The World Is Your Playground: A Bibliometric and Text Mining Analysis of Location-Based Game Research

Chien Lu[1] [0000−0002−3143−4202], Elina Koskinen[0000−0002−4082−6582], Dale Leorke[0000−0001−8741−1758], Timo Nummenmaa[0000−0002−9896−0338], and Jaakko Peltonen[0000−0003−3485−8585]

Tampere University, Tampere, Finland
{chien.lu, elina.m.kostinen, dale.leorke, timo.nummenmaa, jaakko.peltonen}@tuni.fi

Abstract. Location-based games have become mainstream and have been increasingly emphasized in the academic community. However, so far, to our knowledge, no bibliometric analysis of location-based games research literature has been undertaken. We carry out an analysis of 606 publications using bibliometric analysis and text mining. The results reveal prominent researchers, institutions, and countries, as well as the most common research topics and their prevalence over time. The results are useful both to understand the current state of location-based research and for designing future research.

Keywords: Location-based Games, Bibliometric analysis, Text Mining, Topic Modeling

1 Introduction

Location-based games emerged in the early 2000s as a result of artistic, academic and commercial experimentation with Global Positioning System (GPS) and location-based media (LBM, or locative media) technologies [31]. These types of games incorporate physical locations into the gaming experience, with early games employing technologies like text messages and primitive mobile Java [46, 50]. Other, mixed-reality location-based games blended physical and virtual environments to create large-scale, immersive experiences in public space [24,30]. More recently, commercial location-based game apps like *Pokémon GO*, *Harry Potter: Wizards Unite* and *Minecraft Earth* feature augmented reality built on smartphone cameras and sensors, and utilize their ecosystem features including cellular data and the app distribution model to reach potentially millions of simultaneous players [31, 42]. With *Pokémon GO*, location-based games reached a mainstream audience through combination of a casual play style with a recognisable franchise, although no games to date have matched its success [27].

Due to this significant growth, the importance of location-based games has been increasingly emphasized in the academic community. Location-based game

research and design has a strong connection to academia, dating back to its emergence. One major research effort occurred with the IPerG research project (2004-2008) [1, 56], even though consumer hardware was limited at the time. Other early research employed participant observation of specific experimental games, often by designers themselves or with their involvement [6,19,43]. Subsequent research focused on interviews with small samples of players to understand their everyday experiences playing location-based games [32, 33, 36]. Since the advent of smartphones with app stores, location-based games can reach millions of players simultaneously in countless locations around the world [31, p. 101]. This also affords researchers the opportunity to study player experiences on a larger scale. Quantitative and qualitative surveys that compare and contrast players' experiences in international and local contexts have become common – although a majority focus only on *Pokémon GO* [21, 28, 41, 49, 53]. In-depth interviews with players remain the most common method for media and cultural studies research on location-based games [24, 51, 55].

Despite this body of research dating back two decades to the emergence of location-based games, no bibliometric analysis of it has been undertaken. While some critical literature reviews do exist [31, 39, 47], these are either dated or limited to discourse analysis of a selection of publications. This article provides a bibliometric analysis that sheds light on trends in location-based game research. Our research questions are: RQ1 - How has the amount of research grown over time, and what are the countries and institutions contributing the research? RQ2 - What are the main themes of the research, and how do they vary over time?

Using bibliometric analysis and text mining techniques on 606 publications from 1995-2019 sourced from Scopus we provide insight into the frequency of publications on location-based games and the institutions and countries from which these publications most frequently appear, answering RQ1; and the most common keywords and phrases that appear in these publications, co-occurring groups of the keywords, and their prevalence over time, answering RQ2. This research will provide insights into present and previous research on location-based games, as well as providing a framework for further, qualitative research that explores more in depth the approaches used to study these games.

2 Background

Location-based games gained widespread attention after pinpoint GPS technology was made publicly available in 2000. Artists, academics, game designers and amateurs began to experiment with the potential for devices that could trace their users' location to bring play into public, primarily urban, spaces. Practices like geocaching and Alternate Reality Games like *The Beast* (2001) can be seen as precursors to location-based games, as can a whole range of practices and movements aimed at reclaiming urban space for playful behavior - including the Situationist International, New Games Movement, Gutai group, live-action role playing games and parkour [18, 39]. But location-based games more specifically emerged simultaneously as a form of artistic experimentation with "locative"

media technologies, exemplified by the games of artist group Blast Theory [17] and commercial games like *Botfighters* (2001) [31, p. 29]. As a result, location-based games blend a wide range of practices and influences, combining and reconfiguring new media art, locative media, mobile gaming and urban play.

One consequence of location-based gaming's heterogeneity is that no universal definition of location-based games exists. "Location-based game" most commonly refers to a game that uses mobile and/or networked technologies to incorporate their players' location into the game. One way to categorise these games is based on how they track and incorporate their players' location. "Urban games" use city streets as the game's playground or 'board', e.g. in *PacManhattan* (2004), in which players wear costumes of *Pac-Man* characters and chase each other through the streets of Manhattan communicating via "voice chat" on mobile phones. "Location-based mobile games" similarly use the city's streets and public spaces as the game space, but with the added element that player locations are tracked and incorporated into the game through GPS-enabled devices, e.g. Blast Theory's,*Can You See Me Now?* (2001) and *Botfighters*. Lastly, "hybrid reality games" include elements of the previous two types of games, combining them with a 3D visualisation of the city to produce games that blend online and offline play, such as in the treasure hunt game *Mogi, Item Hunt.* [47]

In addition to these genres of location-based games, many other terms have been employed by scholars to describe similar and related games over their near-20-year history. "Mixed-reality game" commonly refers to games that blend the virtual and physical environment, taking place in both simultaneously. "Locative game" is similar to "location-based game", but more specifically connects them to the "locative media movement" of artistic experimentation with location-ware technologies [17,18]. "Pervasive game" refers to any type of game or playful practices that extends the boundaries of play out into the physical environment, and location-based games are one type of pervasive game [39]. In turn, "augmented reality game" refers specifically to games that use augmented reality (AR) technology. Each of these terms refers to a specific, nuanced form of these types of games but they are often synonymous with, or strongly overlap, with location-based games. Other terms - like "mobile urban game", "massively multiplayer mobile game" and "hybrid reality game" - further complicate attempts to clearly delineate research on location-based games and related phenomena [31, p. 36-38].

