

Network Connections and Neighbourhood Perception: Using Social Media Postings to Capture Attitudes among Twitter Users in Estonia

Daniel Baldwin Hess, *University at Buffalo, State University of New York*
Evan Iacobucci, *Rutgers, The State University of New Jersey*
Annika Väiko, *University of Tartu*

Abstract – The residential landscape of a city is key to its economic, social, and cultural functioning. Following the collapse of communist rule in the countries of Central and Eastern Europe (CEE) in the late 1980s and early 1990s, urban residential dynamics and household mobility have been critical to urban change under new economies and political systems. This article explores neighbourhood perception, which is a link in the chain to better explanation of socio-spatial processes (and their interruption by the socialist system). We use a novel data set – opinions expressed on one of social media (Twitter), and a novel empirical method – neural network analysis, to explore people's current attitudes and perceptions about the neighbourhoods and districts in Tartu, Estonia. The findings suggest that Twitter comments about urban neighbourhoods display attitudinal and perceptual commentary, which is subdued compared to other subjects. The socialist goal of homogeneity in neighbourhoods is not reflected in present day perspectives about urban neighbourhoods, 25 years after the disintegration of the USSR. Ambivalence about neighbourhoods persists, but this ambivalence may be in flux. Older, formerly neglected neighbourhoods, the subject of positive perception on social media, are currently experiencing increased investment, and the observed trends in our data support a narrative of neighbourhood transition.

Keywords – Neighbourhood satisfaction; perception; social media; Tartu, Estonia; Twitter; urban planning.

INTRODUCTION

The residential landscape of a city is key to its economic, social, and cultural functioning. Following the collapse of communist rule in the countries of Central and Eastern Europe (CEE) by the early 1990s, urban residential dynamics and household mobility have been critical to urban change under new economies and political systems. Residential dynamics have been in flux since the end of the socialist period, with certain neighbourhoods ascending (and other descending) in prestige [1], [2] based on complex factors. Centre stage in these processes, formerly state-owned apartments were acquired by private owners, new apartment buildings were erected to meet new types of demands, and new sprawling western-style suburbs (unavailable during Soviet times) appeared on formerly undeveloped peripheral urban space [3], [4]. These cityscape changes occurred as CEE countries developed commercial housing markets that were not part of the system for the decades under socialism.

Since the demise of the Soviet Union and the re-introduction of capitalist markets, housing construction is no longer under state

control [5] and, with increasing wealth, those who can afford to do so often seek alternatives to apartment blocks developed under state socialism [6], [7]. For the first time in more than a half century, *consumer choice* is a key dimension of the housing market (in the Soviet Union, housing was allocated and administratively assigned) and people may now freely express their preferences in housing type and housing location. Consequently, views and perceptions about apartment buildings, neighbourhoods, and districts play a role in shaping the attractiveness of places across metropolitan space [8].

This study examines how, 25 years after the end of socialism, residents in CEE urban space express views and perceptions about neighbourhoods. We explore the topic through people's perceptions of neighbourhoods across a modestly sized metropolitan area. Although there was little relation between household income and housing opportunities during socialist times [9], today's market economy demands such a relationship. We contribute in this article to the scholarship about neighbourhood perception, which is a dimension of a better understanding of socio-spatial processes (and their interruption by the socialist system). We use a novel data set – opinions expressed on social media, and a novel empirical method – neural network analysis, to explore people's current attitudes and perceptions about urban neighbourhoods and districts.

The remainder of this article is structured as follows. First, we provide background and context about housing and neighbourhood development in Estonia, especially the massive shifts brought on first by the installation of the Soviet Union and then again after its fall. Secondly, we describe our geographical area of study and explain our methods for collection and analysis of online data. Finally, we present lessons drawn (from a synthesis of two data sources) about people's perceptions of neighbourhoods.

I. RESEARCH STRATEGY

Housing mobility was low in the USSR; many people lived in state-subsidized housing for their entire work careers and lifetimes. The difficulty in changing place of residence – a cornerstone of the rigid housing system, suggests there was a lesser need only a generation ago for residents to possess information about neighbourhoods [10]. In the same way, people were ambivalent



about neighbourhoods and districts because there was little or no control in place of residence in state-provided housing. In addition, it was forbidden to print or distribute maps in Soviet cities for fear of maps falling into enemy hands [11]; consequently, residents of Soviet cities were usually familiar with their immediate neighbourhood and the city centre but had little knowledge about other neighbourhoods.

With censorship during Soviet times, there were few opportunities to publicly discuss opinions. Less than a generation ago, it was uncommon to share (or perhaps even possess) opinions and perspectives about neighbourhoods. Today, with more than two elapsed decades since the end of socialism, we also intend to assess the degree to which people voluntarily express neighbourhood attitudes and perceptions. During Soviet times, such activity was discouraged, if not impossible, due to censorship. We consequently seek to understand what effects the change of socio-political system, as well as the emergence of a digitally connected world, may have on expression of opinions.

to sentiments such as like or dislike; we are interested in broad perspectives and viewpoints. As an extension of this question, we will assess the degree to which neighbourhood perceptions are or are not homogeneous. Second, to what extent do people volunteer such opinions or perceptions? It is one thing for people to give information when asked; it is quite another for people to express their views unbidden. These questions directly relate to the established narrative of the transition from the Soviet neighbourhood development model to the post-socialist model in place today.

We situate our study in Tartu (population 93 700 as of 2016, Statistics Estonia), the second largest city in Estonia, possessing a diverse residential environment (see the city's 16 official neighbourhoods in Fig. 1). During World War II, roughly half of Tartu's then-extant housing stock was destroyed, inducing demand for subsequent replacement construction. During the Soviet period, pre-World War II residential districts, which contained a mixture of small wooden apartment houses and single family homes, were neglected and became the least desirable places [11]. Currently, however, the inner-city wooden residential areas have become coveted destinations, resulting in ongoing gentrification [14], [15]. Now, these gentrifying places have more prestige and residents enjoy historic apartment buildings, good access to the city centre, and convenient services.

We arrange Tartu's neighbourhoods into a typology of 6 neighbourhood types, based on similar characteristics. City centre, or *Kesklinn*, has maintained its historical milieu and added contemporary elements. Low-density urban form typifies Karlova, Tähtvere, and Tammelinn, allowing for greenery, and gardens are common. Vaksali was an elite district earlier in the 20th century due to its location near the railway, but is now mixed. Development in Maarjamõisa and Ropka focuses on the presences of university facilities.

