🏆 *reviewed paper*

Web Data Extraction Systems versus Research Collaboration in Sustainable Planning for Housing: Smart Governance Takes It All

Valerie Dewaelheyns, Isabelle Loris, Thérèse Steenberghen

(Dr. Ir. Valerie Dewaelheyns, KU Leuven Department of Earth and Environmental Sciences, Spatial Applications Division Leuven, Celestijnenlaan 200E, 3001 Heverlee, Belgium, valerie.dewaelheyns@kuleuven.be)
(MSc Isabelle Loris, Ghent University Department of Civil Engineering - Center for Mobility and Spatial Planning, Vrijdagmarkt 10/301, 9000 Ghent, Belgium, Isabelle.Loris@UGent.be, and KU Leuven Department of Architecture - Housing and Urban Studies, Paleizenstraat 65-67, 1030 Brussels, Belgium)
(Dr. Ir. Thérèse Steenberghen, KU Leuven Department of Earth and Environmental Sciences, Spatial Applications Division Leuven, Celestijnenlaan 200E, 3001 Heverlee, Belgium, therese.steenberghen@kuleuven.be)

1 ABSTRACT

To date, there are no clear insights in the spatial patterns and micro-dynamics of the housing market. The objective of this study is to collect real estate micro-data for the development of policy-support indicators on housing market dynamics at the local scale. These indicators can provide the requested insights in spatial patterns and micro-dynamics of the housing market. Because the required real estate data are not systematicly published as statistical data or open data, innovative forms of data collection are needed. This paper is based on a case study approach of the greater Leuven area (Belgium). The research question is what are suitable methods or strategies to collect data on micro-dynamics of the housing market. The methodology includes a technical approach for data collection, being Web data extraction, and a governance approach, being explorative interviews. A Web data extraction system collects and extracts unstructured or semi-structured data that are stored or published on Web sources. Most of the required data are publicly and readily available as Web data on real estate portal websites. Web data extraction at the scale of the case study succeeded in collecting the required micro-data, but a trial run at the regional scale encountered a number of practical and legal issues. Simultaneously with the Web data extraction, the dialogue with two real estate portal websites was initiated, using purposive sampling and explorative semi-structured interviews. The interviews were considered as the start of a transdisciplinary research collaboration process. Both companies indicated that the development of indicators about housing market dynamics was a good and relevant idea, yet a challenging task. The companies were familiar with Web data extraction systems, but considered it a suboptimal technique to collect real estate data for the development of housing dynamics indicators. They preferred an active collaboration instead of passive Web scraping. In the frame of a users' agreement, we received one company's dataset and calculated the indicators for the case study based on this dataset. The unique micro-data provided by the company proved to be the start of a collaborative planning approach between private partners, the academic world and the Flemish government. All three win from this collaboration on the long run. Smart governance can gain from smart technologies, but should not loose sight of active collaborations.

2 INTRODUCTION

The complexity and multi-dimensionality of societal and environmental challenges, such as climate change and resource scarcity, challenge spatial planning towards more collaborative planning in which private and public actors converge towards a collective future (Aarts and Leeuwis 2010; Polk 2015). Spatial planning increasingly seeks to activate stakeholders for knowledge production and for the development of down-to-earth visions and policies.

One way to support such collaborative planning is the evolution towards 'Smart Cities'. The concept 'Smart Cities' refers to cities where digital technology and information is deployed for a more efficient use of resources. The use of digital technologies better equips cities to plan their future, taking into account new forms of governance, financing mechanisms and data exchange. The evolution towards Smart Cities goes together with the rise of an 'Open Data' culture. Data must be (1) available and accessible (e.g. at a reasonable price and in a handy and adjustable format or through download from the internet); (2) presented under conditions that allow the reuse and redistribution (including the merging with other datasets); and (3) data availability is universal, e.g. everyone must be able to use, reuse and redistribute the data (Bauer and Kaltenböck 2012; Khusro, Jabeen et al. 2014).

To make Smart Cities happen and work, we believe that research collaboration is a powerful approach. Transdisciplinary research projects for example aim at the creation of new knowledge on a common question

through collaboration between research and non-research partners (Katz and Martin 1997; Tress, Tress et al. 2005). In this paper we use the European definition of transdisciplinarity that focuses on the involvement of non-academics in research (Darbellay 2015; Zscheischler and Rogga 2015). This involvement can range from including stakeholders in the research as advisors, informants and users; to actual transdisciplinary co-production where solutions to (urban) planning problems and visions for (urban) planning are co-created by different actor groups (including policy-makers, administration and business) (Albrechts 2013; Polk 2015).

Transdisciplinary research is gaining momentum in the realm of sustainable land use management and spatial planning. This seems part of a broader movement towards more collaborative planning, with approaches such as collaborative planning (Healey 1997; Healey 1998), fuzzy planning (De Roo and Porter 2007), adaptive co-management (Olsson, Folke et al. 2004), strategic planning and co-production (Healey 2004; Healey 2007; Albrechts 2013).

