 3 The developed and Implemented Geospatial Query Model for Determining Latest Unique Metric CRS	88
B 4 Query for Measurement variation of Decimal digits among the 100 points Latitude and Longitude	89
B 5 Query Model for Measuring rounding Error using PostGIS geospatial algorithms	90
C Python scripts	93
C 1 Generating the Random bounded point	93
C 2 Developing the optimization algorithm and benchmarking it	93
D Table of Generate Random Geographic Points	99

LIST OF ABBREVIATIONS

ADT	Abstract Data Types
API	Application Programmable Interface
BIM	Building Information Model
CGI	Common Gateway Interface
CRS	Coordinate Reference System
DHCP	Dynamic Host Configuration Protocol
EPSG	(European Petroleum Survey Group) standards for coordinate reference system
EIF	European Interoperability Framework
G2C	Government Service To Citizen
GCI	Geocode Certainty Indicator
GIS	Geographical Information System
GNSS	Global Navigation Satellite System
GPL	General Public License
GPS	Global Positioning System
GUI	Graphical User Interface
HTTP	Hyper Text Transfer Protocol
HTTPS	HTTP secured
ICT	Information and Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IGRQM	Geospatial Relational Query Model
INSPIRE	Infrastructure for Spatial Information in the European Community
IP	Internet Protocol
ISO	International Standardization Organization
ISO/AWI	ISO Approved Working Item

ISO/CD	ISO Committee Draft
ISO/DIS	ISO Draft international Standard
ISO/WD	ISO working Draft
ISO/TC	ISO Technical Committee
ISO/TR	ISO Technical Report
LBA	Location Based Application
LBS	Location Based Services
LD	Linked Data
LDW	Linked Data Web
LCBS	Location and Context-Based Services
LGPL	Lesser General Public License
LOC	Line Of Code
OASIS	Organization for the Advancement of Structured Information Standards
OGC	Open Geospatial Consortium
OMA	Open Mobile Alliance
ORDBMS	Object Relational Database Management System
OS	Operating Systems
OSGeo	Open Source Geospatial Foundation
OSS	Open Source Software
OWA	Ordered Weighted Averaging
OWL	OWL Web Ontology Language
MOSS4G	Mobile Open Source Software for Geoinformatics
MRDB	Multi Representation database
NWIP	New work item proposal
PHP	Hypertext Pre-processor (a web programming Language)
POI	Point Of Interest

RDBMS	Relational Database Management System
RDF	Resource Description Framework
REST	Representational State Transfer
RSO	Rectified Skewed Orthomorphic coordinate system
SDI	Spatial Data Infrastructure
SQL	Structured Query Language
SRID	Spatial Reference IDentifier
SRS	Spatial Reference System
TCP/IP	Transfer Control Protocol / Internet Protocol
TOCG	ISO/ TC211- OGC coordination group
UML	Unified Modelling Language
UPU	Universal Postal Union
URI	Unified Resource Identifier
URL	Unified Resource Locator
US FGDC	United State Federal Geographic Data Committee
UTM	Universal Transverse Mercator coordinate system
VGI	Volunteered Geographic Information
W3C	World Wide Web Consortium
WGS84	World Geodetic System 1984 coordinate system
XML	eXtensible Markup Language
XSD	XML Schema Definition
\forall	for all; for any; for each
\Rightarrow	implies

CHAPTER 1

INTRODUCTION

1.1 Background

Within the last decade, Location Based Services (LBS) have seen rapid development in both web and mobile technologies, (Ahson & Ilyas, 2011; Brimicombe & Li, 2009; Gartner & Ortog, 2011; Gartner & Rehrl, 2008; Karimi, 2013; Kolodziej & Hjelm, 2006; Krisp, 2013; Küpper, 2005; Schiller & Voisard, 2004; Uhler & Mehta, 2013).

To encourage Location-Based Application (LBA) development and utilize location-aware sensors, mobile systems vendors and LBS providers announces competitions hoping for wider adoption of the technology. Nowadays, researchers in LBS have extended the research to include social network and augmented reality hoping to understand users behaviours and to widen applications variates (Brimicombe & Li, 2009; Gartner & Ortog, 2011; Gartner & Rehrl, 2008; Karimi, 2013; Krisp, 2013; Uhler & Mehta, 2013). However as researchers try to find new challenges in augmenting LBS technical aspects and adoption in these days community and technology vendors, there is limited research on quality issues of geospatial web-based data sources of these services have been granted publicly (Brimicombe & Li, 2009; Gartner & Ortog, 2011; Goldberg, 2013; Krisp, 2013; Uhler & Mehta, 2013).

Quality issues and measurement still have ambiguity. The ambiguity includes issues of accuracy, precision, coverage, missing data, in-correctness, fitness-for-user and improper definition metadata. Though it seems these data and metadata are a machine generated, its issues of ambiguity are mostly human error made, and therefore they are known as under-reporting.

Under-reporting is not new to researchers and investigators where a lot multiple disciplines of science and research publications mentioned it as issues in data collection and statistical analysis. It effects statistical analysis that is required to provide an evidence of phenomenon or fact existence. Such evidence assist decision makers, planners and strategists in selecting the right choice, and standardized the resources and the path for a national development. In this context, decision makers and planners that make use of LBS are casual mobile phone and internet users that use theses service in their daily life; therefore, it impacts on the public.

Reverse-Geocoding service is one the core LBS services that is based on the request-response, also known as request-reply and messaging pattern. It requires spatial data with attributes to extract textual address. The data, either licensed or public domain, are being served by web services providers. However, programmers and users use information

from these providers as an authoritative source without knowing the correctness of the information. The problem can be realized when switching from one service provider to another; as result, for a single point, one provider's Reverse-Geocoding service has many differences from the other. The standardization effort of reverse-Geocoding Services was addressed by Open Geospatial Consortium (OGC) as core service of OpenGIS® Location Service (OpenLS); however, the standards cover the service quality aspects and not the data. These cases need a study to provide a standardized solution.

