GEO 620 Master's Thesis

Development, Implementation, Assessment and Analysis of a Real-Time Tile-Based Location-Based Game for Geographic Information Mining Regarding Land Cover Data

Manuel F. Bär

Matriculation Number: 08-923-914

Email: manuel.baer@uzh.ch / manuel.baer@geonet.ch

Biel, 21st of April 2017

Faculty Representative: Prof. Dr. Ross Purves

Advisors: Prof. Dr. Ross Purves and Dr. Flurina Wartmann

GIScience: Geocomputation

Department of Geography

University of Zürich

"The most incomprehensible thing about the world is that it is comprehensible." - Albert Einstein

Acknowledgements

I would like to thank all people involved in my master's thesis.

First, I would like to thank Prof. Dr. Ross Purves and Dr. Flurina Wartmann for several detailed discussions, their crucial feedback and their invaluable input and support throughout the whole process of writing this thesis.

I also especially thank my parents for their unconditional support of my studies and for the valuable input, interesting discussions and tireless proofreading.

Furthermore, I thank Annabelle Jaggi for her undivided attention whilst proofreading this master's thesis and testing the implementation, as well as Chistophe Willemsen for his collaborative efforts in implementing a Neo4j-Plugin for PHP. In addition, many thanks go to Alain Burkhalter, Christophe Mosimann, Lukas Oberli and Tobias Stalder for their time spent on testing the implementation and giving valuable feedback.

Lastly, I thank all persons who registered and played the implemented game, in particular all persons who gave additional feedback or suggested improvements. Without their time, interest and feedback, this master's thesis would not have been possible.

Abstract

Land cover change has been identified as one of the most important variables of global change, but when comparing individual land cover products large disagreements of what land cover class they declare at given locations emerge. Cost efficient in-situ classifications from a large number of non-expert users offer a way of generating large amounts of data for land cover product assessments. By identifying and synthesising several key concepts and methods found in the literature regarding user generated data collection, location-based game design and user motivation, this thesis elaborates the implementation of a full-fledged location-based game for in-situ land cover classification collection by non-expert users. A Neo4j Graph database was combined with a PostGIS raster table and a PHP-framework to successfully implement a browser-based location-based game, allowing for easy access. Data was collected over three months and a considerable amount of land cover classifications were reported by non-expert users. The data was analysed using different visual and statistical analysis approaches. The attributes were analysed individually, in comparison with other contributed data and compared to the official CORINE 2012 land cover dataset of Switzerland. Both, absolute and relative confusion matrices were used to visualise and statistically analyse the data. In addition, various bar and line plots as well as map examples were created to highlight specific characteristics. Since user motivation was found to be a key factor in generating data through gaming and in effective user retention, various types of motivational elements were implemented such as competition elements and a sense of progression. The results reveal that the implemented location-based game offers a plausible and cost efficient way to collect large amounts of data in a short time span. The user contributed data shows agreement rates between the user generated data and an official dataset on par with findings of similar analyses, underlining the success of the implemented application. Observed disagreements are closely inspected and three main sources of error are identified: (1) overrepresentation of classes, (2) difficulties in differentiating classes and (3) spatial autocorrelation of classes. Findings of this thesis suggest that individual or group-related differences in the perception and thus classification of land cover classes are the most prominent source of error. Understanding the different semantics of land cover classes is thus of utmost importance in land cover assessment efforts. The results of this thesis ultimately confirm that the implemented location-based game is less suited for automated land cover curation processes, but rather to detect areas of spatially autocorrelated classification errors and to gain insights into potential semantic issues.

Contents

List of Figures

List of Tables

List of Abbreviations

Term	**Meaning**
CORINE	Coordination of information on the environment
CSV	Comma separated value
GPS	Global positioning system
LBG	Location-based game
LOESS	Locally weighted scatterplot smoothing
POI	Points of interest
UGC	User generated content
UN-ASIGN	Adaptive system for image communication over global networks
UNITAR	United nations institute for training and research
UNOSAT	United nations institute for training and research (UNITAR) operational satellite applications programme
VGI	Volunteered geographic information

1 Introduction

1.1 Context

The rapid development of internet technologies such as cellular internet and display interfaces such as browsers, as well as ever increasing bandwidth allowed for new forms of collaboration, social exchange and data collection. In addition, the significant growth in available digital storage space and its decline in price have led to an abundance of unassessed or even unused digital data within information societies.

Among the large variety of open-source and free software products, high quality remote sensed products are also increasingly becoming freely available. Of these remote sensed products, land cover datasets have "been identified as one of the fundamental variables needed in order to study the morphological and functional changes occurring in the Earth's ecosystems and the environment" (Congalton et al. 2014, p.12071), thus becoming crucial for large scale policy and decision making processes (Congalton et al. 2014; Mallupattu & Sreenivasula Reddy 2013; Lambin et al. 2001). However, according to Fritz et al. (2009) "global land cover datasets still show quite a high degree of disagreement" and See et al. (2013, p.2) agree that "[…] datasets such as GLC-2000, MODIS and GlobCover […] frequently disagree over the land cover they record at any given location." In the scientific community, a particular point of interest is the verification and improvement of land cover products using citizen science and crowdsourcing, especially *user generated content (UGC)* such as *volunteered geographic information (VGI)* (Goodchild 2007; Foody & Boyd 2012; Fritz et al. 2009; See et al. 2013). The sudden interest in research focusing on the use of VGI for land cover product verification, assessment and improvement is partially due to the advances in internet and mobile technologies, allowing users to easily contribute and share geographic information using handheld devices. Fritz et al. (2009) propose the use of "competitive games such as those used for most computer games […] to make the challenge of land cover validation more attractive." *Location-based games (LBGs)* are of particular interest in this respect, because users contribute in-situ geographic information. *LBGs* have seen an unprecedented increase in the number of users in 2016 after the release of the highly popular LBG *Pokémon GO*[1], produced by Nintendo[2] and Niantic labs[3], a Google[4] spinoff. Even though games like *Geocaching*[5] and *INGRESS*[6] existed and were played by millions of users before the release of *Pokémon GO*, LBGs were still unknown to the general public. *Pokémon GO* first introduced the concept of location-based gaming to a wider public, in particular to persons owning

[1] http://pokemongo.nianticlabs.com (accessed: 22.03.2017)
[2] www.nintendo.com (accessed: 22.03.2017)
[3] www.nianticlabs.com (accessed: 22.03.2017)
[4] www.google.com (accessed: 22.03.2017)
[5] www.geocaching.com (accessed: 22.03.2017)
[6] www.ingress.com (accessed: 22.03.2017)

a smartphone with *global positioning system (GPS)* localisation capabilities. Aided by the omnipresence in the media, *Pokémon GO* was the first LBG which gained (almost) global prominence.

Location-based games have become an object of research mainly over the past two decades and the literature agrees on location-based gaming being a useful tool for geospatial data acquisition, geospatial data validation and for edutainment purposes (Matyas 2007; Yoshii et al. 2011; Celino et al. 2012; Ionescu et al. 2013; Davidovic et al. 2013; Matyas et al. 2012; Charsky 2010; Matyas et al. 2011; Avouris & Yiannoutsou 2012; Richter et al. 2012; Matyas et al. 2008; Winter et al. 2011; Celino 2015; Yanenko & Schlieder 2014). Various authors have analysed the usability of location-based games for data mining (Matyas 2007; Matyas et al. 2008; Celino et al. 2012; Winter et al. 2011; Davidovic et al. 2013; Matyas et al. 2012), mainly focusing on *points of interest (POI)* information collection. Location-based games typically only allow certain interactions with a virtual environment if specific real-world location-based criteria are met. A location-based game for geographic information mining is usually characterised by various indicators. The game field structure can be unstructured, semi-structured or structured (Matyas et al. 2012; Matyas 2007), encouraging the collection of specific types of data (e.g. POI, Path, Tiles). Three other key characteristics are the typical duration of a game (Avouris & Yiannoutsou 2012), if the game has a narrative or story-line (Avouris & Yiannoutsou 2012) and if the game is team based or not (Matyas 2007; Matyas et al. 2008; Matyas et al. 2012; Celino et al. 2012; Winter et al. 2011; Yoshii et al. 2011; Davidovic et al. 2013). Most of the LBGs analysed in a scientific context concentrate on collecting POI information and the duration of the games is predominantly short to medium. Furthermore, most of the studied games do not have storylines or narratives at all or only weak ones. Finally, the games are based on the player distribution concept of either "everyone for themselves" or split into teams. Thus, a research gap concerning location-based games becomes apparent. The implementation and analysis of a location-based game focusing on tile-based information collection, incorporating strong narrative characteristics, with a continuous gameplay and faction or team based playing offers the possibility to make land cover validation more attractive and to appeal to a large number of users. This great potential for motivating non-expert users to aid in the task of land cover product validation has already been identified (Fritz et al. 2009), but little research has as yet been done on implementing a location-based game for this purpose. Furthermore, location-based gaming for geospatial data acquisition, geospatial data validation and for edutainment purposes has been widely discussed in the scientific literature, but no literature was found concerning location-based games for tile-based information mining. Thus, the following questions merit consideration: *How* can a tile-based location-based game be implemented to collect data which can be used for land cover validation and assessment and *how* can the generated data be analysed?

1.2 Scope and Overview

This master's thesis addresses these questions on how the collection of data through tile- and location-based games can be implemented and on how the generated data is best analysed through an interdisciplinary approach, merging theories of UGC, VGI, LBGs, game design, citizen sciences, geographic information science and statistical analysis. New cost efficient ways of quality assessment and improvement are called for to assess an increasing array of land cover products. Therefore, besides developing and implementing a real-time tile-based location-based game for geographic information mining, this master's thesis also analyses the generated data in regards to land cover product assessment. The results of the various analyses are then synthesised and discussed in light of the literature. Thus, the aim of this master's thesis is not the implementation of a location-based game for *automated* land cover dataset curation, but to generate a considerable amount of non-expert land cover information data to evaluate the reliability of a location-based game with regards to the assessment of land cover products. Therefore, the overarching research questions can be summarised as:

How can a location-based game with a non-expert target audience be implemented to mine tile-based geographic information, in particular land cover data; and can the generated land cover and semantic data be used in a research context, in particular with regards to the validation or assessment of land cover products?

1.3 Thesis Structure

A comprehensive literature review regarding land cover products, including potential sources of error is given in chapter 2, *State of the Art*. In addition, current efforts in assessing and curating land cover datasets are highlighted. Further, the state of the art of VGI and crowdsourcing as data sources of geographic information are presented with a focus on location-based games as geographic information mining tools. In Chapter 3, *Research Gap*, the identified research gap is presented and chapter 4, *Research Questions*, elaborates the research questions of this master's thesis. Chapter 5, *Methods*, gives a detailed overview of the implementation process of the proposed location-based game as well as the methods and approaches used to analyse the generated data. The results of the analyses are presented in chapter 6, *Results*, focusing on user distribution and attributes, user contributions, analyses of intra-tile agreement rates on the land cover classification of given locations and finally, on the comparison of the user contributed data with the official CORINE 2012 land cover dataset of Switzerland. Chapter 7, *Discussion,* thoroughly discusses the research questions in light of the results obtained from the own research and previous findings from the literature. The *Discussion* is followed by chapter 8, *Limitations,* where major limitations of this thesis are elaborated on. A summary of the main recommendations and suggestions as well as an outlook and a call for further research round off this thesis in chapter 9, *Conclusions and Further Work*.

All diagrams and plots were coded using *R Studio Version 1.0.136* or were drawn in *Microsoft Word 365*. Map material was created using *QGIS 2.18.4 La Palmas*.

2 State of the Art

This chapter presents a detailed overview and the state of the art concepts related to land cover products, land cover product assessment, geographic information mining and location-based gaming. These are fundamental topics that need to be understood in order to implement the intended location-based application. The state of the art also introduces key literature allowing a better understanding of the results and ultimately enabling a synthesis of the results and literature in the discussion.

The chapter is structured in a pyramid or top-down structure, starting with a general overview followed by more and more detailed subsections.

2.1 Land Cover

Researchers agree that land cover is one of the most important variables in understanding and analysing fundamental processes of the earth (Congalton et al. 2014; Foody et al. 2002). Land cover is generally used as a synonym for the characteristics of the earth's surface at a given location, i.e. the "biophysical attributes of the earth's surface" (Lambin et al. 2001, p.262). Furthermore, for the scope of this thesis the following definition of land cover from a remote sensing perspective needs to be taken into account:

> *Land cover is the physical material at the surface of the earth. It is the material which we see and which directly interacts with electromagnetic radiation and causes the level of reflected energy which we observe as the tone or the digital number at a location in an aerial photograph or satellite image.*
>
> (Fisher et al. 2005, p.2)

Issues revolving around land cover and land cover changes have been of increasing importance to the scientific community, especially in regard to climate change and climate change feedback loops. Foody et al. (2002) summarise the findings of various authors by stating that land cover changes have a considerable effect on basic processes of the earth's natural cycle (Skole 1994; Douglas 1999; Vitousek 1994 as cited in Foody et al. 2002). Land cover change has been identified as the "single most important variable of global change affecting ecological systems" (Foody et al. 2002, p.185) and it is presumed that for the next 100 years land cover change will be the most significant variable impacting biodiversity (Chapin et al. 2000).

2.1.1 Land Cover Products

Land cover products are datasets which consist of classifications of the biophysical attributes of the earth's surface at given locations. Land cover products are essentially a generalised representation of the earth's land cover in form of pixels, sometimes referred to as "objects". Raw satellite sensor data is collected and pre-processed to account for various sources of errors including atmosphere related, geometrical and calibration errors. After the raw satellite data is pre-processed, the pixels are classified into a set of classes in supervised or unsupervised classification processes. The unsupervised classification process clusters the pixels according to spectral similarity and defines classes out of the resulting clusters. The supervised classification process requires predefined spectral classes (mostly defined by an expert user) and allocates the pixels according to said predefined classes. The pixels in the resulting dataset can belong to multiple classes (fuzzy) or one class (hard) (Alexis Comber et al. 2005).

The produced products consist of pre-processed and classified datasets, in which each pixel has one or multiple values. The generated product is then provided by the producer and used by a user. The use of the data can vary depending on the intent of the user, most commonly the data is used to describe or analyse the landscape of a particular geographic extend (Alexis Comber et al. 2005).

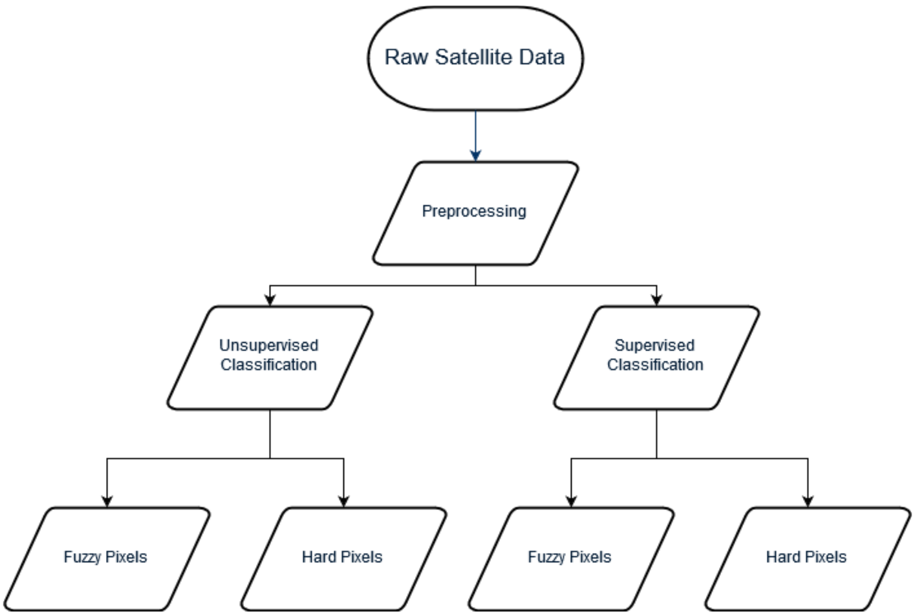

Figure 1 - Remote sensing product workflow
Figure showing the basic workflow of land cover product creation

Land cover products are slowly diffusing into disciplines other than remote sensing and geography, especially with increasing resolutions and decreasing prices of land cover products. Thus, land cover products are not only used in a scientific context, but they have become important tools in various domains within the private sector, such as urban planning, architecture, energy production and water management.

2.1.2 CORINE Land Cover Data

The *Coordination of Information on the Environment (CORINE)* land cover inventory is a database of land cover information, which was initiated on June 27, 1985 (European Environment Agency 1994). The CORINE initiative's objective is the coordination, management and assessment of information regarding the environment and natural resources. Therefore, one of the main goals of the CORINE programme is "to bring together all the many attempts which have been made over the years at various levels (international, Community, national and regional) to obtain more information on the environment and the way it is changing" (European Environment Agency 1994, p.3). The CORINE program provides consistent geographic information on the land cover of 39 different states (as of 19.02.2016) as vectors or as rasters with cell sizes of 100m x 100m to 250m x 250m.

According to the official documentation of the CORINE program (European Environment Agency 1994), a supervised computer-aided photo interpreted classification approach is used to delineate different land cover classes. After completion of the preliminary pre-processing, false-colour images are generated. The generated false-colour images are then delineated and identified by computer-aided photo interpretation, where an expert in photointerpretation manually highlights key areas of homogeneous land cover. The expert classifies an area according to strict rules using a predefined list of classes. The class *green urban areas* is, for example, defined as follows:

"Green urban areas concern all vegetated areas greater than 25 ha that are either situated within or in contact with urban fabrics. Strips of lanes and paths created for recreational use may be found within these areas." (Kosztra et al. 2014, p.29)

After classifying the images, the expert in photointerpretation highlights those areas where additional information is needed in order to correctly classify a particular area in compliance with CORINE standards. Additional information which is viable to be used in the land cover classification includes tourist and topographic maps, aerial photographs, SPOT data and maps of farmland.

Once the process of interpreting the false-colour images has been completed, the national project leader deals with unresolved questions in accordance with the CORINE standards. The expert identifies key problems and divides these into two groups: Areas where ground surveys are needed and areas where aerial photographs meet all necessary requirements. One major problem in this final stage of interpretation is the cost-efficiency-ratio, because high costs are to be expected for only small areas of classification. As a result, the CORINE guidelines state that "it is important to ensure that the cost of ground truth surveys does not exceed 10% of the total budget for the national land cover project" (European Environment Agency 1994, p.66). Figure 2 shows key steps in the process of generating the CORINE land cover dataset.

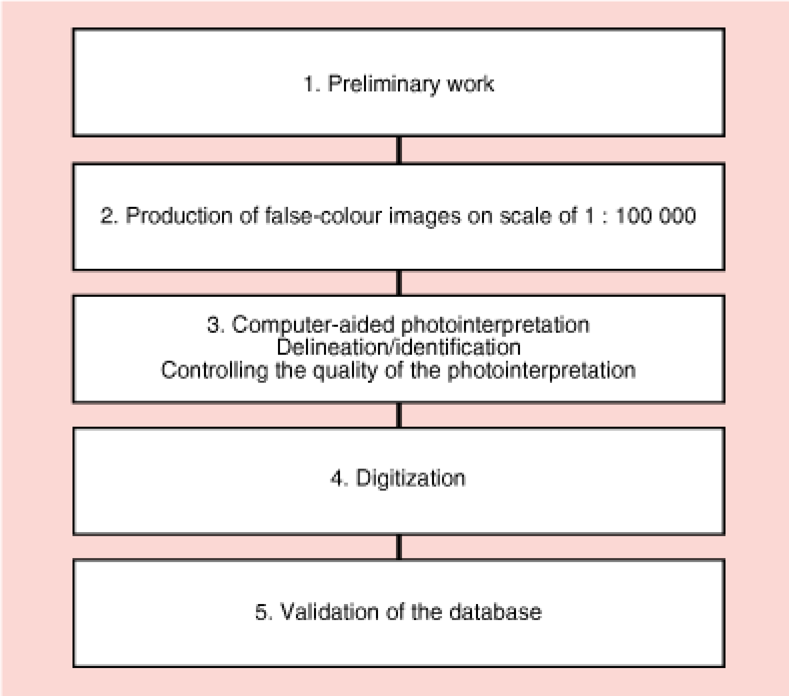

Figure 2 - CORINE land cover product creation. (Reprinted from European Environment Agency 1994)
Figure showing key processes in the creation of the CORINE land cover dataset.

2.1.3 Land Cover Product Errors

Despite land cover being used in various federal, academic and private institutions and being acknowledged as one of the most important environmental variables, there are still major issues in terms of accuracy and data quality regarding land cover products and the implications thereof. The insufficient amount of accurate in-situ land cover data leads to considerable limitations in quality assessment efforts. Land cover datasets show substantial disagreements between individual products, i.e. the most widely used land cover products largely disagree on the land cover data recorded for the same location (See et al. 2013). This poses a major problem seeing that data from (global) land cover products are often used in the context of large scale policy or decision making processes and in understanding large scale environmental processes. Thus, land cover products are also vital in the assessment, analysis and monitoring of climate change and in detecting other large scale morphological or functional changes in the earth's ecosystems (Congalton et al. 2014). Due to the substantial differences between the datasets of individual products, a gradient of possible results of any given analysis using land cover products is inevitable. Therefore, the outcome of an analysis is unavoidably biased according to the land cover product a researcher or institution chooses to base their analysis on. In view of the importance of land cover datasets and the lack of reliable quality thereof, new methods to assess and new technologies to enhance the quality of land cover datasets are called for.

2.2 Land Cover Product Assessment and Curation

Seeing the importance of land cover datasets in making large scale policies or decisions, or detecting large scale environmental changes, a lot of research has been done on the quality assessment of land cover products. Foody et al. (2002) summarise four historical stages for land cover assessment. The first and earliest stage of accuracy and quality assessment was based on the highly subjective question on whether a map looked "right" or "wrong". If the map looked "right", it would be classified as accurate and in contrast, if it looked "wrong", it would be classified as inaccurate. The second stage of land cover product quality assessment was based on the relative quantity of a map class correlating to a ground truth dataset. Although this method does not specifically take the location of a class into consideration, it did, for the first time, quantify the agreement of the relative extent of the mapped classes in comparison to a ground or reference dataset. A major limitation of this method was that a high accuracy could be achieved even if a map class had a similar extent in the reference dataset but completely different spatial attributes (i.e. the same amount of a class can be found in the reference dataset but at different locations). This major limitation called for further quality assessment methodologies for land cover datasets. The third stage saw an extension to the second stage by adding location to the methodology. Certain map classes were compared to the ground truth at specific locations. The relative amount of agreement was then translated to an accuracy score. The fourth stage mentioned by Foody et al. (2002) has been widely accepted and is still being used in accuracy assessments today. The fourth stage is based on a confusion or error matrix from which an accuracy indication is derived. The confusion matrix consists of a table containing mapped classes that are compared to reference data in the form of validation pixels (Carneiro & Pereira 2014). The confusion matrix can be used to calculate the common performance metrics and to make assumptions on the general accuracy of the underlying dataset (Fawcett 2006).

Not only does a large amount of research exist concerning accuracy estimation of land cover products, but also concerning refinement and curation using UGC or VGI. To crowdsource land cover pixel validation, various authors use an interactive internet platform incorporating underlying Google Maps [7]services. The users are tasked with agreeing or disagreeing on the classification of different land cover products when comparing a pixel in the land cover product to the referenced area in Google Maps (Fritz et al. 2009; See et al. 2013).

[7] www.google.ch/maps/

2.3 Volunteered Geographic Information and Crowdsourcing as Data Sources for Geographic Information

User generated content, especially volunteered geographic information and other crowdsourcing efforts have become valuable sources of expert and non-expert geographic information. The following section presents the state of the art in VGI and crowdsourcing.

2.3.1 User Generated Content and Volunteered Geographic Information

UGC refers to all forms of data created by a user and then uploaded to the internet (Neuendorf 2016). It is rapidly becoming an often used source of data for scientific research, specifically when large collections of data are needed. According to the literature, the generated content can be divided into two groups: explicitly and implicitly generated content (Alt et al. 2010; Senaratne et al. 2016). If a user individually and actively contributes content by performing specific actions (e.g. uploading a photograph to Flickr[8]; posting a message on Facebook[9]), the contribution is described as being explicit (Alt et al. 2010). Implicit user generated content encompasses content or data which is generated without any additional effort from the user (e.g. Google[10] scanning and saving a user's search results and emails to individualise ads; web-shop storing information on orders to make user specific product recommendations) (Alt et al. 2010).

There are obvious forms of user generated content, especially since UGC has seen a rapid increase in contributors with the emergence of web 2.0 technologies and broadband cellular internet services (Alt et al. 2010). With the advent of widespread access to accurate GPS location data, there has been a significant increase in user generated content with spatial attributes and location-based user generated data (Alt et al. 2010). VGI is a special form of user generated content, focusing on geographic or spatial aspects of UGC (Goodchild 2007). Similar to UGC, Craglia et al. (2012) argue that VGI can be divided into four groups, each of them represents a combination of the type of explicit or implicit geography which is captured with the type of explicit or implicit volunteering (Alt et al. 2010; Antoniou et al. 2010; Craglia et al. 2012). Spatially explicit content is generated by users knowingly interacting with spatial features (e.g. digitising geometries in a Web-GIS), whereas spatially implicit content is only implicitly associated with a geographic extent or location and can be generated as any type of media (e.g. an article containing toponyms; a geotagged photo) (Senaratne et al. 2016; Antoniou et al. 2010).

[8] www.flickr.com
[9] www.facebook.com
[10] www.google.com

	Geographic	
	Explicit	Implicit
Explicitly volunteered	This is 'True' VGI in the strictest sense. Examples include Open Street Map.	Volunteered (geo)spatial information (VSI). Examples would include Wikipedia articles about non-geographic topics, which contain place names
Implicitly volunteered	Citizen-generated geographic content (CGGC). Examples would include any public Tweet referring to the properties of an Identifiable place.	Citizen-generated (geo)spatial content (CGSC) such as a Tweet simply mentioning a place in the context of another (non-geographic) topic.

Figure 3 - Typology of VGI. (Reprinted from Craglia et al. 2012, p.405)
Figure showing combinations of explicit and implicit geographic information both explicitly and implicitly volunteered

There exists a significant number of applications using VGI. The *Adaptive System for Image communication over Global Networks (UN-ASIGN[11])* for example is an application hosted by the *United Nations Institute for Training and Research (UNITAR) Operational Satellite Applications Programme (UNOSAT)*, through which non-expert users can contribute location-based information and photographs. This application is used for disaster management such as emergency response efforts. One of the best-known collection of volunteered geographic information can be merited to the tireless efforts of the OpenStreetMap (OSM) team, offering quality on par with commercial datasets by leveraging users to voluntarily contribute geographic information (Sehra et al. 2014).

2.3.2 User Generated Content for Land Cover Data Collections

Particular points of interest in the scientific community are the verification and improvement of land cover products using citizen science and crowdsourcing (Fritz et al. 2009; Foody & Boyd 2012; See et al. 2013). Fritz et al. (2009) and See et al. (2013) examine the usability of the crowdsourcing platform Geo-Wiki.org for land cover validation processes. Regarding the quality of data, See et al. (2013) conclude that the overall quality of crowdsourced information is relatively high and that differences between experts and non-experts are small, but vary depending on land cover class and throughout the test period. These findings do not comply with the general assumption that "data produced by volunteers is often considered as being of lesser quality than data produced by experts" as stated by Yanenko and Schlieder (2014, p.1)(Yanenko & Schlieder 2014, p.1). It is also mentioned that the "reliability of the information provided by non-experts improved faster and to a greater degree than experts" (See et al. 2013, p.10), which calls for targeted means of training (i.e. training non-expert users efficiently during or before they generate data should increase the overall quality of the generated data). Research has also been conducted by Hutchison et al. (2012) on how to allocate non-

[11] https://www.unitar.org/unosat/un-asign-crowd-source-photos-mobile-app/ (accessed: 22.03.2017)

expert users to a task when crowdsourcing satellite imagery analysis. The authors argue that "the precision rate of any parallel strategy increases with the number of users (less false identification)" (Hutchison et al. 2012, p.128) and "an iterative strategy improves the spatial coverage (and thus the recall rate) as the iteration goes on" (Hutchison et al. 2012, p.128). It is also stated that "allocating more than 5 volunteers has a low impact on the accuracy and variability, while increasing unnecessary the resources" (Hutchison et al. 2012, p.125). This is somewhat in contrast with the findings of Haklay et al. (2010), who confirm the validity of Linus' Law in volunteered geographic information. Linus' Law is the assumption that the quality of user contributions is directly linked to the quantity of contributions, meaning that the quality of a user generated dataset increases with increasing contributions (Haklay et al. 2010). Another approach is discussed by Leung and Newsam (2014), who analyse and present methodologies of land cover classifications using geo-referenced photos. The authors focus on photo collections on Flickr[12] and on the Geograph Project[13]. They agree that "large collections of geo-referenced ground level photos can be used to derive maps of what-is-where on the surface of the Earth" (Leung & Newsam 2014, p.15). The authors also highlight the "potential for discriminating between land use classes" (Leung & Newsam 2014, p.17), but warn that the intent of the photographer has a high influence on the usability of the geo-referenced photo for land cover or land use classifications. The use of the textual data associated with the geo-referenced images is also discussed and the authors conclude that using text features to classify land cover leads to more accurate results than using image processing techniques, but only for the Geograph Project collection, where the users' intention is to provide typical geographic characteristics of predefined regions. Fritz et al. (2009) show the viability of crowdsourced information for land cover validation but also mention future challenges, in particular how to "attract a wide range of volunteers from all over the world" (Fritz et al. 2009, p.351). The authors propose the use of "competitive games such as those used for most computer games […] to make the challenge of land cover validation more attractive" (Fritz et al. 2009, p.351).

2.3.3 Location-Based Games as Data Sources for Geographic Information

Location-based games are based on a common denominator which only allows an interaction with the virtual environment when specific real-world location-based criteria are met. A location-based game aiming at geographic information mining is characterised by various indicators. The game field structure can be unstructured, semi-structured or structured (Matyas et al. 2012; Matyas 2007), depending on the specific types of data to be collected (e.g. *Points of interest (POI)*, Path, Tiles). Three other key characteristics encompass the duration of a game (Avouris & Yiannoutsou 2012), the presence or absence of a narrative or story-line (Avouris & Yiannoutsou 2012) and the decision on

[12] www.flickr.com (accessed: 25.03.2017)
[13] www.geograph.org.uk (accessed: 25.03.2017)

whether the game is team based or played without teams (Matyas et al. 2008; Matyas et al. 2012; Matyas 2007; Celino et al. 2012; Winter et al. 2011; Yoshii et al. 2011; Davidovic et al. 2013). The following table highlights these key characteristics regarding a selection of the most prominent location-based games analysed in a scientific context.