The addition of new suburban neighbourhoods of single-family detached homes, such as Ihaste, represents one of the most important socio-spatial processes to occur since the end of socialism [9]. In Tartu, modern suburbanization first began in the 1990s, when Soviet-era summer cottages were transformed into year-round living spaces.

68

TABLE I
PERCEPTION OF MOST ATTRACTIVE AND LEAST ATTRACTIVE NEIGHBOURHOODS
[‘TARTU AND ITS RESIDENTS’ MUNICIPAL SURVEY, 2013]

Neighbourhood Type	Neighbourhood Name	Most Attractive Neighbourhood, %	Least Attractive Neighbourhood, %
City Centre	Kesklinn	13.3	5.0
Urban Garden District	Tammelinn	15.8	0.1
	Karlova	9.0	0.8
	Tähtvere	18.5	0.4
	Supilinn	3.5	9.4
Socialist Housing Estate	Annelinn	8.2	42.2
	Jaamamõisa	1.4	9.5
	Ülejõe	4.0	1.3
Mixed	Variku	1.0	1.1
	Ropka	2.5	16.9
	Veeriku	4.6	1.6
	Raadi-Kruusamäe	3.4	2.2
	Vaksali	1.2	3.2
Suburban Campus	Maarjamõisa	2.4	0.2
	Ränilinn	0.9	3.4
Newer Suburb	Ihaste	10.3	2.4
Total		100.0	100.0

II. NEIGHBOURHOOD PERCEPTION IN TARTU: EMPIRICAL RESULTS

We begin by summarizing ‘Tartu and its Residents’, a local panel survey (measuring satisfaction with various aspects of livability) conducted periodically by the Tartu government and employing a district-based random sample reflecting population composition [18]. Of 1500 respondents, 85 percent were ethnic Estonians, 12 percent Russians, and 3 percent other nationalities, mainly Finnish and Ukrainians. The ethnic Russian survey respondents resided mainly in pre-fabricated panel housing (Annelinn and Jaamamõisa). Survey results suggest that 34 percent of respondents have lived in Tartu since birth.

Given the opportunity to choose any style of living space, many respondents expressed preference for living in newer, more isolated neighbourhoods, especially single-family detached garden homes [18], [19]. Consistent with these preferences, the neighbourhoods that ranked highest in desirability were Tammelinn, Ihaste and Tähtvere (see Table I). Fewer people preferred the city centre. Neighbourhood prestige and ethnicity of neighbours were not significant factors in generating these preferences.

The least desired area was the panel housing district of Annelinn. In qualitative survey responses, respondents often cited a lack of privacy and high density of people, and some indicated excessive distance from the city centre. Others suggested that there were too many Russian-speaking residents. Previous studies have shown that people living in such places may have more positive perceptions of the neighbourhood than people living outside of them [20]. The least attractive districts also included the mixed area of Ropka, panel housing district Jaamamõisa, and Supilinn,

a gentrifying area in the inner city [14]. Recently, the prestige of Supilinn has improved due to accelerating gentrification [11].

Tartu residents clearly have positive opinions about garden districts with detached housing. The most favourable districts are located both in the inner city and in suburbs, where there are many owner-occupied single-family houses. The most desirable housing type since the end of the Soviet period is a detached private family house [21], where people feel a strong sense of ownership. The mixed areas and panel housing have equally low perceptions, which could be explained by a lack of information about these places or their connection with the legacy of the Soviet past [20].

III. SCHOLARLY RATIONALE

Use of social media data to augment urban scholarly inquiry has recently emerged in the field of geography. The advent of geotagging, for example, has allowed researchers to investigate a variety of questions using shared user location data. Such endeavours have ranged from using ‘check-in’ data to observe clusters of activity [22] to the identification of distinct groups of users and patterns of socio-spatial activity from Instagram photo locations [23]. Moving beyond geotags, researchers have developed innovative ways to make use of online media itself, such as using Instagram photos to create ‘digital signatures’ for cities [24].

Recently, researchers have argued that effective use of such data sources must go beyond simple, sometimes undertheorized quantitative analyses, emphasizing the role that theory and qualitative methods should play in examining social media data [25]. For instance, an examination of Twitter data surrounding Hurri-

cane Sandy uses small, comprehensible subsets of big data to perform mixed-methods analysis; applying appropriate conceptual and theoretical frameworks to help yield more nuanced and useful understanding of these data [25]. Continued development of such understanding is important: as Ash, Kitchin, and Leszczynski argue, the ‘digital’ realm is not separate from the rest of the world, but a part of it, and one that marks a change in the way that geographies are conceived, studied, and created [26].

A. Using Twitter to Explore Urban Issues

Researchers in the field of urban planning have recently begun to explore social media analysis as a way to observe sentiments, attitudes, and perceptions about urban life. One of the studies employs a technique to scrape Twitter data and then tabulate sentiments about various American public transit agencies [27], ultimately drawing comparisons between those results and opinions about peer services, such as airlines.

Twitter has also been considered a form of ‘micro-participation’ in urban planning processes [28]. Tweets, or ‘microblogs’, are used to gauge public engagement in the Strategic Transportation Mobility Plan in Austin, Texas; while certain obstacles exist between microblogging and influencing decision-making, micro-participation is found to be effective.

Another study, focused in Phoenix, AZ, uses data from open-ended Yelp reviews to assess travel modes that people used to access various non-work destinations [29]. Content analysis is employed to associate transportation with relevant descriptive terms, e.g. transport by light rail tends to be associated with ‘central’, ‘close’, and ‘easy’. Both within and outside of the field of urban planning, research using Twitter content seems to focus on ‘sentiment analysis’, often incorporating various types of machine learning approaches.

Researchers in South Korea using Twitter data surrounding the 2012 Korean Presidential election debates to identify ‘issue networks’ (composed of hubs and connections between them) mapped these networks to the positions of opposing candidates on key issues [30]. Another study investigating contentious issues in Korea found that certain key users such as media outlets often can dominate Twitter space relating to certain events, but that Tweets from these sources are less likely to be recirculated by other users [31].

Most recently, advances in the application of robust textual analysis techniques have allowed researchers to increase the visibility of social media content while minimizing observation bias. For example, an artificial neural network designed to read text [32] was used to uncover patterns and trends in large quantities of unstructured Tweets related to public transport. The technique generates valuable insight into the landscape of the Twitter universe vis-à-vis a targeted subject. In this study, we develop an extension of this technique to advance neighbourhood research.

IV. DATA AND METHODS

A. Twitter Use in Estonia

By 2013, there were more than 30 000 Twitter users in Estonia, generating 40 000 Tweets a day [33]. The vast majority of these are written in Estonian. While this is a significant number of Tweets, most of them is produced by Estonia’s most active 6 000 users.