1.2 Justification

A lot of people these days uses internet web and mobile connected devices to access information that supports their daily life activities. However, they are not aware neither concern of some errors and missing data as they use it by casual means. In contrast, mobile and web application developers link particular parts of data from different sources that present an application model to deliver information of interest to end-users the case of LBS nowadays, the link starts with one of location-awareness enabling technologies connected to reverse geocoding service and further to the application model.

Reverse geocoding services depend on spatial data; and if it is different in content and metadata, it may lead to wrong addresses; within the application model, and it may mislead the end-users with a different direction. Until this study, there are sparse researches and publications on errors in reverse geocoding services this study as under-reporting.

1.3 Problem Statement

Public reverse geocoding web services for LBA and LBS have received acceptance of majority of web and mobile developers, as it is believed that its data quality and current standards (de jure and de facto) is sufficient for LBA and LBS development with its readable representation models (Lisboa-Filho et al., 2013; Krisp, 2013; Yue, 2013; Kresse et al., 2012; ISO/TC 211, 2009).

Though Reverse-Geocoding with its aspects (request and response) was addressed in OGC OpenLS standard with XML schema template, however, the representation models of data were not meant for automated computer processing as in Semantic web (Knapena et al., 2013). In addition, a computer could not resolve address ambiguity it requires more data collection and effort from the programmers, which is inefficient (Gaitanis and Winter, 2014; Garcia-Rojas et al., 2013; Knapena et al., 2013). Aspect of address presentation, abbreviations, incomplete data and granularity of dataset provide an obstacles in providing quality data (Dang et al., 2013; Zandbergen et al., 2012; Goldberg, 2010; Behr, 2010) .

These limitations provide prevent of having seamless integration with other services. Where such technological trends such semantic web and Internet of Things (IoT) requires data and devices to be inter-linked seamlessly (Behkamal, 2014; Fetahu et al., 2014; Assaf and Senart, 2012).

The effectiveness of quality control on request and response were ignored leaving it as user's preference with limited of guidelines, especially for public services till year 2018 (Mabrouk, 2008). For example, coordinates presentation was addressed in ISO 6709:2008, it didn't provide guidelines nor refer to other established standard (ISO/TC 211, 2008a), where geo-precision is either overused and underused by users (Nelson, 2013; Loohuis, 2010), and optimizing its digits needed for tuning data storage and transaction (Nelson, 2013; Harding, 2006; Vauglin, 2002). Though coordinates geo-precision accuracy was determined for distances measurement, (Wikipedia, 2015; Nelson, 2013; Bessel et al., 2010; Cook, 2008), based on WGS84 datum and oblate spheroid earth model (Acree, 2014; Bénard, 2014), local and governmental authorities still prefer using national coordinate system (Wang, 2013; Dima and Covaciu, 2012).

The standardization of geo-precision for public reverse geocoding web service requests, thus, requires accuracy comparison between spheroid model and latest developed national coordinate system, where its development progress varies from one region to another and selection criteria are heterogeneous. The solution to these issues requires improving reverse geocoding web service aspects standards (request and response). The request aspect requires guidelines in tuning geo-precision considering WGS84 and national projected coordinate system; whereas, the response requires dynamic representation models with linked data that consider national jurisdictions divisions.

1.4 Research Aim

From the problem statement point of view, it is clear the main objective is to improve a standard for reverse geocoding services aspects for a street level quality. The standard should be able to address possible aspects of under-reporting issues with concerns of coordinates precision and differences of administrative jurisdiction divisions' representation. To do so, it is required to define research questions that can quantify aspects of reverse geocoding.

1.5 Research Questions

Therefore, the questions are as follows:

1. What is the specific relation between reverse geocoding web services, ISO6709 and OGC OpenLS?
2. What are the possible difference in metric distance measurements between spheroid earth model and any national coordinates system based on geo-precision to justify in reverse-geocoding web service?
3. How to measure decimal degrees digit precision having variation of multiple CRS?
4. What is the relationship between decimal degrees digit precision quality, and data size (storage and internet transition)?

From these questions, and to answer them, specific objectives are addressed in the next section.

1.6 Objectives

The research aims and questions address specific objectives that should be achieved in this research as follows:

- i) To develop and standardized an Initiative Geospatial Relational Query Model (IGRQM), dataset and process, for determining latest unique metric Coordinate Reference System (CRS) for a single geographic point.
- ii) To develop an algorithm for generating random bounded sample points by valid address to be tested with IGRQM, where each point must have one CRS.
- iii) To measure decimal degrees precision accuracy based on Multiple CRS and rounding digits, validate the results against WGS 84 decimal degrees precision and benchmark the data size of the geospatial textual file.

1.7 The Knowledge Gap and Research Contribution

Such research questions may lead to defining standards specifically for Reverse-Geocoding since some of them is based on weak compromise without conformance testing (Cooper, 2012; Kresse et al., 2012). In this regards, OGC OpenLS –Location Utility Service standard have shortage where its basis are made upon use case scenarios; in addition, the technical aspects focus on Extendable Markup Language (XML) and its derivatives.

1.8 Research Scope and Limitation

The research tries to investigate issues of address model in two-dimension coordinates, as well outdoor geocode. Three-dimensional coordinates and indoor geocode requires more detailed data, handling privacy topics and possible integration with Building Information Model (BIM), which is out of the scope of this research.