This definitional elusiveness poses a clear dilemma for a bibliometric analysis of the literature on location-based games, since many publication may be relevant to scholarly work on this topic but employ a different term to describe them. Further, publications may use the term "location-based" and "game" but not in combination, meaning irrelevant combinations would be picked up on a broad search using these terms. As described below, we have used these two broad terms to select our sample for this initial, exploratory study, followed by a curation stage to exclude the irrelevant results. However, future research will be needed to cast a wider net that extracts publications using different terms, with curation needed again to exclude results that may not be relevant to the topic.

3 Method

We use bibliometric and text mining approaches. Bibliometric analysis uses quantitative and statistical techniques to analyze volume, growth and distribution of research literature. To better understand the content of the research works, we also employ a text mining technique called topic modeling. It enables researchers to perform a qualitative analysis in a computational manner [16, 40].

3.1 Data Collection

The Scopus database was chosen because it has indexed a wide variety of significant scientific literature and has been one of the most used sources for game studies [4]. We retrieved the data from the Scopus with the following query:

TITLE-ABS-KEY ("game*" AND "location based") AND (LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "ch"))

where "game*" is set to cover both game and games and the results are limited to conference papers ("cp"), journal articles ("ar") and book chapters ("ch"). The search was executed in November 2019 and 778 items were collected.

3.2 Data Curation

In an initial topic analysis of our sample, topics were skewed due to publications using "location-based" and "game" separately without referring to location-based games. These separate search terms yielded relevant works using variant terms for location-based games, for example, "location-based mobile game", or a "pervasive game" that is location-based, but also yielded irrelevant publications.

To exclude irrelevant publications without excluding relevant ones, a curation stage was done where all publications were inspected. The work was divided equally among 5 co-authors of the present manuscript: each inspected about 155 articles in a spreadsheet, labeling them relevant, irrelevant (i.e. to be excluded), or unsure based on the title, abstract, and keywords; the full article was checked if needed. Articles marked unsure were discussed among co-authors, who then double-checked their exclusions and listed reasons for exclusions. Each co-author then cross-checked 5-10 articles marked for exclusion by another co-author, to ensure agreement and consistent reasoning for exclusions across the full sample. Co-authors then discussed and resolved disagreements on exclusions or reasons. Excluded articles were coded based on the reasons and the consensus reached.

This process excluded 172 (22%) out of the original 778 publications. The most common reasons for exclusion were: 1. Articles applying game theory concepts to analyze a problem involving e.g. "location-based" sensors but without discussion of actual games or related phenomena. 2. Articles on privacy or security in location-based technologies or social networks, with "game" only as an example application or coincidental word. 3. Articles on location-based and location-aware technologies, with "game" only as an example application or in another context. 4. Articles that were not about a location-based game: these

included game projects and interactive art, some using AR and VR, which mentioned "location-based" in the abstract or full paper, but the game itself was not location-based. This broad term was sometimes hard to define; simply tracking participant motion through sensors, similar to Microsoft Kinect, was not deemed location-based, but if the project tracked people's movement through a physical location, including one room or building, we included it. If the distinction was not clear from the full paper, we opted to include it.

The least common exclusion reasons, with only a few examples in total, were: 5. Publications that were not research articles, for example a workshop proposal or news article. 6. Otherwise unrelated articles using "location-based" and "game" separately. 7. Articles that appeared in the search, but were not found during the coding process and seemed irrelevant based on the title.

3.3 Bibliometric Analysis

Bibliometric analysis has been widely used in different disciplines including game studies [8, 14, 35]. It has proven a useful technique to explore impact and popularity of researchers, institutions and also countries in specific research fields. We use the metadata and impact metrics (mainly citations) collected from Scopus to conduct a quantitative analysis of the publications. The analysis puts the emphasis on the temporal trend of publications, and on importance of researchers, keywords, institutions and countries in terms of publications and citations.

3.4 Text mining

We use topic modeling to explore contents of the collected publications. A topic model finds groups (distributions) of frequently co-occurring words as topics, so that content of documents is represented as mixtures of these topics; the resulting topics often have clear semantic meaning which we can then analyze.

A topic model represents the probability $p(w|d)$ of word w in document d as

$$p(w|d) = \sum_k p(k|d)p(w|k)$$

where $p(k|d)$ is the probability a word is generated from topic k out of K possible topics, and $p(w|k)$ is the probability topic k generates word w out of the vocabulary. This can be seen as a multi-step generative process: each document is first generated a distribution of topics and each topic is generated a distribution of words, words are then generated by picking a topic and picking a word from the topic. The model is fitted by maximizing the probability of the observed words in documents, with respect to the underlying probability distributions, by iterative training algorithms. The resulting probabilities $p(k|d)$ describe which topics are most prevalent in each document, and the $p(w|k)$ describe which words are most prevalent in each topic. Unlike e.g. principal component analysis, topic modeling takes into account that documents consist of discrete words from a vocabulary.

Several topic modeling methods have been proposed. In this work a recently developed topic model called Structural Topic Model (STM, [45]) is used to analyze the collected text. Compared to methods such as Latent Dirichlet Allocation (LDA, [7]) and Dynamic Topic Model (DTM, [7]), STM is a more advanced model that is able to take available document-level covariates into account when modeling the text. For each document d, STM models the topic probabilities $p(k|d)$ of all topics $k = 1, \ldots, K$ as a vector $\boldsymbol{\theta}_d = [p(1|d), \ldots, p(K|d)]^\top$ arising from a Logistic Normal distribution,

$$\boldsymbol{\theta}_d \sim LogisticNormal_{K-1}(\boldsymbol{\Gamma}^\top \mathbf{x}_d, \boldsymbol{\Sigma}) \tag{1}$$

so that the expected topic probabilities are a linear transformation of the covariate vector $\mathbf{x}_d$ of the document, followed by logistic (softmax) normalization. The weight matrix $\boldsymbol{\Gamma}$ models interaction of covariates and topics and the covariance matrix $\boldsymbol{\Sigma}$ models variation of topic probabilities around the expected values.

We use the STM model to evaluate the association between the text content of a publication and the citations it has received. In order not to overemphasize older content merely due to accumulation of citations over time, we adjusted the number of citations according to the year of publication, so that

$$adjusted\ citations = \frac{citations}{(2020 - publication\ year) + 1} \tag{2}$$

the adjusted citation count was specified as the covariate in the model.

For each document the abstract, title, and author keywords were concatenated and considered as the document content. 8 publications without abstract were ruled out. Format-related words in the abstract (e.g. purpose, methodology, etc.), copyright text, numbers, punctuation and stop words were removed; the term "location-based" was also considered as a stop word since it appears in almost every document so that it doesn't contain meaningful semantic information. The pruned text was further lemmatized before the model building.