Table 1 - Key characteristics of LBGs in a scientific context
Table presenting key characteristics of the most prominent location-based games in the reviewed literature

Papers	Game	Game Field	Data	Duration	Narrative	Team
(Celino et al. 2012; Celino 2015)	Urbanopoly	Semi-structured	POI Information	Continuous	Weak	Everyone for themselves
(Davidovic et al. 2013)	MapSigns	Semi-structured	POI information, focus on street signs	Short – medium	Weak	Team based, teams made before every round, 2 teams
(Winter et al. 2011; Richter et al. 2012)	Tell Us Where	Unstructured	POI information, focus on place descriptions	Short – medium	None	Everyone for themselves
(Matyas 2007; Matyas et al. 2008)	CityExplorer	Semi-structured	POI information, focus on POIs defined before a game session	Short – long	Weak	Team based, 2 teams
(Matyas et al. 2012; Matyas et al. 2011)	GeoSnake	Structured	Path & POI information	Short – medium	Weak	Everyone for themselves
(Avouris & Yiannoutsou 2012; Neustaedter et al. 2013; Matyas 2007)	Feeding Yoshi	Unstructured	POI Information focus on open and closed WiFi hotspots	Continuous	Weak	Everyone for themselves
(Yanenko & Schlieder 2014)	Alien GeoSpy	Semi-structured	POI information focus on POI categories in area	Continuous	Weak - middle	Everyone for themselves

The literature agrees that location-based gaming can be a useful tool for (geospatial) data acquisition, (geospatial) data validation and for edutainment purposes (Matyas 2007; Yoshii et al. 2011; Celino et al. 2012; Ionescu et al. 2013; Davidovic et al. 2013; Matyas et al. 2012; Charsky 2010; Matyas et al. 2011; Avouris & Yiannoutsou 2012; Richter et al. 2012; Matyas et al. 2008; Winter et al. 2011; Celino 2015; Yanenko & Schlieder 2014). Various authors have analysed the usability of location-based games for data collection (Matyas et al. 2008; Matyas et al. 2012; Matyas 2007; Celino et al. 2012; Winter et al. 2011; Davidovic et al. 2013) focusing primarily on POI information collection. Matyas (2007) and Matyas et al. (2008) assess the game *CityExplorer*, a location-based game, in which players capture a tile when they reach the highest number of markers in a tile. Markers can only be placed on predefined location types (e.g. restaurants, beer-gardens, train stations), thus encouraging the collection of specific points of interest in a predefined region. Similar in some regards is the game *Alien Geospy*, analysed by Yanenko and Schlieder (2014), in which players are tasked with mapping categories of items in pre-defined areas. Regions will change colour depending on how many items from how many categories were mapped by the user. Celino et al. (2012) present the location-based game *Urbanopoly*, where players buy virtual-properties (e.g. restaurants, theatres, shops) and where properties may be snatched away from other players. Unlike *CityExplorer*, *Urbanopoly* allows for continuous gameplay and has no start or end of a game session. It does not only focus on the collection of information on POIs, but also on verifying, correcting and enriching existing information found on OSM. *MapSigns*, presented by Davidovic et al. (2013), is an attempt to use location-based gaming to motivate users to collect niche datasets, usually with low importance for mainstream users (e.g. traffic signs, park benches, trash cans). *Feeding Yoshi* is mentioned in various papers (Avouris & Yiannoutsou 2012; Neustaedter et al. 2013; Matyas 2007) and is a LBG to map open and closed WiFi hotspots. *Feeding Yoshi* also allows for continuous gameplaying and, like all games mentioned above, focuses on the collection of point of interest data. The only location-based game found in the scientific literature which can vaguely be seen as not to only collect point of interest data is *GeoSnake* (Matyas et al. 2012; Matyas et al. 2011), a location-based game adaptation of the highly popular mobile *snake* game. As this LBG involves strategic routing decisions, it could be used to gather route information from different users. Another game, *Tell-Us-Where*, which focuses solely on the collection of place descriptions, is referred to in (Winter et al. 2011) and (Richter et al. 2012). In contrast to the above-mentioned games, *Tell-Us-Where* has no common gameplay or competitive elements. The users get a chance to win a gift voucher by verifying their GPS position and by describing their current location. This may be seen rather as a spatial questionnaire than a location-based game. It can also be argued that the *Geograph Project* is a semi-location-based game (the users have to go to a specific location to take a representative photo of the landscape) with a focus on tile-based geographic information mining, including competitive elements (e.g. list of high scores). The *Geograph Project* however does not use real-time location information of the users to allow or deny certain interactions, thus classifying as an asynchronous location-based game.

Many authors (Matyas 2007; Matyas et al. 2008; Matyas et al. 2012; Celino et al. 2012; Winter et al. 2011; Davidovic et al. 2013) mostly agree that location-based gaming can be used as an effective tool to collect large amounts of spatial data and that the gaming aspects suffice to motivate users to contribute data over a longer period of time. Not only are location-based games viable as data collection tools, but also for data verification and curation purposes (Celino 2015; Yanenko & Schlieder 2014). Celino (2015) presents *Urbanopoly* from the data curation perspective and proposes a methodology to verify and enrich the data of OSM. The author concludes that applying "the power of

Human Computation to Citizen Science" (Celino 2015, p.9) using location-based games "can bring effective tools for geospatial data curation by exploiting the physical presence of the contributors in the environment" (Celino 2015, p.9). Yanenko and Schlieder (2014) primarily focus on the data quality improvement mechanisms of "confirmation" and "retesting". The authors implement both mechanisms in a location-based game designed for assessing the two data quality improvement mechanisms and their findings show that both mechanisms have a positive impact on data quality. Another positive impact of using the dual mechanisms of "confirmation" and "retesting" is, according to the authors, a decrease in the probability of players cheating.

A number of researchers (Charsky 2010; Ionescu et al. 2013; Avouris & Yiannoutsou 2012) have conducted broader research on the topic of (location-based) gaming in a scientific context. Ionescu et al. (2013) propose a multiplatform framework for developing location-based games or transitioning existing games to a location-based game style. In (Charsky 2010), key characteristics of serious games (games which incorporate instructional and video game elements but are not used for entertainment) and edutainment games (games which combine education and gameplay) are presented and discussed. These include competition, goals, rules, choices, challenges and fantasy. Even though these characteristics are discussed as underlying elements of serious or educational games, they also apply to location-based games and games in general. Most noteworthy are the positive effects on motivation by using competitive elements (Lund et al. 2010), and fantasy elements (Kenny & Gunter 2007), to immerse a player in a game and ensure longer and more frequent gameplay. Avouris and Yiannoutsou (2012, p.2121) state that a solid narrative is "a valuable tool for [the] construction of meaning" and that the "narration is a means for combining different heterogeneous parts (actions, events, etc.) into a coherent whole and crafting the relationships between these different parts". In the context of location-based gaming for geographic information mining a strong narrative can thus be used to immerse the player into the game world and create a continuous and coherent story, motivating the player to continue playing.

2.4 User Motivation

Crowdsourcing is a powerful tool to generate large amounts of data through non-expert users. There are many reasons, why users contribute to data collection, including volunteerism and monetary incentives (Hoe et al. 2017), but also personal goals such as advancing one's career, being able to express oneself, working in collaboration with other interested users and getting to know new technologies (Brabham 2012). These incentives have an impact on the users' motivation and enjoyment. Thus, the success of a project collecting volunteered geographic information is highly dependent on the incentives used to ensure non-expert user motivation (Hoe et al. 2017). Virtual rewards are of interest in crowdsourced environments because the data generation task is often not motivating enough to attract many non-expert users. Through virtual reward systems users can be motivated with "extrinsic reward[s] like virtual items while not really enjoy[ing] the play activity itself" (Wang & Sun 2011, p.2).

The vast majority of games use virtual reward systems as a fundamental motivational mechanism to ensure player satisfaction and thus increase the time a user invests in playing a game (King et al. 2010). Virtual rewards can be seen as a virtual proxy of the time and effort a player has invested into a given task in a game and are often comparable and communicable with other players (Hoe et al. 2017). The literature agrees that there are two predominant types of virtual rewards: points and badges. Points are widely used to indicate a player's progression in the game and they encourage competition amongst players (Hamari 2015). Badges, on the other hand, symbolise a goal which can be achieved by multiple users, thus intensifying the feeling of belonging to a specific group (Hoe et al. 2017). Mekler et al. (2015) tested the effect of varying reward systems on user motivation and performance in an experiment. Subjects playing a game which incorporated either a leader board or a level system were found to deliver significantly more content than subjects playing a game with a point system, but they still generated significantly more content than subjects playing a game without any reward system at all (Mekler et al. 2015). Relatedness is also mentioned as a key variable with the potential to increase intrinsic motivation (Hoe et al. 2017). Relatedness is defined by Deci and Ryan (2000, p.231) as "the desire to feel connected to others" and Hoe et al. (2017, p.368) argue that "if an activity allows interaction with others, an individual will likely experience a sense of connectedness thereby increasing intrinsic motivation."

3 Research Gap

The reviewed literature reveals multiple research gaps in the domains of crowdsourced land cover classifications, location-based gaming and the combination thereof. Of particular interest is that the use of games has been identified as having a great potential to make land cover validation more enjoyable and to attract a large amount of users (Fritz et al. 2009). However, to my knowledge, no research has been done on implementing a real-time tile-based location-based game for said purpose. Location-based gaming for (geospatial) data acquisition, (geospatial) data validation and for edutainment purposes has been widely discussed in the scientific literature, but no literature is available concerning location-based games for tile-based information mining.

Most of the games shown in Table 1 concentrate on collecting POI information. Only *GeoSnake* could qualify as a data mining application with which route data could be collected. Furthermore, the duration of a game is predominantly short to medium with only *Urbanopoly*, *Alien Geospy* and *Feeding Yoshi* allowing for a continuous gameplay, hence allowing continuous data collection. All of the studied games have either no storyline or narrative or only a weak one, with the exception of *Alien Geospy*.

Finally, the games focus around the player distribution concepts of "everyone for themselves" or split into two teams. This clearly highlights a research gap in location-based gaming implementation and research, namely, in the implementation and analysis of a location-based game focusing on tile-based geographic information collection with a continuous gameplay duration, incorporating strong narrative characteristics and faction based team play.

The above presented research gaps regarding land cover classification and verification and the research gaps concerning location-based games for geographic information mining can be combined to formulate the overarching goal of this master's thesis:

The development, implementation, assessment and analysis of a tile-based location-based game with continuous gameplay, including narrative as well as competitive elements, focusing on geographic information mining regarding land cover data and the analysis of the generated data.

---

4 Research Questions

The research questions revolve around the two main topics of implementing a location-based game for geographic information mining and the analysis of the generated data:

Research Question One

How can a location-based game with a non-expert target audience be implemented to mine tile-based geographic information, in particular land cover data?

Research Question Two

Can the generated land cover data be used in a research context, particularly in regards to the validation or improvement of land cover products?

---

5 Methods

This chapter outlines the fundamental concepts and methods I used to implement a location-based game and analyse the generated land cover classification data. The first section describes the implementation, including the underlying infrastructure, fundamental game principles and the gameplay. Importantly, the game was improved through an iterative development, and I describe this process and its impacts on the game and users' behaviour. The second section introduces the methods I applied for analysing the generated data in terms of users' distribution and for assessing the agreement between tiles classified by multiple users as well as for the comparison of the land cover classifications generated by users with the official CORINE land cover datasets.

5.1 Implementation

This section highlights concepts and methods used to implement a location-based game for geographic information mining regarding land cover information. Underlying concepts and basic gameplay characteristics are introduced as well as infrastructure choices and key improvements that were made to the game in a participatory process.

5.1.1 Game Concept and Basic Gameplay

The defined goal of the location-based game, which I named "StarBorn", was to generate land cover data from users, exploiting their ability to sense their immediate surroundings through visual and auditory stimuli (cf. Wang & Ben-arie 1996). Therefore, the game only allowed users to classify land cover data of their immediate surroundings by using the built-in GPS functionality of smartphones. I defined the list of possible land cover classes the users could select from based on the official CORINE land cover classification scheme[14]. I chose the second level of detail for the land cover classes in order to assure a balance between the level of detail with which users can classify the observed land cover and the number of selectable classes. In total, the CORINE land cover dataset includes 15 land cover classes, whereas only 13 are applicable for Switzerland (two marine land cover classes are not applicable). To enable a smooth user experience and make the contribution of land cover data attractive for a wide audience without using monetary incentives, I based my location-based game on concepts of highly popular online- and board-games. One inspiring example was the board game *The Settlers of Catan*[15], which features the island of *Catan* made up of hexagonal tiles with different land cover types. Players can acquire specific resources by building settlements on the respective tiles. The resources can then be spent on developing settlements to acquire more resources.

I adapted this basic idea to develop an online game focusing on the collection of land cover information. Raster cells were chosen over hexagonal cells to reduce the complexity of implementation

[14] http://uls.eionet.europa.eu/CLC2000/classes/ (accessed: 02.04.2017)
[15] www.catan.com (accessed: 01.02.2017)

and to better correspond with the CORINE data, with which the generated data was then compared. To encourage competitive playing over tiles, I chose tiles with extents between the 100m x 100m and 250m x 250m cells in the CORINE datasets. The tiles needed an extent large enough so that an effort on behalf of the users is needed to change which tiles they can interact with. Additionally, taking the positional assisted-GPS accuracy of ~9m (Zandbergen 2009) into account, an extent of 200m x 200m was chosen for the tiles. In the game the user is prompted to capture tiles for his or her team by examining the real-world location corresponding to the tile displayed in the game and supplying land cover information using the pre-defined list of land cover classes. The players can destroy tiles captured by an enemy team and recapture the tiles for themselves, allowing multiple users to provide classifications of the same tile and multiple captures of a tile by the same user. The following sections describe the technical solutions I chose for the implementation of the location-based game "StarBorn".

5.1.2 Relational Database and Graph Database

The following section describes the two databases I used in the implementation and explains why a graph database was chosen as the primary data storage and why a PostGIS database was implemented as a complex spatial index.

In recent years there has been an increase in the use of graph databases to model complex relationship structures (Joishi & Sureka 2015). Modern services handling large datasets of information (e.g. Facebook[16], Google[17], Twitter[18]) have reached the limitations of relational databases and have migrated their systems to underlying graph database systems (Joishi & Sureka 2015; Kolomičenko et al. 2013). These database systems have high computational advantages over conventional relational databases, especially for performing complex relationship queries (Vicknair et al. 2010). A graph database is extremely adaptive and excels at scaling with growing applications. Queries to a relational database take longer to process with growing database size, whereas queries to a graph database show more or less constant response times if optimal queries are used, because the query only needs to be executed on a subsection of the graph and not on the whole graph (Joishi & Sureka 2015; Vicknair et al. 2010). Various authors argue that graph databases will become more important, especially with growing spatially relevant data collections which have to be queried (Joishi & Sureka 2015; Kolomičenko et al. 2013; Vicknair et al. 2010).

Keeping the above mentioned literature in mind, I chose Neo4j as the primary data storage, a graph database written in JAVA. In Neo4j all game relevant information is stored in nodes (data points with attributes; Appendix 11.1 for a table of implemented nodes) and relationships (relationships between

---

[16] www.facebook.com (accessed: 22.03.2017)
[17] www.google.com (accessed: 22.03.2017)
[18] www.twitter.com (accessed: 22.03.2017)

the data points with attributes; Appendix 11.1 for a table of implemented relationships) in an interconnected graph. Graph databases excel at complex queries. They do justice to the strongly interlinked nature of the data stored in the database by allowing analytical queries such as: *"Which user has the most buildings on tiles which used to belong to the other team?".*

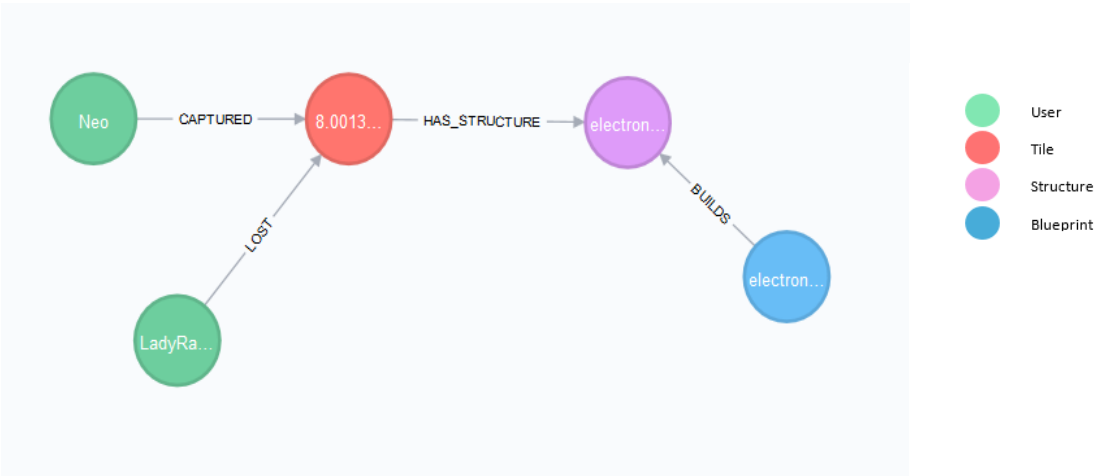

Figure 4 - Node relationship structure in Neo4j
Figure showing an example of the node relationship structure in the implemented Neo4j database.

Figure 4 highlights the useability of a graph structured database for highly related data and gives an example of how the data is structured and stored as nodes (circles) and relationships (arrows) in Neo4j. This example shows a user (screenname: Neo) who has captured a tile. This tile has a structure built onto it and the structure has a connected blueprint, which is used to build the structure. The tile had previously also been captured by another user (screenname: LadyRailbird). The tile was then attacked and the user (screenname: LadyRailbird) lost the tile.

Another advantage of the graph database Neo4j is the possibility to write and efficiently execute complex queries over several relationships. Figure 5 shows the result of such a query. This example query written in text form would be: *"Show me all the users who are in current possession of a tile which used to belong to the user with the screenname LadyRailbird and their relation to other tiles which used to belong to the user with the screenname LadyRailbird".* Not only is the possibility to write and execute complex queries of high importance for running the game and ensuring a satisfactory query performance, but also for exporting and analysing the data.

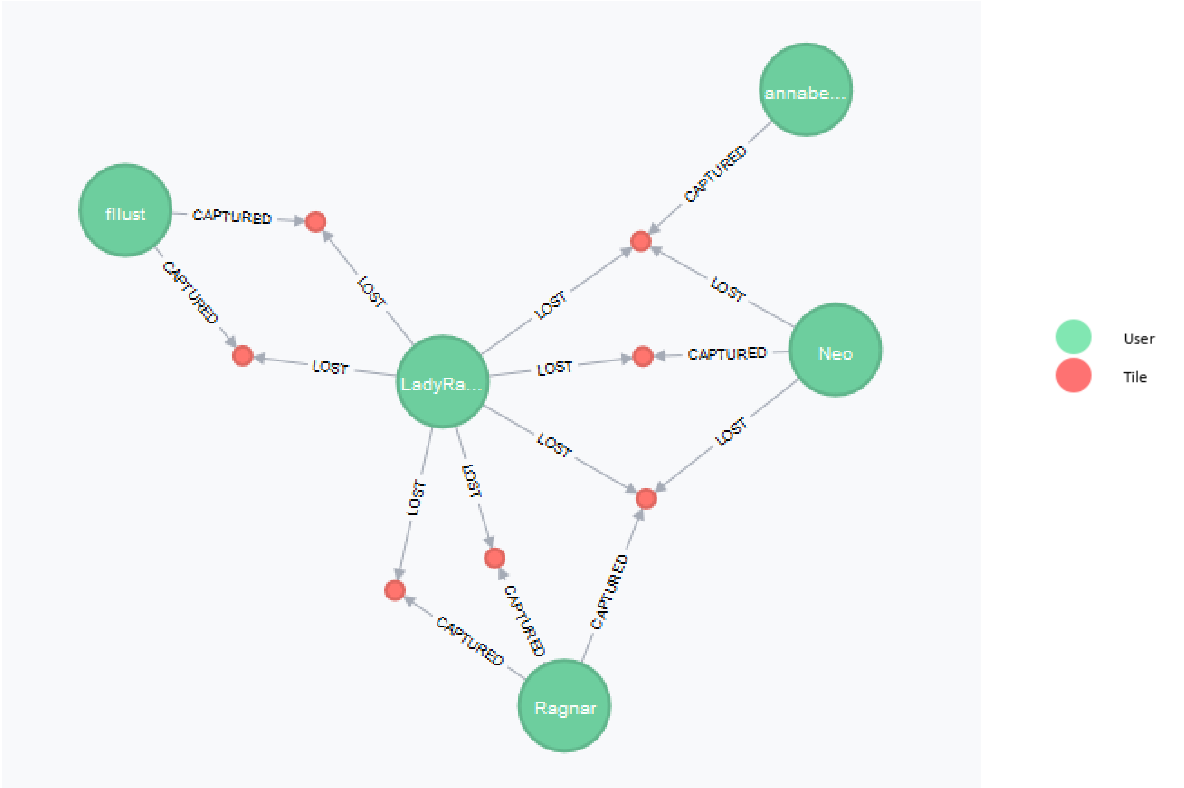

Figure 5 - Complex query result in Neo4j
Figure showing an example of a result of a complex query in Neo4j.

After implementing Neo4j as the primary data storage, a solution was needed to efficiently store and query spatial data. I installed the Neo4j spatial plugin to test its usability for the location-based game. Neo4j spatial provides the possibility to save nodes with geographic coordinates in an indexed rTree, drastically reducing query time for geographic queries. Realising that for the proposed implementation the storage of the entire country (Switzerland) in 200m x 200m cells was of crucial importance, millions of points had to be created. A small test dataset containing 100'000 points was created to test query speed. The tested queries took >20s, which was considered too long for a real-time gaming experience. The preliminary goal query time for spatial queries in the proposed implementation was around 100ms up to a maximum of 500ms. As Neo4j spatial did not offer the necessary speed for spatial queries, I implemented the most stable open source solution available, which was Postgres with the PostGIS extension. I created a raster data table containing a raster of 200m x 200m cells, stored using the new national equal-area projection coordinate system of Switzerland LV95 (EPSG: 2056) and covering the whole area of Switzerland. Using this approach, I achieved response times for spatial queries of under 600ms. After optimising the queries, splitting the raster into smaller raster tiles and adding a spatial index to the raster tiles the query time dropped to around 80ms, which I deemed appropriate for the implementation.

PostGIS raster tables unfortunately only allow storing integer or float types in different bands. For the proposed implementation additional information including timestamps, land cover classifications and bounding boxes needed to be stored. Whilst chosen attributes such as timestamps could easily be transformed to integer or float values, storing a list of strings such as the land cover classifications was more complex. To store the predefined list of land cover classifications the list had to be transformed to an integer between 1 and ($2^{13} - 1$) which, written as a binary number, would then signal the presence or absence of a specific land cover class.

E.G. WITH A TOTAL OF THREE LAND COVER CLASSES "URBAN", "INDUSTRY" AND "WATER", IF A USER REPORTED "URBAN" AND "WATER" AS BEING PRESENT, THIS COULD HAVE BEEN TRANSFORMED TO THE BINARY NUMBER 101, WHICH CORRESPONDS TO THE INTEGER NUMBER OF 5, WHICH COULD EFFICIENTLY BE STORED IN A POSTGIS RASTER BAND.

I deemed this solution too complex for debugging and analysing and therefore implemented a workaround. The implemented solution was to store IDs in the PostGIS raster tables, corresponding to specific node IDs in the Neo4j database or corresponding to specific game related attributes. The PostGIS raster acts as a complex spatial index to identify the needed entry nodes in the Neo4j database or to identify potential special cases in the game. As a PostGIS raster can only store one value per cell per band, multiple raster bands were implemented to allow storage of all the needed information. A total of four bands was used for the implementation of the location-based game:

- Band 1: the ID of the user who captured the tile
- Band 2: the ID of the team the user belongs to
- Band 3: the ID of the tile itself
- Band 4: an indication of whether a treasure is present in the tile or not

The IDs stored in bands 1 – 3 point to specific nodes in the Neo4j graph database, which store vital information regarding the game, whereas raster band 4 stores information which is directly used in the Symfony controller, without retrieving additional information from the Neo4j graph database. I coded the server side application using the Symfony Framework, a PHP server side framework for complex web-application development. Symfony controllers are functions containing parts of the game logic. They transform the users' requests to the game into a viable response according to the inherent game logic coded into the controller. The official Symfony documentation[19] summarises the function of a controller as "execut[ing] whatever arbitrary logic your application needs to render the content of a page" (Symfony n.d.). This pipeline enables efficient querying, even with large datasets, resulting in final page loading times of around 120ms.

[19] http://symfony.com/doc/current/index.html (accessed: 05.03.2017)

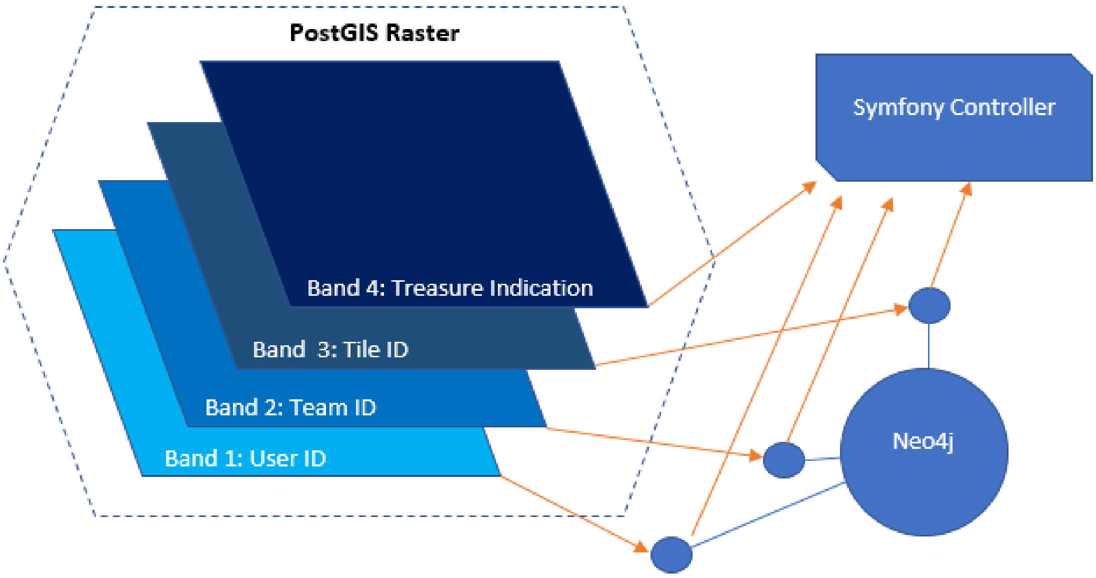

Figure 6 - Database interactions
Figure showing the four implemented PostGIS raster bands and their interactions with Neo4j and Symfony.

Figure 7 presents the underlying pipeline of the key infrastructural elements and their interactions with each other. A user sends a request using the browser and its location capabilities. The request is sent to the PHP server and is then interpreted by the Symfony framework. The Symfony framework routes the request to a Symfony controller, which first queries the PostGIS raster. PostGIS raster queries return a result which includes the ID of the tile, the ID of the team who owns the tile and the ID of the user who owns the tile, all specific to the location from which the browser sent the request. The IDs, now available in the Symfony controller, are then used to identify the needed access nodes in the Neo4j database. The information retrieved from the Neo4j database and the PostGIS database is combined to a response and sent to the browser of the user. The browser then interprets the response using HTML, CSS and JavaScript and displays the information in a visual feedback for the user so that every time a user acts in the game, a request is sent to the server, which then decides on how to react to the action of the user. The reaction is then sent back to the user in a visually pleasing manner.

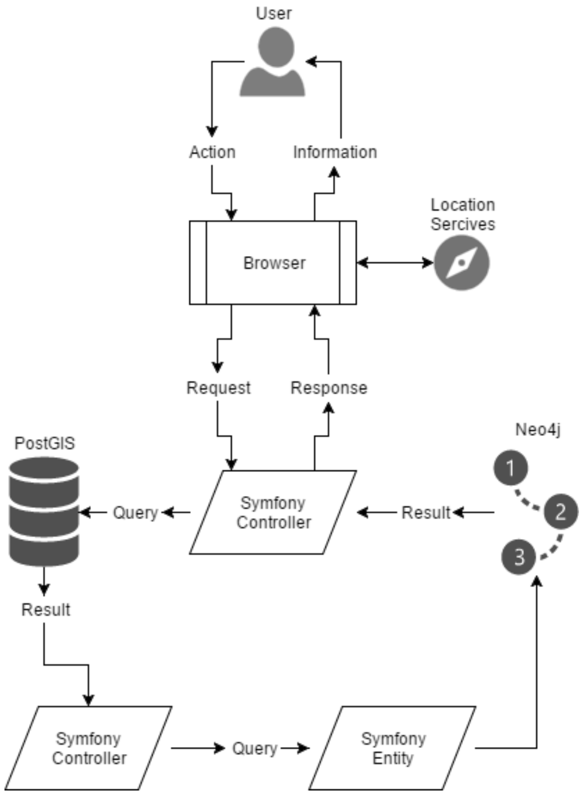

Figure 7 - Key infrastructure elements and interactions
Figure showing the key interconnected infrastructure elements of the implemented location-based game and the corresponding information flow.

5.1.3 Web Infrastructure and Interface

I implemented the location-based game as a browser based web-application, which uses modern browsers' built-in location capabilities. A browser based solution decreases the complexity of implementation and allows all devices with a modern location enabled browser to partake in the game, not limiting access to a specific device model or operating system. I implemented the game in PHP, JavaScript, HTML and CSS. All game logic is either stored in the Neo4j database, PostGIS database or coded in the Symfony framework.