The data collection of addresses are made using public geo-web services providers. Though it is possible to employ Volunteered Geographic Information (VGI), setting up the platform and handling unresolved issues of quality and reliability imposes challenges of validation, correction and verification which is out of the scope.

1.9 Thesis Layout

This thesis is organized in five chapters: Introduction, Literature Review, Methodology, Results and discussion, and conclusion. The objectives of this research were explored with some limitations due the nature of the findings. The objectives were addressed in methodology and its outcome were debated in the results and discussion. The last chapter conclude the finding and achievement.



REFERENCES

- Abdelmoty, A.I., Al-Muzaini, K.M., 2016. Reasoning with Place Information on the Linked Data Web, in: ALLDATA 2016. Presented at the The Second International Conference on Big Data, Small Data, Linked Data and Open Data, IARIA, Lisbon, Portugal, pp. 48–53.
- Acree, E.S., 2014. Method and apparatus of taking aerial surveys. US8693806 B2.
- Aggarwal, N., Panag, T.S., Ryait, H.S., 2013. Optimization of Energy Utilization in Future Internet: A Study of Existing Approaches and a Proposed Conceptual Framework based upon Neural Networks. *IJCA* 66, 31–35.
- Ahlgren, P., Persson, O., Tijssen, R., 2013. Geographical distance in bibliometric relations within epistemic communities. *Scientometrics* 95, 771–784. <https://doi.org/10.1007/s11192-012-0819-1>
- Almeida, P.D., Rocha, J.G., Ballatore, A., Zipf, A., 2016. Where the Streets Have Known Names. Presented at the International Conference on Computational Science and Its Applications, Springer International Publishing, pp. 1–12. https://doi.org/10.1007/978-3-319-42089-9_1
- Assaf, A., Senart, A., 2012. Data quality principles in the semantic web, in: *Semantic Computing (ICSC)*, 2012 IEEE Sixth International Conference On. IEEE, pp. 226–229.
- Ballatore, A., Wilson, D.C., Bertolotto, M., 2013. A Survey of Volunteered Open Geo-Knowledge Bases in the Semantic Web, in: Pasi, G., Bordogna, G., Jain, L.C. (Eds.), *Quality Issues in the Management of Web Information*, Intelligent Systems Reference Library. Springer Berlin Heidelberg, pp. 93–120.
- Behkamal, B., 2014. Metrics-Driven Framework for LOD Quality Assessment, in: Presutti, V., d'Amato, C., Gandon, F., d'Aquin, M., Staab, S., Tordai, A. (Eds.), *The Semantic Web: Trends and Challenges*, Lecture Notes in Computer Science. Springer International Publishing, pp. 806–816.
- Behr, F.-J., 2010. Geocoding: Fundamentals, Techniques, Commercial and Open Services, in: *AGSE 2010*. Presented at the Applied Geoinformatics for Society and Environment, pp. 111–122.
- Beldar, P., Bansode, P., Mane, R., Gaikwad, S., 2014. Traveler Guide using GPS.
- Bénard, P., 2014. An oblate-spheroid geopotential approximation for global meteorology. *Q.J.R. Meteorol. Soc.* 140, 170–184. <https://doi.org/10.1002/qj.2141>
- Bertolotto, M., Giovanni, P.D., Sebillo, M., Vitiello, G., 2014. Standard-Based Integration of W3C and GeoSpatial Services: Quality Challenges, in: Casteleyn, S., Rossi, G., Winckler, M. (Eds.), *Web Engineering*, Lecture

- Notes in Computer Science. Springer International Publishing, pp. 460–469. https://doi.org/10.1007/978-3-319-08245-5_31
- Bessel, F.W., Karney, C.F.F., Deakin, R.E., 2010. The calculation of longitude and latitude from geodesic measurements. *Astronomische Nachrichten* 331, 852–861. <https://doi.org/10.1002/asna.201011352>
- Bill, R., 2008. Spatial Data Infrastructures For Emergency Services And Security Management, in: Coskun, H.G., Cigizoglu, H.K., Maktav, M.D. (Eds.), *Integration of Information for Environmental Security*, NATO Science for Peace and Security Series C: Environmental Security. Springer Netherlands, pp. 233–245.
- Bodbyl-Roels, S., Peterson, A.T., Xiao, X., 2011. Comparative analysis of remotely-sensed data products via ecological niche modeling of avian influenza case occurrences in Middle Eastern poultry. *International journal of health geographics* 10, 1–12.
- Bue, A.L., Machì, A., 2015. Open Data Integration Using SPARQL and SPIN: A Case Study for the Tourism Domain, in: Gavanelli, M., Lamma, E., Riguzzi, F. (Eds.), *AI*IA 2015 Advances in Artificial Intelligence*, Lecture Notes in Computer Science. Springer International Publishing, pp. 316–326. https://doi.org/10.1007/978-3-319-24309-2_24
- Cetl, V., Kliment, T., Jogun, T., 2016. A comparison of address geocoding techniques – case study of the city of Zagreb, Croatia. *Survey Review* 0, 1–10. <https://doi.org/10.1080/00396265.2016.1252517>
- Chen, M., Zhang, X., Green, E., 2015. Analysis of Historical Travel Time Data.
- Chentout, K., Vaisman, A., 2013. Adding Spatial Support to R2RML Mappings, in: *On the Move to Meaningful Internet Systems: OTM 2013 Workshops*. Presented at the OTM Confederated International Conferences “On the Move to Meaningful Internet Systems,” Springer, Berlin, Heidelberg, pp. 398–407.
- Choudhary, P., Anand, M.R., 2015. Determination of Rate of Degradation of Iron Plates due to Rust using Image Processing. *IJER* 4, 76–84.
- Chow, T.E., Dede-Bamfo, N., Dahal, K.R., 2016. Geographic disparity of positional errors and matching rate of residential addresses among geocoding solutions. *Annals of GIS* 22, 29–42. <https://doi.org/10.1080/19475683.2015.1085437>
- Christensen, J.H., 2009. Using RESTful Web-services and Cloud Computing to Create Next Generation Mobile Applications, in: *Proceedings of the 24th ACM SIGPLAN Conference Companion on Object Oriented Programming Systems Languages and Applications, OOPSLA '09*. ACM, New York, NY, USA, pp. 627–634. <https://doi.org/10.1145/1639950.1639958>
- Cook, J.D., 2008. Finding distances using latitude and longitude. *The Endeavour*. URL <http://www.johndcook.com/blog/2008/09/11/finding-distances-using-latitude->

and-longitude/ (accessed 6.4.15).