We select the number of topics using the held-out likelihood criterion, where a proportion of the documents is left out and the remaining part is used as the training data; the trained model is then used to predict the content of the "held-out" documents and their probability is computed as the held-out likelihood. This held-out likelihood value shows the predictive ability of the model, and we choose the number of topics that yields the best held-out likelihood. We further trained models with the selected number of topics from 20 different initializations with the full data set, and the model with the best averaged semantic coherence [38] value is taken as the final model. In brief, the semantic coherence value measures how strongly the top words in each topic co-occur over documents, thus, it can be employed to evaluate the performance of topic models and to choose the best-performing model among several models.

After topic model training, extracted topics were labeled and analyzed. Labeling was done by a common practice similar to Thematic analysis [15] but without reading through all the collected text. The labels were decided with authors of this paper examining the semantics of the top words and the example documents (documents containing high prevalence of the topic) of each topic.

4 Results

We first display the bibliometric analysis results and then the topics extracted by text mining, their prevalence over time, and their citation impact.

4.1 Bibliometric Analysis Results

Number of publications. The distribution of publications of in terms of publication types can be found in Figure 1 (left). The majority type, Conference Paper, contains 421 publications, followed by Journal Article with 158 publications and Book Chapter with 27 publications. Figure 1 (right) shows the number of publications for each year and the porportion of each type of publication. A clearly rising trend can be observed which reached its peak in 2017 with 90 publications and a drop happened next year. This is likely because the game *Pokémon GO* was published in the summer of 2016. Note though that the smaller number of publications in 2019 can be influenced by the delay of publication and processing time of Scopus indexing. Note also that there are 2 publications indexed with a publication date in 2020 which is after the time of our data collection; they are indexed from early released conference proceedings and a book.

Type	Frequency
Conference Paper	421
Journal Article	158
Book Chapter	27

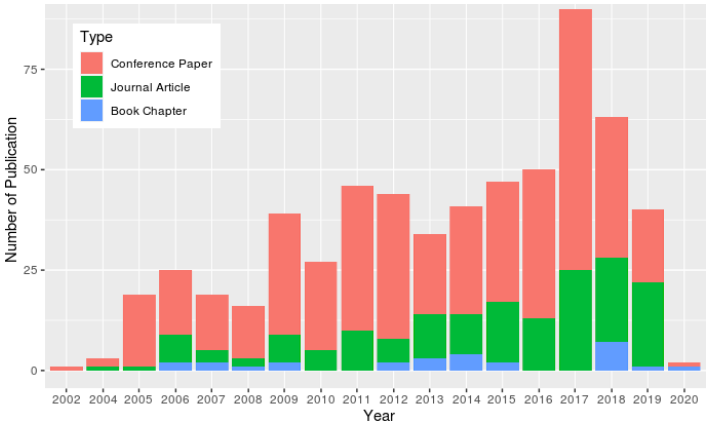

Fig. 1: Publication types. Left: overall counts, Right: counts per year/type.

Researchers. There are 1480 identified authors in the data. Table 1 (left) lists the 10 researchers with highest number of publications in which Christoph Schlieder (head of the Cultural Informatics Research Group at University of Bamberg, Germany) is the most productive researcher with 15 publications. Figure 2 (left) shows the percentage of number of publications for researchers, showing that the majority of researchers have few publications on location-based games so far: around 80% of researchers have only 1 publication and around 91% of researchers have less or equal to 2 publications. We further calculated the ratio

of overall citations to number of publications for researchers having at least 2 publications to identify researchers whose publications are highly cited on average. Table 1 (right) shows the 10 researchers with highest calculated ratio values. The researcher Sanne Floor Akkerman (currently Professor of Educational Sciences at Utrecht University) has the highest ratio, followed by Matt Adams (co-founder of Blast Theory) and Rob Anastasi (member of the Mixed Reality Laboratory at University of Nottingham at the time of the publications). Note that the researchers in Table 1 (right) have at most 3 publications on location-based games in our data but each publication has over 100 citations.

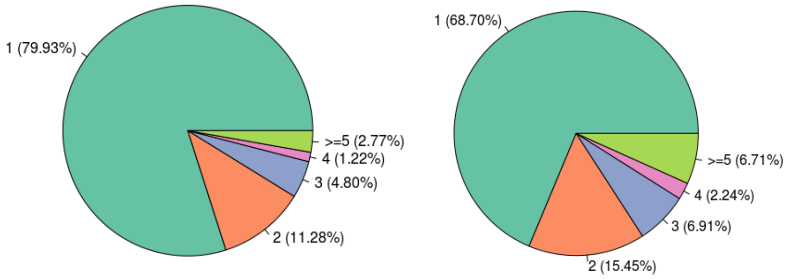

Fig. 2: **Left**: Number of Publications for Researchers. **Right**: Number of Publications for Research Institutions.

Researcher	Publications
Schlieder, Christoph	15
Coulton, Paul	14
Goh, Dion H.	13
Benford, Steve D.	12
Oppermann, Leif	11
Avouris, Nikolaos M.	11
Kiefer, Peter	11
Yiannoutsou, Nikoleta	11
Matyas, Sebastian	10
Sintoris, Christos	9

Researcher	Publications	Cites	Ratio
Akkerman, Sanne Floor	2	338	169.00
Adams, Matt	2	267	133.50
Anastasi, Rob	2	267	133.50
Ten Dam, Geert T.M.	2	258	129.0
Crabtree, Andy	2	247	123.50
Huizenga, Jantina C.	3	339	113.00
Admiraal, Wilfried F.	3	339	113.00
Paxton, Mark	2	214	107.00
Rowland, Duncan A.	3	302	100.67
Hampshire, Alastair	2	197	98.5

Table 1: Researchers. Left: Researchers with highest number of publications. Right: Researchers with highest ratio of citations to publications

Research Institutions. In total there are 492 institutions from 57 countries among the publications. Table 2 lists the 10 institutions with the most publications, showing that University of Nottingham holds the largest count of 28

publications. As shown in Figure 2 (right), most institutions have contributed few publications to location-based game research so far: around 69% of institutions have only 1 publication, around 84% of institutions have 2 or less.