Almost parallel to this evolution towards more collaboration, a 'sustainability-turn' appeared in planning in reaction to the undesirable environmental and societal effects of continuous housing development (Berke 2002; Atkinson-Palombo 2010). Spatial efficiency popped up as a new concept in planning, and the increase of residential densities in both new-growth areas and existing neighborhoods through densification is considered a solution for the space consuming effects of urban sprawl (Gallent 2009; Flemish Government 2012).

Pursuing sustainable planning solely through densification programs will probably lead to strategic gaps. In-fill developments may indeed preserve valuable larger units of agricultural and natural open space from further urbanization. However, often abstraction is made of the importance of smaller open spaces – be it public, semi-private or private - for the environmental quality of life and support of ecosystem services in urban areas (Ståhle 2010; Oktay 2012; Dewaelheyns, Vanempten et al. 2014). So, space efficient strategies in planning should not only focus on urban densification through the development of new housing on (remaining) urban open spaces, but also through the intensification of the existing housing stock.

Housing is one of the main drivers of spatial development and transformation, besides employment and mobility (European Environment Agency 2006; European Environment Agency 2013). While land-use changes and urbanization processes proceeding spatial transformations are widely documented (Engelen, Lavalle et al. 2007), the underlying micro-dynamics of housing are less investigated. Current and future housing requirements reflect changing ambitions, expectations, values and wishes. Property prices for example are a sign of these accumulated desires of individual citizens to live and work in a particular location, and to commute between both (Gallent 2009). Any spatial efficiency strategy focusing on housing requires more quantitative and qualitative insights in the local dynamics of the housing market, and planning should pay greater attention to price signals and imbalances between supply and demand on the housing market (Barker 2004; Gallent 2009).

To date, there are no clear insights in the spatial micro-dynamics of the housing market. Nevertheless, policy-support indicators could measure them. The research objective of this study is the development of a proof-of-concept of two 'open' (e.g. freely available and accessible) policy-support indicators, speed of sale and listing price, that allow insights in the micro-dynamics of the housing market. For the development of these indicators, we focus on micro-data of real estate listings. The research question relates to the methodology: what are suitable methods or strategies to collect data on micro-dynamics of the housing market? We explored a quantitative and a qualitative approach for data collection, being web data extraction and a transdisciplinary research collaboration process initiated through explorative interviews. The proof-of-concept was developed for the case of the greater Leuven situated in Flanders (Belgium).

3 MATERIAL AND METHODS

3.1 Selected indicators

Two indicators are investigated: speed of sale and listing price. Speed of sale is defined as the duration that houses are listed for sale on the market ('time-on-market'), with the time that a listing is published online as a proxy. Filippova and Fu (2011) found that properties in a booming market sold more quickly than properties sold in a declining market. In addition, a prolonged time-on-market reduces sale price. So, speed of sale seems to interact with house price (Clauretie and Thistle 2007; Johnson, Benefield et al. 2007). In addition, Miller & Sklarz (1987) confirmed that a greater degree of overpricing (listing price relative to

value) results in longer marketing time and lower selling price. So, the indicator 'listing price' also provides valuable insights. It should be clear however that there is a difference between the expected (listing price) and realized price.

3.2 Case study Belgium and the greater Leuven

The study is situated in Flanders, the northern region of Belgium in Western Europe. In general, Flanders is currently one of the most densely populated regions in Europe with a population density of 462 inhabitants per km² in 2010.[1] It is known as a strongly urbanized and highly built-up region, that is characterized by urban sprawl, a dense road network (4.5km/km²), fragmentation, and ownership figures far above the European average (European Commission ; Antrop 2004; Bengs, Schmidt-Thmoé et al. 2006; Kasanko, Barredo et al. 2006; De Decker, Ryckewaert et al. 2010; Verbeek, Boussauw et al. 2014).

For the proof-of-concept of both indicators, the research focused specifically on the greater Leuven composed out of the municipalities Leuven and Herent. Leuven itself is a small regional city. It is known for amongst other things its university and related research and developments spin-offs. With 98,376 inhabitants in 2015, it is the 10th most populated city of Flanders.[2] In 2010, the greater Leuven had a population density of 1,686 inhabitants per km²,[3] and the population growth in the past 10 years (2005-2015) was almost 10%.

A combination of arguments makes the greater Leuven an interesting case study for the proof-of-concept of indicators on micro-dynamics of the housing market. First, the average housing price in the city of Leuven in 2010 was 2.5 times (+149 %) as expensive as in 2000. Moreover, the average housing prices further evolved from € 253.002 in 2010 to € 312,162 in 2014.[4] This average housing price in Leuven equals about 124% of the average price for a house in Flanders in 2014. Initial results of a study of Helgers and Buyst (2014) suggest that the price elasticity of supply in Flanders is very inelastic. An increase in price due to increased demand, leads barely to an increase in supply but rather to inflationary effects. This also has the effect that an increase in the demand mainly leads to rising prices.