Cooper, S.C.A.K., 2012. Opportunities for research and innovation from involvement in standards development—Experiences of two researchers.

Crandall, D.J., Backstrom, L., Cosley, D., Suri, S., Huttenlocher, D., Kleinberg, J., 2010. Inferring social ties from geographic coincidences. *PNAS* 107, 22436–22441. <https://doi.org/10.1073/pnas.1006155107>

Cui, Y., 2013. A systematic approach to evaluate and validate the spatial accuracy of farmers market locations using multi-geocoding services. *Applied Geography* 41, 87–95. <https://doi.org/10.1016/j.apgeog.2013.03.016>

Dang, K.T., Trong Phan, N., Chan Ngo, N., 2013. An OpenLS privacy-aware middleware supporting location-based applications. *International Journal of Pervasive Computing and Communications* 9, 311–345.

Danko, D., 2014. Overview of ISO 19115-1 (what is new!). Presented at the ISO/TC 211 Standards in Action Workshop, DIN Burggrafenstraße 6, 10787 Berlin.

Dasgupta, A., 2013. Has OGC Lost its Way? [WWW Document]. URL <http://geospatialworld.net/Paper/Opinion/ArticleView.aspx?aid=30557> (accessed 12.27.15).

Decimal degrees, 2017. . Wikipedia, the free encyclopedia.

Deidda, M., Pala, A., Vacca, G., 2013. An example of a tourist location-based service (LBS) with open-source software. *Appl Geomat* 5, 73–86. <https://doi.org/10.1007/s12518-012-0097-x>

DeLuca, P., Kanaroglou, P.S., 2015. An Assessment of Online Geocoding Services for Health Research in a Mid-Sized Canadian City. *Spatial Analysis in Health Geography* 31.

Demchik, V., 2011. Pseudo-random number generators for Monte Carlo simulations on ATI Graphics Processing Units. *Computer Physics Communications* 182, 692–705.

Dima, D., Covaciu, D., 2012. Improved Methods To Measure And Generate 3D Road Profile In Concerning With Simulation Errors. Presented at the 21th EVU Congress, EVU Grup, Braşov, România.

ESRI, 2017. Geocode coverage—ArcGIS REST API: World Geocoding Service | ArcGIS for Developers [WWW Document]. URL <https://developers.arcgis.com/rest/geocode/api-reference/geocode-coverage.htm> (accessed 1.22.17).

Farrelly, G., 2012. The role of location-based services in shaping sense of place. *Proceedings of the American Society for Information Science and Technology*

- Fetahu, B., Dietze, S., Nunes, B.P., Casanova, M.A., Taibi, D., Nejdl, W., 2014. A Scalable Approach for Efficiently Generating Structured Dataset Topic Profiles, in: Presutti, V., d’Amato, C., Gandon, F., d’Aquin, M., Staab, S., Tordai, A. (Eds.), *The Semantic Web: Trends and Challenges*, Lecture Notes in Computer Science. Springer International Publishing, pp. 519–534.
- Fugard, A.J.B., Potts, H.W.W., 2015. Supporting thinking on sample sizes for thematic analyses: a quantitative tool. *International Journal of Social Research Methodology* 18, 669–684. <https://doi.org/10.1080/13645579.2015.1005453>
- Gaffney, D., 2015. Final Paper: Exploring the Role of Geographic Diffusion in Informational Cascades. Presented at the Computational Social Science Summit 2015 (CSSS15), Kellogg School of Management, Northwestern University, Evanston, Illinois.
- Gaitanis, H., Winter, S., 2014. Is a Richer Address Data Model Relevant for LBS?, in: Liu, C. (Ed.), *Principle and Application Progress in Location-Based Services*. Springer International Publishing, Cham, pp. 121–137.
- Garcia Paje, V., 2015. Improving an Open Source Geocoding Server (Master). Lund University, Lund, Sweden.
- Garcia-Rojas, A., Athanasiou, S., Lehmann, J., Hladky, D., 2013. GeoKnow: leveraging geospatial data in the Web of data. *Open Data on the Web (ODW13)*.
- Geonames, n.d. Reverse Geocoding Webservices [WWW Document]. URL <http://www.geonames.org/export/reverse-geocoding.html> (accessed 8.2.17).
- Goedertier, S., Loutas, N., Peristeras, V., 2014. Linked Location Data as a Service. Presented at the Joint W3C/OGC workshop: Linking Geospatial Data, Google Campus London, Shoreditch, pp. 1–5.
- Goldberg, D.W., 2010. Spatial approaches to reducing error in geocoded data (Doctor of Philosophy). University of Southern California.
- Good Calculators, 2018. Sample Size Calculator - Good Calculators [WWW Document]. Good Calculators. URL <https://goodcalculators.com/sample-size-calculator/> (accessed 3.25.19).
- Google, 2017. Map Coverage Details | Google Maps APIs [WWW Document]. Google Developers. URL <https://developers.google.com/maps/coverage> (accessed 1.12.17).
- Grütter, R., Purves, R.S., Wotruba, L., 2017. Evaluating Topological Queries in Linked Data Using DBpedia and GeoNames in Switzerland and Scotland. *Trans. in GIS* 21, 114–133. <https://doi.org/10.1111/tgis.12196>