Institution	Country	Publications
University of Nottingham	United Kingdom	28
University of Bamberg	Germany	18
Lancaster University	United Kingdom	15
University of Patras	Greece	13
Wee Kim Wee School of Communication and Information	Singapore	13
Aalborg University	Denmark	11
Simon Fraser University	Canada	11
University of Oulu	Finland	10
University of Tampere	Finland	10
University of Porto	Portugal	9

Table 2: Top Institutions Based on Number of Publications

Table 3 shows the 10 institutions with highest ratio of citations to publications, where Utrecht University has the highest value. Note that there are two institutions in the list (HP lab, Bristol and Nokia Corporation) that are not under universities but belong to technology companies. Figure 3 (left) shows the top 10 countries (based on where the institutions are located) according to the number of publications; the top two countries are United Kingdom having 133 publications and Germany having 106 publications. Figure 3 (right) provides a heat map presentation according to the number of publication of each country.

Country	Publications
United Kingdom	133
Germany	106
United States	83
Finland	65
Greece	43
Portugal	38
Taiwan	34
Australia	33
Brazil	24
Japan	24

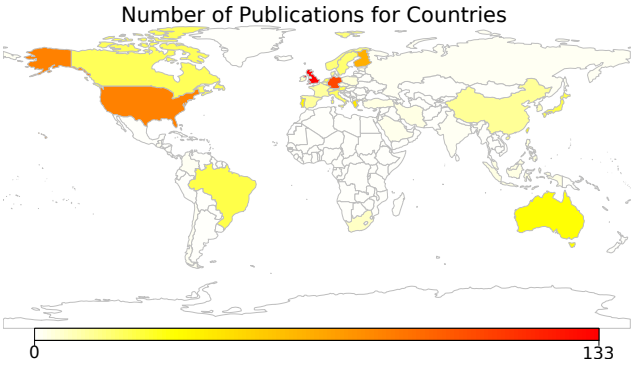

Fig. 3: Countries by number of publications. Left: top countries. Right: heat map.

Author specified keywords. The top 10 author-specified keywords that have occurred the most in the collected publications can be found in Figure 4 (left).

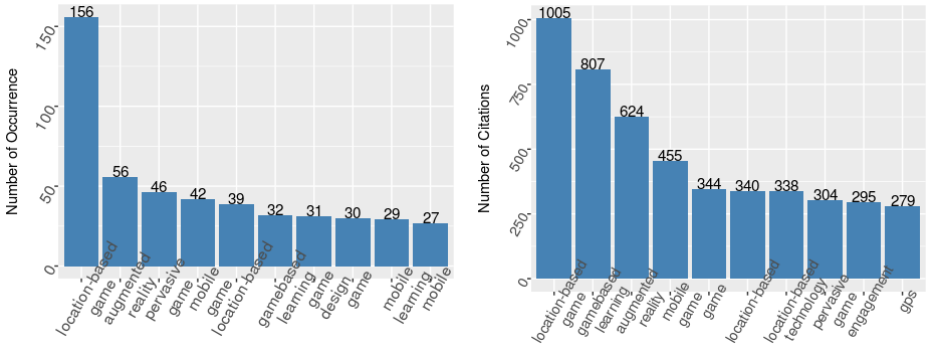

Fig. 4: **Left**: Top 10 Author Specified Keywords. **Right**: Top 10 Cited Author Specified Keywords.

The term "location-based game" has occurred 156 times, followed by game genre related terms such as "augmented reality", "pervasive game" and "mobile game" that have occurred 56, 46, and 42 times respectively. When the terms are ranked by number of citations to the publications, as shown in Figure 4 (right), the term "location-based game" still holds the first place (1005 citations), however, the term "game-based learning" has risen to the second place (807 citations), whereas "pervasive game" has dropped to 8th place.

Institution	**Country**	**Publications**	**Cites**	**Ratio**
Utrecht University	Netherlands	8	383	47.88
North Carolina State University	United States	5	155	31.00
University of Nottingham	United Kingdom	28	779	28.54
Hewlett Packard Laboratories, Bristol	United Kingdom	6	169	28.17
Nokia Corporation	Finland	7	166	23.71
Interactive Institute, Kista	Sweden	6	94	15.67
Nottingham Trent University	United Kingdom	6	92	15.33
University of Oulu	Finland	10	134	13.40
RMIT University	Australia	5	62	14.40
Tampere University of Technology	Finland	5	60	12.00

Table 3: Top Institutions Based on Ratio of Citations to Publications

4.2 Extracted Topics

The STM topic model with 10 topics was the best fitting one according to the held-out likelihood criterion and was selected for analysis. The top word list of each topic can be found in Table 4, and the topics in the table are ordered by

the proportion of all document content that they represent. The most prominent topic in our data, representing 16.71% of all document content, is **Design of Pervasive and Mixed Reality Games**. As the history of location-based games is connected to experimentation with design of pervasive and mobile games and other forms of play, often in an academic context using prototypes and controlled studies of games, it is to be expected many publications will reflect on design, implementation and player experiences of these games. For example, [57] discusses design of immersion in pervasive games whereas [13] uses sound design as an element of augmented reality, which is under the umbrella of mixed-reality.

The second most prominent topic is **Education**, corresponding to the second most commonly cited author keyword game-based learning (GBL). For example, [37] discusses how teachers can be involved in designing these type of games.

The third most common topic is **Mobile Application** which relates more specifically to the technologies employed by location-based games and services. For example, [2] introduces "mobile Real-time Kinematics", a refined solution related to GPS (Global Positioning System).

The topic **Urban Space** further focuses on where location-based games are played - primarily (although not always) the urban environment. For example, [23] has introduced a game called *Placemaking* to facilitate participatory urban planning. Often location-based games are framed by researchers and their designers as encouraging social interaction and bringing communities together.

Since location-based games are still relatively new to mainstream audiences, there is a lot of discussion on game production. **Design and Development** is related to, for example, the authoring tools that support making location-based mobile experiences [9]. **Player Studies** is about how users perceive and experience location-based games and media, and their impact on players. One example is [25] which probes the impact of the gamification design called "user score" of the mobile app *Foursquare*.

Pokémon GO is the only specific game within the topics. No other location-based mobile game has attracted such a large audience of players to date, so it makes sense that it would be prominent in academic scholarship even though it is relatively recent (released in 2016). Articles where this topic was prominent studied *Pokémon GO* from perspectives such as: how it affects players' place attachment [41], social behavior [26], and how players experience the game [5]. **Physical Activity** is an essential component of location-based games, since they usually encourage players to walk and explore the environment around them to accomplish their goal. For example, [52] investigates the exercise and health benefits of playing location-based games and [22] discusses how simulations can be used in exploring the available design options for location-based games.