While several social media platforms may contain information useful to research, Twitter is the most readily accessed and thus provides a significant volume of material from which to draw. Facebook posts, for example, are private and do not lend themselves to capture *en masse*. Tweets, by contrast, are publicly available and provide sufficient variety of content for targeted research. In addition to vastly better accessibility to Tweets, Twitter’s wide adoption makes it a preferable research tool compared with niche social media outlets such as Tumblr.

Why use Twitter to investigate neighbourhood perceptions when survey data already exist? First, Tweets are unprompted and unscripted, as the content is not sought by researchers. We gain insight as to what people think and feel about neighbourhoods without having to ask them, having them answer a survey, or otherwise potentially skewing responses via researcher interaction, survey question phrasing, or research rigidity. The understanding of Twitter commentary and conversation is still limited in an academic sense, but an ability to ‘listen in’ on spontaneous instances of expression is a valuable resource [32]. That people volunteer information in this way is also significant, as it implies that they have relevant thoughts and reason to share. Again, this is unlike a survey or interview, where a question may require someone to form an opinion, whether or not they have already, and even if they otherwise would never have shared.

B. Data Preparation

Data are acquired from Twitter’s website search engine, setting the search parameter to ‘Live’ to capture the largest number of Tweets possible spanning the longest period of time. We use Tartu’s neighbourhood names as search terms, constraining the results only when necessary. For example, when a neighbourhood name is orthographically identical to a word in another language, we constrain the results to only the Estonian language in order to ensure the results are in fact about the target neighbourhood, and not something else entirely.

We attempt to acquire the entire volume of Tweets available via Twitter’s search engine. The first mentions of the target neighbourhoods appear between late 2008 and early 2009, depending on the neighbourhood, with a gradual acceleration in subsequent volume. We collect all available data through August 2016. While it would have been desirable to obtain geographical information about the data, users almost never include it [34]. The result of this process is a text file for each target neighbourhood containing the content of each Tweet, as well as information about likes, retweets, usernames, and other supporting data.

The data are then ‘cleaned’ using regular expressions in a powerful text editor to eliminate all but the content of each Tweet.

TABLE II
DATA DESCRIPTION
[AUTHORS OF THE ARTICLE, COLUMNS (2) AND (3) DERIVE FROM ESTONIAN NATIONAL CENSUS, 2011]

Neighbourhood Type and Name (1)	Population (Share of Total) (2)	Ethnic Distribution, % EST / % RUS (3)	Number of Tweets (Share of Total) (4)	Tweets per Capita (5) = (4)/(2)
<i>City Centre</i>				
Kesklinn	7,768 (8.0 %)	90/10	511 (7.1 %)	0.07
<i>Urban Garden District</i>				
Tammelinn	5,929 (6.1 %)	96/4	206 (2.8 %)	0.03
Karlova	9,629 (9.9 %)	94/6	867 (12.1 %)	0.09
Tähtvere	3,434 (3.5 %)	96/4	481 (6.7 %)	0.14
Supilinn	1,925 (2.0 %)	95/5	828 (11.5 %)	0.43
<i>Socialist Housing Estate</i>				
Annelinn	27,047 (27.7 %)	67/33	1,140 (16.0 %)	0.04
Jaamamõisa	3,400 (3.5 %)	53/47	45 (1.0 %)	0.01
Ülejõe	7,677 (9.3 %)	87/13	503 (7.0 %)	0.07
<i>Mixed</i>				
Variku	1,773 (1.8 %)	94/6	321 (4.5 %)	0.18
Ropka	8,325 (8.5 %)	88/12	110 (1.5 %)	0.01
Veeriku	5,228 (5.4 %)	90/10	380 (5.3 %)	0.07
Raadi-Kruusamäe	4,578 (4.7 %)	83/17	14 (0.0 %)	0.00
Vaksali	3,552 (3.6 %)	89/11	183 (2.6 %)	0.05
<i>Suburban Campus</i>				
Maarjamõisa	1,455 (1.5 %)	88/9	7 (0.0 %)	0.00
<i>Newer Suburb</i>				
Ihaste	2,690 (2.8 %)	86/14	1,382 (19.3 %)	0.51
Total	95,627 (100.0 %)	82/18	7,143 (100.0 %)	0.07

Note: Column (3) provides the share of Estonian residents (EST) and share of Russian-speaking residents (RUS) Ülejõe is not included in the final analysis diagrams due to an unresolvable confound in the data.

This procedure is necessary in order to focus the analysis software on the relevant information contained in the Tweets, rather than related data (i.e., the number of favourites, the number of retweets) that are not relevant to this study. As a final pre-analysis step, data are subdivided into sets of roughly equal numbers of Tweets, to ensure comparability of results between analyses.

C. Empirical Method

Our data, summarized in Table II, include 7 143 tweets. Neighbourhoods with fewer than 400 Tweets (Column (4)) are excluded from further analysis, as this volume is too low to produce significant results. We proceed with 6 neighbourhoods for the remaining analysis, omitting 10 neighbourhoods with too few points (while Ülejõe has more than 400 Tweets, it is not included in the remaining analysis due to a known orthographic data confound).

Wölpak, an artificial neural network-based software package for analysing text, is next employed to identify patterns in each dataset. Wölpak is an iteration of Catpac, a program that

reads unstructured textual data [35] and is useful for analysing large Twitter datasets [32]. While lacking some capabilities of Catpac, particularly in scalability to larger datasets, Wölpak is ideally situated for this project since its text interpreter can read Unicode text. This is essential for allowing Wölpak to process Tweets written in the Estonian language.

Artificial neural networks are computational devices designed to learn in a way that simulates, at least in analogous fashion, the way that human beings learn certain things, particularly patterns [36], [37]. These networks 'learn' patterns by adjusting the connections between 'neurons' to represent them. The way in which Wölpak adjusts its connections, according to a Hebb rule, requires no supervision because 'without external feedback and with only input data, the net structures itself to represent whatever systematicity it can find in the input' [36]. In other words, without any advance input (e.g. coding) from researchers, Wölpak 'learns' patterns that exist in the text it processes.

