Project number:
1.1.1.2/VIAA/1/16/100

TOURISM INTELLIGENCE LATVIA (TOURINTELV) THEORETICAL BASE

PostDoc Latvia

NACIONĀLAIS ATTĪSTĪBAS PLĀNS 2020

EIROPAS SAVIENĪBA
Eiropas Reģionālās attīstības fonds

IEGULDĪJUMS TAVĀ NĀKOTNĒ

for goal-oriented applied research including human geography, complexity science with participatory research methods, ICT and GIS modelling techniques for in-depth analyses of tourism

VIDZEME UNIVERSITY
OF APPLIED SCIENCES

Valmiera, 2019

The aim of the research project is using advantages of ICT development to create GIS solutions-based tourism intelligence platform of market information, providing analyses and forecasts of pan-Baltic big data representing cross-sectoral industry performance in multi-destination levels to improve competitiveness of Latvia in export markets and facilitate regional development.

Andris Klepers, Dr.geogr. post-doc researcher.

Spatial context of tourism flow and local destination systems

Understanding the geographical preferences of international tourists is critical for the tourism planning and marketing. However, it is not an easy endeavour to gather the corresponding information, given the absence of **city-level tourism statistical data** and high costs of participant survey. (cite{Wuhan University[Wuhan University]} (Shiliang, 2016[i]). Understanding the **movement of tourists within a destination has practical applications for destination management**, product development, and attraction marketing. (Lew & McKercher, 2006).

Tourism is, by definition, based on movement, and all **phenomena involving movement are difficult to measure**. There are many different forms of tourism, including holidays and business trips, short and long stays and so forth. The notion of tourism flows has different meanings for those in charge of road, rail or air traffic management, and for those in charge of tourist visits. (Terrier, 2009[ii]).

Lew & McKercher (2006) explored some of the conceptual challenges in understanding tourist **intradestination movement patterns**, to summarize the major influences on such movement, and to model the basic spatial forms that such movement can take. This insight can then be used as a **basis for empirical studies of tourist movements**, which can lead to practical applications for destination planners. Understanding **how tourists move through time and space, and the factors that influence their movements, has important implications** for infrastructure and transport development, product development, destination planning, and the planning of new attractions, as well as management of the social, environmental, and cultural impacts of tourism.

Tourist flow (*tūristu plūsma*, *der Touristenstrom, туристский поток)*: *Tūristu kopums, kas dodas uz noteiktām tūrisma vietām. Izšķir* ***ienākošo****, izejošo un* ***vietējo tūristu plūsmu****. Tūristu plūsmas, to apjomu, virzienu, izmaiņu tendences pēta tūrisma ģeogrāfija*. (Tūrisma un viesmīlības terminu skaidrojošā vārdnīca. — R., 2008)[iii]

Important data issue is the definition of a ‘‘destination’’ especially that on a local level.

The World Tourism Organization defined ‘‘**local tourism destination**’’ as a *physical space that includes tourism products such as support services and attractions, and tourism resources. It has physical and administrative boundaries* *defining its management, and images and perceptions defining its market competitiveness. Local destinations incorporate various stakeholders, often including a host community, and can nest and network to form larger destinations. They are the focal point in the delivery of tourism products and the implementation of tourism policy* (WTO 2002[iv]:np).

This somewhat inelegant description nonetheless provides insights into destination minima and maxima. The intent of this framework was to conceptualize destinations as local entities **that can include cities, towns, or regional areas. This definition excludes, at one end, resort complexes regardless of their size, and at the other end states/ provinces, countries, or multinational agglomerations**. (Lew & McKercher, 2006[v]).

Li & Yang (2016[vi]) demonstrated the **use of big data as an alternative data source to monitor tourist flows at the national scale**, and employed geospatial tools, like geographic information system (GIS) and spatial interaction model (SIM), to rigorously **explain the pattern of tourist movement**.

Tourism modelling and forecasting

A variety of approaches have been used to **forecast tourism demand**. The choice of a certain forecasting technique depends on data availability, time horizons, and research objectives.