The topic of **Gamification of Cultural Experiences** relates to how different cultural experiences can be gamified through location-based mobile technology. For example, [11, 12] enhance museum experiences with location-based games; [29] has developed an authoring platform in order to deliver unique experiences to cultural heritage visitors, and [48] discusses the gamification of tourism through geocaching. The topic **Spatial Data Processing** reflects

the technical aspect of how the information (e.g. player's position and movement, data curation) related to location-based games is dealt with. Some works are about platforms used to process this data, from GIS (Geographic Information System) software to crowdsourced mapping platforms like MapQuest [44] or OpenStreetMap [10]. Other works focused on the exploitation of geographic data within the game itself, such as "location" spoofing' in *Pokémon GO* [10].

Topic	**Pr(%)**	**Top Words**
Design of Pervasive and Mixed Reality Games	16.71	game, player, reality, mobile, pervasive, augment, design, play, experience, virtual
Education	13.14	learn, game, design, mobile, student, educational, use, education, technology, teacher
Mobile Application	13.05	mobile, application, user, system, use, game, service, information, device, paper
Urban Space	8.80	game, urban, city, gaming, mobile, technology, social, community, use, new
Design and Development	8.78	design, tool, author, mobile, support, present, compute, new, will, development
Player Studies	8.69	study, game, social, geocaching, result, share, user, perceive, group, effect
Pokémon GO	8.32	pokémon, place, mobile, player, play, medium, social, model, game, experience
Physical Activity	8.28	game, activity, design, player, physical, use, simulation, child, time, walk
Gamification of Cultural Experiences	7.62	cultural, experience, mobile, museum, tourism, design, technology, game, narrative, heritage
Spatial Data Processing	6.61	data, game, map, use, lbgs, information, spatial, position, movement, object

Table 4: Extracted topics. "Pr(%)" is the proportion of document content represented by a topic, and "Top Words" are its most common words.

4.3 Topic Prevalence over Time

Figure 5 shows the sum of prevalence of topics over time and Figure 6 shows the average prevalence of topics over time. In a topic model, each topic has a proportion (zero to 100%) of prevalence in each document, and the figures are yearly summaries of the prevalences. The sum of prevalence is the sum of a topic's prevalence over all documents of the year, and can represent the overall volume of content published about the topic, thus its change over time reveals the trend of absolute volume of discussion about the topic. On the other hand, the average prevalence of a topic is the average of document-wise prevalences over publications of the year, and reflects the relative popularity of topics among

publications. The average prevalence can be seen as an adjusted topic popularity of the year according to the total number of publications of the year.

The topic **Education** had been popular since 2014 as seen in its average prevalence, the sum of the prevalence had grown and reached a peak in 2017, although due to growing popularity of **Pokémon GO** and other topics, the average prevalence in 2017 is not as high as in some previous years. The average and sum of prevalence of **Education** has a clear decreasing trend since 2017.

Pokémon GO had a growing trend from 2016 and reached a peak in 2017 both in sum of prevalence and average prevalence, likely because *Pokémon GO* was launched in 2016. Even though it probably wasn't discussed in articles before that, other topic words of this category have come up already before 2016.

Design and Development remains present in all years by average prevalence, indicating design of location-based games is not settled and there is continuing need for new design research, in part this may be due to new technologies.

Likewise, **Physical Activity** also has stable average prevalence over time but with a slight increasing trend. The sum of prevalence reaches its peak on 2017 which is corresponding to the trend of total publication volume.

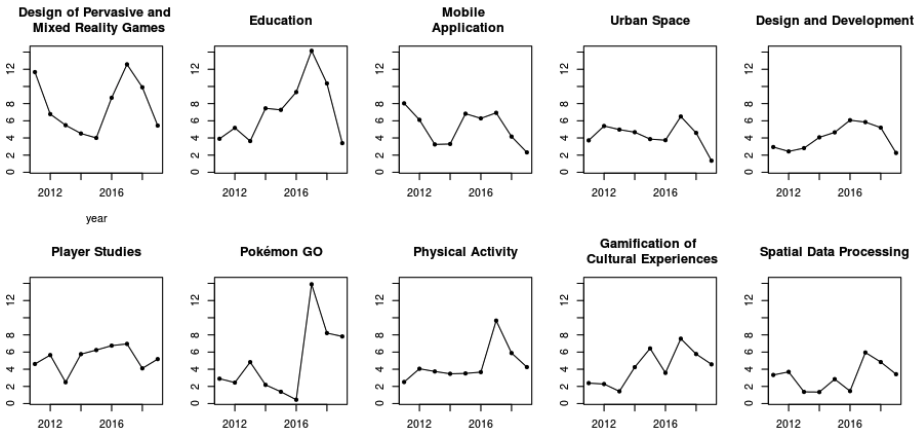

Fig. 5: Sum of Topic Prevalence Over Time

4.4 Relationship between Topics and Citations

Figure 7 shows how the adjusted citation count (according to Equation 2) interacts with topic prevalence, as extracted from the trained topic model. The STM topic model uses adjusted citation count as a covariate when modeling the text content of documents, and the figure depicts this interaction: the horizontal axis displays how much the topic prevalence of a publication increases or decreases when the publication has one more adjusted citation. Thus, compared

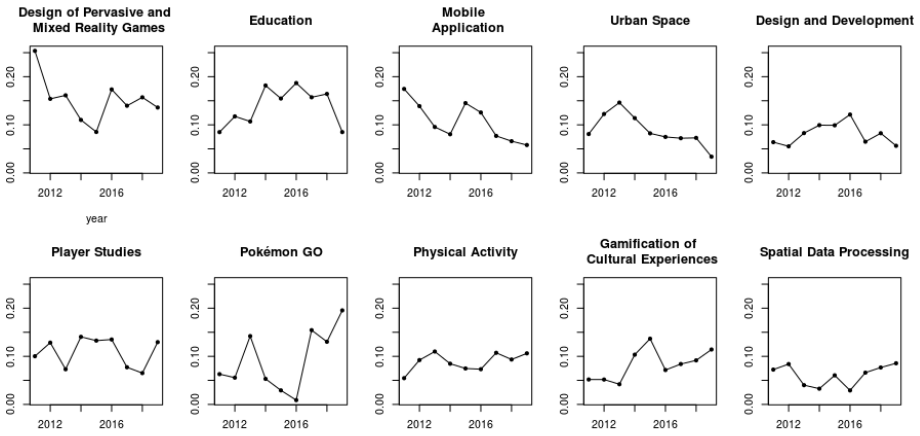

Fig. 6: Average Topic Prevalence Over Time

to an average publication, an increase of one adjusted citation tends to happen for publications having around 1% more content of **Pokémon GO**, 1.1% less content of **Mobile Application** and correspondingly for the other topics.