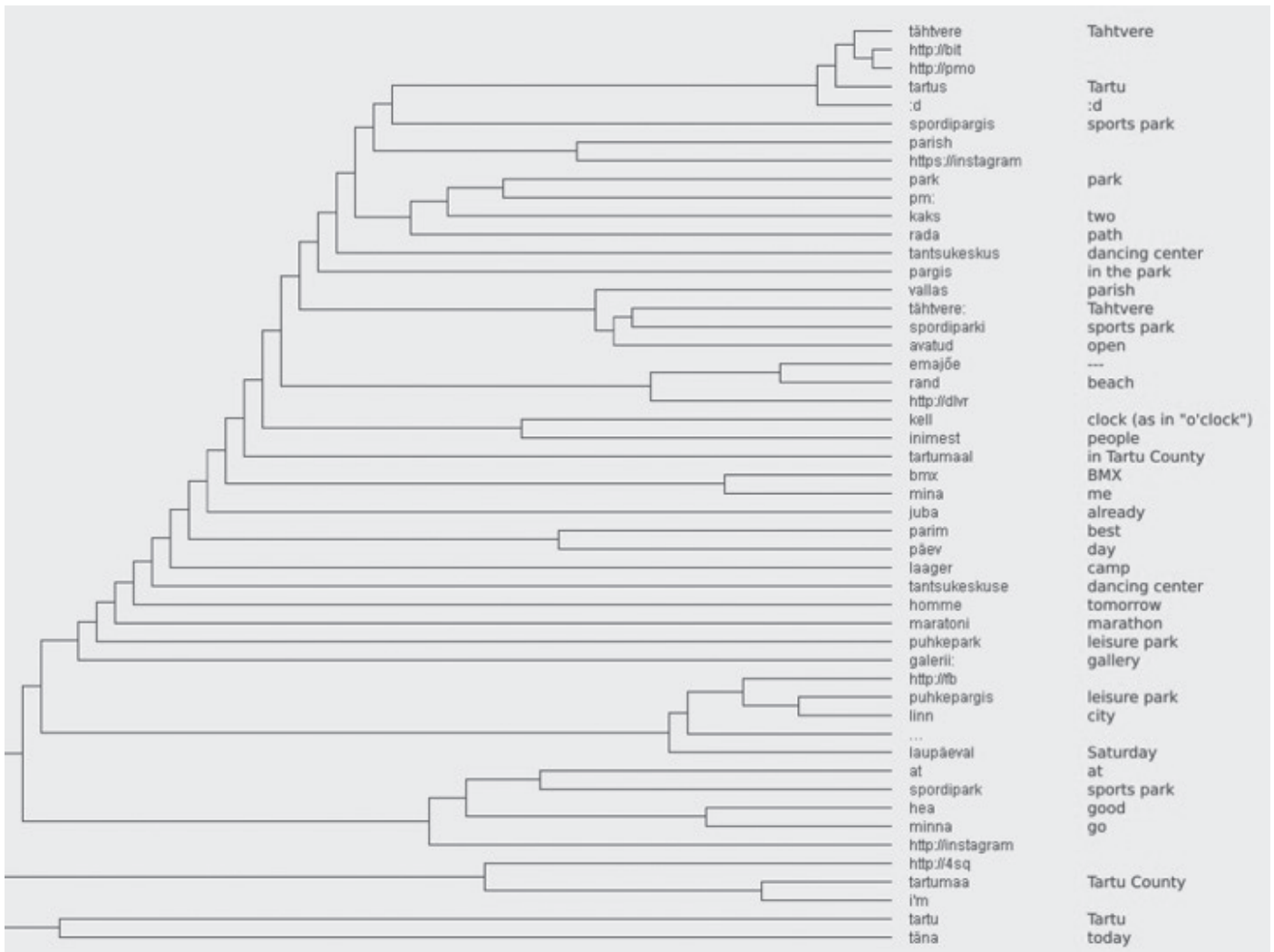


Fig. 2. Cluster analysis diagram for Tähtvere [Figure: D. B. Hess, E. Iacobucci, A. Väiko].

Wölfpak creates outputs demonstrating relevant concepts drawn from the target text in several forms, of which we use hierarchical clusters represented by dendograms. This approach has the advantage of not only extracting the most important components of given inputs but also of demonstrating how concepts are related to one another and how strongly they are interrelated.

Instead of pre-selecting certain themes, e.g. concepts related to positive or negative sentiments, and devising a way to quantify them, Wölfpak processes a data set and identifies the concepts that are most prevalent, whatever they may be. This is accomplished without requiring the coding of concepts, a pre-existing lexicon, or any such instruction as to what the software should look for in a data set, unlike previously employed strategies for analysing such data [27], [38]. This being the case, Wölfpak is always prepared to read text in any form, in any language, and will isolate text patterns [39]. Resulting clusters serve as the basis for interpreting Wölfpak's outputs.

We use the resultant clusters to seek out Tweets in the dataset related to our particular research inquiry. The entire procedure was carried out on all content in its original language (Estonian), and translated into English to present results. The vast majority of Tweets were written in Estonian, with other languages (such as English and Russian) appearing infrequently. Translation was performed by a native Estonian speaker. The end-result of these combined processes is a separation of useful data from noise, so that we may find meaning about attitudes and perceptions that exist within the Twitter universe.

V. FINDINGS

Emergent themes in the data are observed by inspecting Wölfpak's output clusters. Figure 2. presents an example of cluster analysis output for the Tähtvere district. Note that the collection of terms – *spordiparki*, *avatud*, Tähtvere, and *vallas* (translated: sports park, open, Tähtvere, and parish, respectively), form

TABLE III
ANALYSIS OF TWITTER CONTENT [AUTHORS OF THE ARTICLE]

Neighbourhood Type and Name	Search Constraints	Themes
<i>City Centre</i>		
Kesklinn	AND 'Tartu'	Apartments for rent Buildings/architecture Development Events (e.g. festivals) University
<i>Urban Garden District</i>		
Karlova	lang: et	Apartments for rent/sale Beautiful Boat harbour News Karlova Theatre School (Tartu Karlova Kool) 'Neighbourhood' 'District'
Tähtvere	NA	Beach Dance centre Rural News (fire) Sports park
Supilinn	NA	Estonia Green Movement (<i>Eesti Roheline Liikumine</i>) Mika Keränen (children's author) Supilinna Saltaselts (film title)
<i>Socialist Housing Estate</i>		
Annelinn	AND 'Tartu'	Apartment living Buses Emajõgi River News events Residential activity Safety at night Sports matches Stores/restaurants
<i>Newer Suburb</i>		
Ihaste	lang: et	Building of new bridge Bus service News Traffic

a cluster. The proximity of the terms and their interlinking suggests that they are meaningful in combination.

In this figure, the words at right are the terms that occurred most frequently in the Twitter data from Tähtvere. Words appear vertically close to one another either because they occur frequently together, or because they occur frequently in conjunction with a common term or terms. In other words, Wölfpak's algorithm forms connections between concepts that are directly related, but also when they are related through intermediate concepts.