- Guzzetti, F., Pasquinnelli, A., Viskanic, P., 2015. Geographic information for the management of flood risk insurance. Presented at the FOSS4G Europe Como 2015, Como, Italy, pp. 705–716.
- Haran, M., 2011. Gaussian random field models for spatial data. *Handbook of Markov Chain Monte Carlo* 449–478.
- Harding, J., 2006. Vector Data Quality: A Data Provider’s Perspective, in: Devillers, R., Jeansoulin, R. (Eds.), *Fundamentals of Spatial Data Quality*. ISTE, pp. 141–159.
- Harvey, F., 2014. We Know Where You Are. And We’re More and More Sure What That Means, in: *Emerging Pervasive Information and Communication Technologies (PICT)*. Springer, pp. 71–87.
- Here, n.d. Search Response - Geocoder API - HERE Developer [WWW Document]. URL <https://developer.here.com/documentation/geocoder/topics/resource-type-response-geocode.html>
- Heumann, B.W., Walsh, S.J., McDaniel, P.M., 2011. Assessing the application of a geographic presence-only model for land suitability mapping. *Ecological Informatics* 6, 257–269. <https://doi.org/10.1016/j.ecoinf.2011.04.004>
- Holsinger, D., 2016. Does Google license maps from Navteq? Quora.
- Horanont, T., Prapakornpilai, J., Phithakkitnukoon, S., Witayangkurn, A., Shibasaki, R., 2014. Space Profile-Based Reverse Geocoding Service Using Cloud Platform, in: *Services Computing (SCC), 2014 IEEE International Conference On*. IEEE, pp. 842–843.
- Hu, Y., Janowicz, K., Prasad, S., Gao, S., 2015. Metadata Topic Harmonization and Semantic Search for Linked-Data-Driven Geoportals: A Case Study Using ArcGIS Online. *Transactions in GIS* 19, 398–416. <https://doi.org/10.1111/tgis.12151>
- IOGP Geomatics Committee [WWW Document], 2017. URL <http://www.epsg.org/> (accessed 4.29.17).
- ISO, 2016. Publicly Available Standards [WWW Document]. URL <http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html> (accessed 8.7.16).
- ISO/TC 211, 2009. *Standards Guide: ISO/TC 211 Geographic Information/Geomatics*. ISO.
- ISO/TC 211, 2008a. 6709, *Standard representation of geographic location by coordinates*.
- ISO/TC 211, 2008b. ISO 6709:2008 - *Standard representation of geographic point location by coordinates*.

- ISO/TC 211, 2007. ISO 19132:2007 - Geographic information -- Location-based services -- Reference model.
- ISO/TC 211, 2005. ISO 19133:2005 - Geographic information -- Location-based services -- Tracking and navigation.
- ISO/TC 211, 2003. ISO 19115:2003 - Geographic information -- Metadata.
- Jahnke, L., Fleckenstein, J., Wenz, F., Hesser, J., 2012. GMC: a GPU implementation of a Monte Carlo dose calculation based on Geant4. *Physics in medicine and biology* 57, 1217.
- Jana, D., 2014. C++ AND OBJECT-ORIENTED PROGRAMMING PARADIGM. PHI Learning Pvt. Ltd.
- Jelliffe, R., 2007. Where to get ISO Standards on the Internet free [WWW Document]. URL http://archive.oreilly.com/pub/post/where_to_get_iso_standards_on.html (accessed 8.7.16).
- Juhász, L., Rousell, A., Jokar Arsanjani, J., 2016. Technical Guidelines to Extract and Analyze VGI from Different Platforms. *Data* 1, 15.
- Karney, C.F., 2013. Algorithms for geodesics. *Journal of Geodesy* 87, 43–55.
- Knapena, M.J.R., Roosenschoona, O., Lokersa, R., Janssena, S., van Randena, Y., Verweija, P., 2013. Data integration technologies to support integrated modelling, in: *Proceedings of the MODSIM 2013, 20th International Congress on Modelling and Simulation, 1-6 December 2013, Adelaide, Australia*. pp. 845–851.
- Kondo, M.C., Bream, K.D., Barg, F.K., Branas, C.C., 2014. A random spatial sampling method in a rural developing nation. *BMC Public Health* 14, 338. <https://doi.org/10.1186/1471-2458-14-338>
- Kounadi, O., Lampoltshammer, T.J., Leitner, M., Heistracher, T., 2013. Accuracy and privacy aspects in free online reverse geocoding services. *Cartography and Geographic Information Science* 40, 140–153. <https://doi.org/10.1080/15230406.2013.777138>
- Kresse, W., Danko, D.M., Fadaie, K., 2012. Standardization, in: Kresse, W., Danko, D.M. (Eds.), *Springer Handbook of Geographic Information*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 393–565.
- Krisp, J.M. (Ed.), 2013. *Progress in Location-Based Services, Lecture Notes in Geoinformation and Cartography*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Krosse, S., 2010. Spatial Identification of Subjects, in: *Wants to Be a Topic Map. Presented at the Sixth International Conference on Topic Maps Research and Applications, Leipzig, Germany*, pp. 75–83.