Topics **Urban Space**, **Education**, **Spatial Data Processing** and **Design of Pervasive and Mixed Reality Games** lean to higher adjusted citations but not **significantly**. The 95% confidence interval of topics **Gamification of Culturl Experiences**, **Design and Development**, **Physical Activity** and **Player Studies** covers the middle but lean to lower adjusted citations.

Education has a positive but not significant effect although the game-based learning (see Figure 4, right) is the second most cited keyword. The reason might be the accumulation of citations over time, as shown in Figure 5 there were a considerable amount of publications after 2014, and the citations of the keyword "game-based learning" have then accumulated. The equation 2 mitigated this influence, therefore, the effect on adjusted citations is not significant.

5 Discussion

The frequency, prevalence and geographical dispersion of publications in our sample shows the number of publications has been growing over time and it is not only a regional phenomenon. Publications peaked in 2017, although the drop in 2018 might be partly due to delayed indexing as our search was conducted before the end of 2019. The 2017 peak can be partially attributed to the surge of interest in these games following *Pokémon GO*'s release, but there is no clear correlation between this and prevalence of individual topics. Publications in our sample are from a variety of institutions (universities and companies) in different countries, although the U.K. and Germany have a higher proportion of publications than others (and this proportion would likely be higher if weighted by population).

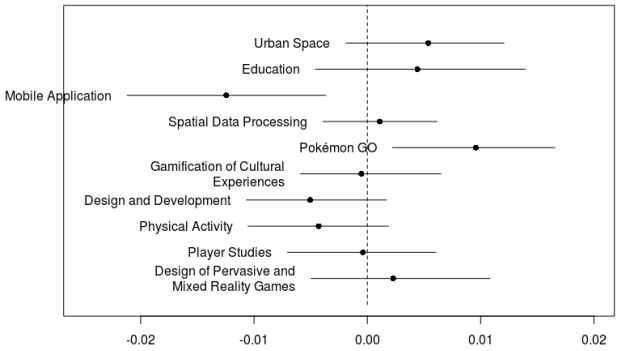

Fig. 7: Graphical Display of the Differences of Topical Prevalence on Adjusted Citations. For each topic the dot shows the mean influence of adding one adjusted citation, and the bars show a 95% confidence interval.

Most topics concern practical aspects of location-based games, such as their design; evaluation of players' experiences; the technologies and data they employ; and their application to real-world situations. This reflects that location-based games have been a highly experimental genre as mentioned in Section 2 [31, p. 29], many early, seminal location-based games, such as *PacManhattan*, *Can You See Me Now?* and *Insectopia*, were developed by or in collaboration with research centres: NYU's Big Urban Games program, Nottingham University's Mixed Reality Lab and the E.U.-funded IPerG project, respectively. Similarly, commercial games like *Botfighters* and *Mogi, Item Hunt* were the subject of numerous player studies and have long been noted for groundbreaking design that inspired future location-based games [37, 38, 60]. This long-standing association between location-based games, research and design and humanities and social sciences research is particularly evident in the position of **Design of Pervasive and Mixed-Reality Games**, **Design and Development**, **Urban Space**, **Mobile Application** and **Player Studies** among the extracted topics. "Pervasive game", a term developed and popularised in an academic context [46], also ranks highly in both top author-specified and cited keywords. Academic interest in the practical application of location-based games and location-based gamification to other fields is evidenced by the presence of **Education** and **Gamification of Cultural Experiences** in the extracted topics; and keywords like "game-based learning", "mobile learning" and "engagement" that appear in top author-specified and cited keywords. Nevertheless, as gamification has been applied to a wider domain (e.g. healthcare, marketing), we expect that a wider-ranged application can be further explored in the future.

Despite the history of location-based games dating back to the early 2000s, *Pokémon GO* is the only topic about a specific game. This may be unsurprising given its unprecedented popularity, but it does indicate a disproportionate number of publications about the game relative to its brief history. *Can You See Me Now?*, *Botfighters* and *Mogi, Item Hunt* were all released by 2003 and have con-

sistently been cited as canonical location-based games [31, 47, 50]. *Pokémon GO* is also not the first commercially successful location-based game on a global scale. It was preceded by *Parallel Kingdom* (2008), *Shadow Cities* (2011) and *Ingress* (2012) which, although far less successful, attracted media attention at their release [31, p. 106-112]. *Pokémon GO*'s presence in the top 10 topics indicates that even though these games had between 4-15 years longer to accumulate mentions in articles, the ubiquity of *Pokémon GO* since its release enabled it to surpass them. This underscores the impact of the game in industries and academia, as well as how quickly the research community has responded to the phenomenon of the game - something existing research has already highlighted [42].

Pokémon GO's prominence be due to interest from fields outside the arts, humanities and social sciences (especially game and media studies) and computer science, where most previous location-based game research was situated, such as interest from health and medicine fields [34]. Some publications do not discuss *Pokémon GO* in most of the full text but use it as an example location-based game in the abstract or introduction [3, 54]. The popularity of *Pokémon GO* in can enhance the visibility and opportunity of receiving citations.

Conversely, when comparing **Mobile Application** and **Design and Development** to other topics, we note their effect on citation is more negative. This can be attributed to more interest being on other aspects of games than design and development processes. Due to the limitations and constraints in creation of early location-based games and the rapidly changing technologies that support them, technical discussions in early papers are unlikely to remain relevant to the wider community. Also, many papers may discuss game development but refer to it as "game design", reducing prevalence of the "game development" topic. Intriguingly, the topic **Urban Space** has decreased in prevalence (both sum and average) over time, trending downwards since 2017. This coincides with the release of *Pokémon GO*, but we could speculate that as location-based games have become more popular and accessible through smartphone apps, and thus able to be played virtually anywhere with cellular data, discussions about them have become less tethered to "the urban". Although commercial location-based games still often privilege densely populated urban locations, they are less restricted to cities compared with pre-smartphone location-based games, which relied on bespoke technologies and were often limited to specific cities, countries or regions, rather than available to potentially billions of players globally [31, p. 102-106].