Terms that occur most frequently *and* are tightly connected, have short lines that quickly connect to others to form evident clusters. Since the length is a product of both the connectedness of terms and their frequency of use, longer lines typically mean that terms occurred relatively infrequently compared to shorter lines, and that their links are less meaningful.

For example, the shortest lines connect two web address prefixes, which occur frequently and indicate that users often employ

such links to share pictures, etc. relating to Tähtvere. By contrast, the longest line, which connects the final term with the final bracket, links all of the terms found in the analysis, indicating nothing other than that they are the top fifty terms in the dataset.

Neighbourhoods and datasets display various patterns of clustering and strength of clusters, as expected. Strong and tight clusters are typically indicative of corresponding strong patterns within datasets, while loose, disparate clustering suggests more conceptual entropy. Most results from our datasets fall somewhere in between these two extremes, with some tendency toward looseness and disparity. This pattern makes sense given the nature of Twitter, especially such localized Twitter data, as many different subjects are expressed, unified only by neighbourhood name. Much of the data is noisy and related only by a single common term, but within that landscape there is discussion of specific subjects in which there is much overlapping terminology.

Additionally, by using the contents of clusters to perform targeted searches of our datasets, we compile collections of Tweets relevant to our original research inquiry, especially entries concerning opinions about neighbourhoods. By examining the qualitative nature of these data, we strive not only to be able to identify attitudes and perceptions expressed via Twitter about the neighbourhoods in question, but also to compare these instances of self-expression with established data from existing survey work. The results of this process are mixed.

The Estonian language has 14 productive ‘cases’, allowing a concept to belong to more than one cluster, unlike in English, which has only two cases (cases are variants of certain words that apply in different grammatical situations). For example, *spordipargis*, *spordiparki*, and *spordipark* (variants of sports park) appear in different clusters.

Each dataset gathered for this project is generated using the Tartu neighbourhood name as a search term. Wölpak analysis provides an accurate summary of *exactly* what is contained in each dataset. The most prominent emergent themes for the analysed neighbourhoods are summarized in Table III.

Several findings emerge from an investigation of key themes. First, local attractions and points of interest figure large in the Twitter discussion of Tartu’s neighbourhoods. People often mention places or events that they have visited, and organizations and news outlets frequently mention notable events or incidents, such as a festival or a major fire.

Examining the themes in Table III, it is clear that explicit expressions of opinions about the neighbourhoods do not dominate the Twitter landscape surrounding Tartu’s neighbourhoods. Instead, neighbourhood points of interest, general descriptive characteristics, news items, and neighbourhood events, particularly those that occur regularly, appear most frequently. There are a few possible reasons why this is the case.

First, events such as sports matches, which occur regularly over the course of a season, are repeatedly advertised and/or commented upon. This pattern was observed especially in Annelinn, where a sports facility is located. Secondly, incidental events, like a fire or an accident, appear to prompt responses from people, as well as to draw coverage from media outlets, who Tweet links to news stories. This phenomenon was observed about a fire in Tähtvere and about the construction and opening of a new bridge in Ihaste.

Finally, general points of interest and descriptions figure large in the comments surrounding various neighbourhoods. Often, these appear to be generated by people sharing comments about something they see or do, like visiting a beach. This pattern is corroborated by the frequent appearance of web addresses linking to Instagram, Facebook, and other ways of sharing media, such as Tweets containing photos. These patterns often, but not always, relate to leisure time activity, which is presumably generated by people who do not necessarily live in an area about which they are tweeting.

Conspicuous absences are noted in our cluster analysis. Expression of opinions, attitudes, and perceptions about Tartu’s districts does not dominate discussion. Events and generalities seem to overshadow small-scale expressions of opinion on Twitter,

at least in the samples we acquired. This, however, does not suggest that such expressions of opinion do not exist in the datasets, and through targeted searching of the material, we are able to uncover various instances of people commenting about Tartu neighbourhoods.

A. In Their Own Words

Although Tweets describing opinions about where people live do not make up a large share of Twitter activity, we detected many instances of such expressions, especially in the highly populated Soviet-era housing estate in Annelinn, the recently gentrifying area of Karlova, and the suburban district of Ihaste. These Tweets provide micro-level insight as to how people view places.

1. Annelinn, a dense development consisting of Soviet-era prefabricated panel apartment blocks, has a high proportion of ethnic Russian residents and a reputation as being less desirable than other areas in Tartu.

- Life in a panel housing estate in Annelinn: I don’t know my neighbours but I can hear them every time they go to the bathroom or make other noises.
elu annelinna paneelmajas ehk naabreid ei tunne aga kuulen neid iga kord kui vetsus käivad või häält teevad #район
- When I look at old pictures of Tartu I see that a lot has changed but Annelinn has stayed the same, lol.
vaatan vanu tartu pilte ja päris palju on muutunud. annelinn on ikka sama lol
- Annelinn is full of gunshots and mafia families.
Annelinn on täis püssipauke ja maffia perekondi #elu #tartu
- It’s so beautiful to look out of the kitchen window in the evenings. #annelinn
Küll on ilus õhtuti köögi aknast välja vaadata! #Tartu #annelinn @â€¹
- The street art in Annelinn. So pretty.
Annelinna streetart. Nii kaunis <3

A resident suggests that his/her neighbours are audible in the panel housing apartments, but that they do not know their neighbours, and that this is typical. Some have concerns about safety and crime, and others express enjoyment of underappreciated elements of neighbourhood beauty. Interestingly, one person suggests that the district is stuck in the past. Each of these to some degree matches established neighbourhood notions. Apartment buildings in Soviet-era housing estates were often made of cheap materials, rapidly constructed, and are densely populated, which is reflected in some of the comments.

2. Karlova is an older neighbourhood, consisting mostly of wooden apartment buildings and houses constructed before

World War II. Like Supilinn, it has shown the signs of gentrification in recent years.

- It's nice to be back in the hood, Karlova 4life.
ütilahe on tagasi hoodi elama tulla. Karlova 4 lyf.
- I just met a fox on Kalevi Street in Karlova.
Kohtusin just rebasega keset Kalevi tänavat #karlova #tartu
- Karlova is an interesting place in which to live: in the evening you hear gunshots and in the morning bikers start racing past your windows.
Karlova on ikka huvitav paik elamiseks - õhtul kuuled aknast püstolilaske, hommikul hakkavad ratturid maja eest läbi kihutama #karlova
- In love the streets of Karlova.
armunud karlova tänavatesse
- Karlova in the dark is a very creepy place :(
pime karlova on nii kuradi creepy koht :(
- Spoke to a scientist who studies air pollution effects on public health and Karlova is one of the most dangerous areas in Tartu.
Vestlusel õhu mõju inimeste tervisele uuriva teadlasega ilmes üllatav tõsiasi: õhukvaliteedilt on Tartu üks ohtlikumaid piirkondi Karlova.