- Küpper, A., 2005. Location-Based Services: Fundamentals and Operation. John Wiley & Sons.
- Lawrence, C., Riezler, S., 2016. NLmaps: A Natural Language Interface to Query OpenStreetMap., in: COLING (Demos). pp. 6–10.
- L'Ecuyer, P., Simard, R., 2007. TestU01: AC library for empirical testing of random number generators. *ACM Transactions on Mathematical Software (TOMS)* 33, 22.
- L'Ecuyer, P., Simard, R., 2005. Testu01. A software library in ANSI C for empirical testing of random number generators. University of Montreal, Canada, 2005.
- Lemmen, C.H.J., Van Oosterom, P.J.M., Uitermark, H.T., Zevenbergen, J.A., Cooper, A.K., 2011. Interoperable domain models: The ISO land administration domain model LADM and its external classes, in: UDMS 2011: 28th Urban Data Management Symposium, Delft, The Netherlands, September 28-30, 2011. Presented at the 28th Urban Data Management Symposium (UDMS 2011), ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Delft, The Netherlands, pp. 31–39.
- Libkin, L., Martens, W., Vrgoč, D., 2013. Querying graph databases with XPath, in: Proceedings of the 16th International Conference on Database Theory. ACM, pp. 129–140.
- Lisboa-Filho, J., Nalon, F.R., Peixoto, D.A., Sampaio, G.B., de Vasconcelos Borges, K.A., 2013. Domain and Model Driven Geographic Database Design, in: Domain Engineering. Springer, pp. 375–399.
- Loohuis, J., 2010. Precision of geographical coordinates for use in Google Maps [WWW Document]. Hackvalue. URL http://www.hackvalue.nl/en/article/106/precision_of_geographical_coordinates_for_use_in_google_maps (accessed 5.29.15).
- Ma, M., Zhong, Z., Guo, N., Jing, N., Xiong, W., 2016. An efficient reverse geocoding method based on Global Subdivision Model, in: 2016 24th International Conference on Geoinformatics. Presented at the 2016 24th International Conference on Geoinformatics, pp. 1–9. <https://doi.org/10.1109/GEOINFORMATICS.2016.7578966>
- Mabrouk, M. (Ed.), 2008. OpenGIS location services (OpenLS): core services.
- Markey, P., Clynych, G., 2013. A performance analysis of WS-*(SOAP) and RESTful Web Services for Implementing Service and Resource Orientated Architectures.
- Marrara, S., Pasi, G., 2016. Fuzzy Approaches to Flexible Querying in XML Retrieval. *International Journal of Computational Intelligence Systems* 9, 95–103. <https://doi.org/10.1080/18756891.2016.1180822>

- Matsumoto, M., Nishimura, T., 1998. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 8, 3–30.
- Maturi, D.A., Ullah, M.Z., Ahmad, S., Ahmad, F., 2014. An Efficient Computation of Effective Ground Range Using an Oblate Earth Model. *Abstract and Applied Analysis* 2014, e459790. <https://doi.org/10.1155/2014/459790>
- McCullough, B.D., 2006. A review of TESTU01. *J. Appl. Econ.* 21, 677–682. <https://doi.org/10.1002/jae.917>
- McKenzie, D., 2015. The Open Geospatial Consortium (OGC). *GeoInformatics* 18, 14–17.
- McKenzie, G., Janowicz, K., 2015. Where is also about time: A location-distortion model to improve reverse geocoding using behavior-driven temporal semantic signatures. *Computers, Environment and Urban Systems* 54, 1–13. <https://doi.org/10.1016/j.compenvurbsys.2015.05.003>
- Moura, T.H., Davis, C.A., Fonseca, F.T., 2015. Reference data enhancement for geographic information retrieval using linked data. *Transactions in GIS*.
- Nelson, J., 2013. Silly Geographic Precision. *UXBlog | IDV Solutions' User Experience*. URL <http://uxblog.idvsolutions.com/2013/11/silly-geographic-precision.html> (accessed 5.29.15).
- Noel, K., 2016. Convergence analysis of random generators in Monte Carlo simulation: Mersenne Twister and Sobol. *Social Science Research Network (SSRN)* 34. <https://dx.doi.org/10.2139/ssrn.2734013>
- O’Cathain, A., 2018. *A Practical Guide to Using Qualitative Research with Randomized Controlled Trials*. OUP Oxford.
- Okada, K., Okuda, T., Yamaguchi, S., 2013. Design of Geographically Aggregatable Address and Routing toward Location Based Multicast, in: *Advanced Information Networking and Applications Workshops (WAINA)*, 2013 27th International Conference On. pp. 1307–1312.
- O’Leary, Z., 2017. *The Essential Guide to Doing Your Research Project*, Third edition. ed. SAGE Publications Ltd, Thousand Oaks, CA.
- Paredes-Valverde, M.A., Alor-Hernández, G., Rodríguez-González, A., Valencia-García, R., Jiménez-Domingo, E., 2015. A systematic review of tools, languages, and methodologies for mashup development. *Softw. Pract. Exper.* 45, 365–397. <https://doi.org/10.1002/spe.2233>
- Phan, T., Baek, A., Guo, Z., 2014. Reducing the cloud cost of mobile reverse-geocoding, in: *Proceedings of the Fifth International Workshop on Mobile Cloud Computing & Services*. ACM, pp. 15–22.