As with most web-applications, the implemented game has a main page from which the user can access various subpages. In order to increase security and to identify what user-specific information must be displayed, the main page could only be accessed when logged in. It consisted of a large map area, in which the tiles in the immediate vicinity of the user are shown and coloured according to which team owns the tile. The user's exact location is shown with a small icon coloured in the colour of the team the user chooses to affiliate him or herself with. At the bottom of the screen, a bar shows the user's current level and the experience points needed to reach the next level in form of a proportionally coloured bar. In the middle of the screen, a menu button can be found. On clicking the menu button a pop-up appears showing the various menu options (cf. Figure 10 left).

5.2 The Implemented Location-Based Game and its Iterative Development

The following section sheds light on the basic gameplay characteristics and how key improvements were identified and implemented over the course of implementing and running the location-based game "StarBorn".

5.2.1 Basic Gameplay

In the following section, I describe a typical experience of a user who signs up for the game as a way to illustrate the general idea of the gameplay. The game was designed to have a static procedure when new users register to facilitate an easy entry into the game. A newly interested user can sign up using an online form. Basic user data is needed to be able to sign up, including a desired username, a screenname, a password and the user's email address. Optional data can also be provided including the indication of gender and age of the user.

Once the user has submitted the needed information, the user is redirected to a page asking the user to check his or her email to confirm the email as valid. The user receives an email with an invitation to visit the "Star-Forge" and create an avatar. This confirmation email has a multi-facetted purpose. On the one hand, it is used to verify if a user submitted real contact information minimising the risk of spammers or cheaters. On the other hand, the email is a first introduction into the game's virtual fantasy world. Once the user clicks on the link in the mail, the user is asked to create an avatar (a self-designed visual representation of him- or herself; cf. Figure 8), which serves as the user's profile image in the in-game profile.

Figure 8 - Email confirmation and avatar creation pages
Figure showing examples of in-game screens. A screenshot (left) shows the page a user is directed to after confirming his or

her email address and another screenshot (right) shows an example of the avatar creation screen, where users can create a visual representation of their in-game character, which serves as a profile image.

After the user has created an avatar, the user must choose which faction or team he or she would like to affiliate him- or herself with. Both factions are introduced with a brief background story (cf. Figure 9).

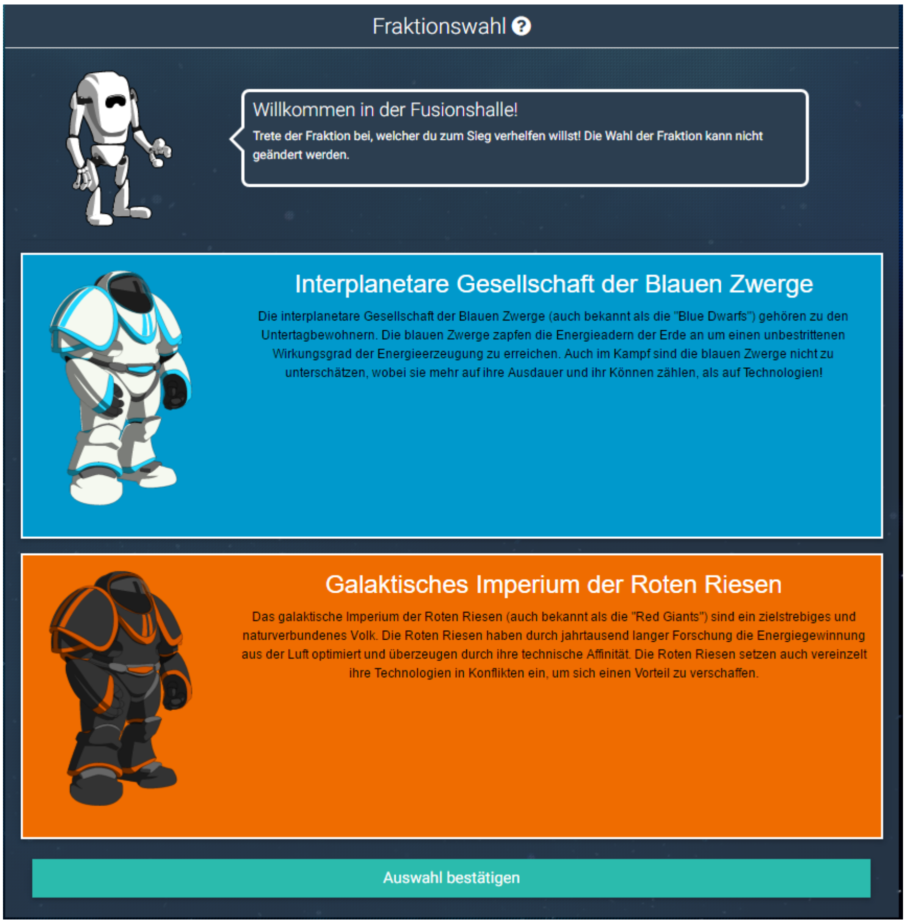

Figure 9 - Team selection page
Figure showing the team selection page on which new users must choose which team they would like to affiliate themselves with.

After successfully selecting a team, the user is directed to a tutorial page, explaining the various functions of the game and how the gameplay works. The tutorial includes a link leading the user to the game's homepage depicting a map of the location of the user and the user's immediate surroundings. The map-view shows the position of the user with a small icon, which is coloured according to the user's chosen team affiliation (cf. Figure 10 left). In a radius of 700m around the position of the user, colour-coded tiles are shown. The tiles correspond to the colour of the team which is currently in possession of the tile. Grey indicates that the tile is currently not in possession of either team.

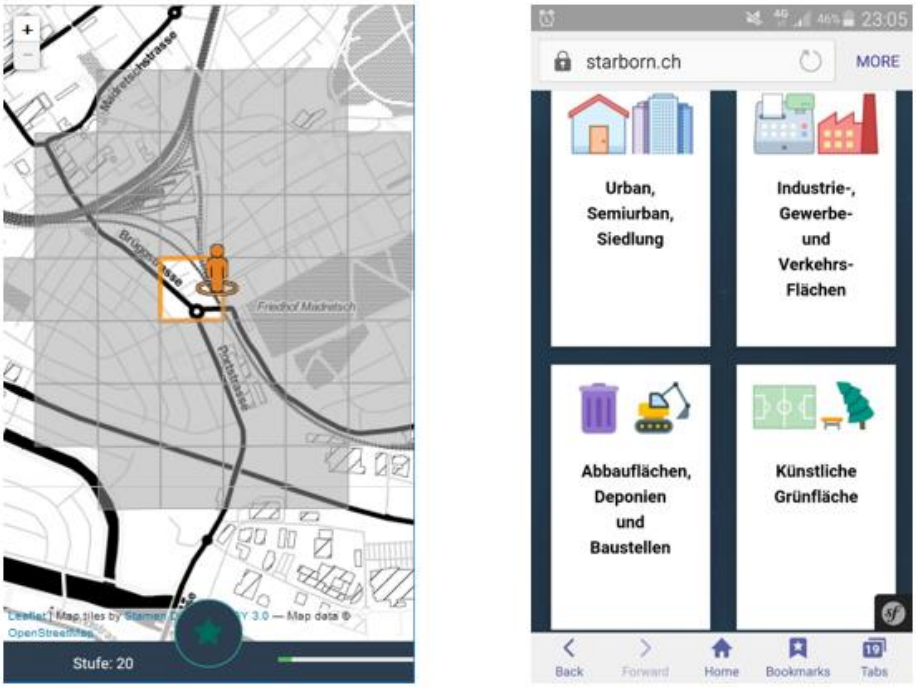

Figure 10 - Main map and capture interface
Figure showing the main map page (left) and the capture interface (right) with which a user reports land cover classifications.

As presented in the concept, the key functionality of the implemented game revolves around users generating land cover data in a playful and engaging way, with the game being the primary motivation or incentive to generate data. Users can sign up and play the game for free and without being confronted with any kind of advertisements. When registering, the user agrees that the generated data can and will be used for scientific purposes. Once signed in, the users can virtually capture real-world areas of 200m x 200m by being in said area (or at least, the location in which the browser reports the user to be) and reporting the land cover classes the user recognises in the area from a predefined land cover classification scheme (Figure 10; right image). In addition to providing the users with the name of the different land cover classes in linguistic terms, each land cover class is also visualised through two icons. The icons were implemented to help the user differentiate between the different land cover classes, which are not always easily distinguishable for non-experts. My choice fell on icons rather than real images in order to underline that the abstract visual aids provided should merely be a hint at what the land cover class could look like.

Once a user has successfully captured an area, he or she is rewarded with two virtual currencies, the common "stardust" and the rare "ethertokens". The user also has the option to collect resources his or her captured tiles are generating in a daily interval. Once a user has successfully captured an area, the area is allocated to the team of the user. Now any user of the same team has the option to build shields, structures or drones onto the area, all of which can be built using blueprints. These blueprints can be bought in an in-game shop for specific combinations of the two in-game currencies. Different shields, structures and drones have varying specifications such as health points or bonus effects (e.g. generating bonus resources). Users can attack areas of the opposing team by attacking the shields, structures and drones built onto the area. Every user has two different attack possibilities, a primary attack, which does low damage but can be used frequently (every ten seconds), and a secondary attack, which does considerable damage but can only be used from time to time (every three minutes). Once all shields, structures and drones have been destroyed, the area is marked as uncaptured and may be recaptured by any team. The basic decision making processes are depicted in Figure 11. The presented flowchart also highlights key gameplay elements.

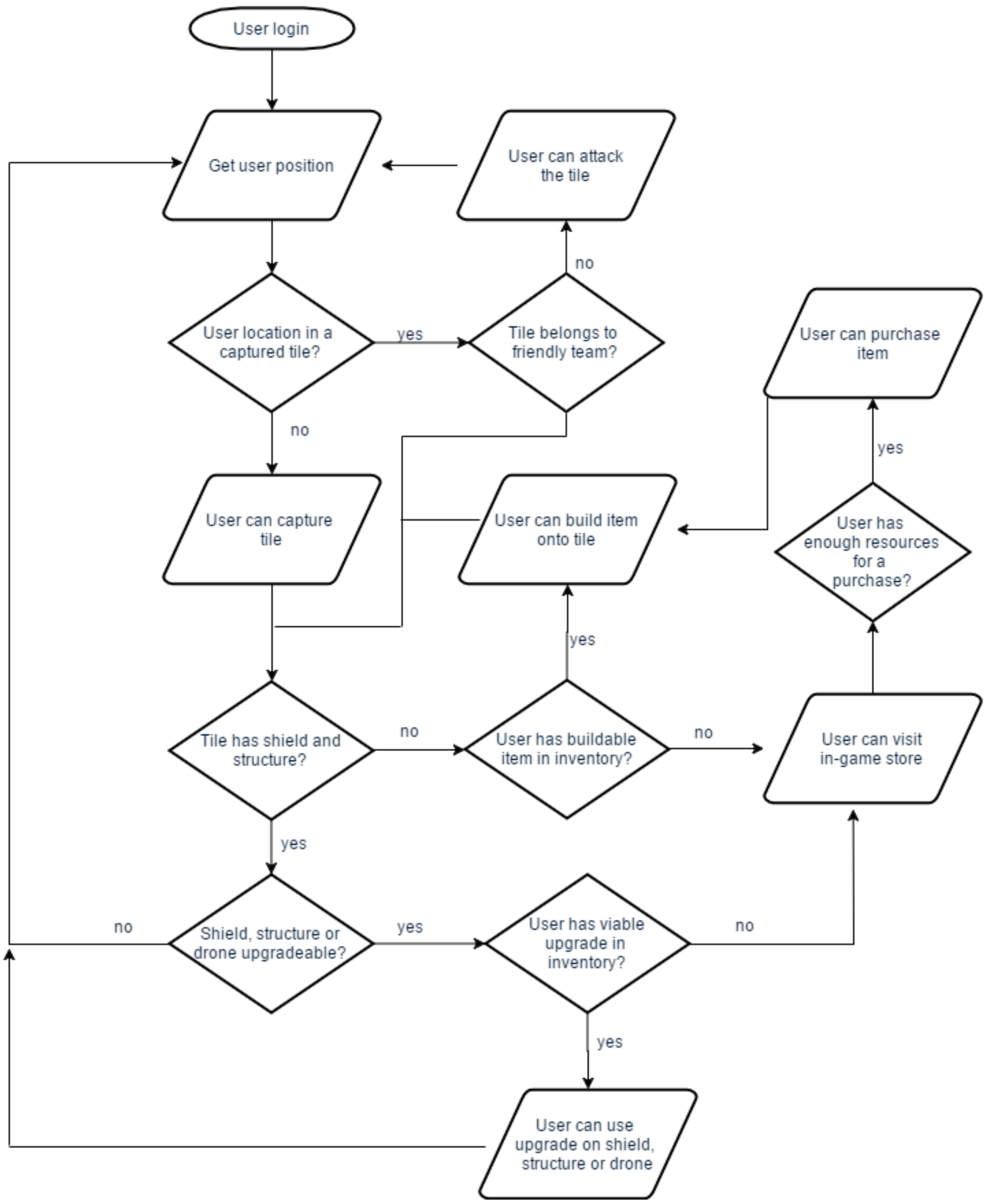

Figure 11 - Key decision flowchart
Figure showing a flow diagram of simplified general decisions in the implemented game representing the gameplay.

5.2.2 Key Improvements

Following a closed implementation period, the implemented location-based game was beta-tested with a chosen audience of five users to eliminate any major bugs and to test performance. After this initial test phase, I presented the game to the public and actively promoted the game verbally and on social media platforms including Facebook[20] and Google+[21], which resulted in users beginning to play the game and promoting the game themselves. The users were actively encouraged in verbal exchanges and in a Google+ community to give feedback, report bugs and propose improvements and ideas. Many users used this opportunity and reported various bugs and potential improvements. The improvements were implemented according to the reports or ideas of the users in an iterative process whilst the game was in the beta- or live stage.

In this section, I present the key improvements that I implemented due to user feedback. The aim of this section is to present key changes and improvements to the game and the reasoning behind them.

Design Theme: Light vs. Dark – Two theme proposals were individually presented to 13 persons belonging to different genders, ages and with or without experience in gaming or location-based gaming. A core part of a game involves the overall design of the game, which should then be consistent throughout the game. Two different versions of the overall design were created. One a light-coloured version, and the other a dark-coloured version.

Figure 12 - Light versus dark themes
Figure showing two initial design ideas: a light theme (left) and a dark theme (right).

[20] www.facebook.com
[21] https://plus.google.com

The users strongly favoured the dark theme. Additionally, users with a background of gaming argued that the dark theme fits more into the game fantasy of a futuristic game with the name "StarBorn". As a result, the dark theme was chosen for further work.

Resource Management – The initial idea was to base the implemented location-based game on the board game *Settlers of Catan*. A major part of the gameplay of this board game is the collection and management of various resources (e.g. stone, wood, food). I originally planned to implement a game having six resources to manage (stone, wood, water, food, work force, overwatch), which the users earned depending on the land cover of the captured areas. Three users who have had long lasting gaming experiences were asked for their opinion on the proposed resource system. In short interviews, in which I asked the users' opinion about the proposed resource system, all users stated that the intended resource system was too complicated and that there were too many resources to manage. The interviewed persons all agreed that they would enjoy the game more if there were less resources to manage and if the land cover of the area they captured had no influence on the gained resources. Two of the interviewees elaborated on the second point and stated that they might feel inclined to falsify the land cover information when capturing an area if the rewarded resource is dependent on the land cover recorded.

Based on the information gained from these interviews, I drastically simplified the resource system. In the ultimately implemented game, users receive a virtual currency called "stardust" when performing various actions including capturing or destroying a tile. To act as an incentive to capture areas and not only destroy areas, a secondary, rarer currency called "ethertokens" was introduced, which users only receive when capturing areas. Users can buy virtual items or upgrades with a specific combination of these currencies from an in-game store.

Level System and Experience Points – Whilst conducting the small interviews for the resource system, one user pointed out that due to the lack of a level system he felt no sense of progression. Various papers (Wang & Sun 2011; King et al. 2010; Deeds 2016; Adams & Dormans 2012) reveal that experience points or level systems constitute a vital part of a user's sense of progression. I thus implemented a user level and experience point system. The users' level is the floored square root of the users' experience points. Giving a user equal experience points independent of the level of the user for specific actions results in a non-linear player progression in terms of levels. Thus, the users' progress is fast in the beginning but with increasing levels, more experience points are needed to reach the next level. Adams and Dormans (2012) identify this non-linear level progression as a desirable negative feedback mechanism. To enforce a player's sense of progression, buying different items available in the in-game store requires a specific minimum player level. In addition, the items with higher level and/or currency requirements are more desirable than cheaper items obtainable at lower levels. More expensive items or items unlocked at higher levels generally have more health points or have special effects such as generating a supplementary daily extra production of one of the two mentioned currencies.

Treasure Hunt – I presented the implemented game in a citizen science lecture at the University of Zürich where the students were asked to play the game for two weeks and give feedback about their experiences. One of the major points raised in the feedback was that the users had no incentive to collect tiles which were not on their direct path of travel. Thus, the users agreed that they would play the game as a form of distraction whilst traveling, but that they had no incentive to deliberately change their route of travel only for the purpose of playing the game. The students who gave feedback helped devise a new game feature to increase the users' motivation to deviate from planned routes. According to the feedback I implemented an in-game treasure hunt system, which led to additional resources as a reward. The treasure hunt system incorporated three different types of treasures: common or small, seldom or medium and rare or large treasures. The common or small treasure is depicted on the map as an icon showing a small pot of gold and grants the user who captures the tile containing the treasure 5 "ethertokens" and 1500 "stardust". The seldom or medium treasure is illustrated by an icon showing a treasure chest full of gold and grants the user who captures the tile which contains the treasure 10 "ethertokens" and 3000 "stardust". The rare or large treasure depicted by a diamond grants the most resources (15 "ethertokens" and 6000 "stardust") but is also the rarest type of treasure. Once a user captures a tile which contains a treasure, the user earns bonus resources and the treasure disappears. The treasure icons are visualised in Figure 13.

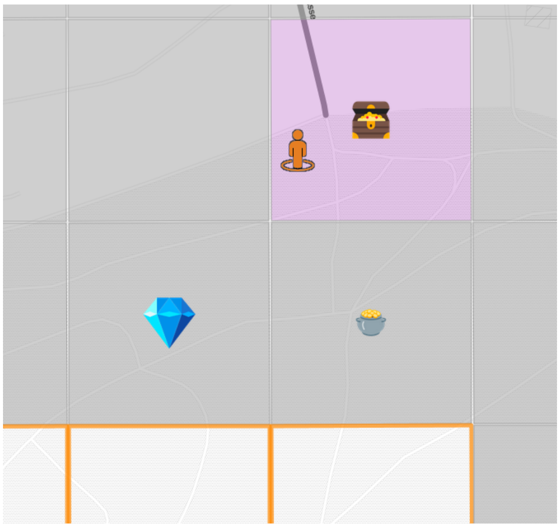

Figure 13 - Treasure hunt system
Figure showing the implemented treasure hunt icons.

The treasures were randomly distributed over the area of Switzerland. The goal of the treasure implementation was that users deviate from their intended path of travel to actively collect treasures and thus generate more resources for themselves. Additionally, the treasure hunt system had the goal of increasing the number of tiles for which users reported a land cover classification. Another potential impact is that the treasures in common areas, where many users play the game, are found and collected faster, thus motivating users to play in more remote areas hoping to find and collect treasures which are still collectable in the virtual game world. The user-treasure interactions (e.g. when did which user find a treasure, where was it and what was the size of the found treasure) were logged in a Postgres table, but were not evaluated or analysed in the scope of this thesis.

Special In-Game Events – To promote the game and to motivate players to come back and continue playing the game due to curiosity, I created and hosted special in-game events (as hypothesised in Cantallops & Sicilia 2016). In the duration of the game, the users witnessed two in-game events lasting from Thursday till Sunday. The first special event was dubbed "The double damage days" and the damage caused by all attacks was doubled. This effectively increased the competition between the two opposing factions. The event also increased the number of tiles with multiple classifications since in the event period, more tiles were destroyed and recaptured.

The second in-game event was dubbed "Resources are plentiful" and as the name suggests, the amount of resources allocated for capturing tiles was increased. The goal of this event was to motivate users to capture more tiles which had not previously been captured by other users. This also led to a slight increase in captures.

Catch-Up Mechanics for New Users – Two users mentioned the difficulty for new players to generate resources because more experienced players would already have constructed large in-game buildings on many tiles, which were hard to destroy. After brainstorming ideas to counteract this problem, I introduced a catch-up mechanic for new players, through which new players would deal more damage than experienced players. This improvement ensured that the gaming experience was more rewarding for newer players and effectively increased the number of destroyed tiles whilst increasing the number of tiles with multiple captures.

Competition – Regarding the elements of competition in the game, one user stated that a direct comparison between the teams and the users would be appreciated. The initial implementation incorporated no way of comparing the two opposing teams or the individual users of each team. As stated in the literature (Hoe et al. 2017; Hamari 2015; Senaratne et al. 2016), badges and points can be a useful tool for users to compare their achievements. This cultivates an in-game sense of competition, which can lead to greater user motivation. Thus, I created a new page showing the users ranked by the number of tiles each user has classified (cf. Figure 14).

Figure 14 - Rankings page
Figure showing the implemented rankings page (Screenshot saved: 27.02.2017).

In addition, the page shows the cumulative number of classified tiles of each of the two teams, fostering the competitive elements of the game. The newly created ranking-page had the goal of motivating users to capture more tiles as to benefit the whole team in the statistics. Furthermore, a title-and-icon-reward system was introduced. At certain predefined thresholds, users would be rewarded with a new title and icon, which is shown in the ranking-page and in the profiles of the individual users. The rewarded titles and icons range from the "discoverer" depicted by an icon of a compass and rewarded for capturing 20 or more tiles, to "extra-terrestrial" depicted by a spacecraft and rewarded for capturing 800 or more tiles.

5.3 Analysis of Data Generated by a Location-Based Game

This chapter revolves around the analysis of the generated data collected through the implemented location-based game. It highlights methods and underlying theories used to generate key results. All key used methods, pipelines, data flows, algorithms and temporary files generated are shown in Figure 15. This flowchart should act as a visual representation of how the generated data of the implemented game was analysed. In addition, the created flowchart also sheds light on the different interlinkages of methods and the varying in- and outputs which are used as underlying data for the generated results. A detailed description of each of the different workflows and algorithms can be found in the subsequent sections of this chapter. In combination with the flowchart, this section gives an extensive and comprehensive overview of the used methodologies.

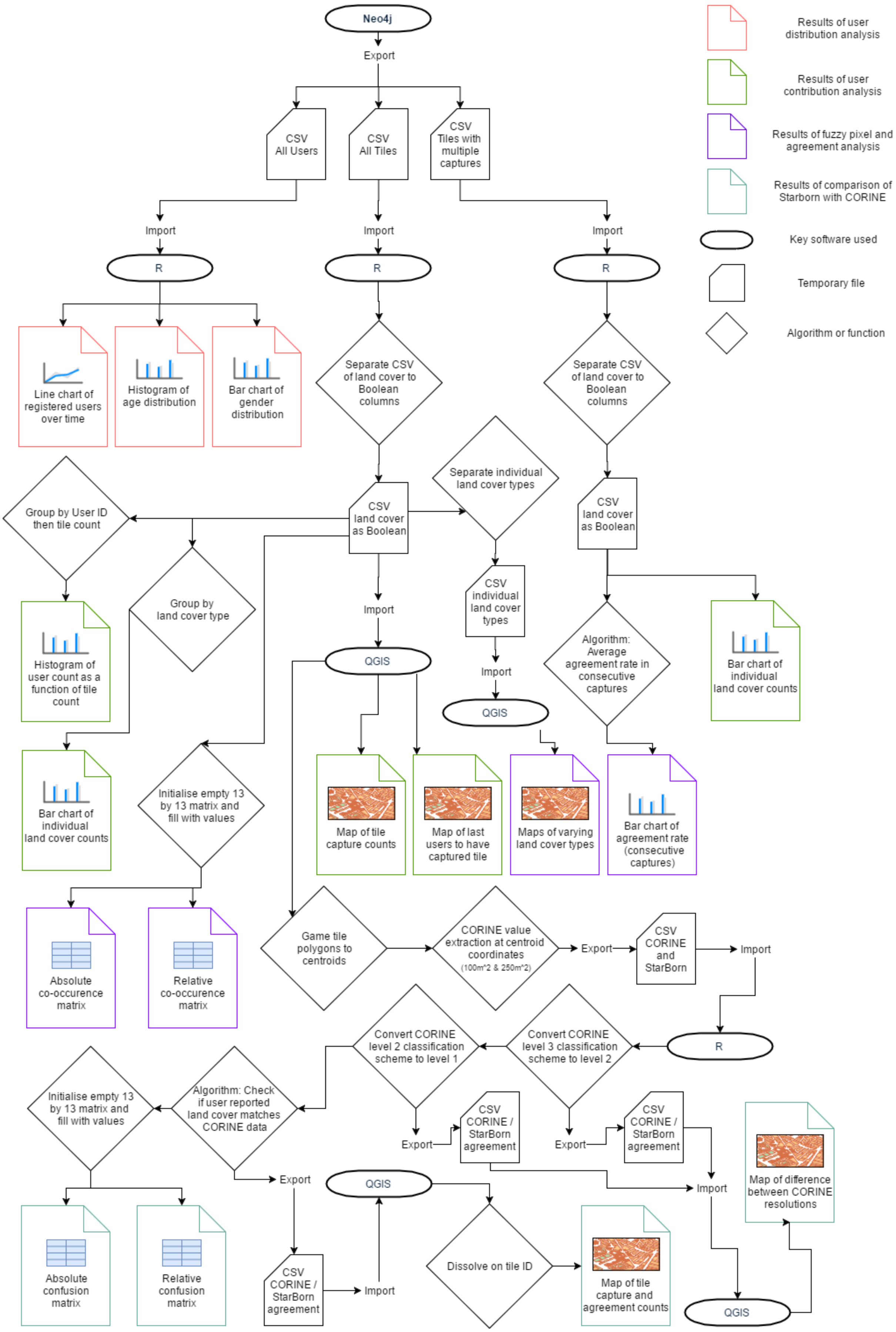

Figure 15 - Key analysis flowchart

Figure showing a flowchart depicting key analysis processes and their interconnections

5.3.1 User Distribution

After collecting the user generated data, the user distribution was analysed with a specific focus on the user distribution regarding temporal variations and the reported attributes of the users. A *comma separated value (CSV)* file was exported from the Neo4j database containing a list of all the users, the date showing when users registered for the game, the reported age and gender of the user and the team the user chose upon registration. To analyse the data I used R and inspected the data both quantitatively through calculations and visually through plots and diagrams. To visualise and summarise the data I produced various plots including the total number of registered users over time and age and gender distribution of the registered users.

5.3.2 User Contributions

After analysing different attributes of the registered users, the contributions themselves were analysed.

As with the user distribution analysis, the data for the user contribution analysis was exported from the database in multiple CSVs. One exported CSV contained all the individual tile classifications including the corresponding attributes such as the timestamp, user ID and the spatial information of the classified tile. The second exported CSV-file consisted of all the individual tile classifications and the corresponding attributes of tiles which were captured more than once.

Because the reported land cover classes are stored as comma separated values in a single string, the values were separated into a machine-readable format. The values were divided into separate Boolean columns for each land cover class. In other words, a column was created for every land cover class and if the land cover class was present in the string of comma separated values, the value in the column with the corresponding land cover class as column name was set to 1, otherwise to 0.

E.G.: THE STRING "URBAN,FOREST,WATER" WAS CONVERTED TO:

URBAN	INDUSTRY	ARABLE	FOREST	PASTURE	AGRICULTURE	WATER	NOVEG	SHRUB	…
1	0	0	1	0	0	1	0	0	…

After converting the strings of comma separated values of the reported land covers to individual columns, the count of tiles containing each land cover class could efficiently be queried. The count of tiles containing a specific land cover class is the sum of all the rows where the value of the specific land cover class column is equal to 1.

The CSV containing all user classifications was used as underlying data to generate a histogram of the number of users as a function of the number of tiles classified. I achieved this by first grouping the classifications according to the user and then counting the classified tiles for each user in R. The resulting data was stored in a temporary table, where each row represents a user and the number of tiles collected by said user. The temporary table was grouped again according to a predefined range of the number of classified tiles. Then the number of users in each group was counted.

In order to analyse how many times specific tiles were classified (i.e. spatial clusters of increased classification behaviour) the CSV containing all classifications was grouped by tile ID in R. The resulting table contained each tile, its spatial attributes and the number of times one or several users classified the tile. The resulting table was imported into QGIS for visualisation. In QGIS, a map visualisation was created to visually inspect the results. To further quantify and underline the results, I created a bar plot showing the number of tiles containing each land cover class, using only the tiles which were captured more than once.

In addition to analysing the tile counts from different perspectives, I also analysed the tiles in regard to individual users and the areas in which individual users classified tiles. Again, the CSV containing all the tiles, the username of the user who classified the tile and the spatial information of the tiles were imported into QGIS. The tiles were then classified according to the username, which resulted in a map showing the tiles in colours based on the user who last classified a given tile. Even though this method does not take multiple captures of the same tile into account and only shows the last user to have captured a tile, it suffices to visualise examples for the key spatial characteristics of different users, such as potential user movement behaviours.

5.3.3 Intra-Tile Agreement and Fuzzy Pixels

The analysis of the user contributions focused on the inter-tile variations and the fundamental characteristics of the dataset. This section shows how the intra-tile variations of reported land cover classes and the fuzzy nature of the tiles regarding land cover classifications were analysed. Therefore, this section describes methods used to identify the intra-tile agreement rates regarding the different land cover classes and the approaches used to examine land cover co-occurrences in the data generated by the implemented location-based game.

The data of tiles that were captured multiple times was analysed by building pairs of [n, n+1] consecutive captures and analysing the rate of agreement within each of the pairs. As an underlying dataset, a CSV was exported from the Neo4j database, in which each row contains a pair of classifications and the respective users. Therefore, each row contains a pair of consecutive captures of a given tile. I argue that analysing consecutive captures rather than averaging all captures of a given tile can potentially shed light on changes over time in the land cover at a specific location. However, the timespan in which data was collected was too short to make relevant judgements on land cover changes over time and I do not further explore this aspect in my thesis.