One person comments that there is crime at night and bicycles during the day. Another notes creepiness after dark. Two of the comments express affection for Karlova.

3. Ihaste, a newer and growing area of Tartu, consists primarily of suburban-style single-family homes and has a reputation for safety and family-friendliness.

- I do not understand Ihaste, I get lost 10 times before I can get out of here
Mis asi see Ihaste on, suuda 10 korda ära eksida, enne, kui siit välja saan
- 2.40am... I look out of the window and see a random giant rabbit running past my house ... #Ihaste
2.40 am... vaatan mina aknast välja ja mida ma näen? mingi random hiiglaslik jänes jookseb mu majast mööda..... 0.o #Ihaste
- Now I'm at home, saw some paranormal things in the dark woods of Ihaste, creepy : D
kodus nüüd :) ihaste pimedates metsades nägime suht palju paranormaalseid asju, creepy :D
- I look at the 300,000€ houses they sell in Ihaste and cry.
vaatan ihaste 300 000 euriseid maju ja nutan

Tweets about Ihaste help to illustrate the differences that set it apart from other Tartu neighbourhoods. Curvilinear suburban streets are more difficult to navigate than older urban grids. Homes are situated in rural settings as opposed to more urbanized Tartu. The lifestyle that Ihaste affords does not come cheap. As one of the few neighbourhoods that newly developed in Estonia's post-socialist economy (following its humble beginning as a *dacha* neighbourhood), pronounced differences exist between Ihaste and older Tartu places.

B. Comparison with Survey Data

Our findings are largely compatible with the data presented in 'Tartu and its Residents', but provide additional insights not evident in the survey alone. The trends we observed are commensurate with the perceptions of most and least attractive neighbourhoods according to the survey. Commentary we observed about Karlova is indicative of a hip if rough-around-the-edges neighbourhood. Annelinn is mixed, as expected. We do not detect overwhelming negativity, however, that may have been predicted by the high rating of Annelinn as 'Least Attractive' in the survey (see Table I). Mention of activities associated with Tähtvere also support its perception as an attractive place, particularly due to its recreational facilities (a beach and fitness park). Ihaste and Supilinn both received positive attention. While Ihaste also scored high in the survey, favourable mention of Supilinn is a departure from the earlier survey results.

While the survey provides attitudinal data, it does so in the context of people being asked for this information. Were respondents not surveyed, it is possible that they would never have shared, or perhaps even explicitly formed, opinions about the various areas referenced in the questions. In contrast, while perhaps lacking the comprehensiveness of the survey, the information we find via Twitter is voluntarily offered. We contend that the voluntary nature of this information may suggest genuine sentiment behind the observed comments.

C. Limitations and Social Media Effects

Chief among our research limitations is an undeveloped understanding of the population (or populations) using Twitter. The research suggests that younger demographics are overrepresented on Twitter [34]. Overrepresentation of youth is one part of a larger difficulty in ascertaining the degree to which people who use Twitter can or cannot be generalized to wider populations. We observed signs of this trend in some of our data, particularly in comments that appeared associated with a local high school.

Additionally, the majority of our analysis involves translating the material originally written in Estonian into English. Elements of nuance and meaning may be lost in this translation. Also, the nature of Twitter removes context behind comments. We use Tweets for a purpose other than that for which they were intended, and in which a 140-character statement is the greatest achievable granularity. Therefore, we do not assume the same stringency as we might for a survey in which the import of respondents' answers could be imparted to them.

Further, posting a Tweet is not necessarily linked to specific actions or choices. For example, it is not possible to determine in most cases whether people decided to visit, shop, or seek housing in a neighbourhood based on an opinion expressed on Twitter. The expressed neighbourhood perceptions do not necessarily translate to housing preferences.

While we focus on defined geographic space, there are a few things to be borne in mind about the results. First, as geolocation data for Tweets either did not exist or was not available, we are unable to determine whether Tweets were made inside each neighbourhood, or simply about it. Similarly, Tweets that mention a neighbourhood may not necessarily be about that neighbourhood but rather events occurring there. Should significant geo-locational data become available in the future, spatial analyses are to be linked to our research process.

We observe a significant degree of conceptual disparity in our data, despite being unified around common neighbourhood names. Within the pool of content, the proportion of data relevant to our specific research aim is lower than would be ideal. This is a challenge that will continue to be a part of research using social media (or any large sets of unstructured online data) used for analysis outside of its original intent.

Finally, while the neural network analysis functioned as intended, a relatively low volume of data available from Tartu, combined with a high noise-to-content ratio, presented challenges for finding specific neighbourhood perceptions using only cluster analysis. In other words, while we were able to discover much about where people share content on social media, what landmarks people discuss in different neighbourhoods, and events that take place in neighbourhoods, neighbourhood perceptions related specifically to housing and quality of life were not as clear or rich as anticipated. We attempted to mitigate this difficulty through rigorous manual analysis, with a significant degree of success.

CONCLUSION

The research area that relates social media behaviour with urban planning and policy is nascent. We offer new insights into how people and organizations communicate (in Estonian) about neighbourhoods on Twitter and advance the developing field of research relying on social media data.

Our findings provide evidence to address our research questions. First, it is apparent in our datasets that what stands out in discussion about neighbourhoods are points of interest and events. This is informative in multiple ways. The Soviet goal of homogeneity in neighbourhoods is not reflected in present-day conditions, despite approximately half of the residents of a typical post-socialist city living in pre-fabricated buildings inherited from the Soviet years. Each neighbourhood that we observed displays a distinct character, albeit to a varying degree. It is clear in our analysis that while attitudinal and perceptual commentary is present in the landscape of Twitter surrounding Tartu, it is subdued in contrast with other subjects. We cannot say for certain that this finding points to a lack of Estonians wanting to share

their views, but it does not suggest an overwhelming desire to do so, at least using Twitter.

Combined with existing evidence of low residential mobility in post-socialist cities our data seem to suggest, or at least not refute, the idea that Soviet-era ambivalence about neighbourhoods to some degree still persists, at least as a determinant in housing choice. Findings also suggest, however, that this state of affairs is likely in flux. Neighbourhoods like Ihaste, Karlova, and Supilinn are currently experiencing new or renewed investment and interest related to positive or improved neighbourhood perceptions. Trends in our data support this narrative of transition.