- Posada-Sánchez, M., Bischof, S., Polleres, A., 2016. Extracting Geo-Semantics About Cities From OpenStreetMap., in: SEMANTiCS (Posters, Demos, SuCCESS). Presented at the 12th International Conference on Semantic Systems, Leipzig, Germany.
- Pozueco, L., Tuero, A.G., Paneda, X.G., Melendi, D., Garcia, R., Paneda, A.G., Rionda, A., Diaz, G., Mitre, M., 2015. Adaptive learning for efficient driving in urban public transport, in: Computer, Information and Telecommunication Systems (CITS), 2015 International Conference On. IEEE, Gijon, pp. 1–5. <https://doi.org/10.1109/CITS.2015.7297724>
- Renteria-Agualimpia, W., Lopez-Pellicer, F.J., Lacasta, J., Muro-Medrano, P.R., Zarazaga-Soria, F.J., 2015. Identifying geospatial inconsistency of web services metadata using spatial ranking. *Earth Sci Inform* 8, 427–437. <https://doi.org/10.1007/s12145-014-0172-4>
- Saito, M., Matsumoto, M., 2013. Variants of Mersenne twister suitable for graphic processors. *ACM Transactions on Mathematical Software (TOMS)* 39, 12.
- Salas-Zárate, M. del P., Alor-Hernández, G., Valencia-García, R., Rodríguez-Mazahua, L., Rodríguez-González, A., López Cuadrado, J.L., 2015. Analyzing best practices on Web development frameworks: The lift approach. *Science of Computer Programming* 102, 1–19. <https://doi.org/10.1016/j.scico.2014.12.004>
- Saxena, P., Kamal, R., 2014. An Information Dissemination Framework for Location Based Services. *IJETTCS* 3, 210–216.
- Schiller, J., Voisard, A., 2004. Location-based services. Elsevier.
- Sehra, S.S., Singh, J., Rai, H.S., 2014. A systematic study of OpenStreetMap data quality assessment, in: Information Technology: New Generations (ITNG), 2014 11th International Conference On. IEEE, pp. 377–381.
- Sehra, S.S., Singh, J., Rai, H.S., 2013. Assessment of OpenStreetMap Data - A Review. *International Journal of Computer Applications* 76, 17–20. <https://doi.org/10.5120/13331-0888>
- Shorter, C., 2013. Cameron Shorter: OGC heed community pressure regarding “GeoServices REST API.” Cameron Shorter. URL <http://cameronshorter.blogspot.my/2013/05/ogc-heed-community-pressure-regarding.html> (accessed 12.27.15).
- Sierra, J.P.G.-B., Callejo, M.A.M., Garrido, R.A., 2015. Volunteering assistance to online geocoding services through a distributed knowledge solution, in: Workshop at 18th AGILE Conference On.
- Silva, V., Malheiro, B., 2011. Multi-agent system for personalization of location-based services, in: Information Systems and Technologies (CISTI), 2011 6th Iberian

Conference On. IEEE, pp. 1–7.

Sim, J., Saunders, B., Waterfield, J., Kingstone, T., 2018. Can sample size in qualitative research be determined a priori? *International Journal of Social Research Methodology* 21, 619–634. <https://doi.org/10.1080/13645579.2018.1454643>

Tamayo, A., Granell, C., Huerta, J., 2012. Measuring complexity in OGC web services XML schemas: pragmatic use and solutions. *International Journal of Geographical Information Science* 26, 1109–1130. <https://doi.org/10.1080/13658816.2011.626602>

Tanasescu, V., Smart, P.D., Jones, C.B., 2014. Reverse Geocoding for Photo Captioning with a Meta-gazetteer, in: *Proceedings of the 22Nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, SIGSPATIAL '14*. ACM, New York, NY, USA, pp. 509–512. <https://doi.org/10.1145/2666310.2666492>

Tawofaing, A.F., Sallaberry, C., Le Parc-Lacayrelle, A., Moal, T., 2016. Extraction of Business Information on the Web to Supply a Geolocated Search Service, in: *ALLDATA 2016*. Presented at the Second International Conference on Big Data, Small Data, Linked Data and Open Data, IARIA, Lisbon, Portugal, pp. 82–85.

Teske, D., 2014. Geocoder Accuracy Ranking, in: Lamprecht, A.-L., Margaria, T. (Eds.), *Process Design for Natural Scientists, Communications in Computer and Information Science*. Springer Berlin Heidelberg, pp. 161–174.

Tian, X., Benkrid, K., 2009. Mersenne Twister random number generation on FPGA, CPU and GPU, in: *Adaptive Hardware and Systems, 2009. AHS 2009. NASA/ESA Conference On. IEEE*, pp. 460–464.

Tiwari, S., Kaushik, S., 2014. User Category Based Estimation of Location Popularity using the Road GPS Trajectory Databases. *Geoinformatica: An International Journal (GIJ)* 4, 20–31.

Trammell, B., 2014. Textual Representation of IP Flow Information Export (IPFIX) Abstract Data Types. Internet Engineering Task Force (IETF), Request for Comments (RFC) 14.

Tredak, P., JR, J.C.W., 2016. Parallelization of random number generators. US20160162262 A1.

van Rossum, G., 2016. random - Generate pseudo-random numbers - [WWW Document]. Python 2.7.13 documentation. URL <https://docs.python.org/2/library/random.html> (accessed 8.22.17).

Vauglin, F., 2002. A practical study on precision and resolution in vector geographical databases, in: *Spatial Data Quality*. Taylor & Francis, London, pp. 130–143.