Second, about 34% of the inhabitants of the city of Leuven stated in 2014 that they want to move within five years. Of these 34%, slightly more than 15% wanted to move to a different city or municipality. About 64% of the emigrants between 0-9 year and 25-39 years of Leuven moved to Herent in 2014. The reverse movement of Herent to Leuven, over 18%, happened mostly in the age group 20-24 years.

Third, 27% of the households in Leuven spend more than 30% of the total household expenses on housing.[5] Fourth, the owner occupies just fewer than 53% of the houses in het city of Leuven. For the suburb Herent, this figure is almost 80%. Finally, the greater Leuven is part of a region with a high potential for sustainable (re-)development of structurally underused detached housing (Bervoets, van de Weijer et al. 2015).

3.3 Used methods

Information on the 'speed of sale' and 'listing price' of properties listed for sale is not readily available in official censuses and databases in Flanders or Belgium. Nor does a housing pressure indicator exists in Belgium, his regions or municipalities. Nevertheless, Flemish and Belgian real estate agencies have large databases with these data. Therefore, we used two approaches to collect listings information from real estate portal websites. The first approach was a technical approach using a web data extraction system. The second approach was a collaborative approach in which research collaboration was initiated through explorative interviews.

3.3.1 Web data extraction

Real estate portal sites on the World Wide Web do publish most of the required information on speed of sale and listing price publicly and readily available. To be able to use these Web data, they need to be collected from the web and structured in a database using a "web data extraction" system. A web data extraction

[1] http://statbel.fgov.be/nl/statistieken/cijfers/bevolking/structuur/woonplaats/oppervlakte_dichtheid/
[2] http://statbel.fgov.be/nl/statistieken/cijfers/bevolking/structuur/woonplaats/groot/
[3] http://statbel.fgov.be/nl/modules/publications/statistiques/bevolking/Bevolking_nat_geslacht_opp_bevolkingsdichtheid.jsp
[4] https://www.immotheker.be/Immotheker3/Site_v3/Default.aspx?WPID=323&MIID=169&MSIID=186&L=N&Regio=Leuven&ChartType=0&Jaar=2014
[5] http://www.stadsmonitor.be/betaalbaarheid-van-het-wonen-woonquote

ISBN 978-3-9504173-0-2 (CD), 978-3-9504173-1-9 (print)
Editors: Manfred SCHRENK, Vasily V. POPOVICH, Peter ZEILE, Pietro ELISEI, Clemens BEYER

system makes use of a software application that collects and extracts unstructured or semi-structured data that are stored or published on Web sources (Laender, Ribeiro-Neto et al. 2002; Sarawagi 2008; Ferrara, De Meo et al. 2014). This data can then be further processed in a semi-automatic or fully automatic way: data can be converted into workable and structured data, merged and unified for further processing and saved for further use (Ferrara, De Meo et al. 2014). Database building is one of the known applications of web data extraction tasks, besides opinion mining and sentiment analysis, customer care and context-ware advertising (Ferrara, De Meo et al. 2014).

Among the available tools for web data extraction, we used the free desktop application 'import.io' (https://import.io/). It works well on websites based on templates or regular structures, and uses information that is provided by users in the form of labeled example pages to build a training set (Ferrara and Baumgartner 2011). Import.io offers several advantages. Because of its Graphical User Interface (GUI), simple extractions do not require users to code, so a non-programmer can use the tool. The desktop application does not require a local server since it uses an online server hosted by import.io. The programming codes behind the Web Extraction use a standard Application Programming Interface (API) structure in multiple formats, promoting the sharing of the code with other developers. In addition, failed updates of the extraction are notified. Finally, import.io allows to download the extracted data in four different formats: Excel, HTML, JavaScript Object Notation (JSON) and as comma-separated values (csv).

We used two web data extraction features of import.io: the Extractor and the Crawler. Both are semi-automatic tools that need to be guided by the user through a training session on a minimum of five Web pages. For template elements (e.g. data or information published at fixed places in the Website template) and well-structured web pages, a machine learning approach can be used. This approach requires the user to highlight the required pieces of information and to identify their datatype (further called highlighting). Data or information that has no fixed position in the Web page template or that is published on less structured webpages has to be addressed through the XML Path Language (XPath). XPath is a syntax for defining fragments of an XML document, and is part of the W3C's XSLT standard.[6] The XPath syntax uses path expressions to select nodes or node-sets in an XML document. The difference between the Extractor and the Crawler is the depth at which data will be extracted. An Extractor only extracts data from the indicated Webpage, while the Crawler goes to Webpages of the same website at a deeper hierarchical level.[7]

To decide which website(s) to extract, we composed a set of screening criteria based on website evaluation checklists from the universities of Berkely[8], Leicester[9], Maryland[10] and Wisconsin[11]. The composed criteria included a suitable goal and content of the webpage (being the publication of real estate listings) and the availability of the required information (type of property; market; listing price; address; date since when the house is listed or speed of sale). Also, it should be clear who is the owner of the website; the website needs to be maintained and updated frequently; it should be user-friendly; and its relevance has to be clear (e.g. website of a local real estate agent versus a real estate portal site that offers listings of different real estate agents and listing providers).