6 Conclusions, Limitations and Opportunities

In this research, bibliometric and text analysis were conducted. We found that "location-based games" as a term has been broadening in the research community and entangling with different notions from different fields. This yields a wide-ranged result of keywords and extracted topics. The analysis answered our research questions: for RQ1, the analysis revealed a rising trend of publications up to 2017, and differing contributions in terms of number of publications and citations across researchers, institutions and countries. For RQ2, the text analy-

sis extracted topics from the collected text content across the publications, and revealed clear themes, their temporal trends, as well as their interactions with citations, which were further analyzed and discussed. This research provides an overview of the current state of development and can help researchers gain a general understanding and identify potential future research directions.

We used the Scopus indexing system to gather the articles. We chose Scopus as it allows detailed Boolean queries and assures results represent peer-reviewed work and not for example technical reports. However, there might be some relevant research works not available through Scopus search and not analyzed in this article. This can be solved for instance by combining with other digital libraries such as Google Scholar to expand the result set; however, as this would greatly expand the workload at our data curation stage, we focus here on Scopus and consider an expanded search future work. It is possible to extend the corpus also by using additional search terms such as "mixed-reality game," "locative game, or "augmented-reality game" and merging these with our existing data. Besides, our sample focuses only on English-language publications. We acknowledge that there is an abundance of relevant research works in other languages. The inclusion of literature in other languages is considered a future direction.

The qualitative analysis of text content was assisted by a computational technique (topic modeling). The results were both interesting on their own and supported human interpretation and efficient reading of the documents. We expect such computational methods will become ever more crucial in similar studies.

Our results and discussion illuminated and revealed trends in location-based game research, focusing on quantitative findings. Further qualitative research, such as coding for common themes and terms, could be conducted to reveal other trends: such as the top-mentioned specific games, genres and types of games, technologies they employ and disciplines where they are discussed. An even richer discourse analysis could examine how the discourse around these games has evolved over time and the balance of positive versus negative framing about their impact. This type of analysis is beyond the scope of this paper, however, given the aforementioned resource constraints. Instead, we signal it here as an avenue for future research building on the findings of this paper.

ACKNOWLEDGMENTS

This work was supported by Academy of Finland decisions 312395, 313748 and 327352.

References

1. Iperg. http://web.archive.org/web/20141028144114/http://www.pervasivegaming.org/index.php, accessed: 2019-11-29
2. Alanen, K., Wirola, L., Kappi, J., Syrjarinne, J.: Inertial sensor enhanced mobile rtk solution using low-cost assisted gps receivers and internet-enabled cellular phones. In: 2006 IEEE/ION Position, Location, And Navigation Symposium. pp. 920–926. IEEE (2006)

3. Alavesa, P., Pakanen, M., Voroshilov, A., Kukka, H., Samodelkin, A., Ojala, T., Pouke, M.: City knights: Spatial realism and memorability of virtual game scenes in pervasive gameplay. In: 2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games). pp. 71–78. IEEE (2017)
4. Alha, K.: The imbalanced state of free-to-play game research: A literature review. In: DiGRA Conference (2019)
5. Alha, K., Koskinen, E., Paavilainen, J., Hamari, J.: Why do people play location-based augmented reality games: A study on pokémon go. Computers in Human Behavior **93**, 114–122 (2019)
6. Benford, S., Anastasi, R., Flintham, M., Greenhalgh, C., Tandavanitj, N., Adams, M., Row-Farr, J.: Coping with uncertainty in a location-based game. IEEE pervasive computing **2**(3), 34–41 (2003)
7. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. Journal of machine Learning research **3**(Jan), 993–1022 (2003)
8. Bragge, J., Storgårds, J.: Profiling academic research on digital games using text mining tools. In: DiGRA Conference (2007)
9. Brundell, P., Koleva, B., Wetzel, R.: Supporting the design of location-based experiences by creative individuals. In: 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP). pp. 112–116. IEEE (2016)
10. Celino, I.: Geospatial dataset curation through a location-based game. Semantic Web **6**(2), 121–130 (2015)
11. Cesário, V., Coelho, A., Nisi, V.: Enhancing museums' experiences through games and stories for young audiences. In: International Conference on Interactive Digital Storytelling. pp. 351–354. Springer (2017)
12. Cesário, V., Trindade, R., Olim, S., Nisi, V.: Memories of carvalhal's palace: Haunted encounters, a museum experience to engage teenagers. In: IFIP Conference on Human-Computer Interaction. pp. 554–557. Springer (2019)
13. Chatzidimitris, T., Gavalas, D., Michael, D.: Soundpacman: Audio augmented reality in location-based games. In: 2016 18th Mediterranean Electrotechnical Conference (MELECON). pp. 1–6. IEEE (2016)
14. Çiftci, S.: Trends of serious games research from 2007 to 2017: A bibliometric analysis. Journal of Education and Training Studies **6**(2), 18–27 (2018)
15. Cooper, H.E., Camic, P.M., Long, D.L., Panter, A., Rindskopf, D.E., Sher, K.J.: APA handbook of research methods in psychology, Vol 1: Foundations, planning, measures, and psychometrics. American Psychological Association (2012)
16. Evans, M.S.: A computational approach to qualitative analysis in large textual datasets. PloS one **9**(2), e87908 (2014)
17. Farman, J.: Mobile interface theory: Embodied space and locative media. Routledge (2013)
18. Flanagan, M.: Critical play: Radical game design. MIT press (2009)
19. Flintham, M., Benford, S., Anastasi, R., Hemmings, T., Crabtree, A., Greenhalgh, C., Tandavanitj, N., Adams, M., Row-Farr, J.: Where on-line meets on the streets: experiences with mobile mixed reality games. In: Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 569–576. ACM (2003)
20. Foursquare: Foursquare City Guide (2009)
21. Fragoso, S., Reis, B.M.S.: Ludic re-enchantment and the power of locative games: A case study of the game ingress. In: International Conference on Culture, Technology, and Communication. pp. 131–148. Springer (2016)
22. Heinz, T., Schlieder, C.: Addressing spatio-temporal geogame relocation issues using design evaluation heuristics and agent-based simulation. In: 2019 11th Inter-