- Vitolo, C., Elkhathib, Y., Reusser, D., Macleod, C.J., Buytaert, W., 2015. Web technologies for environmental Big Data. *Environmental Modelling & Software* 63, 185–198.
- Wang, G., 2013. Teaching High-Accuracy Global Positioning System to Undergraduates Using Online Processing Services. *Journal of Geoscience Education* 61, 202–212.
- Wikipedia, 2015. WikiProject Geographical coordinates. Wikipedia.
- Winterford, B., 2013. Open source crusade blocks geospatial standard. iTnews. URL <http://www.itnews.com.au/news/open-source-crusade-blocks-geospatial-standard-345493> (accessed 12.27.15).
- Woyke, E., 2009. Google's Navigation Bombshell [WWW Document]. Forbes. URL <http://www.forbes.com/2009/10/23/android-navigation-internet-technology-wireless-google.html> (accessed 1.6.17).
- Xu, S., Flexner, S., Carvalho, V., 2012. Geocoding Billions of Addresses: Toward a Spatial Record Linkage System with Big Data, in: *GIScience in the Big Data Age*. Columbus, Ohio, USA., pp. 17–26.
- Yahoo, n.d. Key Concepts - YDN [WWW Document]. URL <https://developer.yahoo.com/geo/geoplanet/guide/concepts.html#woeidsplanet> (accessed 1.13.17).
- Yana, Y., Lia, G., Xiea, J., Guoa, L., 2011. a Comparative Study of Spatial Data Harvest Standards. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 3825, 49–55.
- Yates, A., Beal, K., Keenan, S., McLaren, W., Pignatelli, M., Ritchie, G.R., Ruffier, M., Taylor, K., Vullo, A., Flicek, P., 2014. The Ensembl REST API: Ensembl Data for Any Language. *Bioinformatics* btu613.
- Yue, P., 2013. Automatic Composition of Geospatial Web Service, in: *Semantic Web-Based Intelligent Geospatial Web Services*, SpringerBriefs in Computer Science. Springer New York, pp. 65–85.
- Zandbergen, P.A., Hart, T.C., Lenzer, K.E., Camponovo, M.E., 2012. Error propagation models to examine the effects of geocoding quality on spatial analysis of individual-level datasets. *Spatial and Spatio-temporal Epidemiology, Special Issue on Geocoding in the Health Sciences* 3, 69–82. <https://doi.org/10.1016/j.sste.2012.02.007>
- Zarem, M., Vuillermet, E., DeAguiar, J., 2007. Intelligent reverse geocoding. *US20070191029 A1*.
- Zhou, T., 2013. Understanding user adoption of location-based services from a dual perspective of enablers and inhibitors. *Inf Syst Front* 1–10. <https://doi.org/10.1007/s10796-013-9413-1>

Zipf, A., Jöst, M.M., 2012. Location-Based Services, in: Kresse, W., Danko, D.M. (Eds.), Springer Handbook of Geographic Information. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 417–421.



BIODATA OF STUDENT

His first name is Mohammed, His Father's name is Mustafa and his family name is Al-Habshi. He was born in Aden, Yemen (1974), and raised in Saudi Arabia with his parents. He grew with interest for video games which led him to home computers and Basic programming. At his intermediate education, he got a Personal Computer where he practiced system maintenance. After the second gulf war, he returned to Yemen to continue his academic education.

He entered the Engineering Faculty, Sana'a University where he joined Architecture and Urban Design department due to limitation (at that time) of the electronic engineering and popularity of Architecture in Yemen. After graduation, he tried to find a job; however, the unemployment was the common case in Yemen. He gave computer course in private institutes.

He tried to obtain scholarship through Ministry of Planning and Development (MPD) in Yemen, where he found a vacancy of contractor supervisor for computer network in-building implementation. He worked with MPD where he also worked as short-term consultant for the German Technical Cooperation (GTZ) and later he works for advisory Service for the ministry (AS-MPD), a project funded and supported by UNDP.

At his work with MPD, he obtained two certificate of enterprise networking. Also, his awareness has grew of finding free and cheap alternatives of commercial software for developing countries information technology infrastructure implementation have raised, where he found open-source software provide the suitable solution for such global problem while restricting expenditure internally within the country.

He approaches to continue his Master studies by obtaining the Islamic Development Bank Scholarship, where he found GIS is a new field suite his background in Urban Development. However, his first initial proposal for M.Sc found obstacle in find data resource from Yemen for GIS application, where he found LBS application possible to accomplish for M.Sc. He worked in Malaysia as tutor, software development and integrator in cooperation with RSRC, MIRO, CampusCity institute, CTH technologies, Sarawak Forestry, Wildlife.

He interests lays in 3D modelling and visualization, computer languages, web application and services, databases, distributed mobile computing, engineering computerized methods, communication, e-learning, wireless sensing and creative design.

LIST OF PUBLICATIONS

- Al-Habshi, M. M., & Shafri, H. Z. 2019. A Query Model of Determining latest unique metric Coordinate Reference System. *Journal of Recent Technology and Engineering*, 8(4).(accepted)
- Suhaili, Affendi, M. Al-Habshi, Jukka Okkonen, and A. Rodzi Mahmud. 2010. An Integrated Geospatial System For Forest Monitoring And Enforcement, In *World Engineering Congress 2010*, 91–95. Kuching, Sarawak, Malaysia.
- Barakat, O.L., RSA, S.H., Ramli, A.R., Hashim, F., Samsudin, K., Al-baltah, I.A. and Al-Habshi, M.M., 2013. Scarecrow: Scalable malware reporting, detection and analysis. *Journal of Convergence Information Technology*, 8(14), p.1.
- Shaharum, N.S.N., Shafri, H.Z.M., Ghani, W.A.W.A.K., Samsatli, S., Al-Habshi, M.M.A. and Yusuf, B., 2020. Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms. *Remote Sensing Applications: Society and Environment*, 17, p.100287.



UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT

ACADEMIC SESSION : _____

TITLE OF THESIS / PROJECT REPORT :

NAME OF STUDENT : _____

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (v)

☐

CONFIDENTIAL

(Contain confidential information under Official Secret Act 1972).

☐

RESTRICTED

(Contains restricted information as specified by the organization/institution where research was done).

☒

OPEN ACCESS

I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for :

☐

PATENT

Embargo from _____ until _____
(date) (date)

Approved by:

(Signature of Student)
New IC No/ Passport No.:

Date :

(Signature of Chairman of Supervisory Committee)
Name:

Date :

[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentiality or restricted.]