We did a preliminary web data extraction test for the case of greater Leuven. After checking seven major real estate portal websites that offer real estate listings, we decided to focus on one real estate portal website that offered information on the 'listing date', e.g. date since when a property was listed. We used the import.io Extractor to extract the required information from a listings overview page of the considered website. Since the required data were well structured or published as fixed elements on the webpage, the Extractor was trained through highlighting. For this prototype extraction round, the number of pages was limited to less than 10 pages; and page selection was random.

We did a second trial run for the case of Flanders to test (1) if web scraping could be used to collect the required information at a larger scale and (2) if we could use the import.io Crawler to collect additional information from the pages of individual listings (like lot size, Energy Performance Certificate (EPC) value

[6] For the specifications see: http://www.w3.org/TR/xpath/
[7] http://support.import.io/knowledgebase/
[8] www.lib.berkeley.edu/TeachingLib/Guides/Internet/EvalForm.pdf
[9] http://www2.le.ac.uk/offices/ld/resources/study/eval-web
[10] www.lib.umd.edu/tl/guides/evaluating-checklist
[11] http://www.education.wisc.edu/docs/WebDispenser/soe-documents/aics-lesson3-webpagecredibilitychecklist.pdf?sfvrsn=2.

and building year). The website used for the first trial run blocked crawling, so we focused on two new websites that offer complementary information. One site also offers the required information on the date of listing, while the second site provides additional information on the type of sale (including private owner and auction). The Crawler was trained by highlighting for structured data and by XPath syntax for non-structured data (Table 1).

	Fixed elements: training by highlighting	Variable elements: training through XPath syntax
Website 1	*Type:* house, apartment *Market:* for sale, for rent *Listing price* *Address* (sometimes limited to the municipality) *Listed for sale since*: date *Lot size*	*Energy Performance Certificate (EPC)* .//*[.="EPC value"]/following-sibling::div *Building year* .//*[.="Building year"]/following-sibling::div *Housing type* .//*[.="Building placement"]/following-sibling::div *Price history* .//*[@class="table price-history"]
Website 2	*Type_market:* house, apartment; for sale, for rent, sales by auction *Listing price* *Address* *Municipality* *Seller:* agency, by auction, by the owner	*Energy Performance Certificate (EPC)* //*[.="Energieverbruik"]/following-sibling::td *Building year* //*[.="Bouwjaar"]/following-sibling::td *Subtype:* house, villa, bel-étage: //*[.="Subtype property"]/following-sibling::td *Number of façades:* proxy for housing type (detached, semi-detached, terraced) //*[.="Aantal gevels"]/following-sibling::td *Lot size* //*[.="Oppervlakte terrein"]/following-sibling::td

Table 1 Web data extraction by training of the application Import.io through highlighting and XPath syntax of the two extracted websites

3.3.2 Initiating transdisciplinary research collaboration

A transdisciplinary research collaboration is the second approach we explored for the development of policy-supporting indicators on the micro-dynamics of the housing market. The involved partners were university researchers on geosciences and –technology, the spatial planning policy department of the Flemish government, and private businesses active on the real estate market. This collaboration focused on the sharing of 'system knowledge'. System knowledge is knowledge on the current state of given system and its ability to change (Hirsch Hadorn, Bradley et al. 2006; Brandt, Ernst et al. 2013). As such, the research collaboration is situated at the 'consultation' level, but aims at initiating a process that can evolve towards the level of 'collaboration' where participants are partners in a policy project and jointly decide about issues with policy makers, (Pretty 1994; Leeuwis 2000; Aarts and Leeuwis 2010; Lang, Wiek et al. 2012; Brandt, Ernst et al. 2013).

Simultaneously with the web data extraction, the dialogue between stakeholders was started through explorative interviews. The interviews served three goals: (1) initiate the transdisciplinary research collaboration through a dialogue between stakeholders; (2) raise understanding, e.g. gain insights in the context of the housing market; and (3) explore the development of housing dynamics indicators, e.g. (i) whether there is a need of an indicator on dynamics in the housing market, and (ii) possibilities for the development of such indicator in terms of the availability of data, experiences from within the real estate market etc. Related to the latter goal, there were four research questions: is the housing pressure indicator a relevant indicator; when is a sale slow or fast; is web data extraction a suitable, reliable, efficient, feasible way to collect the required data to calculate the indicator; and are there other or better ways available or developable to collect the required data.