Most studies that are devoted to the quantitative analysis of tourism flows can be divided into three groups (Furmanov, K. et al., 2012[vii]):

1) aimed at tourism demand forecasting in a particular country or a group of countries **based on time series models**;
2) causal econometric models that are used to explain the dynamics of tourism flows and elucidate the **relationships that exist between the demand for tourism and the different factors that possibly affect it**. The constructed models can also be used for forecasting purposes;
3) comparative studies whose purpose is to **determine the forecasting methods that outperform others** in most cases.

There has been dramatic growth in Internet usage, especially for travel information search purposes. Each time an indi- vidual uses a website he or she leaves traces on that site. These traces can be collected and used for different purposes such as tracking user behavior, recommending products to the customer on their next visit to the website and optimizing website usability. Google Analytics accounts make it possible for businesses to collect these traces from their websites. Although many destination management organizations (DMOs) are collecting these types of information from their websites, this information is usually not used for making managerial decisions, but merely by IT departments to enhance website usability. Often, the interpretation of website traffic indicators such as Google Analytics is not clear to DMO managers: such as what it means to the DMO to have one million website visitors. However, website traffic data can be very informative: showing, for instance, from which countries users originate.

This information can then be combined to see if there is a correlation between the country of origin of website visitors and the country of origin of the actual arrivals to the destination. (Gunter & Önder, 2016[viii]). That can be used to make build-in forecasting tool in the DMO web platform stakeholder area.

Data sources for tourism forecasting

For measuring and analysing tourism economics, the international consensus represented by United Nations approved International Recommendations with concepts, definitions, classifications and the **basic set of data** and indicators that should be part of any national System of Tourism Statistics (WTO, 2019[ix]). For local destinations that includes **particularly the data about domestic and international tourist overnight stays** and dynamic of tourist accommodations (CSB, 2019[x]). These data are standardized by EU in Latvia and other countries framed by the EU Tourism Statistics Directive (1995[xi]). These statistics are mainly at national or regional levels. There are no direct obligations on local authorities or on tourism businesses – thus the data-driven planning of local tourism destinations systems stay under own initiatives. Even more: such limitations eliminate demand of **same-day visitors** (*Tourism demand: domestic and outbound tourism (excluding day trips))*, analysing just part of local destination performance.

In case of large proportion of tourism demand related to same-day visitors this causes a major problem to estimate, plan and forecast targeted tourism activities for local tourism destinations.

A visitor is a traveller taking a trip to a main destination outside his/her usual environment, for less than a year, for any main purpose (business, leisure or other personal purpose) other than to be employed by a resident entity in the country or place visited. A visitor (domestic, inbound or outbound) is classified as a tourist (or overnight visitor), if his/her trip includes an overnight stay, or as a same-day visitor (or excursionist) otherwise. (WTO, 2019[xii]).

Some studies attempt to collect the information through **participant survey**, but the small sample size and biased interviewee choices greatly impede the analysis at broad level (Xu & Zhang, 2015[xiii]). **Alternative data sources** and more sophisticated techniques should be taken advantage to characterize geographical preferences of international tourists at broad level. In search for alternative data sources there are some large-scale EU initiatives, e.g. Eurostat (2017[xiv]) attempts to emphasize ***big data*** for regular statistic purposes with aim to understand the tourism development dynamics.

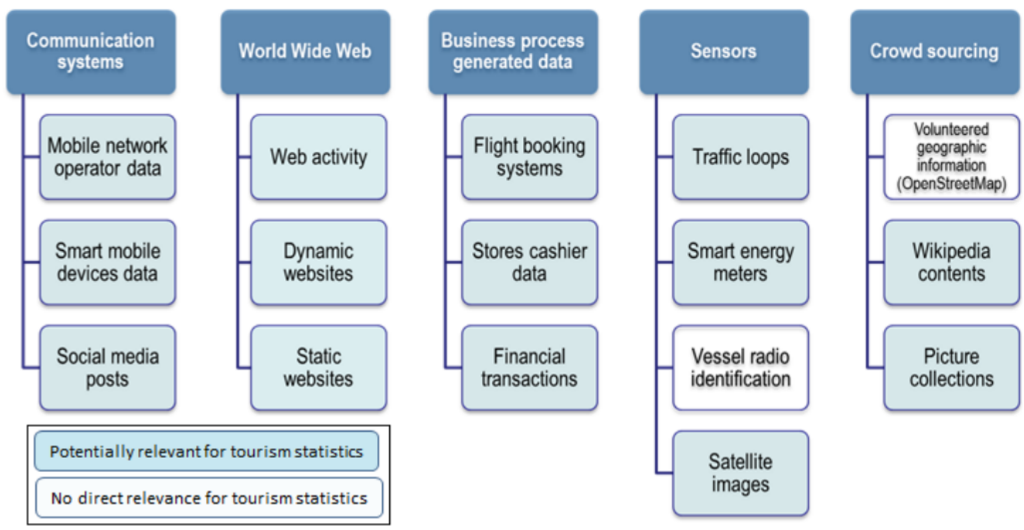
Data is everywhere; and it is revolutionising the world of official statistics. People and businesses are leaving behind a constant flow of digital footprints, voluntary or unintended.