To analyse the agreement rate between consecutive users who classified a tile, an algorithm (cf. chapter 5.3.2) split the comma separated land cover classifications into separate columns. All 26 new columns stored the corresponding land cover information in a Boolean format. The digit 0 indicates that the user did not classify the tile as containing said land cover class and the digit 1 indicates that the user did classify the tile as containing said land cover class. To analyse the average agreement rate between consecutive captures, I created an algorithm in R, in which the sum of agreeing consecutive captures of a given land cover class was divided by the sum of all tiles containing said land cover class. This effectively resulted in the formula:

$$\frac{Sum(LC(Ua) = 1\ AND\ LC(Ub) = 1)}{Sum(LC(Ua) = 1\ OR\ LC(Ub) = 1)}$$

"$Sum(LC(Ua) = 1\ AND\ LC(Ub) = 1)$" includes all the tiles in which both the preceding (Ua) **and** following (Ub) users agreed that a given land cover class was present in the tile and "$Sum(LC(Ua) = 1\ OR\ LC(Ub) = 1)$" includes all of the tiles in which either the preceding **or** following user reported said land cover class. Running the algorithm on all land cover classes resulted in a decimal value between 0 and 1 indicating the percentage of how many consecutive users agreed on a given land cover class.

The implemented game allowed users to not only report one land cover class per tile, but all land cover classes which a user believed to be contained in the tile. This resulted in many tiles which were reported as containing multiple land cover classes. To analyse the number of times a specific land cover class was reported in conjunction with another land cover class, a co-occurrence matrix of land cover classes was created and plotted in R. An initial empty 13 by 13 matrix was created, in which each column and each row correspond to one of the 13 land cover classes. An algorithm fills in the co-occurrence matrix with values: A double for-loop effectively iterates through every land cover class in comparison with every other land cover class and counts the number of rows containing both land cover classes. The generated co-occurrence matrix shows absolute values of co-occurrences and thus has a redundancy axis, which is the diagonal of the matrix where the land cover class of the row and column are equal. The redundant values were deleted to ease comprehension. In addition, the mentioned diagonal is omitted as a single land cover class cannot be reported twice in a single classification of a tile.

The first analyses resulted in a co-occurrence matrix containing absolute counts of the co-occurrence of each land cover class with every other land cover class. Because the data showed large differences in the number of times each land cover class was reported by a user, I calculated an additional relative co-occurrence matrix as to normalise the results. The relative co-occurrence matrix was generated using the same algorithm as with the absolute co-occurrence matrix, but with a small addition: The algorithm additionally divided every generated absolute value of the co-occurrence of a land cover class by the total number of tiles containing the land cover class represented by the row in the matrix. This resulted in a relative co-occurrence matrix, in which every cell represents the percentage of the land cover of a given row co-occurring with the land cover of the corresponding column.

Finally, to visualise the distribution of all tiles with specific reported classifications, I visualised the results in QGIS, using the broad extent of Zürich city and the surrounding areas as an example. This area was chosen to visualise the data because all land cover classes were reported to be contained by multiple tiles within this area.

5.3.4 Comparison with CORINE 2012

After focusing on the analysis of user contributions and agreement rates between user generated land cover classifications, this section describes the methods used to compare the generated data with the official CORINE dataset. The first step consisted in extracting the CORINE raster values at the locations of the individual game tiles and in adding the CORINE values as attributes to the game tiles. As an approximation, I reduced the polygons of the game tiles to their respective centroids. The generated centroids were then used as points of extraction and the respective CORINE raster values at the location of the points were stored as an additional attribute of the centroids. This approximation entails a loss of data, but using other methods of extraction would go beyond the scope of this master's thesis. To potentially identify the impact of using the presented approximate method to extract the raster cell data, all presented analyses were done using the CORINE datasets with a 100m x 100m and with a 250m x 250m cell size. The differences between the two resolutions regarding the land cover classifications at the location of the game tile centroids was analysed.

In addition, temporal aspects of agreement rates between the user contributed land cover classifications and the CORINE land cover dataset were analysed. This was achieved by plotting the total number of agreeing and disagreeing tiles over time as a line diagram depicting the data collection period. In addition, the weekly agreement rate was calculated as to shed light on average agreement rate change over time.

The CORINE classes first had to be summarised to the second level classes. This was necessary because CORINE classifies land cover types into 44 third level classes, which are contained in 15 second level classes, which are again contained in five first level or top level classes (excluding the three "No Data" classes). The CORINE third level classes were therefore summarised to CORINE second level classes in R. The CORINE third level classes were also summarised to CORINE first level classes for additional analyses to shed light on potential difficulties of users in differentiating between semantically similar land cover classes.

The CORINE land cover class names were abbreviated in all processing steps and in the resulting visualisations, results and discussions. Table 2 summarises the abbreviations applied to the CORINE second level classes and Table 3 summarises the abbreviations applied to the CORINE first level classes. The abbreviated names were used to describe the respective classes in the following chapters of this master's thesis. If an abbreviation was used, the land cover class name is written within quotation marks.

Table 2 - Used abbreviations and official CORINE land cover class names (second level)

Abbreviation	Official CORINE land cover second level class name
Urban	Urban fabric
Industry	Industrial, commercial and transport units
Mine	Mine, dump and construction sites
Greenarea	Artificial, non-agricultural vegetated areas
Arable	Arable land
Permacrop	Permanent crops
Pasture	Pastures
Agriculture	Heterogeneous agricultural areas
Forest	Forests
Shrub	Scrub and/or herbaceous vegetation associations
Noveg	Open spaces with little or no vegetation
Wetland	Inland wetlands
Water	Inland waters

Table 3 - Used abbreviations and official CORINE land cover class names (second level)

Abbreviation	Official CORINE land cover first level class name
Artificial	Artificial surfaces
Agriarea	Agricultural areas
Forestsemi	Forest and semi natural areas
Wetland	Wetlands
Water	Water bodies

The described pre-processed CSV then contained the CORINE land cover classifications (level one and two) as well as Boolean land cover columns from the user generated data for each tile (level one and two). To identify those tiles including the land cover class defined by the CORINE dataset in the user reported classifications, a new Boolean column was added to the dataset indicating if an agreement of the CORINE class with one of the user-contributed classifications could be observed, or if a disagreement of all user-contributed classifications with the CORINE classification was present. The resulting output was imported into QGIS for visualisation and used in R to generate confusion matrices.

The above described CSV contains a Boolean value containing the information on whether an agreement between CORINE and any of the user reported land cover classifications exists or does not exist. I visualised this data by generating agreement maps in QGIS. To address the problem of displaying the agreement rate of tiles with multiple captures, the imported data was dissolved and the individual land cover and agreement values were summed up. I visualised the data as a map in which two-dimensional styling was incorporated using the cylindrical coordinate based colour definition HSV (which stands for Hue, Saturation and Value). Various authors have studied colouring bivariate statistical maps and they concluded that using a hue-saturation-value-matrix can be an effective tool for displaying such two-dimensional data (Trumbo 1981; Wainer & Francolini 1980). The created hue-saturation-value-matrix was then uploaded to Coblis[22] to check for colour blind compatibility. The resulting map allows efficient visual analysis of multiple important characteristics of user agreements in and between the different game tiles. A resulting map was generated with the broad extent of Zürich as an example, to be discussed in further detail.

As the generated map serves the purpose of making the data available for visual inspection, the data was additionally processed in R to generate statistical results. The methodology of generating a confusion-matrix was chosen, which corresponds to the used methods in the literature about similar research (Comber et al. 2016). I created an empty 13-by-13 matrix to incorporate all land cover classes of the CORINE level two classification scheme. The rows of the matrix represent the user generated land cover classifications, with each row summing up to the total number of tiles in which the land cover class of said row is reported to be contained in. The columns represent the land cover classes from the CORINE dataset. The diagonal represents the number of tiles in which the user reported land cover class agrees with the CORINE land cover class. The remaining cells represent how many tiles of a specific land cover class reported by users were defined as belonging to another specific class in the CORINE dataset. These cells then contain the number of confusions between the user generated data and the CORINE land cover dataset. Due to the fuzzy nature (multiple classifications) of the user reported land cover classifications in comparison with the hard nature (single classification) of the CORINE dataset, the resulting matrix cannot be considered as a true confusion matrix. In the generated confusion matrix, the rows fulfil the typical characteristics of a confusion matrix (adding up to the total number of user generated tiles containing a given land cover class). However, the columns do not add

[22] http://www.color-blindness.com/coblis-color-blindness-simulator/ (accessed: 02.04.2017)

up to the total number of tiles. I therefore refer to the generated confusion matrix as a "pseudo-confusion-matrix", as to emphasise the similarity but not equality of the generated confusion-matrix to standard confusion matrices. I further generated relative "pseudo-confusion-matrices", where each cell contains the percentage of user reported land cover classifications which are classified as corresponding to the CORINE land cover classes. Naturally, the rows of the relative confusion-matrix add up to 100%, since a row incorporates all tiles which users reported as containing a given class (minor errors occur due to rounding).

Finally, to be able to calculate the confusion matrix metrics mentioned in the literature (Fawcett 2006; Beleites et al. 2013; Olson & Delen 2008), I generated a confusion matrix using the CORINE level two classes. I only included the tiles where users reported exactly one land cover class. This resulted in an initial dataset which allowed hard user reported land cover classifications to be compared with hard CORINE land cover classifications.

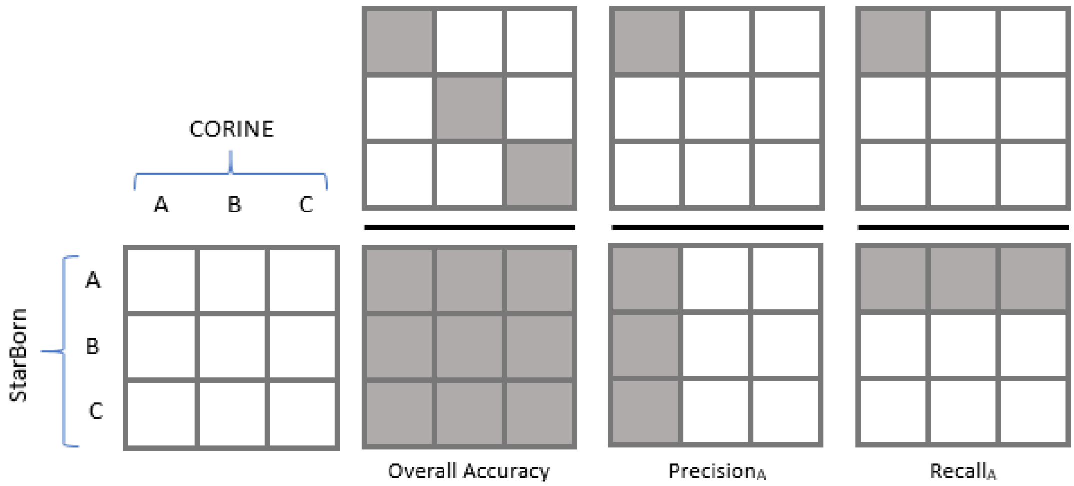

Figure 16 - Algorithm to compute overall accuracy, precision and recall
Figure visualising how the overall accuracy and the precision and recall for each class were computed.

Using the generated confusion matrix, I calculated the precision and recall of each land cover class. The calculation algorithms are visualised in Figure 16. In addition, I calculated the overall accuracy of the user contributed data and F-scores for every land cover class. The confusion matrix metrics were calculated with the user generated data as the reference data and the CORINE land cover dataset as the predicted values. Since I compared the two datasets and because none of them can be viewed as being the ground truth, the dimensions could also be swapped. Adapted from the literature (Beleites et al. 2013; Olson & Delen 2008; Fawcett 2006; Klotz et al. 2016) the mentioned metrics for this thesis are defined as follows:

- The overall accuracy describes the percentage of tiles where an agreement between the user reported classifications and the CORINE land cover dataset can be observed.
- The precision, also referred to as the positive predictive values, defines what percentage of the tiles classified by CORINE as being of a specific land cover class correspond to the user reported classification.
- The recall, also referred to as the sensitivity, describes the percentage of user contributions reported as containing a specific land cover class which corresponds to the CORINE land cover classification.
- The F-score represents the harmonic mean of the precision and the recall and can be used as a quality indicator. The F-score is calculated as $\frac{2}{1/Precision+1/Recall}$. An F-score of 1 consequently indicates complete agreement between the user generated content and the CORINE land cover dataset, whereas an F-score of 0 indicates no agreement.

6 Results

This chapter presents the results of this master's thesis. First temporal and attribute statistics of the generated data regarding the users of the implemented location-based game "StarBorn" are analysed and presented, followed by an analysis of the spatial attributes and classifications of user contributions. Next, the contributions are analysed focusing on the agreement and co-occurrences between individual captures of the same tiles. Last, the generated data is compared to the CORINE 2012 land cover dataset and agreements and disagreements are described.

6.1 User Distribution – Temporal Variations and User Attributes

First, the generated data was analysed regarding the users, taking both temporal variations and user attributes into account. This is an important first step of analysing the generated data as to make assumptions about the non-expert users, who generated the data. It also sheds light on the question of if the game promotion efforts unintentionally reached and appealed only to particular groups of persons and if these groups share common attributes.

The game was made available to the general public on 09.11.2016 and the data used in the analysis was collected on 23.02.2017. In the period in which the game was online, 138 users registered, of which 84 captured at least one tile. The rate of registering users can be divided into four distinct periods: pre-live and testing, beta testing, live and promotion, live without promotion. Figure 17 highlights the total number of registered users over time and thus the rate of registering users.

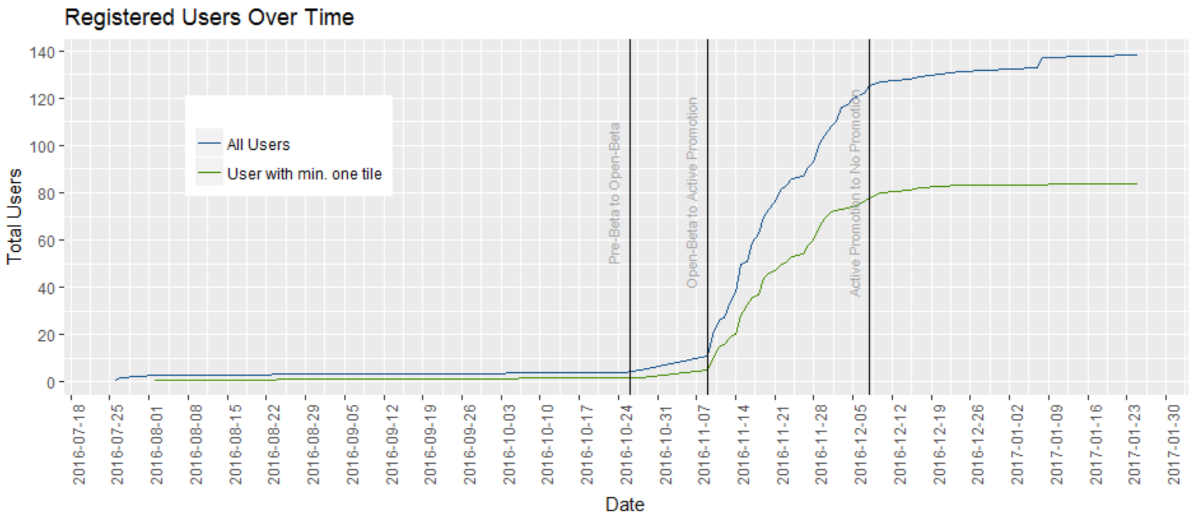

Figure 17 - Registered users over time
Figure showing the total number of users over the extent of the game period. Three lines indicate key transitional moments.

The registered users from 25.07.2016 till 26.10.2016 are test-users, which I created to actively test the implementation using different accounts. A slight increase can be observed in registered users between the 26.10.2016 and the 09.11.2016, the two weeks before opening the game to the public. This was the "open-beta" period, in which the game was shared amongst a chosen audience to test server stability and ease of use. In the period between 09.11.2016 and 08.12.2016 a massive increase in users can be observed. In this period, the game was actively shared and promoted on various platforms, including Facebook[23] and Google+[24]. In addition, a Google+-Community[25] was created to share news and enable users to easily give feedback and report bugs or problems they may encounter. In the mentioned period, the game was also promoted at the University of Zürich in various lectures. After the 08.12.2016, the game was not actively promoted and a rapid decline in new registering users can be observed.

The registered users are of a wide range of age groups and both genders are present. The age and gender declared by the users when registering were however not confirmed. The data, visualised using a bar plot (Figure 18), shows a spike in users with birth years between 1985 and 1995. Another spike can also be observed at the value of 2000, which coincides with the default year of birth in the registration form. The peak at 2000 suggests a high number of users not changing the year of birth from the default. It must be mentioned that this diagram excludes seven users who registered before the 10.11.2016 at 17:00. The reason behind this exclusion is that the default year of birth of the registration form was set to 1895. The first new users to register noted that such a low default year of birth resulted in a high amount of scrolling and consequently reporting the true year of birth was perceived cumbersome. I thus changed the default year of birth to 2000.

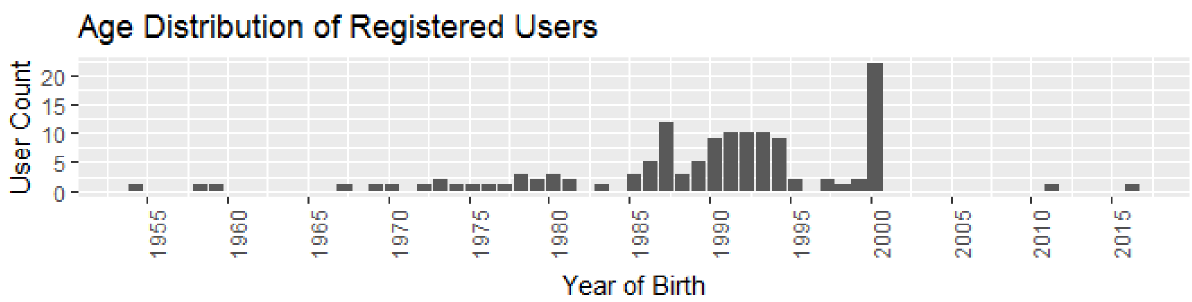

Figure 18 - Age distribution
Figure showing the age distribution of registered users by showing the year of birth in comparison with the user count.

The cluster of registered users who reported being born between 1985 and 1995 can be interpreted as showing the average age range of persons who are interested in the technology of location-based gaming and are able to comprehend the mechanics of such a game. The cluster could also be interpreted as a proxy of the average age range of the social environments in which the implemented location-based game was actively promoted, both on online platforms and in offline interactions.

[23] www.facebook.com
[24] https://plus.google.com
[25] https://plus.google.com/communities/102063937147156703895

The gender distribution is heavily sided towards the male gender. The reliability of the reported gender was not confirmed and must be treated as potentially inaccurate. The users were given multiple options: male, female, "not reported" and "NULL". Figure 19 shows that registered users who declared their gender are comprised of: 94 male users, of which 62 captured at least one tile, 27 female users, of which 16 captured at least one tile, 15 users who declared "not reported" as their gender, of which five captured at least one tile, and two test users having no reported gender.

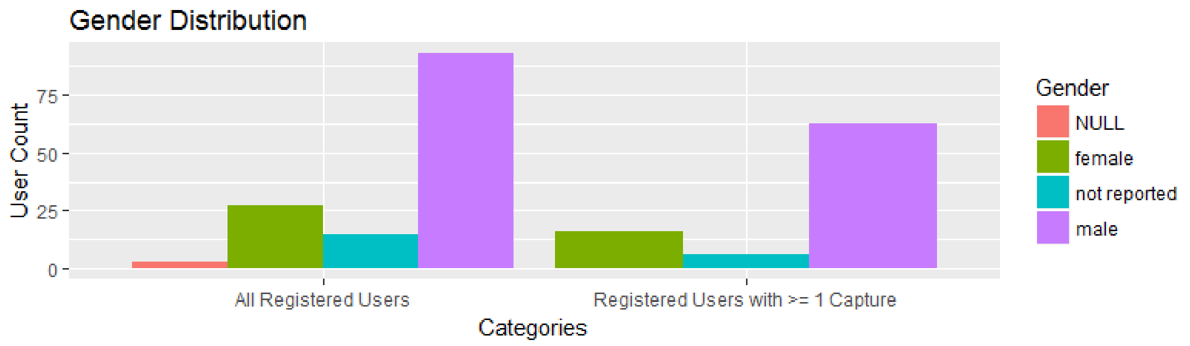

Figure 19 - Gender distribution
Figure showing the gender distribution of all registered users (left) and all registered users who captured at least one tile (right)

Figure 19 is divided into two categories: all registered users and all registered users who captured at least one tile and thus contributed at least one classification. Comparing the two categories shows that 65.96% of the male and 59.26% of the female users who registered captured at least one tile. Male users were not only observed to register more frequently, but also, if registered, were also slightly more likely to participate in the game. The higher number of male users could be interpreted as a higher probability of males of playing video games. The gender distribution of the audience of my promotion and advertising efforts could also be a potential reason for the overrepresentation of males.

6.2 User Contributions

After having analysed user registration patterns and user age and gender distributions, the user generated content was inspected in light of the underlying spatial attributes and the reported classifications. These results give insight into the viability of implementing a location-based game to generate geographic information, in particular land cover classifications. The results in this section also shed light onto the variations of the user generated data in terms of single and multiple captures and thus classifications of the in-game tiles and the variations in classified land cover class frequencies.

From the game going live (09.11.2016) until the data collection date (23.02.2017), the users classified a total of 13'338 tiles in 11'380 unique locations. Each tile has an extent of 200m x 200m, equalling a total of 533'520'000m^2 of classified area, of which 455'200'000m^2 represent the extent without multiple classifications. The 13'338 classified tiles are comprised of 10'444 tiles which were captured once and 936 tiles which were captured at least twice. The clear majority of the tiles are classified as containing "urban" (n=8010) or "industry" (n=7724) land cover classes. These are followed by "pasture" (n=2716), "forest" (n=2418), "arable" (n=2221), "water" (n=1166) and "greenarea" (n=1065), which are followed by the land cover classes with the least amount of classifications: "agriculture" (n=494), "noveg" (n=432), "permacrop" (n=360), "wetland" (n=246) and "mine" (n=150). The land cover classes and the number of times each land cover class was reported was visualised using a bar plot (Figure 20). This results in a total of 28'633 reported land cover classes contained in the 13'338 tiles, averaging 2.15 reported land cover classes in each tile.

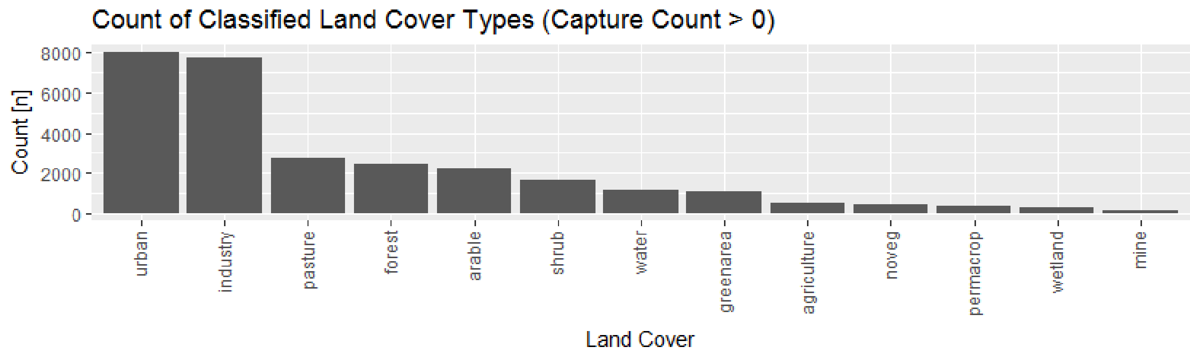

Figure 20 - Count of classified land cover classes (capture[n]>0)
Figure showing the count of reported land cover classes for tiles which were captured at least once

The overrepresentation of the land cover classes "urban" and "industry" indicates that players spend a large portion of their time playing the game in perceived urban or industrial landscapes.

Figure 21 reveals the typical long tail characteristics of user generated content, where the majority of the users captured a few tiles and a few users captured many tiles. The generated diagram shows that the majority of users (n=35) captured 1 - 20 tiles. A major decrease is visible with only nine users capturing 20 - 40 tiles and four users capturing 40 - 60 tiles.

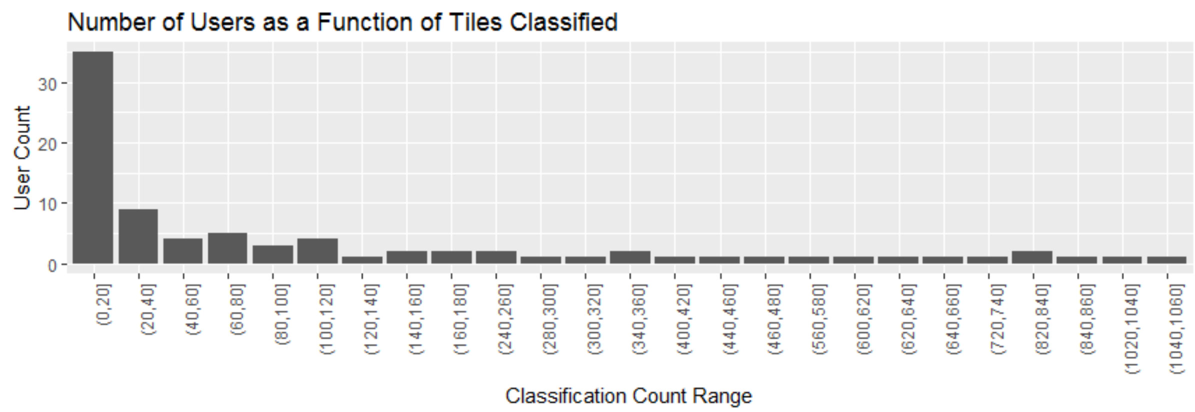

Figure 21 - Number of users as a function of tiles classified
Figure showing the number of users who classified tiles in predefined ranges of tile capture counts. A tile capture count range of 20 tiles was chosen. Ranges without any users contributing the corresponding number of tiles were omitted.

The mentioned large difference in the number of tiles which were captured once and the number of tiles which were captured at least twice hints at small hotspots of multiple captures. This is highlighted with a map of the area of Zürich, where the city centre shows clusters of multiple captures (Figure 22). Some clusters themselves have linear spatial characteristics and are assumed to correlate with most frequently used public transportation lines or other frequented routes of travel. With increasing distance to the city centre, the number of times a tile was captured decreases.

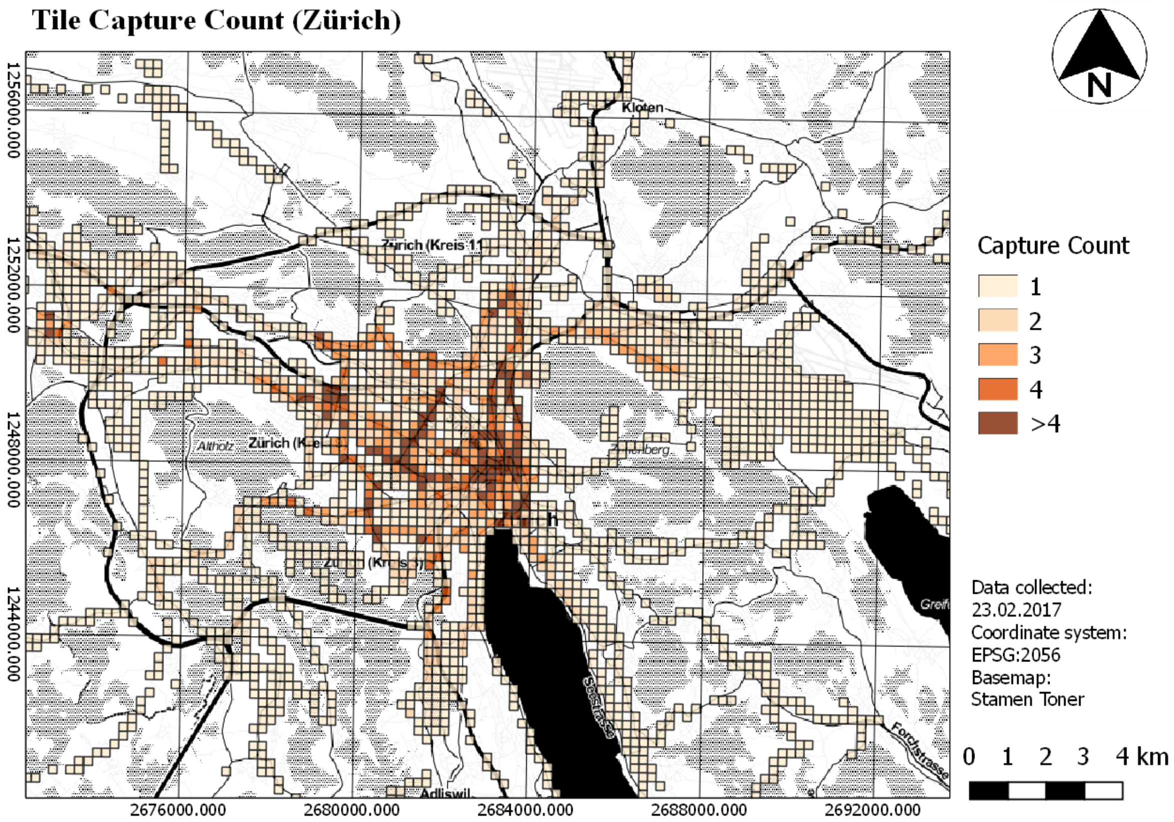

Figure 22 - Tile capture count (Zürich)
Map of the broad area of Zürich showing the number of times each tile was captured.

To quantify these results, a bar plot (Figure 23) was generated showing the number of tiles containing each land cover class, only using the tiles which were captured at least twice. The generated bar plot underlines the assumption that the game was primarily played in an urban environment. The bar plot also shows that the majority of the tiles which were captured more than once were also classified as being "urban" or "industry".