While the Tartu municipal survey shows that people sometimes *do have* attitudes and perceptions about where they live or may like to live, there is apparently little motivation for people to volunteer these views on Twitter. This lack of discussion may speak to the priority these opinions occupy in the consciousness of Tartu residents who Tweet. To succinctly relate this characterization to our research aims, we suggest that perhaps the opinions people possess are not unusually strong, or perhaps they simply prefer to keep these opinions to themselves.

ACKNOWLEDGEMENTS

This research project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Marie Skłodowska-Curie grant agreement No. 655601. The authors thank Alex Bitterman, Ingmar Pastak, and Janika Raun for helpful contributions to the research.

REFERENCES

1. *Residential Change and Demographic Challenge : The Inner City of East Central Europe in the 21st Century* [A. Haase, A. Steinführer, S. Kabisch, K. Großmann, R. Hall, eds.]. Edition ed. Farnham: Ashgate, 2011. 376 p.
2. Kovács, Z. Social and Economic Transformation of Historical Neighbourhoods in Budapest. *Tijdschrift voor economische en sociale geografie*, 2009/09, Vol. 100, Issue 4, pp. 399–416. <https://doi.org/10.1111/j.1467-9663.2009.00549.x>
3. Kährrik, A., Leetmaa, K., Tammaru, T. Residential Decision-Making and Satisfaction among New Suburbanites in the Tallinn Urban Region, Estonia. *Cities*, 2012, Vol. 29, Issue 1, pp. 49–58. <https://doi.org/10.1016/j.cities.2011.07.005>
4. Tammaru, T., Leetmaa, K. Suburbanisation in Relation to Education in the Tallinn Metropolitan Area. *Population, Space and Place*, 2007, Vol. 13, Issue 4, pp. 279–292. <https://doi.org/10.1002/psp.444>
5. Borén, T., Gentile, M. Metropolitan Processes in Post-Communist States : An Introduction. *Geografiska Annaler, Series B: Human Geography*, 2007/06, Vol. 89, Issue 2, pp. 95–110. <https://doi.org/10.1111/j.1468-0467.2007.00242.x>
6. Hirt, S. Suburbanizing Sofia : Characteristics of Post-Socialist Peri-Urban Change. *Urban Geography*, 2007/11, Vol. 28, Issue 8, pp. 755–780. <https://doi.org/10.2747/0272-3638.28.8.755>
7. Kährrik, A., Tammaru, T. Population Composition in New Suburban Settlements of the Tallinn Metropolitan Area. *Urban studies*, 2008/05, Vol. 45 (5–6), pp.1055–1078. <https://doi.org/10.1177/0042098008089853>
8. Liepa-Zemeša, M., Hess, D. B. Effects of Public Perception on Urban Planning: Evolution of an Inclusive Planning System During Crises in Latvia. *Town Planning Review*, 2016/01, Vol. 87, Issue 1, pp. 71–92. <https://doi.org/10.3828/tpr.2016.5>
9. Ruoppila, S., Kährrik, A. Socio-economic residential differentiation in postsocialist Tallinn. *Journal of Housing and the Built Environment*, 2003, Vol. 18, Issue 1, pp. 49–73.
10. Hegedüs, J., Tosics, I. Housing Classes and Housing Policy : Some Changes in the Budapest Housing Market. *International Journal of Urban and Regional Research*, 1983/12, Vol. 7, Issue 4, pp. 467–494.