We used semi-structured interviews and purposive sampling, i.e. respondents from which we expected to receive as much information as possible (Maxwell 1997; Patton 2002; Guarte and Barrios 2006; Teddlie and Yu 2007; Creswell 2008). Qualitative research does not aim at collecting data from a random selection of a large number of data points to obtain statistical information about the opinions of an entire population. Instead, the aim is to choose a small number of respondents that will give in-depth data (Koontz 2003; Messely 2014). The seven major real estate portal websites in Flanders were contacted as well as the associations of notary and real estate agents. Only two real estate portal websites agreed to cooperate. Both

interviews lasted about 45 minutes. The interviews were transcribed ad verbatim using the online tool "otranscribe"[12].

The qualitative data were analyzed according to the grounded theory approach, organizing the interpretation of raw data into a theoretical explanatory scheme (Strauss and Corbin 1998). We used the open and axial coding phases of inductive coding. In the open coding phase, the data were broken down in discrete objects such as ideas, phenomena and feelings, and given a name. These objects are further called concepts. Since this is an exploratory study all objects were included. These concepts were then further analyzed and aggregated into distinct categories. In the axial coding phase we re-assembled the concepts and categories by identifying links and cross-cuts (Rogge, Dessein et al. 2011; Kerselaers, Rogge et al. 2013; Messely, Rogge et al. 2013). This gained a more profound and comprehensive understanding of the data.

3.4 Micro-data of real estate listings and data cleaning

The proof-of-concepts are based on a 10 year database of listings provided by a major real estate portal website after the explorative interview. This company scrapes small(er) independent real estate agents websites across Belgium and has more than 110,000 listings on the Web. The standardization of practice and procedures within the company insures a high enough level of accuracy of the data. The presented analysis specifically focuses on market dynamics in the year 2011. Data exchange required two rounds, since the researchers needed a first understanding of the database to be able to formulate the concrete variables needed for analysis. We received the data in a .csv format file of 2.5 GB. A selection was made of the case municipalities (Leuven and Herent) and required variables using Qlikview[13] and exported to .csv format.

Data clean up consisted of two main processes, being the selection of relevant and sufficiently documented records (houses and apartments for sale, municipalities of Leuven and Herent), and the transformation of data to usable formats (listing price, housing numbers, dates, geographic coordinates, etc.). We used the open source data cleaning tool "OpenRefine"[14] (Verborgh and De Wilde 2013). Only the required variables were kept, with data on listing price, speed of sale, address, type of market (rental and buyers), and context variables.

First, we made a selection for the greater Leuven which resulted in a set of 20,179 records out of more than 3 million. As we want to study houses and apartments for sale, rental properties and commercial buildings, building plots and garages were left out, as well as student rooms, holiday houses and 'other'. This resulted in a selection of 6,766 records. Second, we cleaned up the data skipping unrealistic or blank values for some key variables such as listing price, date first for sale (result: 5,419 records). Addresses were also not always correctly spelled, this was corrected using the text faceting tool of OpenRefine.

Third, not all listings with an address in Leuven or Herent were geographically located in Leuven or Herent. These were removed from the database by a spatial join between the dataset of the listings and a shapefile of municipalities in a Geographical Information System (GIS). A number of 103 records appeared to be geographically located outside Leuven and Herent, and were removed from the dataset. As a result, 5.271 records remained.

We focused on the year 2011 for the calculation of the proof-of-concepts. This is the most recent year in which census data is available concerning the existing building stock, population characteristics etc., what would allow comparison with official census data. Finally, spatial analyses were performed on a dataset of 390 records using Geographical Information System (GIS) techniques (Figure 1).

3.5 Calculating the indicators

The proof-of-concept indicators 'speed of sale' and 'listing price' were calculated for the year 2011. Speed of sale was calculated as the difference in number of days between the first and last date of publication on the portal website. Some processing was needed to calculate this speed of sale. The appropriate variables were "creatie_datum" (date of creation), "is_gearchiveerd" (listing is archived or not), "r_datum" (date of archiving), "laatstewijziging_datum" (date of last change), and the sequence of "Pub1 start", "Pub1 stop", "Pub2 start", "Pub2 stop", "Pub3 start", "Pub3 stop" etc. (start and end date of the publication period). For

[12] http://otranscribe.com/
[13] http://global.qlik.com
[14] http://openrefine.org

example, there seems to be some inconsistencies in the dates of archiving (e.g. when a real estate agent archives a listing) and the end date of the last publication period. This is due to the fact that a real estate agent him- or herself decides and assigns when a listing is archived. For the proof-of-concept, we used "Pub1 start" and "Pub1 stop".

The calculating of the indicator 'listing price' was straightforward. The database provided both the initial asking price and the corrected asking price. Since the corrected asking price is seldom completed, the initial asking price was used. This was compared with the average selling price tracked by Belgian censuses. For the greater Leuven, the average listing price was additionally compared with the average realized price[15] for the entire dataset (2005-2014) to check the consistency in pricing difference between the expected and the realized price over the past 10 years. Pricing difference (%) was calculated as [(realized price – listing prize) / listing prize *100].