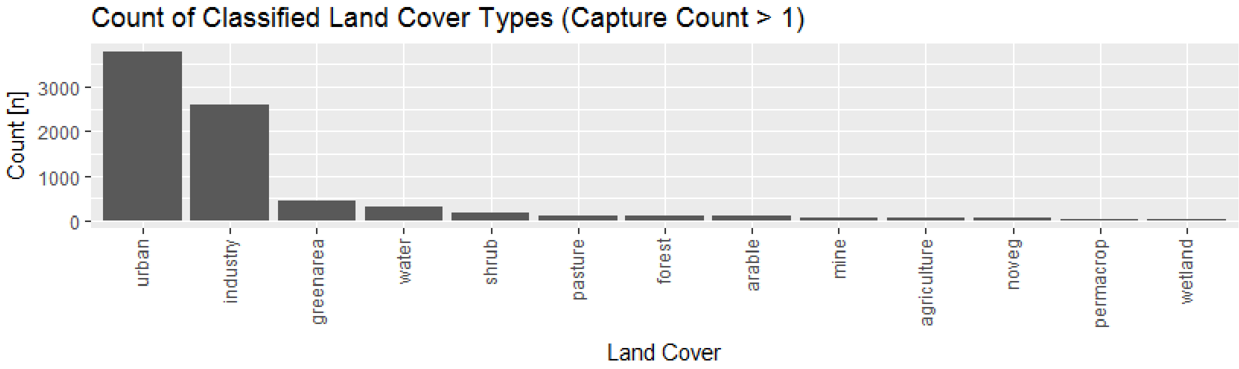

Figure 23 - Count of classified land cover classes (capture[n]>1)
Figure showing the count of reported land cover classes for tiles which were captured at least twice

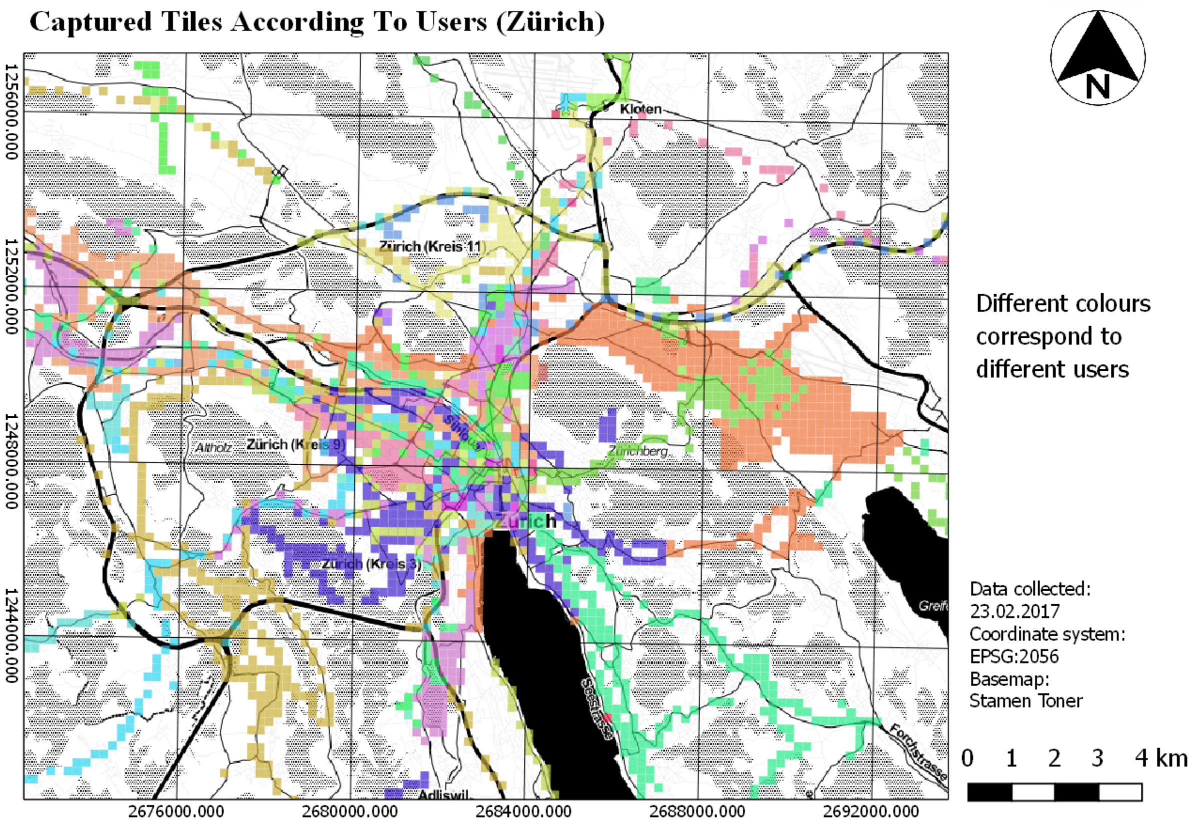

Figure 24 - Captured tiles according to users (Zürich)
Map of the broad area of Zürich showing captured tiles. Colours correspond to the username of the user who captured the tile last.

The data clearly shows that data was primarily collected in urban environments, but it can also be observed that the data shows a potential correlation with road or public transportation networks. It can be observed that a substantial number of users collected data whilst traversing the road network and thus the data often shows linear or line like collection patterns. A few exceptions can also be observed where users made an effort to collect tiles in a larger coherent surface. This duality of user collection behaviour can be seen as a proxy of the way users play the game (cf. Figure 24).

6.3 Agreements and Fuzzy Pixels

In this section, results are presented regarding the analysis of the agreements between multiple captures and the co-occurrences of pairs of land cover classes.

The data of tiles with multiple captures and thus multiple classifications from different users over time was analysed and the results were visualised in a bar plot (Figure 25). The analysis involved building pairs of consecutive captures and analysing the rate of agreement within each of the pairs (see chapter 5.3.3, *Intra-Tile Agreement and Fuzzy Pixels*). Analysing the data of multiple captured tiles shows a high rate of agreement for the land cover class "urban". In the consecutive classification pairs containing "urban" as the classified land cover class in either the first or second classification, 80.6% of the classification pairs both agree that "urban" is present. 62.6% of the classification pairs agree that "industry" is present. Interestingly, this considerable difference in the rate of agreement between "urban" and "industry" can be observed even though both land cover classes show a similar number of classifications. 36.8% agree that "water" is present and 22% agree that "forest" is present in the classified tile. The rest of the land cover classes analysed with classification pairs show an agreement rate of under 20%.

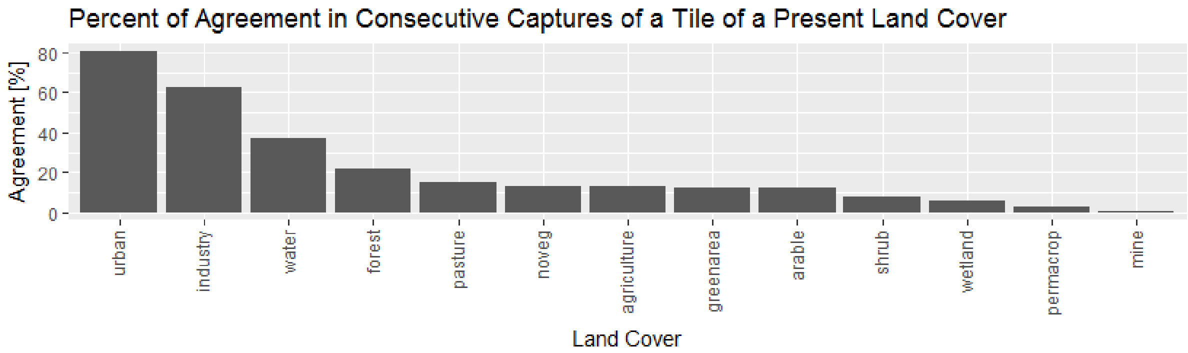

Figure 25 - Consecutive agreement rates
Figure showing the agreement rates of consecutive captures of each land cover class

As to generate a higher concentration of information, the users could classify a tile by selecting multiple land cover classes. The users were urged to report every land cover class they perceived to be present. These multiple classifications per capture resulted in a dataset of fuzzy pixels, where most tiles contained a combination of the mentioned land cover classes. This information is plotted in a land cover class co-occurrence matrix (Figure 26).

The matrix shows the absolute count of tiles which share the corresponding two land cover classes. This is a useful tool to display the sheer number of different combinations and the number of times these can be observed in the data. Most notable is that "industry" and "urban" are mentioned 4'660 times in the same tile, "pasture" and "industry" are found together in 1'440 tiles and "pasture" was counted 934 times in combination with "arable".

These high numbers of co-occurring land cover classes can either indicate high spatial autocorrelation of the land cover classes or difficulties for non-experts to distinguish one from the other and thus reporting both. This becomes especially evident when looking at the relative values of the relative co-occurrence matrix (Figure 27), indicating to what proportion tiles with one land cover class contain another.

The relative co-occurrence matrix of land cover classes gives valuable insights into the characteristics of the underlying data. The presented relative co-occurrence matrix indicates three main characteristics of the data generated in the implemented location-based game: The overrepresentation of specific land cover classes in terms of absolute numbers of classifications, the potential difficulties of semantically differentiating between similar land cover classes and the spatial autocorrelation of specific pairs of land cover classes.

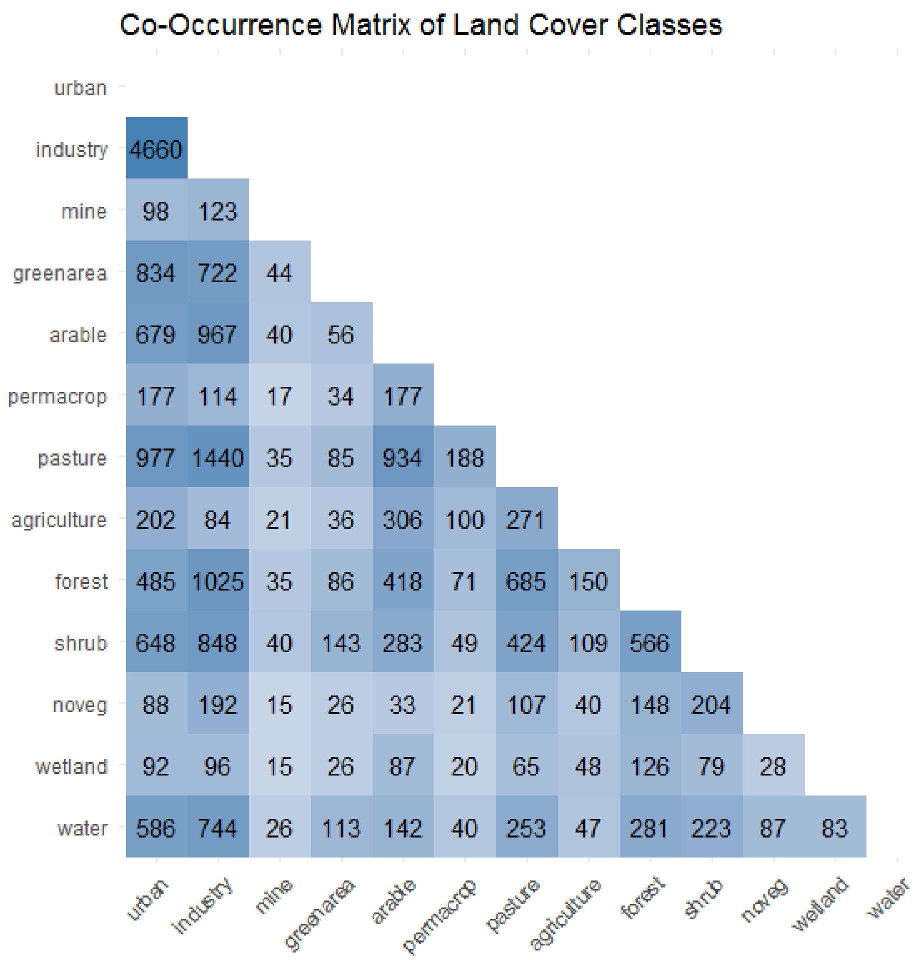

Figure 26 - Land cover co-occurrence matrix
Figure showing the number of times a given land cover class was reported in combination with each of the other land cover classes.

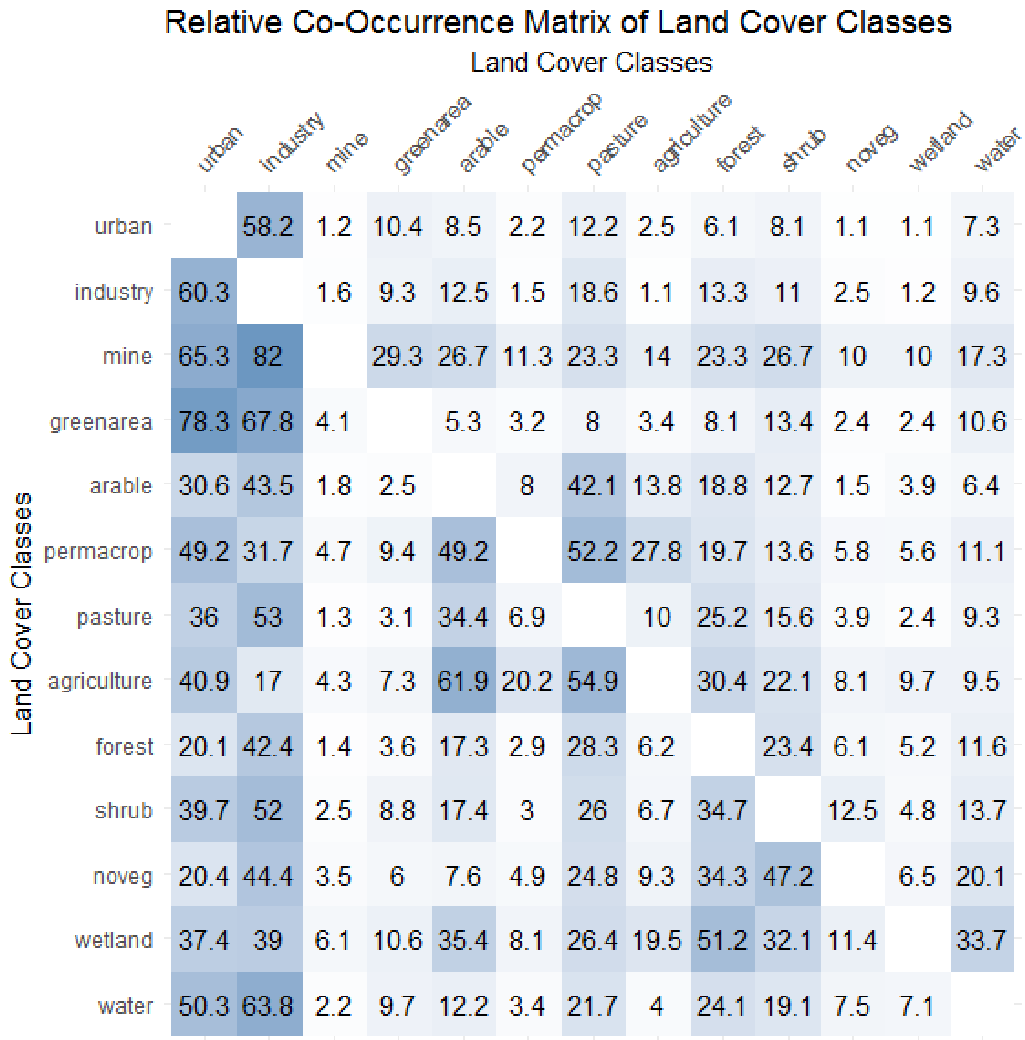

Figure 27 - Relative land cover co-occurrence matrix
Figure showing a land cover co-occurrence matrix with relative values. The relative values indicate what fraction of tiles reported as containing one land cover class were also reported containing another land cover class.

The first two columns of the relative co-occurrence matrix show an above average number of high values. I argue that the first two columns containing "urban" and "industry" indicate an over-representation of the mentioned land cover classes. This assumption is further cemented by the presented diagram showing the number of instances the land cover types were classified (Figure 20). The chance of another land cover class co-occurring with "urban" or "industry" is thus naturally higher. The relative co-occurrence matrix further shows 58.2% of the tiles which were classified as containing "urban" land cover elements were also classified as containing "industry", and 60.3% of the tiles classified as containing "industry" were also classified as containing "urban". This relatively small difference indicates possible perceived spatial autocorrelations between the classes. Of particular interest are the occasional large differences between the relative percentages of a pair of land cover classes depending on which perspective the data is inspected from. Noteworthy is that 78.3% of the tiles classified as containing "greenarea" were also classified as containing "urban", but only 10.4% of the tiles classified as containing "urban" were also reported as containing "greenarea". This can be an indicator that the users predominantly perceived green-areas as being part of the urban landscape and that the land cover class “urban” was reported significantly more frequently than “greenarea”, again highlighting that users played the game predominantly in urban environments. Another noteworthy particularity is that 51.2% of the tiles which were classified as containing “wetland” were also classified as containing “forest”, but only 5.2% of the tiles classified as containing “forest” were also classified as containing “wetland”. This indicates a possible natural spatial autocorrelation of the two classes and could be an indication that most of the wetlands in the landscape of Switzerland are found in forest-like areas. Interesting is also the observed high co-occurrence of “water” with “urban” and “industry”. 50.3% of tiles classified as containing “water” were also classified as containing “urban” and 63.8% of the tiles classified as “water” were also classified as containing the land cover type “industry”. This could indicate potential urban waterbodies. I assume that most of the reported tiles containing waterbodies are tiles with urban waterbodies such as urban rivers or lakeshores in an urban environment, again in line with the observation that the game was played most intensively in urban areas. This is underlined by comparing the two land cover class maps of Zürich, in which Figure 28 shows the tiles in Zürich which were classified as containing “urban” and Figure 29 shows the tiles in Zürich which were classified as containing “water”. The mentioned maps confirm that the tiles which users reported as containing the land cover class “water” trace urban rivers and urban lakeshore lines, whilst mostly co-occurring with the land cover class “urban”.

Finally, the matrix also indicates possible difficulties in semantically distinguishing one class from the other and potential spatial autocorrelations. An example of this characteristic is the high co-occurrence rates of “agriculture” (61.9%), “pasture” (34.4%) and “permacrop” (49.2%) with “arable”. I argue that these co-occurrences, on the one hand, indicate that the users could not identify the more likely class and chose to report both classes instead of choosing one. On the other hand, the mentioned characteristics of the high co-occurrence rates could also originate from highly detailed classifications of users who took their time and investigated the whole tile before reporting a contribution. The latter would then indicate high spatial autocorrelations between classes with relatively high co-occurrence rates.

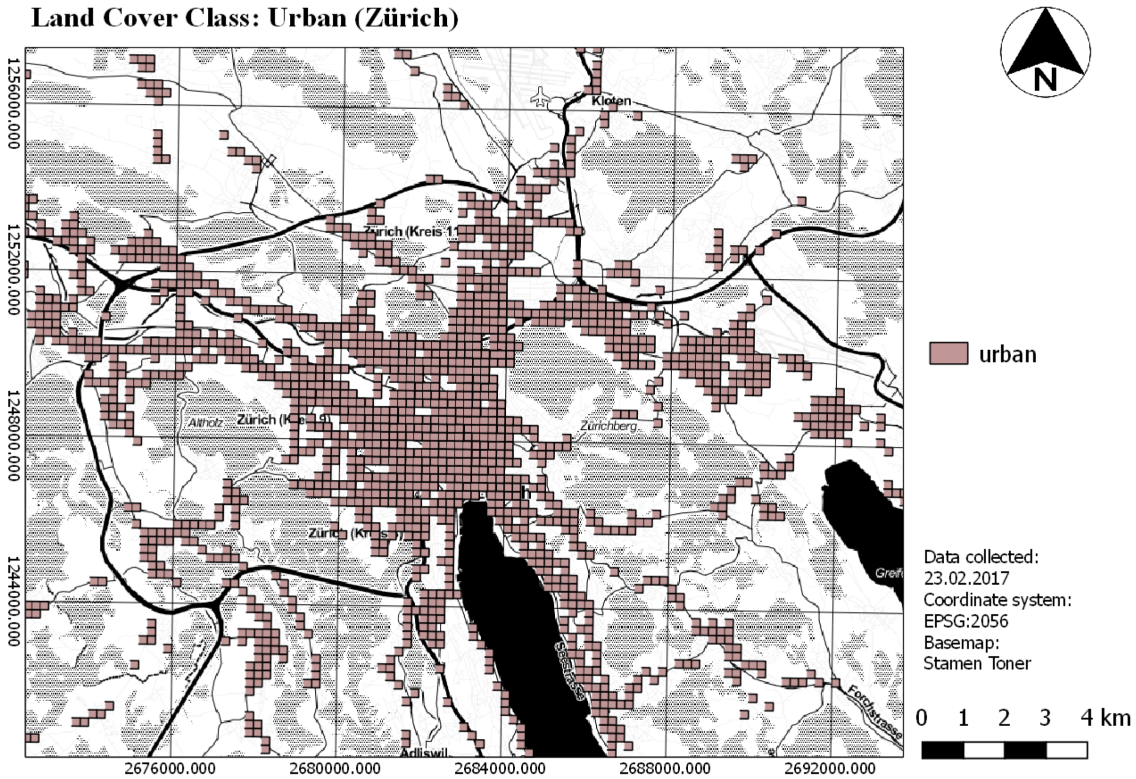

Figure 28 - Land cover class urban (Zürich)
Map of the broad area of Zürich showing captured tiles where at least one user reported the land cover class "urban".

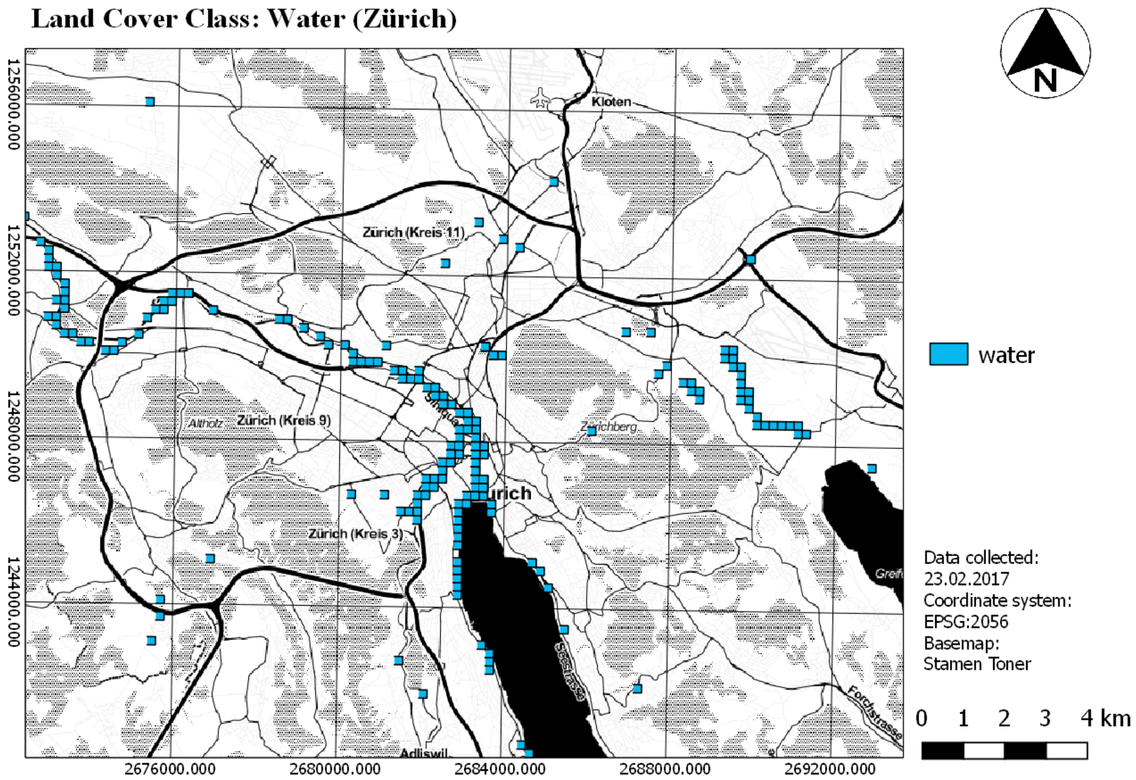

Figure 29 - Land cover class water (Zürich)
Map of the broad area of Zürich showing captured tiles where at least one user reported the land cover class "water".

6.4 Comparison with CORINE 2012

This final section revolves around analysing the user generated content in comparison with the official CORINE 2012 dataset of Switzerland.

First, the approximation of the raster value extraction using the tile centroids and not the tile polygons was analysed. As described in the methods (chapter 5.3.4, *Comparison with CORINE 2012*), the difference of the agreement values when using the CORINE dataset with a resolution of 100m x 100m in contrast to when using the CORINE dataset with a resolution of 250m x 250m was calculated. It can be observed that out of the 11'380 individual tiles (some with multiple captures), 10'537 have at least one user reported classification which agrees with the CORINE land cover dataset in both the 100m^2 and the 250m^2 raster resolutions. The agreement rate between the two tested resolutions thus corresponds to 92.6%. However, 843 tile locations yield different results based on the chosen resolution. Having approximately 7.4% disagreement between the two tested resolutions called for further inspection. A map showing the locations of the disagreeing tiles was generated and visually inspected in QGIS. No visual patterns could be identified. An example map of the broad area of Zürich was generated to show this distribution of varying results based on the chosen CORINE land cover dataset resolutions (Figure 30).

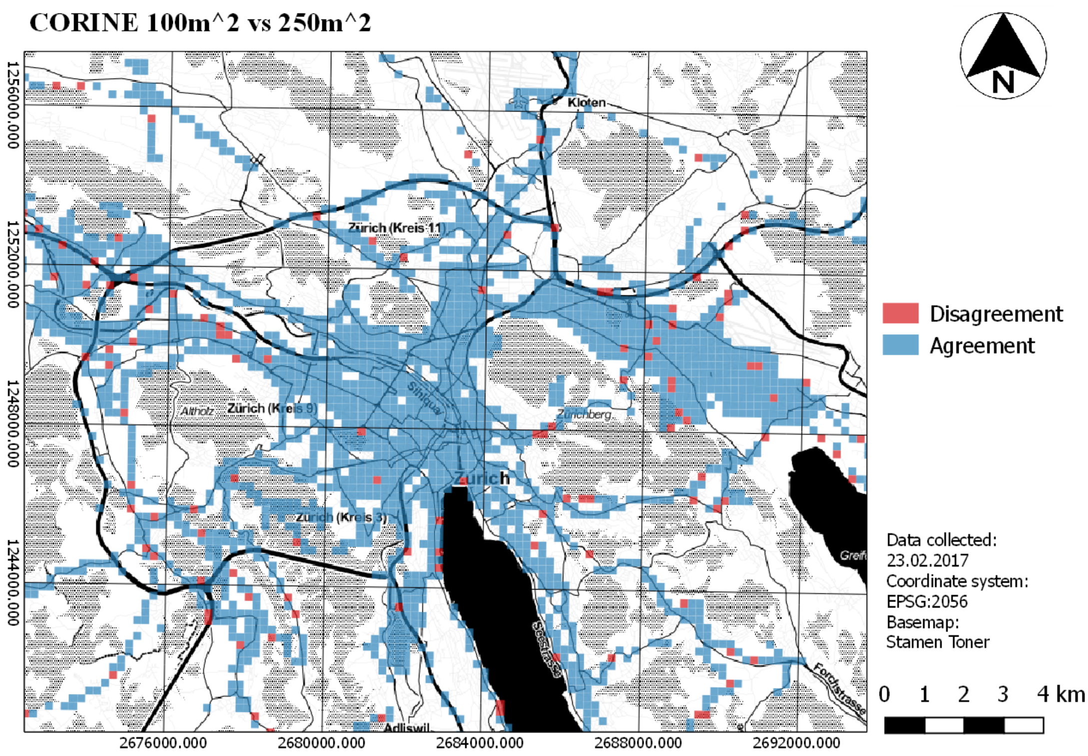

Figure 30 - Comparison between agreements depending on CORINE resolution
Map of the broad area of Zürich showing agreements and disagreements between the CORINE datasets with a resolution of 100m^2 and 250m^2.

Naturally, disagreements can only occur in locations of land cover change and are impossible to observe if the 100m^2 raster cells covering the area of a given 250m^2 raster cell all report an identical land cover class. Accordingly, disagreements always show a change in land cover types and could be analysed further focusing on why the observed disagreements occur. This analysis shows that using different resolutions has an impact on the results and I thus deduce that using the approximation of extracting the raster cell values at the centroids of the game tiles and not using the game tile polygons influences the results and must therefore be kept in mind whilst interpreting the results.

Users contributed the classification of a total of 13'338 tiles (also counting multiple captures) with the possibility to report multiple land cover classes per tile. Of these 13'338 classifications, 10'157 classifications showed an agreement between the CORINE land cover dataset in at least one user reported class. Therefore, in 76.15% of the tile-classifications one of the user reported land cover classes corresponds to the CORINE land cover class. The same calculation using the 250m^2 CORINE land cover raster resulted in 9'994 classified tiles for which an agreement was found between one of the user reported land cover classes and the CORINE dataset, equalling 74.93% of the tiles. These results show a slight difference in the overall agreement rates depending on which resolution of the CORINE land cover dataset is taken for the analysis.

To analyse potential variations in agreement rates over time, the total number of agreeing and disagreeing classifications over time was plotted (cf. Figure 31). A significantly higher increase in captured tiles with agreeing classifications can be observed than where the reported land cover classes do not match the land cover class of the CORINE dataset.

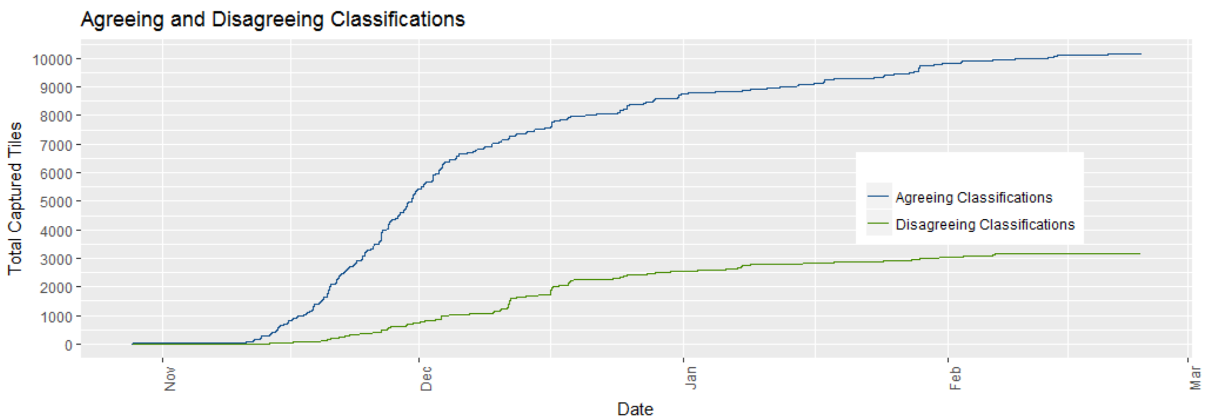

Figure 31 - Total agreeing and disagreeing classifications over time
Figure showing the total number of agreeing and disagreeing classifications during the data collection period.