This data deluge can make it difficult to see the wood for the trees; yet big data undeniably has huge potential for many areas of statistics (Eurostat, 2017).

The arrival of big data is also changing the working environment for statisticians. They no longer hold a monopoly on producing statistics, but now compete with a wide range of data producers. Ignoring innovation will push statistical authorities out of the information market — a development that could jeopardise the critical role of independent, official statistics in any democratic debate. Many sources of big data measure flows or transactions. Within the wide range of statistical domains, tourism statistics are on the frontline of big data-related innovations of sources and methods. Tourism statistics try to capture physical flows of people — as well as the accompanying monetary flows; big data provides promising new sources of data and previously unavailable indicators to measure these flows (and stocks). The different sources of big data and their potential relevance in compiling tourism statistics. It discusses the opportunities and risks that the use of new sources can create: new or faster data with better geographical granularity; synergies with other areas of statistics sharing the same sources; cost efficiency; user trust; partnerships with organisations holding the data; access to personal data; continuity of access and output; quality control and independence; selectivity bias; alignment with existing concepts and definitions; the need for new skills. The global dimension of big data and the transnational nature of companies or networks holding the data call for a discussion in an international context, even though legal and ethical issues often have a strongly local component (Eurostat, 2017).

The diagram in Figure 1 outlines the most commonly discussed sources of big data. Just like any other classification, individual items can be allocated to different groups, depending on the viewpoint. The same is true for this taxonomy, as sources are interrelated and multifaceted. For instance, social media posts can be filed under both 'communication systems' and 'world wide web'; Wikipedia is web-based but also crowd-sourced.

Figure 1: Taxonomy of big data sources



Based on this framework of alternative data categorization, for the local tourism destination system development purposes following data will be obtained in *TourInteLV* web platform for further analysis.

1.Communication systems.

(A). A1. Mobile network operator data (place reserved for synergy with alternative project of Integrated Design of Techno-Social Systems: Next Generation of Tourism Monitoring in Latvia: Project number: 1.1.1.2/VIAA/1/16/110[xv]).

A2. *Data about visiting of tourist attraction from Google profile including "popular times" visual information within a week*

2.World Wide Web.

(B). Web activity: digital demand collected from two main sources:

B1. *Google trends* key-word search dynamic.

B2. Local tourism destination landing-page digital demand statistics.

3.Business process generated data.

(C). Demand data from:

C1. tourist attractions (based on comparison – in real numbers and estimated in 5 grades). *Important issue here is to separate local demand from same-day visitor and tourist demand.*

C2. demand of public events with data set from market leaders in ticket sales of public events (Bilesuparadize.lv & Bilesuserviss.lv).

C3. Accommodation quality monitoring throughout the destinations using publicly available *Booking.com* customer ratings (summarized for recent 2 years).

C4. Demand of railway tickets towards seaside from passenger train[xvi]

4. Sensors.

(D). Tourist flow data from sensors – visitor counters:

D1. Specially managed mobile sensors to measure visitor demand in certain public space areas within certain time – in particular tourist attractions incl. public events.

D2. Third party data extracted and integrated from permanently settled-up sensors (or counting devices in compilation with video-cameras) at gateways of major tourist attractions.

D3. Tourism directly non-related sensor data that are helping to identify time & space dimensions of tourism flow throughout the year.

D3.1. Traffic loops: traffic intensity[xvii] measured by Latvian State roads.

D3.2. Smart electric energy consumption meters
- will be applied to the seaside destinations to characterize volume of seasonality.

5.Crowd sourcing.