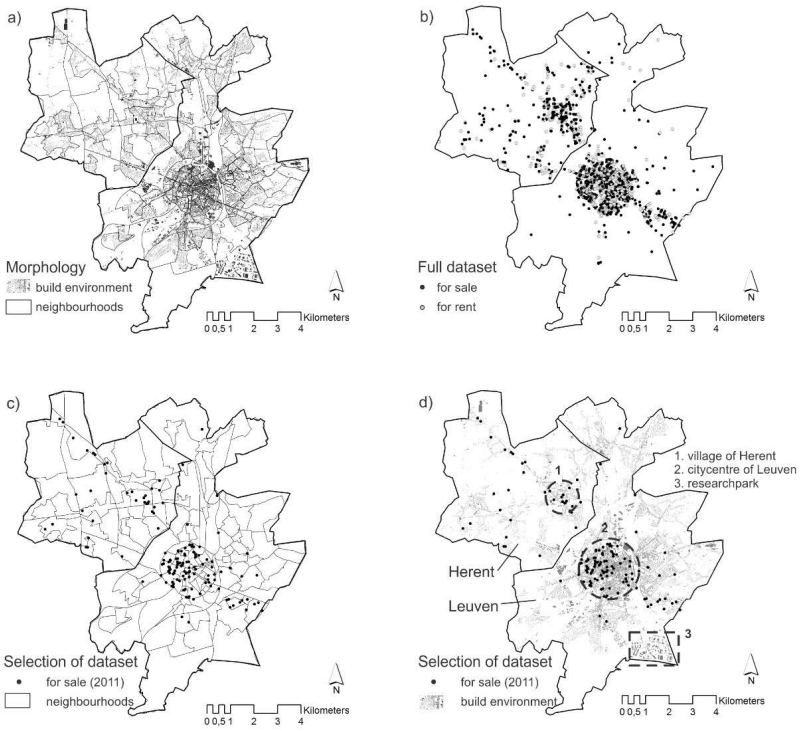

Figure 1 Characterization of the case study of the greater Leuven. Morphology (a) indicates the neighborhoods and built environment. The full dataset (b) gives an overview of all the listings, while the selected dataset (c) focuses on houses and apartments for sale for the year 2011. Map (d) indicates the three main urban cores being the city center of Leuven, the village of Herent and the research park of Haasrode in relation to the selected listings.

4 RESULTS

4.1 The use of web data extraction

At the scale of the greater Leuven, the web data extraction system proved a valuable method for exploring ideas and concepts for new indicators. This is especially the case when data is not readily available, as is the case for micro real estate data.

The web data extraction trial run at the scale of Flanders however encountered a number of practical and legal issues. Practical issues included an extensive duration of 44 hours for extracting one website. The legal issues are related to the grey legal zone in which web data extraction is situated. In Flanders, the act of web data extraction itself is not illegal. Yet, the terms of use of the screened real estate websites all disallowed on the one hand to use information published on the site by third parties and on the other hand to copy and/or store published information in databases. Some real estate websites could allow it on the precondition of a written authority; others literally prohibited web scraping. Besides the individual terms of use, the intellectual property rights and the database right (like European Union Directive No. 96/9/EC of 11 March

[15] Calculations based upon data from Stabel.be: http://statbel.fgov.be/nl/statistieken/cijfers/economie/bouw_industrie/vastgoed/gemiddelde_prijs_woonhuis/ retrieved on 2/1/2016.

ISBN 978-3-9504173-0-2 (CD), 978-3-9504173-1-9 (print)
Editors: Manfred SCHRENK, Vasily V. POPOVICH, Peter ZEILE, Pietro ELISEI, Clemens BEYER

1996; and UK Copyright and Rights in Databases Regulations 1997) apply. Due to these technical and legal issues, web data extraction is not considered an optimal technique to collect the required data for purposes beyond the exploring of a concept for an indicator.

4.2 Insights in the real estate market from two explorative interviews

The two explorative interviews provided preliminary insights in three main issues: (1) housing market indicators and organization of the housing market; (2) characteristics of the interviewed real estate portal sites; and (3) possibilities of data collection through web scraping and cooperation.

4.2.1 Housing market indicators and organization of the housing market

In general, both companies indicated that the development of indicators about housing market dynamics was a very good and relevant idea, yet a challenging task. The respondents referred to their existing price estimation indicators, and suggested four additional possible indicators: the share of listings published online compared to the total supply of listings; the share of listings that goes offline without being sold; the selling price; and the influence of insulation of the property on the price.

"[the development of an indicator about housing market dynamics is] a very good idea. We are very interested in that and tried to develop such indicators ourselves. Of course, that is not always possible because we are in a commercial dynamic. Sometimes we need to make choices. But it has always been the goal of our company to bring this kind of information to the private buyer." Company A

About the validity of indicators that are calculated based on real estate listings, the respondents stressed that (1) the indicator 'speed of sale' only indicates how fast a property is from the market, not how fast it is sold, and that (2) a share of the listings goes offline without actually being sold. This has to be taken into account when interpreting the indicator 'speed of sale' based on the duration that a listing was online. Benchmarking whether a property is slow or fast from the market can be done based on their data. The respondents also stated that 'blind spots' will remain, for example no data is available on the share and transactions of private sales, e.g. sales that happen "under the shower of the local soccer clubs". To check the validity, the respondents suggested involving the notaries, since they have data on actual sales and selling prices. Nevertheless, they also announced that this would be a challenging task.