In addition, the percentage of disagreeing classifications was aggregated to a weekly average and plotted as a line diagram. Figure 32 shows high fluctuations in the average weekly disagreement rates. The reported land cover classes of the users who started playing the implemented location-based game between 09.11.2017 and 13.11.2017 had an average disagreement rate of 6.34% - 12.52%. Users reporting land cover classifications in the early stages of the data collection period thus contributed data displaying higher relative agreement rates compared to the rest of the period. After 27.11.2016, high fluctuations can be observed for the remainder of the data collection period ranging from 13.21% disagreement rate in the week of 19.02.2017 to 38.87% disagreement rate in the week of 08.01.2017. The week before data was collected shows a significant decrease of the average weekly disagreement rate to 13.21%. However, the last week of data collection only consists of a relatively small number of new tile captures (n=46) compared to other weeks. No correlation between the number of classifications and the rate of agreement can be observed when summarising the data according to the week.

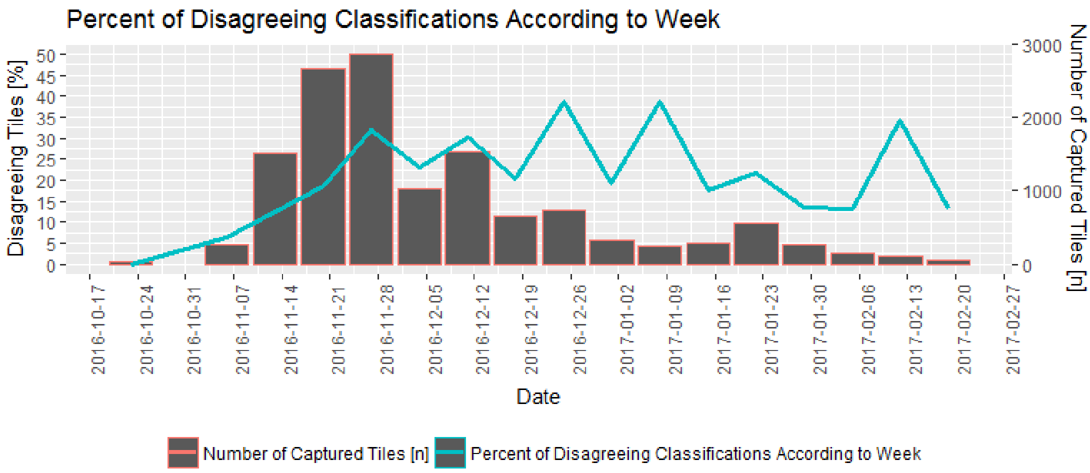

Figure 32 - Percentage of weekly disagreeing classifications
Figure showing the percentage of disagreeing classifications (comparing the generated data with the CORINE land cover dataset) according to the week in combination with the number of captured tiles according to the week

Visually inspecting the agreement counts and tile capture counts using bivariate colouring techniques allows useful insights into the potential accuracy of the user generated data. In addition, patches of spatially autocorrelated disagreements can be identified and further inspected in QGIS. The resulting map output of the comparison analysis between the user generated data in the implemented location-based game and the CORINE 2012 dataset shows that substantial agreement can be observed in the area of cities and less high rates of agreement can be observed in peripheral regions. Not only does the map show high agreement rates in metropolitan areas, but also a high number of captures. Achieving a high number of captures with a high agreement rate is the most desirable outcome. Again, the extent of Zürich was chosen to visualise these results at a scale at which the individual tiles can easily be distinguished. The presented results apply for the whole of the game extent. It becomes clear that in metropolitan areas such as cities, the intra-tile agreement rate as well as the agreement between the user generated content and the CORINE dataset are relatively high. The peripheral, more remote areas on the other hand are subject to a higher disagreement rate.

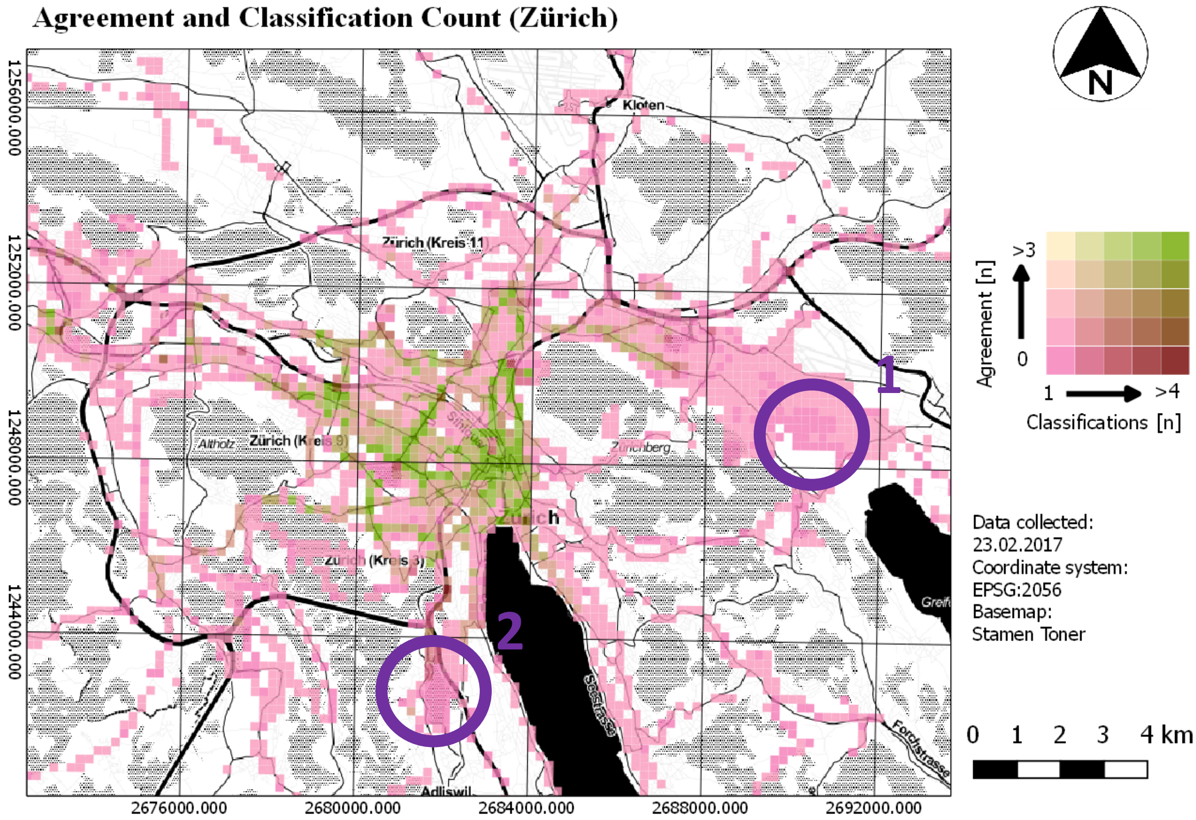

Figure 33 - Agreement and classification count (Zürich)
Map of the broad area of Zürich showing the capture counts and the agreement counts using bivariate colouring. Two areas of spatially autocorrelated areas are highlighted with circles and the numbers one and two.

Of particular interest are the patches of spatially autocorrelated disagreement values, of which two are examined in more detail. These are highlighted and numbered in Figure 33 and displayed in Figure 34.

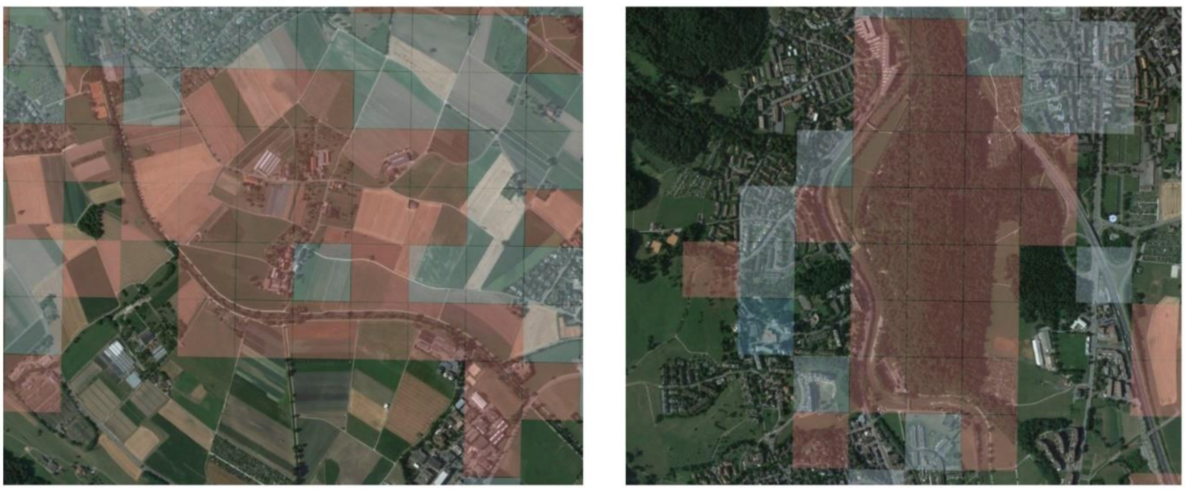

Figure 34 - Detailed examples of disagreements
Satellite images with an overlay (transparency = 30%) indicating agreement (blue) or disagreement (red) between the user generated data and the CORINE land cover dataset. The left image corresponds to the highlighted circle number one in Figure 33; The right image corresponds to the highlighted circle number two in Figure 33.

The first area examined in greater detail is an area dominated by rural characteristics comprised of a mosaic of different agricultural areas and rural infrastructure. The area also includes residential and industrial areas. This becomes clear when viewing a satellite image of the area in question whilst indicating the agreements and disagreements by overlaying semi-transparent coloured polygons.

One of the most prominent differences between the CORINE dataset and the user generated classifications becomes evident when inspecting the data in a tile by tile approach. Most of the area is classified as being "arable" in the CORINE dataset, whereas the user who classified the tiles in question reported them as containing pasture. The tiles containing the small river indicate another issue of high importance regarding potential sources of disagreements. All tiles containing the small river were reported as containing "forest", "shrub", "wetland" and "water". These land cover classifications may be accurate for the area of the river and the immediate surroundings. The area of the small river and its immediate surroundings is however only a relatively small percentage of the area of the whole tile.

The second example additionally highlights potential issues regarding the classification strategy of CORINE, the difficulty for non-local expert photo-interpreters to take into account the local semantics of an area and the potential problems of having a rule and priority based classification pipeline. The second example image shows a large area of spatially autocorrelated disagreement in an area of a visually homogeneous land cover class. The disagreement results from the CORINE classes reporting the tiles as containing the land cover class "greenarea", whereas the user reported the area as containing "forest".

To quantify and analyse further possible reoccurring issues of land cover classification disagreements, I calculated absolute and relative "pseudo-confusion-matrices". Both the results of the comparison of the data from the implemented location-based game with the CORINE land cover datasets in $100m^2$ and $250m^2$ resolution, as well as the results of the comparison of the data using the summarised CORINE level one classes were computed as "pseudo-confusion-matrices". Seeing that the differences of the results between the two mentioned resolutions of the CORINE land cover datasets are small, only the results using the $100m^2$ resolution are discussed. The same conclusions can however be drawn when using the $250m^2$ resolution raster in the extraction pipeline.

The first "pseudo-confusion-matrix" (Figure 35) shows all land cover classes from the implemented game (rows) in comparison with all the land cover classes of the CORINE land cover dataset (columns). The land cover classes are summarised to the level two classes of the CORINE classification scheme. The numbers show how many percent of the user reported classifications for each of the land cover classes were classified as which land cover class in the CORINE dataset.

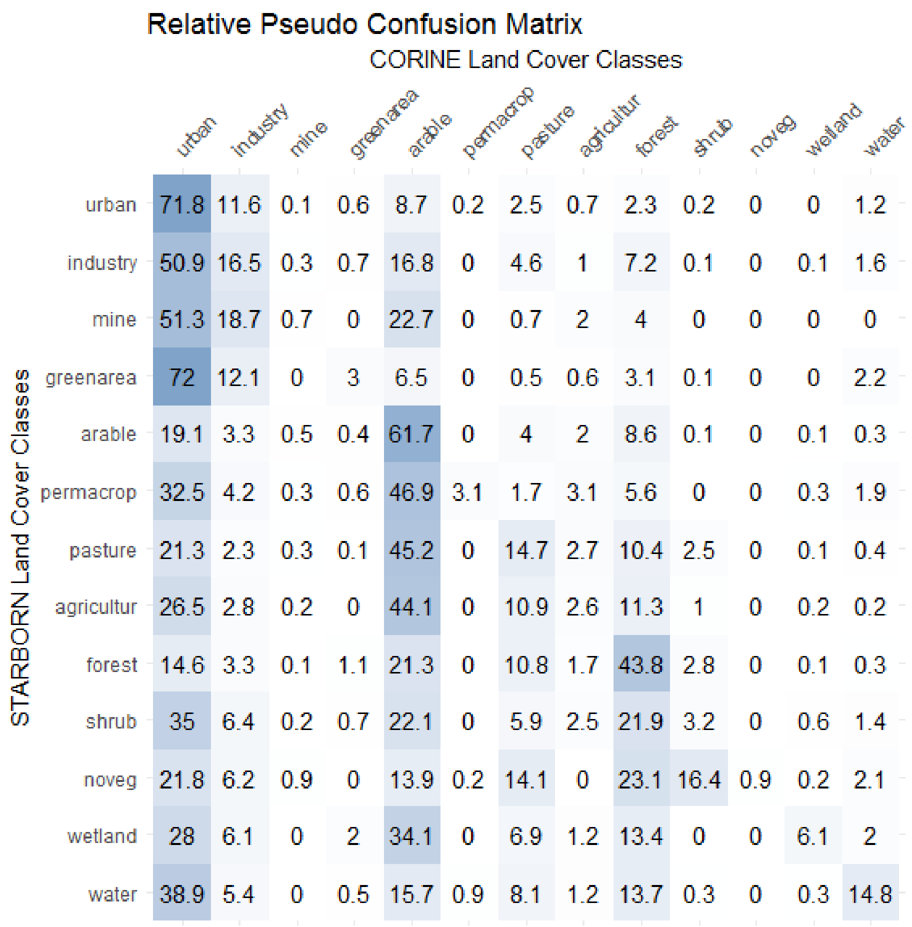

Figure 35 - Relative "pseudo-confusion-matrix" (CORINE level two)
Figure showing the relative "pseudo-confusion-matrix" of the data generated with the location-based game in comparison with the CORINE land cover dataset (level two).

The "pseudo-confusion-matrix" confirms that the land cover classes "urban", "arable" and "forest" show the highest relative agreements between the reported land cover classes and the corresponding land cover classes in the CORINE dataset with 71.8%, 61.7% and 43.8% respectively. The results also show that the highest confusions are found in the reported land cover classes "greenarea", "mine" and "industry", which were classified to be "urban" in the CORINE land cover dataset with 72%, 51.3% and 50.9% respectively. Overall, the results show a high rate of disagreement in the first column, where the CORINE land cover dataset classifies a large number of tiles as "urban" whereas users reported otherwise. These insights in combination with the data from the previous analysis (*Chapter 6.3 Agreements and Fuzzy Pixels*) suggest that these percentage values must be interpreted with caution. Similar to the co-occurrence matrices, this "pseudo-confusion-matrix" incorporates differences in absolute numbers of classifications of each land cover class, spatial autocorrelations of land cover classes and multiple classifications due to difficulties in semantically differentiating specific land cover classes.

The matrix shows that in 72% of the classifications in which a user reported "greenarea", CORINE defines the land cover as "urban". An agreement between CORINE and the user generated data concerning the land cover class "greenarea" was only noticed in 3% of the tiles users reported to contain "greenareas". A possible hypothesis for this large discrepancy between the user generated data and the official CORINE dataset is that the user generated data allows for multiple land cover classifications for each tile whereas the CORINE dataset only encompasses single values. Thus, if a small urban green area is present in a tile, CORINE will classify the tile as being "urban" and the users of the implemented game might classify the tile as containing "urban" as well as "greenarea". This example shows that due to the possibility of having multiple user classifications for one tile, an individual tile can be counted as agreeing with CORINE (e.g. reported "urban" = agreement with CORINE "urban") and simultaneously disagreeing with CORINE (e.g. a user classifying the same tile as also containing "greenarea"; Reported "greenarea" = disagreement with CORINE "urban"). The "pseudo-confusion-matrix" confirms a similar phenomenon in regards to agricultural areas. Tiles in which users reported the land cover classes "agriculture", "pasture" and "permacrop" are classified in the CORINE dataset as "arable" (44.1%, 45.2% and 46.9% respectively). Seeing that the same mentioned land cover classes show a high co-occurrence rate, this again underlines the assumption that users had difficulties in differentiating between the classes and chose to report many land cover classes instead of one, or indicates a potential spatial autocorrelation of the mentioned land cover classes.

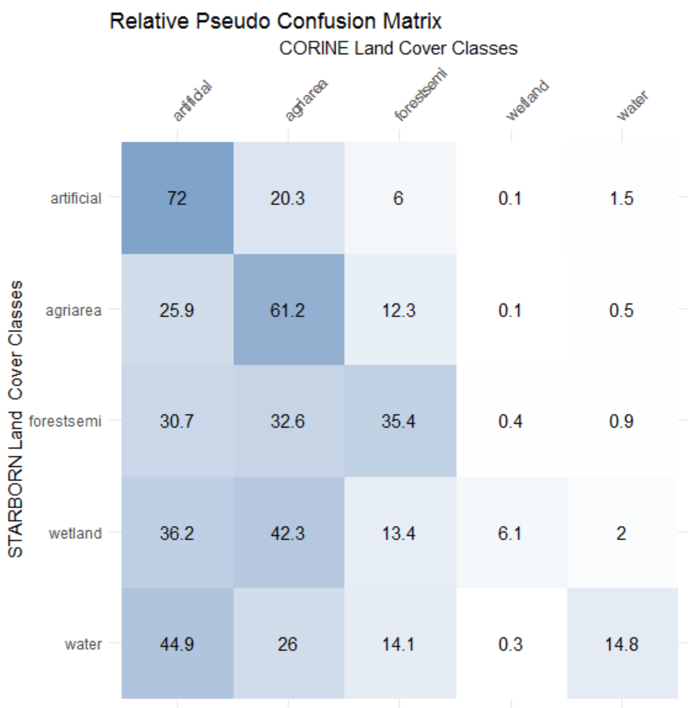

Figure 36 - Relative "pseudo-confusion-matrix" (CORINE level one)
Figure showing the relative "pseudo-confusion-matrix" of the data generated with the location-based game in comparison with the CORINE land cover dataset (level one).

The resulting "pseudo-confusion-matrix" after aggregating the CORINE level two classes to CORINE level one classes (Figure 36) highlights the effects of the difficulties in semantically differentiating different land cover classes. In the matrix containing the relative agreement values of CORINE level one classes, "artificial" now has a 72% agreement rate between the user reported land cover classes and the CORINE land cover classes. The CORINE level one land cover class "artificial" encompasses "urban", "industry", "mine" and "greenarea". Most of the mentioned classes show an increase in agreement by aggregating the level two classes to level one classes. The same effect is visible with the CORINE level one class "agriareas", encompassing the CORINE level two classes "arable", "permacrop", "pasture" and "agriculture". Using the CORINE level one class "agriarea" shows a 61.2% agreement rate between what the users reported and the CORINE land cover dataset. By aggregating "arable", "permacrop", "pasture" and "agriculture" to one class, difficulties of semantic differentiation are dissolved. With the resulting agreement rates being significantly higher using the CORINE level one classifications than the average agreement rates of the classes contained within a said CORINE level one class, this shows that semantic uncertainty can lead to large uncertainties in the data, if not properly addressed or aggregated.

To be able to compute valid confusion matrix metrics, a confusion matrix was generated incorporating only user classifications with exactly one reported class (Figure 37). A total of 3'704 tiles were reported to contain exactly one land cover class, which was deemed a sufficient amount to compute mentioned confusion matrix metrics. The computed metrics are summarised in Table 4. Ideally, the results of a land cover class should show high precision in combination with high recall. The generated metrics show that "urban" has a high precision (0.8124) and recall (0.8462). Using the definitions of precision and recall illustrated in chapter 5.3.4, *Comparison with CORINE 2012*, the mentioned precision translates to 81.24% of the tiles which were classified as "urban" in the CORINE dataset were also classified by the non-expert users of the implemented location-based game as containing "urban". In addition, the mentioned recall translates to 84.65% of tiles that the users reported to contain the land cover class "urban" were found to be classified as "urban" in the CORINE land cover dataset. Similar results were observed with the land cover class "forest" with a precision of 0.7455 and a recall of 0.7553.

Both land cover classes "industry" and "arable" show a high discrepancy between precision and recall ("industry": precision = 0.6796, recall = 0.4496; "arable": precision = 0.3799, recall = 0.8361). 67.96% of the tiles which were classified as "industry" in the CORINE land cover dataset were also classified as "industry" by the users of the implemented location-based game. However, only 44.96% of the user-contributed classifications reported as containing "industry" were also classified as "industry" in the CORINE land cover dataset. Hence, the user-contributed data shows a potential overrepresentation of the land cover class "industry" or the CORINE land cover dataset could show a potential underrepresentation. This is confirmed in the results showing 854 user-contributed tiles which were reported to contain "industry" in comparison with 565 tiles in the CORINE land cover dataset classified as being "industry". Whilst "industry" shows a higher precision than recall, "arable" shows opposite characteristics.

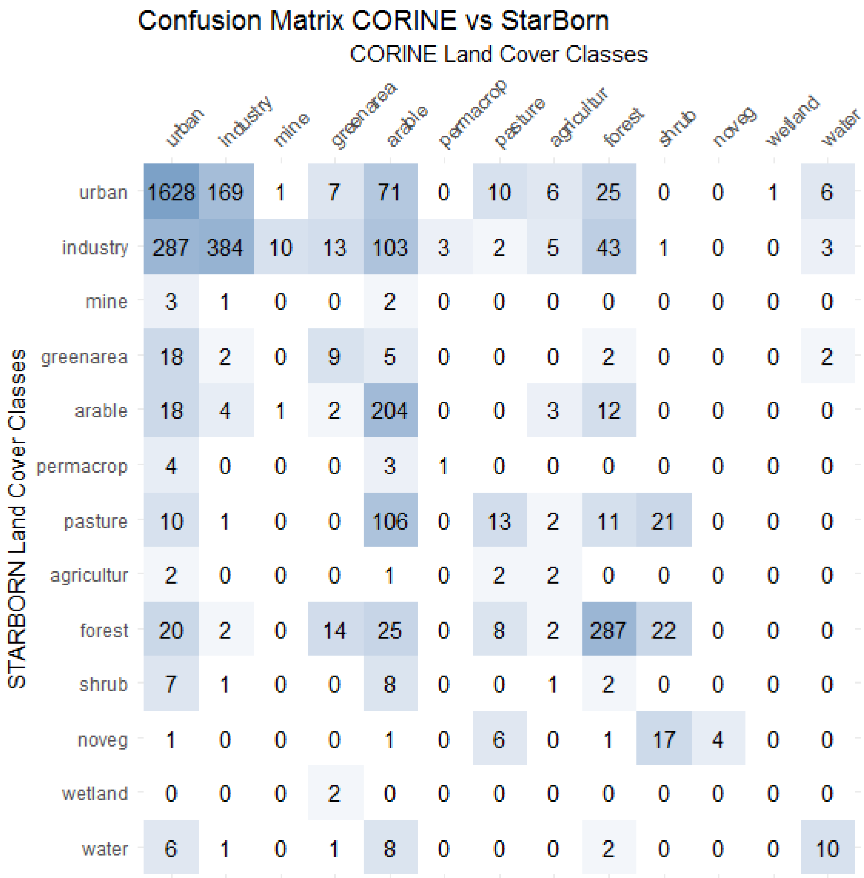

Figure 37 - Confusion matrix (hard classifications)
Figure showing the confusion matrix of the data generated with the location-based game in comparison with the CORINE land cover dataset (level two) using only tiles where exactly one land cover class was reported.

Table 4 - Land cover classes and their respective precision, recall and F-score

Land Cover Class	Precision	Recall	F-Score
Urban	0.8124	0.8462	0.8289
Industry	0.6796	0.4496	0.5412
Mine	0	0	0
Greenarea	0.1875	0.2368	0.2093
Arable	0.3799	0.8361	0.5224
Permacrop	0.25	0.125	0.1667
Pasture	0.3171	0.0793	0.1268
Agriculture	0.0952	0.2857	0.1429
Forest	0.7455	0.7553	0.7503
Shrub	0	0	0
Noveg	1	0.1333	0.2353
Wetland	0	0	0
Water	0.4762	0.3571	0.4081

Only 37.99% of the tiles which were classified as "arable" in the CORINE land cover dataset were also classified as "arable" by the users of the implemented location-based game. In contrast, 83.61% of the user contributed classifications containing "arable" were also classified as "arable" in the CORINE land cover dataset. Thus, mentioned discrepancy between the precision and recall of the land cover class "arable" shows a potential underrepresentation of said land cover class in the user generated data or might indicate a potential overrepresentation in the CORINE land cover dataset. Again, this is underlined by the large difference in number of tiles containing the land cover class "arable" depending on the dataset, with the CORINE land cover dataset showing 537 tiles and the user reported classifications showing 244 tiles. While precision and recall differ largely in "industry" and "arable", the calculated F-scores show similar values of 0.5412 regarding the land cover class "industry" and 0.5224 regarding the land cover class "arable". Consequently, the overall agreement rate of the land cover classes "industry" and "arable" when comparing the user generated data to the official CORINE land cover dataset was found to be similar.

The confusion matrix shows an overall accuracy of 68.63%, which can be attributed to high agreement rates between user generated land cover classifications and the CORINE dataset in land cover classes which were reported frequently, namely "urban" (F-score = 0.8289), "industry" (F-score = 0. 5412), "arable" (F-score = 0.5224) and "forest" (F-score = 0.7503). Considering these findings, a Spearman-correlation test was performed using the F-score and the number of user generated classifications for each land cover class. The results show a significant correlation of rho = 0.8453 (p-value: 0.0003) between the number of user generated classifications of a given land cover class and the F-score.

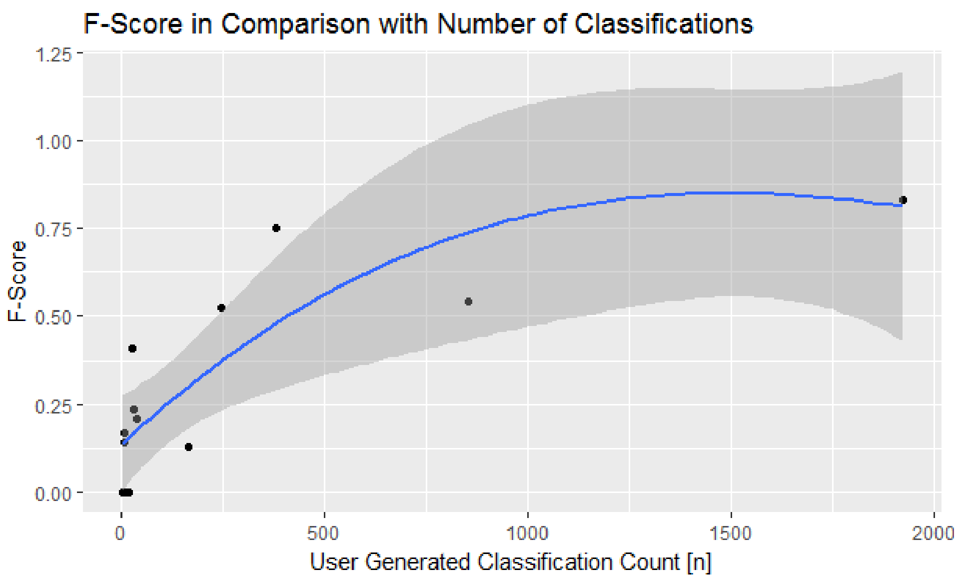

Figure 38 - F-score in comparison with number of classifications
Figure showing the number of classifications of land cover classes in respect to their F-score. A general trendline was added using locally weighted scatterplot smoothing with a span of 2.

Figure 38 visualises the F-score in dependence of the number of classifications, including a general trendline using locally weighted scatterplot smoothing (LOESS; span = 2). In spite of a high calculated correlation, the results must be interpreted with caution as the cluster of points showing low classification counts and low F-scores in combination with few points showing high classification counts and high F-scores could strongly bias the correlation test in favour of a significant correlation.

It must be kept in mind that the mentioned results concerning the confusion matrix that only includes hard classifications take only a fraction of the generated data into account (3'704 tiles with hard user generated classifications compared to 13'338 total classified tiles). While mentioned results shed light on interesting characteristics of the generated hard classifications subset, no founded statements regarding the entire dataset can be made.

7 Discussion

In this chapter, I discuss the results in light of the research questions and in conjunction with literature findings. Overall, the implemented location-based game successfully generated a high amount of land cover data in a short time span. The generated data was analysed and the results reflect various characteristics and issues related to user generated content with regards to land cover classifications. The user generated data was successfully compared with the official CORINE 2012 land cover dataset and statements regarding the usability of a location-based game for the generation and assessment of land cover are presented.

7.1 Research Question One

RQ1: CAN A LOCATION-BASED GAME WITH A NON-EXPERT TARGET AUDIENCE BE IMPLEMENTED TO MINE TILE-BASED GEOGRAPHIC INFORMATION, IN PARTICULAR LAND COVER DATA?

7.1.1 General Suitability of a Location-Based Game for Land Cover Data Validation

Land cover change has been identified as one of the most important variables of global change (Skole 1994; Douglas 1999; Vitousek 1994; Foody et al. 2002), but, as See et al. (2013) point out, there are large disagreements between major land cover products and thus new methods of land cover product assessment are called for. The literature further identifies a "lack of in-situ environmental data for the calibration and validation of remotely sensed products" (See et al. 2013, p.1) while Fritz et al. (2009) suggest implementing a location-based game for land cover validation. The data generated as part of this thesis showed that a location-based game could be successfully implemented and played, generating 13'338 classifications of 200m x 200m extents during the time of data collection. This resulted in an overall extent of 533'520'000 m^2 (some 1.29% of the total area of Switzerland) of classified area contributed by non-expert users between the 09th November 2016 and the 23rd February 2017.