(E). These algorithms of analysing geo-tagged pictures uploaded by travellers and other related content has been described in several scientific papers, but they are not foreseen to include in this *TourInteLV* web platform. However, part of data sharing from main stakeholders as a motivation form to contribute with a content is foreseen.

Currently, many pilot projects are ongoing and statistical authorities have begun releasing 'experimental statistics' based on innovative sources. However, the ultimate aim is to transform the tourism statistics system (or statistics in general) into a **data factory using many input sources to serve many output needs simultaneously**. (Eurostat, 2017).

Therefore, the *TourInteLV* web platform will serve demand for local tourism destination system stakeholders to deliver knowledge based on various data sources about demand of tourism combining official tourism statistics in merge with alternative data allowing to monitor, interpret, plane and forecast tourism development.

When discussing risks and constraints, new sources are typically in a 'defensive' position. The results of pilots are compared with existing data — somewhat arrogantly labelled 'the ground truth'. To fully adhere to the scientific method, statisticians need to make a critical assessment of the current methodology (and even use new sources to do so). A mobile phone penetration rate of only 90%

(with use falling even further when travelling?), is an issue that needs to be assessed and solved — but what about the tourism demand surveys using phone interviews (CATI) on the basis of landline registers, where less than half of the population has a landline nowadays? Or what about the dramatically falling response rates in surveys, sometimes below 50 %, or the significant bias due to the memory effect?

---

[i] Shiliang S. et al. (2016). Characterizing geographical preferences of international tourists and the local influential factors in China using geo-tagged photos on social media. *Applied Geography, 73*, 26-37.

[ii] Terrier, C. (2009). Tourist Flows and Inflows: On Measuring Instruments and the Geomathematics of Flows, in Patrick Bonnel , Martin Lee-Gosselin , Johanna Zmud , Jean-Loup Madre (ed.) Transport Survey Methods, pp.219 - 241

[iii] Tūrisma un viesmīlības terminu skaidrojošā vārdnīca. (2008). Rīga.

[iv] WTO. (2002). WTO Think Tank Enthusiastically Reaches Consensus on Frameworks for Tourism Destination Success. Madrid: World Tourism Organization, Madrid.

[v] Lew, A., & McKercher, B. (2006). Modeling Tourist Movements. A Local Destination Analysis. *Annals of Tourism Research, 33, (2), 403–423.*

[vi] Li, D., & Yang, Y. (2017). GIS Monitoring of Traveler Flows Based on Big Data. In *Analytics in Smart Tourism Design* (pp. 111-126). Springer International Publishing.

[vii] Furmanov K., Balaeva, O., & Predvoditeleva, M. (2012). Tourism flows from the Russian Federation to the European Union, Anatolia: An International Journal of Tourism and Hospitality Research, 23 (1), 17- 31

[viii] Gunter, U., & Önder, I. (2016). Forecasting city arrivals with Google Analytics. *Annals of Tourism Research, 61*, 199–212

[ix] UN WTO. (2019). Understanding Tourism: Basic Glossary.

[x] Centrālā Statistikas pārvalde. (2019). Transports un tūrisms. http://data1.csb.gov.lv/pxweb/lv/transp_tur/transp_tur__turisms__izm__ikgad/?tablelist=true

[xi] The EU Tourism Statistics Directive. (1995). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/88423/EU_Tourism_Statistics_Directive.pdf

[xii] UN WTO. (2019). Understanding Tourism: Basic Glossary.

[xiii] Xu, Z., & Zhang, J. (2015). Antecedents and consequences of place attachment: A comparison of Chinese and western urban tourists in Hangzhou, China. Journal of Destination Marketing & Management. http://dx.doi.org/10.1016/ j.jdmm.2015.11.003.

[xiv] Tourism statistics: Early adopters of big data? (2017). Eurostat. Luxembourg: Publications Office of the European Union, 2017

[xv] Bērziņa, I. (2019). Integrated Design of Techno-Social Systems: Next Generation of Tourism Monitoring in Latvia. Post-doc research project. Vidzeme University of Applied Sciences. Retrieved from: https://va.lv/en/research/research/integrated-design-techno-social-systems-next-generation-tourism-monitoring-latvia

[xvi] https://www.pv.lv/

[xvii] https://lvceli.lv/informacija-un-dati__trashed/#satiksmes-intensitate