The respondents claimed the real estate sector to be a difficult and complex sector. Also, there is an important role for power over data flows. Respondents believed that the government should take an official position concerning this, since they consider it a necessity for private buyers to know as much as possible about the properties that are for sale. So, concerning the organization of the housing market both companies stressed the need for more transparency on the housing market in Flanders.

4.2.2 Characteristics of the interviewed real estate portal sites

Both companies have a similar dream, being the persuasion of a transparent housing market where the end users are put first. They also put forward the development of their corporate real estate database. The business models of both companies differ however. Where one company works partly according to the general routine of buying listings from software providers that are completed with listings from free publications and available government data, the other company collects listings itself through own software and through web data extraction (called 'web scraping') of small initiatives and real estate agents.

"Our ambition is to offer the end-user the most complete housing supply. It (e.g. the API) is also a platform that we are developing specifically for the end-user. We really try to maintain this point of view. […] Our business model functions completely different. Basic functionalities are for free. We generate income through cooperation with the real estate agent. To make life easier for the real estate agent, to have better integrations, and services" Company B

"The past years, many people started [with initiatives to open up the real estate sector], from different stakeholders. Anyway, we want to support each initiative, because we too have a big dream to open up all of that [e.g. data] for everyone" Company A

Concerning the exchange of data, there is an ambition to provide data for free, with respect for the license restrictions of software providers. One company is currently developing an API for providing free data as a way to solve some commercial problems, since it would allow them to collect data as well. In general, both

companies referred to the need for a better organization of the exchange of data, on the one hand between the three regions Flanders, Brussels and Walloon region; and on the other hand between the real estate sector and the government (for example about the Energy Performance Certificate (EPC)).

4.2.3 Possibilities of data collection through web scraping and cooperation

Both companies were familiar with web data extraction systems (e.g. web scraping) and considered it a suboptimal technique to collect real estate data for the development of housing dynamics indicators. They mentioned two reasons for this, being the inefficient workflow of translating data to html to data; and the fact that only current listings are published online and that a long period of scraping is needed to collect time series data, e.g. labor-intensive. Moreover, the technique of web scraping was considered computer intensive, placing a load on the servers of the real estate portal websites. Legal issues encompassed the obligatory protection of the listings bought from software providers. For the company using web scraping as a business strategy, indexing of the website is not allowed and their website is protected against scraping. The company takes the position that the website of a real estate agent is a publication channel, but they respect it if an agent doesn't want to be scraped. Both companies stated that collaboration [with them] is a better option than web scraping.

"We don't prefer it [web scraping], because it is a very suboptimal way to transfer data. We translate our data to html, made for browsers, and then you translate through html this information into data again. Many data gets lost and is very computer intensive. We have to generate all those web pages at our server side, you have to collect all the data and process it. There are just better ways to collect data." Company B

"Web scraping, I don't think it is realistic for you to develop it yourself. I think it is much more interesting to collaborate with us." Company A

So, both companies preferred an active collaboration instead of passive web scraping in the frame of the development of indicators about housing market pressure. A range of concerns and preconditions was mentioned, related to the fact that they are commercial companies. The goals of the indicators are important criteria, although it was also mentioned that it is difficult to tell now where the development of these indicators will end. If there are commercial goals attached to it, it would be a problem to cooperate. But if the developed indicators will be available for everyone, it is okay. There are concerns about advantages for competing companies, and everybody who joins needs to collaborate under equal conditions. The companies' data can be used to develop a proof-of-concept. Specifically concerning the exchange of data, one company would provide their complete database, while the other company would first run a trial query. Once the proof-of-concept of the indicators is developed, a new consultation is requested to further discuss future cooperation.

4.3 Transdisciplinary research collaboration: from stakeholder engagement towards future co-production

The explorative interviews provided a basis for the determination of a shared goal. Both companies share our challenge on developing open (e.g. freely accessible) indicators on micro-dynamics of the housing market. The ambition of both companies is to pursue a transparent housing market where the end users are put first, so our challenge is part of their 'dream'. They also share our success, as they gain when we succeed in our challenge. Finally, they also feel responsible to work on this shared challenge, based on commercial motivations. The recognition of a shared goal opened up the way for co-creation of the project. A users' license was formulated and signed by one company and the research partner.

The process of an ongoing longtime collaboration between the real estate portal site and the Flemish government is now being developed. The formal transfer of the cooperation from the academic research group to the government was covered in a transfer meeting. A mutual engagement for exploring future collaboration between the company and the governmental party was pronounced. A win-win was identified for both parties. For the company, involvement in housing policy support research and insights in their own data were mentioned. Communication about the results and publication of the indicators could happen jointly. Also access to and/or involvement in developing additional information was considered a mutual gain.