The generated data was analysed and compared to the CORINE 2012 land cover dataset, making differences between the generated data and the CORINE 2012 land cover dataset visible. This thesis successfully demonstrates the viability of implementing a location-based game for the task of land cover validation as proposed by Fritz et al. (2009). This is of particular importance for the validation and assessment of the CORINE land cover dataset seeing the official CORINE report states that not more than 10% of the total budget for a national land cover project should be spent on ground truth surveys (European Environment Agency 1994). A cost efficient and engaging new approach of in-situ data collection is thus needed. This master's thesis confirms the viability of using game elements as underlying motivation to generate cost efficient in-situ observations.

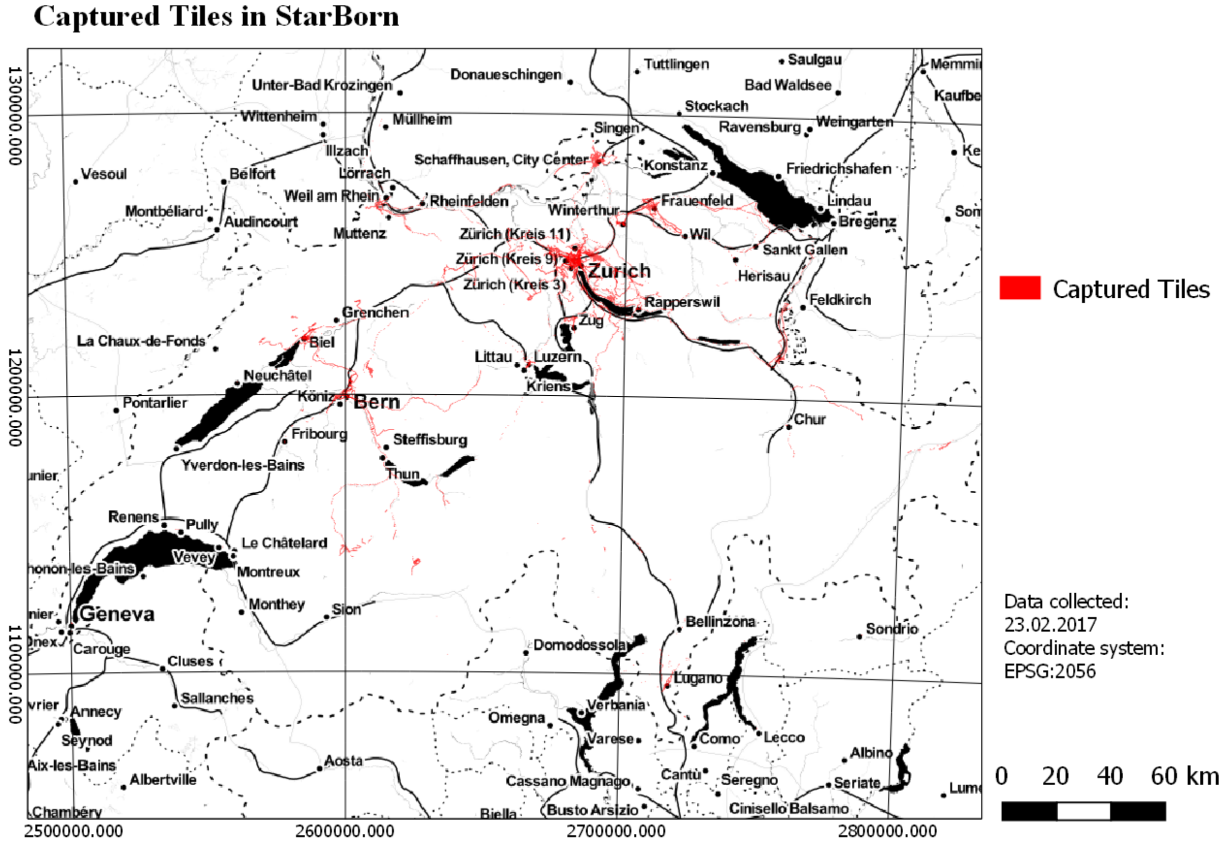

Figure 39 - Captured tiles in StarBorn
Map of Switzerland showing all tiles classified in the implemented location-based game during the data collection period.

7.1.2 Implementation

The implemented infrastructure proved to run stable and efficient throughout the game period. Key improvements were made to the game including updating query notation to increase performance, implementing new features to foster user motivation (user ranking-page; user activity-page; level and experience points system) and introducing new game elements to diversify land cover classifications (treasure hunt system; special in-game events). The overall performance of the game and the response times of 120ms were presumed to be satisfying as no user reported the response time discouraging them from effectively playing the game. The implemented graph database Neo4j proved to be an efficient and well-suited solution for storing data with high numbers of relationships, which corresponds to literature findings (Joishi & Sureka 2015; Vicknair et al. 2010). However, for dealing with explicit spatial data, a spatially enabled database, in this case PostGIS, was more appropriate. This master's thesis shows a clear performance increase when using PostGIS as an advanced spatial index to identify entry nodes in the Neo4j graph database, as opposed to also storing and querying spatial data in Neo4j. The spatial PostGIS query speed was optimised by introducing raster tiles and a spatial index on the tiles, which resulted in a final spatial query time of 80ms.

Symfony handled the implemented game code with appropriate response times of 120ms. This thesis has proven that a web-based solution is widely accessible, practical to implement and maintain and easy to share on internet platforms, which corresponds to the literature (Espada et al. 2012; Charland & Leroux 2011). However, implementing a native location-based game would have had advantages, including direct access to a device's hardware including GPS, compass and accelerometer and more evolved 3D graphics capabilities (Espada et al. 2012; Charland & Leroux 2011).

7.1.3 Classification Selection Page

I put considerable thought into the design of the land cover classification selection page. I argue that using the CORINE level two classification scheme with 15 classes (13 viable for Switzerland) has proven to be a balanced option between the number of options presented to the user and the level of detail of the contributions. Using the CORINE level three classification scheme with 44 classes would be an overload in choices for non-expert users. The literature verifies an increase in user frustration or dissatisfaction if confronted with too many choices (Herbig & Kramer 1994; Iyengar & Lepper 2000). Furthermore, the results show an increase of land cover confusions with an increase in number of land cover class choices, which was identified by comparing the results of CORINE level one and CORINE level two classes. Therefore, the observed difficulties of non-expert users in differentiating semantically similar land cover classes would arguably increase if the CORINE level three classification scheme was used, whereas the CORINE level one land cover classification scheme with five classes does not offer sufficient detail in the land cover class choices. The results show that using the CORINE level one classification scheme can effectively reduce the confusion rates of specific land cover classes, but also greatly reduces the level of detail of the contributions. Consequently, the CORINE level two classification scheme was incorporated.

7.1.4 User Gender and Age Distribution

A higher number of male users (n = 94) in comparison with female users (n = 27) shows the gender specific affinity towards gaming in general, which is confirmed in the 2016 JAMES report (Willemse et al. 2016). The age distribution shows a distinct target group of users born between 1985 and 1995. It remains uncertain, whether this uneven distribution is due to a generally higher affinity of this target group to video games or whether it is due to the specific promotional efforts directed primarily at potential users of this age category.

7.1.5 User Behaviour Patterns

The results suggest two distinct user behaviour patterns. Firstly, users who play the game as a secondary activity whilst performing another activity with a higher subjective priority. This primarily includes users who played the location-based game as a means of distraction whilst traveling or commuting, which corresponds to similar results in the literature (Bell et al. 2006; Zhang et al. 2015). Playing a location-based game whilst traveling can arguably increase the users' in-game performance if not addressed through game mechanics, as a larger in-game area can be covered compared to walking. This hypothesis corresponds to the assumptions of Bell et al. (2006, p.422), who analysed user behaviour in another location-based game, stating: "Journeys may have been good times to play, as players naturally move through different locations […]." This behaviour resulted in users contributing route information potentially allowing valuable insights not only into land cover classification behaviour but also into specific user movement behaviours throughout the period of the game. The second distinct user behaviour was identified as users who play the game as their primary activity, thus adjusting their route of travel to maximise in-game performance. A similar user behaviour pattern was identified by Colley et al. (2017, p.8), pointing out that "*Pokémon GO* might be successfully incenting people to […] substantially change where they choose to go". Bell et al. (2006, p.422) also confirm that users "would take a different route to their destination, either for work or leisure, in order to play […].". This second group of users primarily contributed tiles of coherent extents. I hypothesise that both behaviours can make valuable contributions to a game focusing on generating land cover classifications by effectively increasing the total spatial extent of classified tiles, with the first behaviour contributing route information and the second adding coherent areas. The quality of contributions and agreement rates with the CORINE land cover dataset was not analysed in regards to the user specific behaviour patterns and would be a valuable addition to the presented research. In particular, if the characteristics of contributions (e.g. fuzziness, quality, agreement rates with the CORINE land cover dataset) significantly correlate with a specific user behaviour pattern.

7.1.6 User Motivation

The success of the implemented geographic information mining application with a non-expert target audience was highly dependent on the implemented motivational incentives, which corresponds to the findings of Hoe et al. (2017). These motivational incentives were important, seeing that no monetary or material incentives were used as was the case in similar applications focusing on crowdsourcing land cover assessment or validation tasks mentioned in the literature (See et al. 2013; Bayas et al. 2016). Not only was it vital to motivate users to register, but also to play the game, ideally recurringly. The literature states that key elements such as user competition (Charsky 2010; Lund et al. 2010), character creation (Bessiere et al. 2007; Mcarthur et al. 2015; Ducheneaut et al. 2009) and game fantasy (Charsky 2010; Kenny & Gunter 2007) are vital in motivating users to continuously play a game. All mentioned key elements were successfully incorporated into the game. However, although the literature agrees that purposefully designed elements can be implemented into a game to increase user motivation, a few users stated they felt compelled to help with my scientific research. The subjective moral obligation or the perceived norm of reciprocity (i.e. returning a favour for having received my help in the past) were stated as key motivational elements. The mentioned phenomenon of increased perceived obligation to act according to social norms and reciprocate past received help is confirmed in the literature (Greenglass 1969; Gouldner 1960).

The iterative development of the location-based game allowed the users to participate in the development of the game and this could also be an important motivational aspect. The described key improvements in chapter 5.2.2, *Key Improvements*, demonstrate that developing or improving an application in an iterative and participatory process can have positive effects on user motivation and can lead to new features. Van Rijn and Stappers (2008, p.1) also support these assumptions and further state that a user is more willing to participate if the user "feels respected or trusts the intentions of the designer." Not only was the iterative development important as a motivational element, but also for identifying key improvements to the game, which increased user satisfaction and game performance.

Using various platforms to promote the implemented location-based game resulted in a high number of users registering. The clear decrease in the number of registrations after discontinuing active promotion of the game (cf. Figure 40) and the decrease in the tile capture rate (cf. Figure 41) verify that promotion is a vital part of game development and should be pursued parallel to running an application. Advertising and promotion efforts on social media platforms have been identified as a viable marketing strategy (Wright et al. 2010). In retrospect, a higher devotion to the advertisement of the implemented location-based game could have increased user contributions. This assumption is confirmed in the literature, stating that "marketers are encouraged to keep the lines of communication open with consumers in order to create real value for their customers" (Wright et al. 2010, p.78). On the other hand, Bayas et al. (2016) used a multitude of advertising platforms including newspapers, radio, television and websites, which resulted in 12'278 geo-tagged photographs being collected in 1699 unique locations in a similar timespan. Comparing the mentioned results of Bayas et al. (2016) to the results of this thesis (13'338 classifications in 11'380 unique locations) begs the question if large scale advertising and promotion efforts are an efficient approach to user motivation. Studies suggest that the ability of the advertising party to personalise and deliver messages has a positive impact on the success of a promotion or advertising campaign (Wright et al. 2010), which corresponds to the results of this thesis.

7.1.7 Short and Long Term User Retention

Choosing to implement a location-based game to collect non-expert land-cover classifications proved to generate a high amount of data in a short time. However, the question remains if a location-based game can also ensure long term data collection. The results show that after ceasing all promotional efforts, the rate of newly registered users decreased significantly (cf. Figure 40), with the rate of contributions also gradually slowing down (cf. Figure 41).

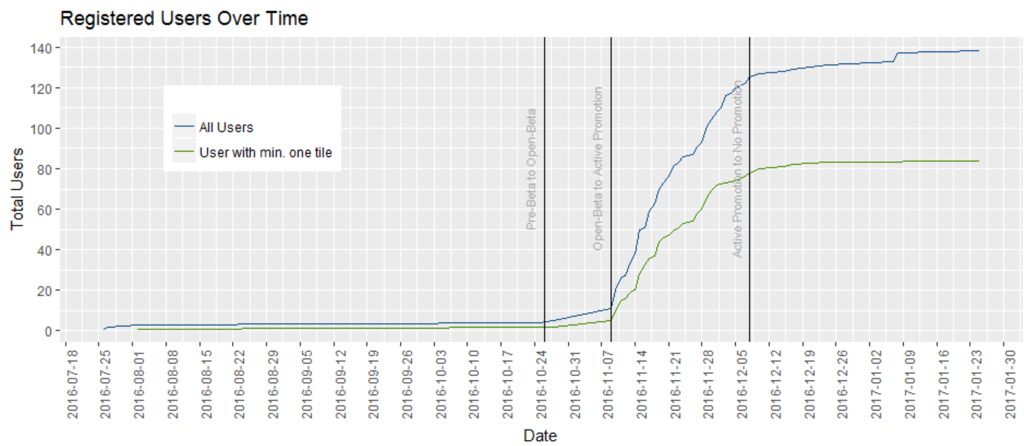

Figure 40 - Registered users over time (discussion)
Figure showing the total number of users over the extent of the game period. Three lines indicate key transitional moments.

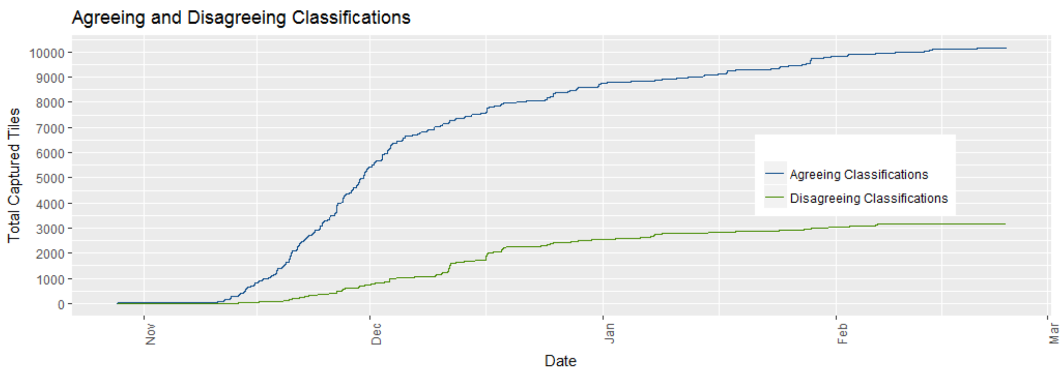

Figure 41 - Total agreeing and disagreeing classifications over time (discussion)
Figure showing the total number of agreeing and disagreeing classifications during the data collection period.

These findings coincide with LeBlanc and Chaput (2016), who mention the unsustainable initial interest levels and significant drop offs of active users shortly after the start of *Pokémon Go*, a widely known location-based game. I argue that for long term user retention, a location-based game might not be the best approach. Targeting interested users with a crowdsourcing platform with competitive or progressive elements emphasising the scientific or professional purposes instead of game fantasy might attract fewer users in the beginning, but could prove to be better suited for long term data collection. An example of a long term crowdsourcing site targeting a professional audience is Stack Overflow[26] as is mentioned in the literature (Reiners & Wood 2015).

7.1.8 Summary RQ1

The implemented location-based game was aimed at a mostly non-expert target audience who contributed many land cover classifications. The first question whether a location-based game can be implemented using key game elements as underlying motivation is thus confirmed by the findings of this thesis.

7.2 Research Question Two

RQ2: CAN THE GENERATED LAND COVER DATA BE USED IN A RESEARCH CONTEXT, IN PARTICULAR WITH REGARDS TO THE VALIDATION OR IMPROVEMENT OF LAND COVER PRODUCTS?

7.2.1 Contributions Over Time

In terms of user contributions over time and their agreement rates, the results show considerable differences to other similar geographic information mining efforts. The total number of user contributions over time shows a characteristic sigmoid curve with the rate of new user contributions decreasing significantly towards the end of the collection period. In contrast, the results of an analysis of the crowdsourced land cover validation project Geo-Wiki[27] show a rapid increase in user contributions towards the end of the data collection phase (See et al. 2013). This major difference in user contribution rate changes over time may be attributed to the different forms of user motivation. In Geo-Wiki, the users with the most points at the end of the data collection period enjoyed a chance to win Amazon[28] vouchers or paper co-authorships (See et al. 2013). Thus, the user contribution rate increased significantly towards the end of the collection period with users wanting to maximise their chances of winning. Unlike Geo-Wiki, the results of this thesis show a significant decline in the user contribution rate towards the end of the test period. After an initial interest period, users were found to contribute less. This may be explained by the fact that the implemented location-based game incorporated no incentive of having the highest score at a specific moment in time. The findings of See et al. (2013) in comparison with the results of this thesis suggest that fixed collection periods with major rewards being distributed after a fixed date could increase the total number of contributions.

[26] www.stackoverflow.com (accessed: 10.04.2017)
[27] www.geo-wiki.org (accessed: 10.04.2017)
[28] www.amazon.com (accessed: 10.04.2017)

7.2.2 Quality Assessment and Fuzzy Classifications

A key advantage of the generated game over the CORINE land cover dataset is the fuzzy nature of the generated tile classifications versus the hard nature of the CORINE raster cell classifications. The advantages of using fuzzy classifications over hard classifications have been identified in the literate as being of higher informational quality and allowing improved uncertainty estimates (Gopal et al. 2016). Of the 13'338 classified tiles, 72.09% were reported to contain more than one land cover class and 30.49% were reported to contain over two land cover classes.

However, the fuzzy nature of the user generated land cover classifications makes it difficult to assess the quality of the CORINE land cover dataset. According to literature findings (Muchoney & Strahler 2002; Carneiro & Pereira 2014; Foody et al. 2002), a common and accepted method for calculating accuracy and quality metrics when validating land cover datasets consists in creating a confusion matrix. The confusion matrix can later be used to calculate precision, recall, accuracy, false positive rates and true positive rates (Fawcett 2006). Seeing that the mentioned metrics are only meaningful if the underlying confusion matrix is used to compare hard pixels, a subset of the entire dataset containing only tiles which users classified as containing exactly one tile was used to compute the confusion matrix metrics.

The high number of fuzzy classifications is a potential indicator of spatial autocorrelations of land cover classes, difficulties of users differentiating between classes or the heterogeneous nature of the landscapes within the classified areas.

7.2.3 Agreement Rates of Consecutive Captures

A high fluctuation of intra-tile agreement rates is illustrated in Figure 25. The results show that the land cover class "urban" has the highest consecutive agreement rate of 80.6% agreement between a preceding and following classification. Following the land cover class "urban", the class "industry" was found to have a consecutive agreement rate of 62.5% and the class "water" a rate of 36.8%. The visible exponential decline in user agreement rates suggests a high uncertainty in most of the land cover classes. "Urban" is arguably a land cover class, of which different users agree on the semantics of what an "urban" landscape consists of. With "industry", the agreement rate and therefore arguably the inter-user agreement on the semantics is considerably lower, declining further in the land cover class "water". One might be inclined to think the land cover class "water" should have large inter-user agreement rates. However, seeing that multiple classifications were allowed for a given tile, low consecutive agreement rates of "water" could have been caused by varying visibility of the manifestation of the land cover class in the natural environment. Local knowledge of an environment and the land cover class not being visible from the position of the user when classifying a tile could therefore be a source of low intra-tile agreement rates of uncommon features or land cover classes with typically small extents.

The high fluctuations of intra-tile agreement rates suggest considerable differences in the intersubjective perception of the semantics of a land cover class by different users. Similar findings were reported by See et al. (2013), who support the assumption that the quality of user generated content varies depending on land cover class and throughout the test period. Differences in users' perception of the semantics of land cover classes have been investigated in the literature (Comber et al. 2015; Comber et al. 2016) and are discussed in further detail in chapter 7.2.5.

7.2.4 Agreements with CORINE

Overall, the results of this thesis show agreement rates of user generated classifications in comparison with the CORINE land cover dataset of 76.15% and 74.93% (100m^2 and 250m^2 CORINE land cover dataset cell size respectively). This corresponds with results of similar research, e.g. Bayas et al. (2016) found an overall agreement rate between crowed sourced land cover data and official land cover data of 70% and See et al. (2013) observed an agreement rate between 66% and 76%. Comparing the agreement rates of this master's thesis with the mentioned literature suggests a higher than average agreement rate, which can be viewed as a proxy of higher data quality. However, the resulting agreement rates of this master's thesis must be interpreted with caution. The conducted analysis implemented an algorithm which outputs true or false (agree or disagree) depending on whether the CORINE land cover classification matches any one of the user contributed classifications or not. Consequently, the user contributed fuzzy classifications were more likely to result in an agreement than the user contributed hard classifications, thus increasing overall agreement.

An additional confusion matrix using only tiles which were reported as containing exactly one class yielded a slightly lower overall agreement rate of 68.63%, which is more in line with the above-mentioned results of similar research. This slightly lower agreement rate also corroborates the assumption that the algorithm used to compare the fuzzy user generated land cover classifications with the hard CORINE land cover classifications results in over-optimistic agreement rates. The calculated confusion matrix metrics show four characteristic combinations of precision (what fraction of tiles with a specific CORINE land cover class agrees with the user generated classification) and recall (what fraction of tiles with a specific user reported land cover class shows an agreement with the CORINE land cover classification). The results indicate that land cover classes with high precision and high recall show high agreement rates between the user generated classifications and the CORINE classifications (e.g. "urban": precision = 0.8124, recall = 0.8462, f-score = 0.8289). The results suggest that land cover classes with a high precision but low recall are overrepresented (e.g. "industry": precision = 0.6796, recall = 0.4496, f-score = 0.5412) and classes which show a low precision but a high recall are underrepresented (e.g. "arable": precision = 0.3799, recall = 0.8361, f-score = 0.5224) in the user contributed data. Despite having fundamentally different precision and recall values, "industry" and "arable" show similar F-score values, consequently indicating similar overall agreement rates. Finally, land cover classes with low precision and low recall also show low agreement rates. Figure 42 visualises the relationship between precision and recall as found in the results of this thesis.

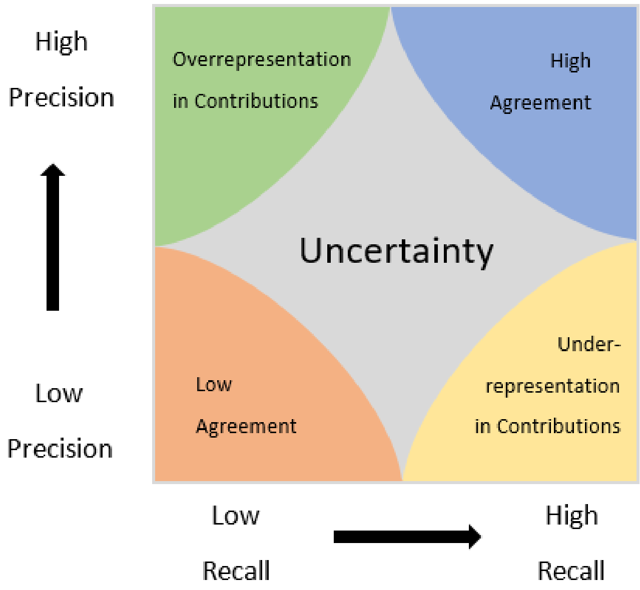

Figure 42 - Precision and recall in user generated land cover classifications
Figure showing the hypothesised relation between precision and recall values and the resulting characteristics of the user contributed land cover classifications.

Over- and underrepresentation as well as over- and underreporting of classifications has been identified as a source of error in user generated or crowdsourced content (Kosmala et al. 2016; Gardiner et al. 2012). Although literature findings confirm the importance of identifying potentially under- or overrepresented classifications, Gardiner et al. (2012) argue that rarer classifications were found to be overrepresented and thus may show a higher grade of uncertainty. This coincides with the findings of this thesis, which show a significant correlation between the F-score and the corresponding number of user generated classifications of each land cover class. However, the distribution of values used in the correlation analysis could be the cause of a high bias towards a significant correlation. Furthermore, the detailed examples presented in chapter 6.4, *Comparison with CORINE 2012*, show a potential tendency to over-classify underrepresented objects as a means of artificially increasing the importance thereof. I argue that in highly spatially autocorrelated landscapes with only small changes over larger distances, irregularities are perceived to be of higher importance than the otherwise perceived monotone surrounding landscape. This again coincides with the arguments of Gardiner et al. (2012) showing that rarer classifications were found to be overrepresented.

The results indicate no significant trend in agreement rate change over time, made visible in Figure 32. The weekly relative disagreement rates varied considerably throughout the test period and the highest percentage of disagreements were recorded in the week of the 08.01.2017 (151 classifications, 96 disagreements) followed by week of the 25.12.2016 (457 classifications, 287 disagreements) and the week of the 12.02.2017 (75 classifications, 39 disagreements). These results show three considerable peaks of relative disagreements in the data collection period. These results contradict the literature stating that the reliability and quality of user contributions increases over time (See et al. 2013; Touya et al. 2017) but the findings do support the assumption that the quality varies depending on land cover class and throughout the test period as mentioned in the literature (See et al. 2013). However, it needs to be pointed out that the applied method did not consider individual users and their contributions independent of the whole dataset. I argue this could be the source of the high fluctuations in the weekly agreement rates as new users registered and started to contribute classifications at different times throughout the data collection period. If all users and their contributions were analysed individually, an increase in agreement rate over time might have been observed.

7.2.5 Disagreements with CORINE

Of arguably higher importance than agreements between the user generated data and the CORINE land cover dataset are the disagreements and their origins. The results show three potential sources regarding the disagreements between the user-contributed data and the CORINE land cover dataset:

- Overrepresentation of classes caused by predominant land cover classifications
- Difficulties of users in differentiating between semantically similar land cover classes
- Spatial autocorrelations and multiple classifications

Overrepresentation of classes – The results reveal users reporting the land cover classes "urban" and "industry" with a higher frequency. Naturally, this uneven distribution of where the game was played led to an overrepresentation of the mentioned land cover classes. I assume the overrepresentation of the land cover classes "urban" and "industry" originates from:

1. a high urban sprawl in Switzerland (Weilenmann et al. 2017) resulting in less crisp borders of the mentioned land cover classes,
2. users choosing to play the location-based game as a distraction whilst commuting through urban areas, which corresponds to the findings of Bell et al. (2006),
3. a higher in-game performance of users in urban areas, coinciding with the results of Colley et al. (2017),
4. or unstable cellular broadband connection in remote areas resulting in slow internet speeds, potentially turning the game-playing into an unsatisfactory experience, as suggested in literature findings (Huang et al. 2013) and further supported by taking into account the

official cellular coverage maps of the three major cellular providers in Switzerland: Salt[29], Sunrise[30] and Swisscom[31].

Colley et al. (2017) noticed a higher in-game performance of users playing the location-based game *Pokémon GO* in urban areas. The authors attribute this increase to the higher number of possible in-game activities in urban areas which grant experience points. This corresponds to the results of this thesis as enemy classification hotspots, typically found in urban areas, imply less travel distance for a user looking to destroy those enemy tiles with the aim of gaining additional experience points.

The overrepresentation of the land cover classes "urban" and "industry" calls for a careful interpretation of the co-occurrence matrices as to not confuse naturally higher co-occurrences between land cover classes with potential spatial autocorrelations of land cover classes.

Difficulties in differentiating classes – The potential subjective differences in regards to the semantics of land cover classes must be taken into consideration and addressed. The results show considerable difficulties of users in differentiating between specific land cover classes. This is arguably one of the most prominent sources of disagreement between the data generated with the implemented location-based game and the CORINE land cover dataset. The confusion rate of specific pairs of land cover classes (e.g. "permacrop" and "pasture" with "arable") presented in Figure 35 is significantly higher than others (e.g. "forest" with "water"), indicating difficulties in semantically differentiating between specific land cover classes. Comparing the "pseudo-confusion-matrix" of the CORINE level two classification scheme (Figure 35) with the "pseudo-confusion-matrix" of the CORINE level one classification scheme (Figure 36) illustrates these semantic issues of land cover classifications. The implemented algorithm aggregates CORINE level two classes to CORINE level one classes resulting in a major decrease of confusions within specific CORINE level one classes (e.g. agricultural areas). The literature confirms that the presented issues regarding the ontological and semantic aspects of land cover classes must be taken into consideration and that they can present a major source of errors or disagreements between different users or disciplines (Comber et al. 2005). Comber et al. (2005, p.226) state that "[l]and cover is perceived differently by different disciplines, and their perceptions inform their assessment of the data and their analyses." Further research has been done on different conceptualisations of landscape features depending on nationality and level of expertise (Comber et al. 2016). The authors conclude "that it is important to consider and test for potential variations in the way that landscape features are labelled and conceptualised by different groups of contributors when analysing crowdsourced data" (Comber et al. 2016, p.16). The results of this thesis agree with the

[29] www.salt.ch/en/coverage/ (accessed: 10.04.2017)
[30] www.sunrise.ch/en/residential/mobile/mobile-network/network-coverage/network-coverage-map.html (accessed: 10.04.2017)
[31] https://scmplc.begasoft.ch/plcapp/pages/gis/netzabdeckung.jsf?lang=en (accessed: 10.04.2017)

literature (Comber et al. 2015; Comber et al. 2016) on the finding that different perceptions of land cover classes can be observed between different users.

Figure 43 - Example image of a classification showing disagreement
Image showing an in-situ photograph of an area where a spatially autocorrelated patch of disagreement between the user generated land cover classifications and the CORINE land cover dataset can be observed. Source: Google Maps[32].