national Conference on Virtual Worlds and Games for Serious Applications (VS-Games). pp. 1–4. IEEE (2019)
23. Innocent, T.: Play about place: Placemaking in location-based game design. In: Proceedings of the 4th Media Architecture Biennale Conference. pp. 137–143 (2018)
24. Innocent, T., Leorke, D.: Heightened intensity: Reflecting on player experiences in wayfinder live. Convergence **25**(1), 18–39 (2019)
25. Jin, L., Zhang, K., Lu, J., Lin, Y.R.: Towards understanding the gamification upon users' scores in a location-based social network. Multimedia Tools and Applications **75**(15), 8895–8919 (2016)
26. Kari, T., Arjoranta, J., Salo, M.: Behavior change types with pokémon go. In: Proceedings of the 12th International Conference on the Foundations of Digital Games. pp. 1–10 (2017)
27. Keogh, B.: Pokémon go, the novelty of nostalgia, and the ubiquity of the smartphone. Mobile Media & Communication **5**(1), 38–41 (2017)
28. Koskinen, E., Leorke, D., Alha, K., Paavilainen, J.: Player experiences in location-based games: Memorable moments with pokémon go. In: Augmented Reality Games I, pp. 95–116. Springer (2019)
29. Kotsopoulos, K.I., Chourdaki, P., Tsolis, D., Antoniadis, R., Pavlidis, G., Assimakopoulos, N.: An authoring platform for developing smart apps which elevate cultural heritage experiences: A system dynamics approach in gamification. Journal of Ambient Intelligence and Humanized Computing pp. 1–17 (2019)
30. Lantz, F.: Big urban game. Space Time Play pp. 390–391 (2007)
31. Leorke, D.: Location-based gaming: play in public space. Springer (2018)
32. Licoppe, C., Inada, Y.: Emergent uses of a multiplayer location-aware mobile game: The interactional consequences of mediated encounters. Mobilities **1**(1), 39–61 (2006)
33. Licoppe, C., Inada, Y.: Geolocalized technologies, location-aware communities, and personal territories: the mogi case. Journal of Urban technology **15**(3), 5–24 (2008)
34. Ma, B.D., Ng, S.L., Schwanen, T., Zacharias, J., Zhou, M., Kawachi, I., Sun, G.: Pokémon go and physical activity in asia: multilevel study. Journal of medical Internet research **20**(6), e217 (2018)
35. Martí-Parreño, J., Méndez-Ibáñez, E., Alonso-Arroyo, A.: The use of gamification in education: a bibliometric and text mining analysis. Journal of computer assisted learning **32**(6), 663–676 (2016)
36. McGonigal, J.E.: This might be a game: ubiquitous play and performance at the turn of the twenty-first century. Ph.D. thesis, University of California, Berkeley (2006)
37. Melero, J., Hernández-Leo, D., Blat, J.: Teachers can be involved in the design of location-based learning games. In: Proceedings of the 6th International Conference on Computer Supported Education. vol. 3, pp. 179–186 (2014)
38. Mimno, D., Wallach, H.M., Talley, E., Leenders, M., McCallum, A.: Optimizing semantic coherence in topic models. In: Proceedings of the conference on empirical methods in natural language processing. pp. 262–272. Association for Computational Linguistics (2011)
39. Montola, M., Stenros, J., Waern, A.: Pervasive games: theory and design. CRC Press (2009)
40. Nikolenko, S.I., Koltcov, S., Koltsova, O.: Topic modelling for qualitative studies. Journal of Information Science **43**(1), 88–102 (2017)

41. Oleksy, T., Wnuk, A.: Catch them all and increase your place attachment! the role of location-based augmented reality games in changing people-place relations. Computers in Human Behavior **76**, 3–8 (2017)
42. Paavilainen, J., Korhonen, H., Alha, K., Stenros, J., Koskinen, E., Mayra, F.: The pokémon go experience: A location-based augmented reality mobile game goes mainstream. In: Proceedings of the 2017 CHI conference on human factors in computing systems. pp. 2493–2498. ACM (2017)
43. Palmer, S., Popat, S.: Dancing in the streets-a design case study. Interactions (New York) **15**(3), 55–59 (2008)
44. Pulsifer, P.L., Caquard, S., Taylor, D.F.: Toward a new generation of community atlases—the cybercartographic atlas of antarctica. In: Multimedia Cartography, pp. 195–216. Springer (2007)
45. Roberts, M.E., Stewart, B.M., Airoldi, E.M.: A model of text for experimentation in the social sciences. Journal of the American Statistical Association **111**(515), 988–1003 (2016)
46. de Souza e Silva, A.: Alien revolt (2005-2007): A case study of the first location-based mobile game in brazil. IEEE Technology and Society Magazine **27**(1), 18–28 (2008)
47. de Souza e Silva, A., Hjorth, L.: Playful urban spaces: A historical approach to mobile games. Simulation & Gaming **40**(5), 602–625 (2009)
48. Skinner, H., Sarpong, D., White, G.R.: Meeting the needs of the millennials and generation z: gamification in tourism through geocaching. Journal of Tourism Futures (2018)
49. Sobel, K., Bhattacharya, A., Hiniker, A., Lee, J.H., Kientz, J.A., Yip, J.C.: It wasn't really about the pokémon: parents' perspectives on a location-based mobile game. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. pp. 1483–1496. ACM (2017)
50. Sotamaa, O.: All the world's a botfighter stage: Notes on location-based multi-user gaming. In: CGDC Conf. Citeseer (2002)
51. Spallazzo, D., Mariani, I.: Location-Based Mobile Games: Design Perspectives. Springer (2018)
52. Stanley, K.G., Livingston, I., Bandurka, A., Kapiszka, R., Mandryk, R.L.: Pinizoro: a gps-based exercise game for families. In: Proceedings of the International Academic Conference on the Future of Game Design and Technology. pp. 243–246 (2010)
53. Stokes, B., Dols, S., Hill, A.: Cities remix a playful platform: prominent experiments to embed pokémon go, from open streets to neighborhood libraries. american university, washington, dc (2018)
54. Tregel, T., Raymann, L., Göbel, S., Steinmetz, R.: Geodata classification for automatic content creation in location-based games. In: Joint International Conference on Serious Games. pp. 212–223. Springer (2017)
55. Vella, K., Johnson, D., Peever, N., Cheng, V.W.S., Davenport, T., Mitchell, J.: Motivating engagement with a wellbeing app using video games and gamification (2017)
56. Waern, A., Benford, S., Goetcherian, V., Åkesson, K.P., Söderlund, T., Björk, S., Mäyrä, F., Holopainen, J., Schäfer, L., Ghellal, S.: IPerG position paper. In: Workshop on Gaming Applications in Pervasive Computing Environments, Second International Conference on Pervasive Computing: Pervasive 2004 (Apr 2004)
57. Waern, A., Montola, M., Stenros, J.: The three-sixty illusion: designing for immersion in pervasive games. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp. 1549–1558 (2009)