11. Hess, D. B., Hiob, M. Preservation by Neglect in Soviet-Era Town Planning in Tartu, Estonia. *Journal of Planning History*, 2014, Vol. 13, Issue 1, pp. 24–49. <https://doi.org/10.1177/1538513213512254>
12. Hess, D. B. Transport in Mikrorayons. *Journal of Planning History*, 2017/06. <https://doi.org/10.1177/1538513217707082>
13. Metspalu, P., Hess, D. B. Revisiting the Role of Architects in Planning Large-Scale Housing in the USSR: The Birth of Socialist Residential Districts in Tallinn, Estonia, 1957–1979. *Planning Perspectives*, 2017/07/30, pp. 1–27. <https://doi.org/10.1080/02665433.2017.1348974>
14. Hess, D. B. Early 20th-Century Wooden Tenement Buildings in Estonia: Building Blocks for Neighborhood Longevity. *Town Planning and Architecture*, 2011/01, Vol. 35, Issue 2, pp. 110–116. <https://doi.org/10.3846/tpa.2011.13>
15. Kährik, A., Temelová, J., Kadarik, K., Kubeš, J. What Attracts People to Inner City Areas? The Cases of Two Post-Socialist Cities in Estonia and the Czech Republic. *Urban studies*, 2015/01/22, Vol. 53, Issue 2, pp. 355–372. <https://doi.org/10.1177/0042098014567444>
16. Leetmaa, K., Tammaru, T., Hess, D. B. Preferences toward Neighbor Ethnicity and Affluence: Evidence from an Inherited Dual Ethnic Context in Post-Soviet Tartu, Estonia. *Annals of the Association of American Geographers*, 2014/11/14, Vol. 105, Issue 1, pp. 162–182. <https://doi.org/10.1080/00045608.2014.962973>
17. Roose, A., Noorkõiv, R., Gauk, M. Suburbanization in Tartu. Quarterly bulletin of Statistics Estonia. *Tallinn: Statistics Estonia*, 2015, (1), pp. 80–110.
18. *Tartu Housing Survey “Tartu and Its Residents 2013”*. Edition ed. Tartu: Tartu City Government, 2013. 80–110 p.
19. Hess, D. B., Tammaru, T., Leetmaa, K. Ethnic Differences in Housing in Post-Soviet Tartu, Estonia. *Cities*, 2012/10, Vol. 29, Issue 5, pp. 327–333. <https://doi.org/10.1016/j.cities.2011.10.005>
20. Hastings, A., Dean, J. Challenging Images: Tackling Stigma through Estate Regeneration. *Policy & Politics*, 2003/04/01, Vol. 31, Issue 2, pp. 171–184. <https://doi.org/10.1332/030557303765371663>
21. Permentier, M., Van Ham, M., Bolt, G. Same Neighbourhood ... Different Views? A Confrontation of Internal and External Neighbourhood Reputations. *Housing Studies*, 2008/11, Vol. 23, Issue 6, pp. 833–855. <https://doi.org/10.1080/02673030802416619>
22. Cranshaw, J., Schwartz, R., Hong, J.I., Sadeh, N. The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. *Proceedings of the Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media (ICWSM 2012)*, Dublin, Ireland, 2012.
23. Boy, J. D., Uitermark, J. How to Study the City on Instagram. *PLoS ONE*, 2016, Vol. 11, Issue 6, e0158161.
24. Hochman, N., Manovich, L. Zooming into an Instagram City: Reading the Local through Social Media. *First Monday*, 2013, Vol. 18, Nr. 7. <https://doi.org/10.5210/fm.v18i7.4711>
25. Shelton, T., Poorthuis, A., Graham, M., Zook, M. Mapping the Data Shadows of Hurricane Sandy: Uncovering the Sociospatial Dimensions of ‘Big Data’. *Geoforum*, 2014/03/01, 52, pp. 167–179. <https://doi.org/10.1016/j.geoforum.2014.01.006>
26. Ash, J., Kitchen, R., Leszczynski, A. Digital Turn, Digital Geographies? *Progress in Human Geography*, 2016/08/24. <https://doi.org/10.1177/0309132516664800>
27. Schweitzer, L. Planning and Social Media: A Case Study of Public Transit and Stigma on Twitter. *Journal of the American Planning Association*, 2014, Vol. 80, Issue 3, pp. 218–238. <https://doi.org/10.1080/01944363.2014.980439>
28. Evans-Cowley, J. S., Griffin, G. Microparticipation with Social Media for Community Engagement in Transportation Planning. *Transportation Research Record: Journal of the Transportation Research Board*, 2012, (2307), pp. 90–98. <https://doi.org/10.3141/2307-10>
29. Mondschein, A. Five-Star Transportation: Using Online Activity Reviews to Examine Mode Choice to Non-Work Destinations. *Transportation*, 2015/03/27, Vol. 42, Issue 4, pp. 707–722. <https://doi.org/10.1007/s11116-015-9600-7>
30. Park, S. J., Park, J. Y., Lim, Y. S., Park, H. W. Expanding the Presidential Debate by Tweeting: The 2012 Presidential Election Debate in South Korea. *Telematics and Informatics*, 2016/05/01, Vol. 33, Issue 2, pp. 557–569. <https://doi.org/10.1016/j.tele.2015.08.004>
31. Hsu, C., Park, S., Park, H. Political Discourse among Key Twitter Users: The Case of Sejong City in South Korea. *Journal of Contemporary Eastern Asia*, 2013, Vol. 12, Issue 1, pp. 65–79. <http://dx.doi.org/10.17477/jcea.2013.12.1.065>
32. Iacobucci, E. *Understanding Attitudes and Perceptions of Public Transport: Investigation through Social Media and Conceptual Analysis*. University at Buffalo, State University of New York, 2016. 97 p.
33. Ideon, A. Estonia Tweets 40,000 Times a Day [online]. *Postimees*, 2013 [cited 10.08.2017]. <https://news.postimees.ee/1244182/estonia-tweets-40-000-times-a-day>
34. Sloan, L., Morgan, J. Who Tweets with Their Location? Understanding the Relationship between Demographic Characteristics and the Use of Geoservices and Geotagging on Twitter. *PloS one*, 2015/11/06, Vol. 10, Issue 11, pp. 142–209. <https://doi.org/10.1371/journal.pone.0142209>
35. Woelfel, J. Artificial Neural Networks in Policy Research: A Current Assessment. *Journal of Communication*, 1993, Vol. 43, Issue 1, pp. 63–80. <https://doi.org/10.1111/j.1460-2466.1993.tb01249.x>
36. Churchland, P., Sejnowski, T. *The Computational Brain*. Edition ed. Cambridge: MIT Press, 1992. 560 p.
37. Rumelhart, D. E., McClelland, J. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Edition ed. Cambridge: MIT Press, 1986. 567 p.
38. Pak, A., Paroubek, P. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *Proceedings of the LREC 2014, Ninth International Conference on Language Resources and Evaluation*, European Language Resources Association. Reykjavik, Iceland, 2010.
39. Evans, C., Chen, H., Battleson, B., Wölfel, J. K., et al. Unsupervised Artificial Neural Networks for Pattern Recognition in Multilingual Text. In *Proceedings of the International Network for Social Network Analysis (INSNA) Sunbelt Conference*, St. Pete Beach, FL2008, 2008.



Daniel Baldwin Hess, PhD, earned a doctoral degree in urban planning from the University of California, Los Angeles, USA, a master of urban planning from the University at Buffalo, State University of New York, USA, and a bachelor degree in civil and environmental engineering from Clarkson University.

Dr. Hess is currently Professor in the School of Architecture and Planning at the University at Buffalo, State University of New York, USA. He is also Visiting Scholar and Director of the Centre for Migration and Urban Studies at the University of Tartu, Estonia,

where his academic stay is funded by a Marie Skłodowska-Curie Fellowship from the European Union's Horizon 2020 Research and Innovation Programme. His scholarship addresses interactions between transport, land use, and metropolitan form, and his recent research projects explore how access to transport and housing foster social and economic functions in cities. Recent research projects explore the influence of market-led urban transformation in post-socialist urban space on the legacy of town planning and local and national planning systems. Dr. Hess won a Fulbright Scholar Award, an Eisenhower Fellowship from the U.S. Department of Transportation, and a Fellowship from the National Endowment for the Humanities.



Evan Iacobucci obtained a Master of Urban Planning degree from the State University of New York at Buffalo, a Master of Arts in philosophy from the State University of New York at Buffalo, and a bachelor degree in French from the State University of New York College at Geneseo.

Evan is a doctoral student at the Edward J. Bloustein School of Public Planning and Policy at Rutgers, The State University of New Jersey. His primary research interests include transportation behaviour, the use of information technology for novel research approaches, the history of cities,

and the role of science in planning research. His particular interest is in better understanding the factors and relationships that influence travel behaviour so that effective interventions may be designed toward a more sustainable future. His recent work focuses on determinants of transportation mode use, particularly the significance of attitudinal factors, and the effects that access to and use of different modes has on equity and accessibility issues.



Annika Väiko received a Master degree in human geography and regional planning from the University of Tartu, Estonia. She is a Research Assistant with the Centre for Migration and Urban Studies of the University of Tartu. Her scholarly interests include migration, residential mobility, and segregation.

CONTACT DATA

Daniel Baldwin Hess

E-mail: dbhess@buffalo.edu

Evan Iacobucci

E-mail: evaniacobucci@gmail.com

Annika Väiko

E-mail: annika.vaiko@ut.ee