Two patches of spatially autocorrelated disagreements were presented in detail (chapter 6.4, *Comparison with CORINE 2012*). Even though the presented examples were analysed individually and no holistic statements should be made, they show potential problematic land cover classes in terms of their semantics. The official CORINE documentation for example defines green urban areas as "all vegetated areas greater than 25 ha that are either situated within or in contact with urban fabrics. Strips of lanes and paths created for recreational use may be found within these areas" (Kosztra et al. 2014, p.29). The presented example (Figure 34 left) shows a spatially autocorrelated area of disagreeing tiles of which 14 were reported by a user as containing "forest" which CORINE classifies as being "greenarea". Figure 43 shows an in-situ photograph of the area in question and highlights the difficulty of classifying certain land cover classes. The presented spatially autocorrelated patch of disagreements can arguably be attributed to the different definitions of "greenarea" and "forest" depending on the classifying user, again underlining the importance of analysing the semantics and spatial autocorrelations within land cover datasets.

[32] www.google.ch/maps

Semantics and how human beings perceive and construct reality and scientific knowledge is no topic specific to land cover product assessment, but has been a fundamental question of philosophy and cognitive science throughout the past millennias. Centuries of study and philosophical discussions have led to a plethora of theories ranging from Plato's allegory of the cave, which argues that "everyday life is sunk in illusion because of its dependence on the senses" (Buckle 2007, p.327), over symbolic interactionism, which focuses on "the interpretation of subjective viewpoints and how individuals make sense of their world from their unique perspective" (Carter & Fuller 2016, p.932), to radical constructivism, which states "all understanding and all communication is a matter of interpretative construction on the part of the experiencing subject" (Olssen 1996, pp. 276-277). Research in geographic information science has identified key cognitive issues regarding semantics in geographic information and the literature has discussed cognitive aspects of geographic information systems and the implications thereof (Mark & Freundschuh 1995; Mennis 2003). Prominent are discussions of cognitive categorisation indicating the interdependence of spatial entities with the conceptual categories which an observer of an entity perceives said entity to belong to (Mennis 2003). Mennis (2003, pp.458–459) further states that "[h]umans typically use prior knowledge about objects to interpret new visual input." The literature underlines the hypothesis that non-expert and expert users alike categorise visual inputs of varying landscapes according to subjective prior knowledge and cognitive processes. Thus, different users might assign the same landscape encompassing the same physical and spectral properties to different land cover classes. This hypothesis is not only corroborated by philosophical theories but also by the findings of this thesis and by similar research findings (Comber et al. 2016).

Understanding the perceived semantics of land cover classes is of utmost importance in crowdsourcing approaches of land cover data generation, which coincides with the argumentation of Comber et al. (2016). In Addition, I argue that discrepancies of the perceived land cover classes between experts and often local non-experts may lead to major problems and potential conflict situations, especially when it comes to large-scale top-down policy or decision making processes. Therefore, individual or group-related differences in the perception and thus classification of land cover classes should be addressed in more detail.

Spatial autocorrelation of classes – A discrepancy in the perceived semantics of land cover classes between expert users who create the CORINE classification scheme and non-expert users who played the implemented location-based game were found in spatially autocorrelated patches of disagreement (cf. Figure 33 and Figure 34). The literature (Pugh & Congalton 2001; Muchoney & Strahler 2002) agrees that the issue of spatial autocorrelation is an important factor to be considered when analysing and interpreting remotely sensed products or land cover datasets. The results of this master's thesis suggest the presence of spatial autocorrelations. A perceived spatial autocorrelation is evident in the agreeing and disagreeing land cover classes. This confirms that land cover classifications can be spatially autocorrelated including potential errors in the CORINE land cover data. These findings correspond to the literature confirming the presence of clusters of spatially autocorrelated classification errors in land cover data (Campbell 1981).

I argue that the presented issue of spatial autocorrelation delivers valuable information for future assessments of CORINE land cover datasets, as spatially autocorrelated errors were found to show major semantic disagreements at specific geographic locations.

7.2.6 Summary RQ2

The user contributed data of the implemented location-based game showed above average agreement rates with the CORINE land cover dataset in comparison to similar research. However, the high agreement rates must be interpreted with extreme caution due to major differences in the two compared datasets and limitations in the used methods. Disagreements were found especially due to comparing fuzzy pixels with hard pixels, having different tile sizes in the implementation and the CORINE land cover dataset, overrepresentations of specific land cover classes, spatial autocorrelations of land cover classes and difficulties of users differentiating between semantically similar land cover classes.

The results of this thesis confirm that the implemented location-based game is less suited for automated land cover curation processes, but rather to detect areas of spatially autocorrelated classification errors and to gain insights into potential semantic issues. Nevertheless, the generated data was found to be of higher informational value than the CORINE land cover dataset because of the fuzzy nature of the contributions. The results have also shown that the implemented location-based game can generate large amounts of data which can be used for quality estimations, especially if analysing consecutive classifications.

The second question is therefore only partially answered as the data can be used to assess the CORINE land cover dataset but mainly to identify problematic areas which need further manual inspection.

8 Limitations

As in every study, the results of this master's thesis have to be seen in the light of several limitations spanning from limitations in the design and implementation of a location-based game to limitations of the analyses presented in this work.

To begin with, even though the possibility of multiple classifications of one tile effectively increases the informational value of a tile, it also comes with major limitations. Firstly, the fuzzy data holds no information about the relative extents of the different land cover classes in one tile. Even though this could be addressed by prompting the user to report the percentage of each land cover class perceived to be present in each tile, such additional tasks would most probably have a negative impact on user motivation. Secondly, comparing a fuzzy dataset with a hard dataset is a potential source of errors or major uncertainties as "conventional measures of classification accuracy cannot be used as they are appropriate only for 'hard' classifications" (Foody 1996, p.1317). Various approaches of defuzzification strategies for fuzzy remotely sensed data have been addressed in the literature (Hofmann 2016; Foody 1996). Artificial neural networks have been identified as a promising defuzzification strategy when dealing with cells with fuzzy memberships (Song & Bortolan 1994; Foody 1996; Beleites et al. 2013). The user contributed data generated by the implemented location-based game could eventually be defuzzified using an artificial neural network, yet the question arises on what data should be used to train these artificial neural networks. Using CORINE land cover data to train an artificial neural network to defuzzify user contributed land cover classifications seems pointless, if the defuzzified dataset is then compared with the CORINE land cover dataset, because the results will be heavily biased.

Another major limitation of multiple classifications is that the provided classifications contain no information on whether every reported land cover class is present or if the user could not differentiate between classes and thus chose to mention all at once. If a user is not sure which class to report he or she might feel inclined to report all potential classes. Therefore, users knowingly report false classifications, not knowing which classification should be reported. Consulting the literature reveals this issue may be addressed by incorporating some form of confidence declaration (See et al. 2013).

Another limitation is reducing the game tiles from polygons to centroids and extracting the CORINE raster values at mentioned centroids. Even though the effect was analysed in this thesis, higher accuracy in the results could be achieved if the raster values were extracted over the extent of the game tiles. A possible solution to this limitation could be an analysis of the user contributed data by intersecting the contributed game tiles with the CORINE 2012 vector land cover dataset.

9 Conclusions and Further Work

9.1 Conclusions

This thesis investigated various research gaps concerning the implementation of a tile-based location-based game for geographic land cover data mining and the analysis of the user contributed data. This was achieved by designing, developing and implementing a location-based game, advertising and promoting this game and improving the location-based game in a participatory and iterative process. The implemented location-based game successfully generated substantial quantities of user generated tile-based land cover classifications. These classifications were subsequently analysed in terms of user attributes, user contributions, intra-tile agreement rates compared with official CORINE 2012 land cover data. The overarching goal of this master's thesis was defined as:

> *The development, implementation, assessment and analysis of a tile-based location-based game with continuous gameplay, including narrative as well as competitive elements, focusing on geographic information mining regarding land cover data and the analysis of the generated data.*

All aspects of the research goal were achieved, including the proposed implementation and analysis.

The implementation of a location-based game for tile-based land cover data mining revealed that choosing adequate infrastructure, facilitating participatory improvement iterations, addressing user motivation and integrating sustained promotional efforts are key variables in user registration, user retention and contribution rates. However, allowing multiple classifications without recording corresponding subjective confidence values or relative extent proportions was identified as a major limitation.

Analysing the generated user contributed data revealed that a location-based game is a viable option to generate a large number of non-expert land cover classifications. The generated data was found to be of equal overall accuracy as presented in similar research. The results revealed three major sources of disagreement between the user generated land cover classifications and the CORINE land cover dataset: (1) overrepresentation of classes, (2) difficulties in differentiating classes and (3) spatial autocorrelation of classes. Understanding the semantics of land cover classes and the perception thereof was identified as being of utmost importance in land cover product assessment and validation. However, the generated data proved difficult to compare with the official CORINE land cover dataset because of the fuzzy nature of the generated data as opposed to the hard nature of the CORINE land cover dataset.

Modern computational capacities and freely available open source frameworks allow for efficient and cost effective implementation of new approaches to geographic information mining. Identifying the strengths and weaknesses of underlying infrastructural elements is essential to be able to detect and eliminate bottlenecks and increase user satisfaction. The implemented Neo4j graph database achieved adequate performance in storing non-spatial data. PostGIS showed exponentially higher performance rates in dealing with explicitly spatial data and was thus incorporated as a complex spatial index to identify Neo4j entry nodes through the utilisation of four custom raster bands. In conclusion, I stress the importance of carefully using infrastructural elements according to their individual strengths, identified by vigorous testing.

Monetary incentives are not essential in motivating users to contribute data or aid in some underlying crowdsourced task. Motivational elements of games such as fostering competition, allowing self-presentation and providing a sense of progression suffice as motivational stimuli to encourage non-expert users to get involved. However, sustained user motivation and effective user retention depend on additional features. Frequent updates, active promotion, widescale advertisement, participatory iterative development and personalisation of community messages can be effective tools to maintain user motivation and to assure a sustained rate of new user registrations.

9.2 Further Work

While this thesis has addressed previously identified research gaps and contributed to the current scientific discussion by using a real-time location-based game for tile-based non-expert user generated land cover data collection, many opportunities remain for extending the scope of the thesis.

First of all, the results of this master's thesis show major difficulties and potential uncertainties when comparing a fuzzy dataset with a hard dataset. An artificial neural network could be implemented to potentially defuzzify the fuzzy user contributed dataset. However, defining a training set would need careful attention. The location-based game could also be enhanced prompting users to contribute additional confidence values and land cover proportion estimations. This additional data could be used to defuzzify the dataset using a predefined rule based approach or it could lead to more accurate artificial neural network predictions. Another possible solution would be to only allow the users to report one classification. This will ultimately result in fuzzy pixels if multiple users classify the same tile with differing classifications, but it could lead to more solid intra-tile user accuracy predictions and potentially easier defuzzification processes.

The results show that the semantics of land cover classifications is one of the major sources of disagreements. The results also indicate a potential correlation between the quality of the generated data and the inter-subjective agreement rates on the semantics of the different land cover classes. Further analysis regarding the different perceptions of land cover classes between and within expert and non-expert users could produce additional valuable results regarding the semantics of different land cover classes and may prove to be vital information for future land cover product quality assessments. The location-based game could be enhanced to also collect in-situ photographs, which might shed light on how users perceive a given land cover.

Even though user movement data was indirectly collected, this was not explicitly analysed in the scope of this thesis. Analysing user movements could provide crucial insights into user behaviour and movement patterns. Analysing user movement patterns regarding the quality or accuracy of user contributions could result in important quality-velocity or fuzziness-velocity correlations, which would then need to be addressed in future implementations.

Finally, the effects of using different types of gamified applications on short- and long term user motivation are of great interest and should be studied in more detail. Sustaining long term motivation would be necessary if the generated data were to be used to detect changes in long term developments. Aimed at an expert group, where long term motivation is more easily achieved, a similar application could be implemented with less gamified elements that might be more appealing to a scientific or niche community.

10 Bibliography

Adams, E. & Dormans, J., 2012. *Game Mechanics - Advanced Game Design*, New Riders.

Alt, F. et al., 2010. Location-based crowdsourcing: Extending Crowdsourcing to the Real World. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10*, pp.13–22.

Antoniou, V., Morley, J. & Haklay, M., 2010. Web 2.0 geotagged photos: Assessing the spatial dimensions of the phenomenon. *Geomatica*, 64(1), pp.99–110.

Avouris, N. & Yiannoutsou, N., 2012. A Review of Mobile Location-based Games for Learning across Physical and Virtual Spaces. *Journal of Universal Computer Science*, 18(15), pp.2120–2142.

Bayas, J.C.L. et al., 2016. Crowdsourcing in-situ data on land cover and land use using gamification and mobile technology. *Remote Sensing*, 8(11), pp.1–18.

Beleites, C., Salzer, R. & Sergo, V., 2013. Validation of soft classification models using partial class memberships: An extended concept of sensitivity & co. applied to grading of astrocytoma tissues. *Chemometrics and Intelligent Laboratory Systems*, 122, pp.12–22.

Bell, M. et al., 2006. Interweaving Mobile Games With Everyday Life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'06)*, pp.417–426.

Bessiere, K., Seay, A.F. & Kiesler, S., 2007. The ideal elf: identity exploration in World of Warcraft. *Cyberpsychol Behav*, 10(4), pp.530–535.

Brabham, D.C., 2012. Motivations for Participation in a Crowdsourcing Application to Improve Public Engagement in Transit Planning. *Journal of Applied Communication Research*, 40(3), pp.307–328.

Buckle, S., 2007. Descartes, Plato and the Cave. *Philosophy*, 82(2), pp.301–337.

Campbell, J.B., 1981. Spatial correlation effects upon accuracy of supervised classification of land cover. *Photogrammetric Engineering and Remote Sensing*, 47(3), pp.355–363.

Cantallops, M.M. & Sicilia, M.-A., 2016. Motivations to read and learn in videogame lore : the case of League of Legends. In *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality,* pp.585-591.

Carneiro, J.T. & Pereira, M.J., 2014. Beyond the confusion matrix: geostatistical error assessment for Landsat landcover maps of the Portuguese landscape. In *2014 IEEE International on Geoscience and Remote Sensing Symposium (IGARSS)*, pp.4902–4905.

Carter, M.J. & Fuller, C., 2016. Symbols, meaning, and action: The past, present, and future of symbolic interactionism. *Current Sociology*, 64(6), pp.931–961.

Celino, I., 2015. Geospatial dataset curation through a location-based game: Description of the Urbanopoly linked datasets. *Semantic Web*, 6(2), pp.121–130.

Celino, I. et al., 2012. Urbanopoly - A social and location-based game with a purpose to crowdsource your urban data. In *Proceedings - 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust and 2012 ASE/IEEE International Conference on Social Computing, SocialCom/PASSAT 2012*. pp.910–913.

Chapin, F.S. et al., 2000. Consequences of changing biodiversity. *Nature*, 405(6783), pp.234–42.

Charland, A. & Leroux, B., 2011. Mobile application Development: Web vs. native. *Communications of the ACM*, 54(5), pp.49–53..

Charsky, D., 2010. From Edutainment to Serious Games: A Change in the Use of Game Characteristics. *Games and Culture*, pp.177–198.

Colley, A. et al., 2017. The Geography of Pokémon GO: Beneficial and Problematic Effects on Places and Movement. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Chi 2017)*.

Comber, A. et al., 2015. Comparing national differences in what people perceive to be there: Mapping variations in crowd sourced land cover. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 40(3W3), pp.71–75.

Comber, A. et al., 2016. Crowdsourcing: It matters who the crowd are. The impacts of between group variations in recording land cover. *PLoS ONE*, 11(7), pp.1–19.

Comber, A., Fisher, P. & Wadsworth, R., 2005. What is land cover? *Environment and Planning B: Planning and Design*, 32(2), pp.199–209.

Comber, A., Fisher, P. & Wadsworth, R., 2005. You know what land cover is but does anyone else?…an investigation into semantic and ontological confusion. *International Journal of Remote Sensing*, 26(1), pp.223–228.

Congalton, R.G. et al., 2014. Global land cover mapping: A review and uncertainty analysis. *Remote Sensing*, 6(12), pp.12070–12093.

Craglia, M., Ostermann, F. & Spinsanti, L., 2012. Digital Earth from vision to practice: making sense of citizen-generated content. *International Journal of Digital Earth*, 5(5), pp.398–416.

Davidovic, N. et al., 2013. Using Location Based Game MapSigns to motivate VGI data collection related to traffic signs. Presented at *International Workshop on Action and Interaction in Volunteered Geographic Information (ACTIVITY)*, AGILE 2013.

Deci, E.L. & Ryan, R.M., 2000. The "What" and "Why" of Goal Pursuits Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, 11(4), pp.37–41.

Deeds, D.W., 2016. *Gamification in Education and Business*, Available at: http://link.springer.com/content/pdf/10.1007/978-3-319-10208-5.pdf.

Douglas, I., 1999. Hydrological investigations of forest disturbance and land cover impacts in South-East Asia: a review. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 354(1391), pp.1725–1738.

Ducheneaut, N. et al., 2009. Body and mind: a study of avatar personalization in three virtual worlds. In *BT - Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, pp.1151–1160.

Espada, J.P. et al., 2012. Extensible architecture for context-aware mobile web applications. *Expert Systems with Applications*, 39(10), pp.9686–9694.

European Environment Agency, 1994. *CORINE Documentation*, Available at: http://www.eea.europa.eu/publications/COR0-landcover/page001.html.

Fawcett, T., 2006. An Introduction to ROC Analysis. *Pattern Recogn. Lett.*, 27(8), pp.861–874.

Fisher, P., Comber, A. & Wadsworth, R., 2005. Land Use and Land Cover: Contradiction or Complement. In P. Fisher & D.J. Unvin, eds. *Re-Presenting GIS* West Sussex: Wiley, pp.85–98.

Foody, G.M. & Boyd, D.S., 2012. Using volunteered data in land cover map validation: Mapping tropical forests across West Africa. In *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp.6207–6208.

Foody, G.M., Giles M, F. & Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), pp.185–201.

Foody, G.M.G., 1996. Approaches for the production and evaluation of fuzzy land cover classifications from remotely sensed data Foody, G. M. *International Journal of Remote Sensing*, 17(1996), pp.1317–1340.

Fritz, S. et al., 2009. Geo-wiki.org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, 1(3), pp.345–354.

Gardiner, M.M. et al., 2012. Lessons from lady beetles: Accuracy of monitoring data from US and UK citizenscience programs. *Frontiers in Ecology and the Environment*, 10(9), pp.471–476.

Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal*, 69(4), pp.211–221.

Gopal, S. et al., 2016. Characterizing urban landscapes using fuzzy sets. *Computers, Environment and Urban Systems*, 57, pp.212–223.

Gouldner, A.W., 1960. The Norm of Reciprocity : A Preliminary Statement. *American Sociological Review*, 25(2), pp.161–178.

Greenglass, E.R., 1969. Effects of prior help and hindrance on willingness to help another: Reciprocity or social responsibility. *Journal of Personality and Social Psychology*, 11(3), pp.224–231.

Haklay, M. (Muki) et al., 2010. How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. *The Cartographic Journal*, 47(4), pp.315–322.

Hamari, J., 2015. Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior*.

Herbig, P.A. & Kramer, H., 1994. The Effect of Information Overload on the Innovation Choice Process. *Journal of Consumer Marketing*, 11(2), pp.45–54.

Hoe, D. et al., 2017. Perceptions of virtual reward systems in crowdsourcing games. *Computers in Human Behavior*, 70, pp.365–374.

Hofmann, P., 2016. Defuzzification strategies for fuzzy classifications of remote sensing data. *Remote Sensing*, 8(6), 467.

Huang, C. et al., 2013. Quantifying User Satisfaction in Mobile Cloud Games. In *Proceedings of Workshop on Mobile Video Delivery*, 4.

Hutchison, D. et al., 2012. Crowdsourcing Satellite Imagery Analysis: Study of Parallel and Iterative Models. In N. Xiao et al., eds. *Geographic Information Science*. *7th International Conference, GIScience 2012, Columbus, OH, USA, September 18-21, 2012. Proceedings*. Berlin, Heidelberg: Springer, pp.116–131.

Ionescu, G., De Valmaseda, J.M. & Deriaz, M., 2013. GeoGuild: Location-based framework for mobile games. In *Proceedings - 2013 IEEE 3rd International Conference on Cloud and Green Computing, CGC 2013 and 2013 IEEE 3rd International Conference on Social Computing and Its Applications, SCA 2013*. pp.261–265.

Iyengar, S.S. & Lepper, M.R., 2000. When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79(6), pp.995–1006.

Joishi, J. & Sureka, A., 2015. Vishleshan: Performance Comparison and Programming Process Mining Algorithms in Graph-Oriented and Relational Database Query Languages. In *International Database Engineering & Applications Symposiums (IDEAS 2015)*.

Kenny, R.F. & Gunter, G.A., 2007. Endogenous fantasy-based serious games: intrinsic motivation and learning. *International Journal of Social Sciences*, 1(11), pp.591–596.

Kiefer, P. & Schlieder, C., 2012. Changing the Rules : Acquiring Quality Assured Geospatial Data With Location-based Games. In *Proceedings of the 4th annual Asian GAME-ON Conference on Simulation and AI in Computer Games (GAME-ON ASIA'2012), pp.44–51.*

King, D., Delfabbro, P. & Griffiths, M., 2010. Video game structural characteristics: A new psychological taxonomy. *International Journal of Mental Health and Addiction*, 8(1), pp.90–106.

Klotz, M. et al., 2016. How good is the map? A multi-scale cross-comparison framework for global settlement layers: Evidence from Central Europe. *Remote Sensing of Environment*, 178, pp.191–212.

Kolomičenko, V., Svoboda, M. & Mlýnková, I.H., 2013. Experimental Comparison of Graph Databases. *Proceedings of International Conference on Information Integration and Web-based Applications & Services - IIWAS '13*, pp.115–124.

Kosmala, M. et al., 2016. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, 14(10), pp.551–560.

Kosztra, B. et al., 2014. Proposal for enhancement of CLC nomenclature guidelines. Available at: http://land.copernicus.eu/user-corner/technical-library/CLC2006_Nomenclature_illustrated_guide_enhanced_final.pdf.

Lambin, E.F. et al., 2001. The causes of land-use and land-cover change:moving beyond the myths. *Global Environmental Change*, 11, pp.261–269.

LeBlanc, A.G. & Chaput, J.-P., 2016. Pokémon Go: A game changer for the physical inactivity crisis? *Preventive Medicine*, 2015, pp.2015–2017.

Leung, D. & Newsam, S., 2014. Land cover classification using geo-referenced photos. *Multimedia Tools and Applications*, 74(24), pp.11741–11761.

Lund, K., Lochrie, M. & Coulton, P., 2010. Enabling Emergent Behaviour in Location Based Games. In *Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media*

Environments. pp. 78–85.

Mallupattu, P.K. & Sreenivasula Reddy, J.R., 2013. Analysis of land use/land cover changes using remote sensing data and GIS at an urban area, Tirupati, India. *The Scientific World Journal*, 2013, 268623.

Mark, D.M. & Freundschuh, S.M., 1995. Spatial Concepts and Cognitive Models for Geographic Information Use. In T. L. Nyerges et al., eds. *Cognitive Aspects of Human-Computer Interaction for Geographic Information Systems*. Dordrecht: Springer Netherlands, pp.21–28.

Matyas, S. et al., 2008. Designing location-based mobile games with a purpose: collecting geospatial data with CityExplorer. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology*. pp.244–247.

Matyas, S., 2007. Playful Geospatial Data Acquisition by Location-based Gaming Communities. *The International Journal of Virtual Reality*, 6(3), pp.1–10.

Matyas, S. et al., 2011. Wisdom about the crowd: Assuring geospatial data quality collected in location-based games. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. pp. 1–6.

Mcarthur, V., Teather, R.J. & Jenson, J., 2015. The Avatar Affordances Framework : Mapping Affordances and Design Trends in Character Creation Interfaces. *CHI Play*, pp.231–240.

Mekler, E.D. et al., 2015. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, 71, pp.525–534.

Mennis, J.L., 2003. Derivation and implementation of a semantic GIS data model informed by principles of cognition. *Computers, Environment and Urban Systems*, 27(5), pp.455–479.

Muchoney, D.M. & Strahler, A.H., 2002. Pixel- and site-based calibration and validation methods for evaluating supervised classification of remotely sensed data. *Remote Sensing of Environment*, 81(2–3), pp.290–299.

Neuendorf, K.A., 2016. *The Content Analysis Guidebook*. Thousand Oaks, London, New Delhi, Singapore: SAGE Publications.

Neustaedter, C., Tang, A. & Judge, T.K., 2013. Creating scalable location-based games: Lessons from Geocaching. *Personal and Ubiquitous Computing*, 17(2), pp.335–349.

Olson, D.L. & Delen, D., 2008. *Advanced data mining techniques*, Berlin, Heidelberg: Springer.

Pugh, S. & Congalton, R., 2001. Applying spatial autocorrelation analysis to evaluate error in New

England forest-cover-type maps derived from Landsat Thematic Mapper Data. *Photogrammetric Engineering & Remote Sensing*, 67(5), pp.613–620.

Reiners, T. & Wood, L.C., 2015. *Gamification in Education and Business*. Cham, Heidelberg, New Yor, Dordrecht, Londin: Springer International Publishing.

Richter, D. et al., 2012. How People Describe their Place: Identifying Predominant Types of Place Descriptions. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information*, pp.30–37.

van Rijn, H. & Stappers, P.J., 2008. Expressions of ownership: motivating users in a co-design process. *Participatory Design*, pp.181–184.

See, L. et al., 2013. Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts. *PLoS ONE*, 8(7), pp.1–11.

Sehra, S.S., Singh, J. & Rai, H.S., 2014. A Systematic Study of OpenStreetMap Data Quality Assessment. *2014 11th International Conference on Information Technology: New Generations*, pp.377–381.

Senaratne, H. et al., 2016. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 31(1), pp.139–167.

Skole, D.L., 1994. Data on global land-cover change: acquisition, assessment, and analysis. In W.B. Meyer & B.L. Turner, eds. *Changes in Land Use and Land Cover. A Global Perspective.* Cambridge: Cambridge University Press, pp.437–471.

Song, Q. & Bortolan, G., 1994. Some properties of defuzzification neural networks. *Fuzzy Sets and Systems*, 61(1), pp.83–89.

Olssen, M., 1996. Radical Constructivism and Its Failings: Anti-Realism and Individualism. *British Journal of Educational Studies* , 44(3), pp.275-295.

Symfony, Symfony Documentation - Controller. Available at: http://symfony.com/doc/current/controller.html [Accessed March 5, 2017].

Touya, G. et al., 2017. Assessing Crowdsourced POI Quality: Combining Methods Based on Reference Data, History, and Spatial Relations. *ISPRS International Journal of Geo-Information*, 6(3), 80.

Trumbo, B.E., 1981. A Theory for Coloring Bivariate Statistical Maps. *The American Statistician*, 35(4), pp.220–226.

Vicknair, C. et al., 2010. A comparison of a graph database and a relational database. In *Proceedings of the 48th Annual Southeast Regional Conference on ACM SE 10*, 42.

Vitousek, P.M., 1994. Beyond global warming: Ecology and global change. *Ecology*, 75(7), pp.1861–1876.

Wainer, H. & Francolini, C.M., 1980. An Empirical Inquiry Concerning Human Understanding of Two-Variable Color Maps. *The American Statistician*, 34(2), pp.81–93.

Wang, H. & Sun, C.-T., 2011. Game Reward Systems: Gaming Experiences and Social Meanings. In *Proceedings of DiGRA 2011 Conference: Think Design Play*.

Wang, Z. & Ben-arie, J., 1996. Conveying Visual Information with Spatial Auditory Patterns. *IEEE Transactions on Speech and Audio Processing*, 4(6), pp.446–455.

Weilenmann, B., Seidl, I. & Schulz, T., 2017. The socio-economic determinants of urban sprawl between 1980 and 2010 in Switzerland. *Landscape and Urban Planning*, 157, pp.468–482.

Willemse, I., Waller, G. & Süss, D., 2016. *JAMES. Jugend, Aktivitäten, Medien. Erhebung Schweiz. Ergebnisbericht zur JAMES-Studie 2016.* Zürich: HAW Zürcher Hochschule für Angewandte Wissenschaften, Departement Angewandte Psychologie.

Winter, S. et al., 2011. Location-Based Mobile Games for Spatial Knowledge Acquisition. In K. Janowicz, M. Raubal, A. Krüger, C. Keßler, eds. *Cognitive Engineering for Mobile GIS.* Belfast: ME. Workshop at COSIT'11., pp.1–8.

Wright, E. et al., 2010. The Lasting Effects Of Social Media Trends On Advertising. *Journal of Business & Economics Research*, 14(3), pp.73–80.

Yanenko, O. & Schlieder, C., 2014. Game Principles for Enhancing the Quality of User - generated Data Collections. In *AGILE 2014 Workshop on Geogames and Geoplay.*

Yoshii, A. et al., 2011. IDetective: A location based game to persuade users unconsciously. In *Proceedings - 17th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications, RTCSA 2011*, pp.115–120.

Zandbergen, P.A., 2009. Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning. *Transactions in GIS*, 13(SUPPL. 1), pp.5–25.

Zhang, N. et al., 2015. GameOn: P2P Gaming On Public Transport. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. MobiSys '15. New York: ACM, pp.105–119.

11 Appendix

11.1 Nodes and Relationships in Neo4j

Table 5 - Table of node labels and node attributes stored in the implemented Neo4j graph database

Labels	Attributes
Resources	Name, name_DE, icon, iconColour
Team	Tid, name, img, title, txt
Role	roleType
Structure	Type, sid, name, name_DE, hp, img
Blueprint	Bid, type, name, minlvl, ilvl, name_DE, img
Inventory	Capacity
User	Uint, uid, primary, secondary, xp, registrationDateTime, password, isAccountNonExpired, isCredentialsNonExpired, isEnabled, isAccountNonLocked, usernameCanonical, screenname, isActive, emailCanonical, email, username, profileImage
Tile	tLat, tLng, bBox, collected, rid, tid

Table 6 - Table of nodes and relationships between the nodes stored in the implemented Neo4j database

Node A	Relationship	Node B	Attributes
User	HAS_RESOURCE	Resource	Amount
User	HAS_ROLE	Role	
User	IN_TEAM	Team	
User	HAS_INVENTORY	Inventory	
User	CAPTURED	Tile	landcover, captured, collected
User	LOST	Tile	Landcover, lost, captured, collected
Blueprint	COSTS	Resource	Amount
Blueprint	BUILDS	Structure	
Inventory	CONTAINS	Blueprint	Amount
Tile	HAS_STRUCTURE	Structure	hp

12 Personal Declaration

Personal declaration: I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

(Place, Date and Signature): ______________________________