For the government, collaboration with the private company allows to gain detailed insights in the housing market processes through data that is not available elsewhere. This would allow the effective development of

ISBN 978-3-9504173-0-2 (CD), 978-3-9504173-1-9 (print)
Editors: Manfred SCHRENK, Vasily V. POPOVICH, Peter ZEILE, Pietro ELISEI, Clemens BEYER

indicators for the real estate market, i.e. speed of sale, housing pressure, listing and sale prices, attractiveness of neighborhoods, and friction inoccupation of houses.

Both parties formulated specific points of interest. For the company, the reason for collaborating in the frame of this explorative study was the clear endpoint of the project in the near future. This ensured them that the collaboration would be manageable. This relates to the fact that they are a commercial business. Discussed preconditions for the governmental party included the possibility to put the company's database at the disposal of third parties in case of outsourcing (part) the analyses. Also the criteria of scientific integrity should be acknowledged and agreed upon for the collaboration.

4.4 Proof-of-concepts of the indicators

Based on the micro real estate data from the database provided by a major real estate portal, we calculated the proof-of-concepts for the indicators speed of sale and listing price.

4.4.1 Speed of sale

For the greater Leuven, the indicator 'speed of sale' provides detailed insights at the micro-level (Figure 2 and Figure 3). For listings submitted in 2011, we see that the two largest groups are those listings sold within two weeks, and those sold within three months (both 19%). Based on the data, almost half of the listings (47%) can be assumed to be sold within three months (Figure 3). Most of the listings are situated in de city center; the listings the most peripheral tend to be longer on the market. The listing price of those premises are lower than the inner city premises.

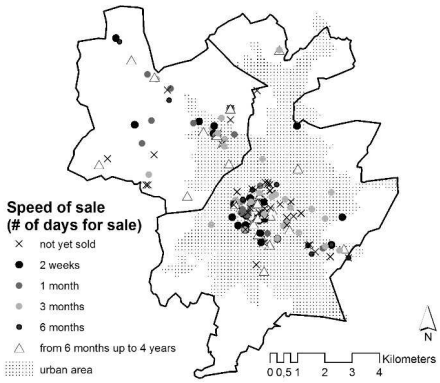

Figure 2 Map of the proof-of-concept indicator 'speed of sale' for the greater Leuven. The map is based on data received from one real estate portal site.

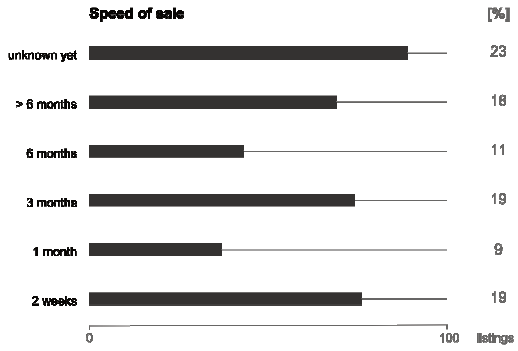

Figure 3 Distribution of listings for the year 2011 per arbitrary time category of speed of sale. The graph is based on data received from one real estate portal site.

4.4.2 Listing price

The indicator 'speed of sale' also provides detailed insights at the micro-level for the greater Leuven. It allows geographical differentiation, for example by regrouping by neighborhoods, municipalities, real estate markets, etc. (Figure 4). For the time series 2005-2014, comparisons were made between the listing prices from the database and the realized prices as reported by Statistics Belgium[16] (Figure 5). Stressing the

[16] http://statbel.fgov.be/nl/statistieken/cijfers/economie/bouw_industrie/vastgoed/gemiddelde_prijs_woonhuis/ retrieved on 2/1/2016

exploratory nature of our study, this first exercise clarifies for a limited time period (2005-2014) that listing prices for houses and apartments are more volatile than the realized price for the case of greater Leuven, for both houses as for apartments. Analyzing further the pricing difference (Figure 6), volatility appears bigger for apartments than for houses.

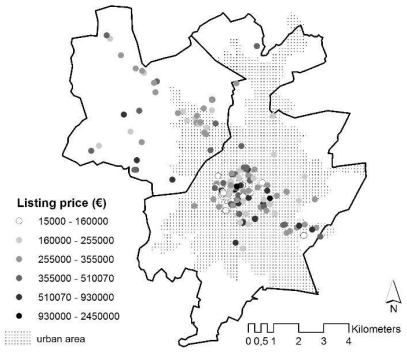

Figure 4 Map of the proof-of-concept indicator 'listing price' for the greater Leuven. The map is based on data received from one real estate portal site.

Figure 5 Evolution of the listing prices of houses and apartments, based on the dataset, versus there realized prices from Statistics Belgium over the period 2005-2014.

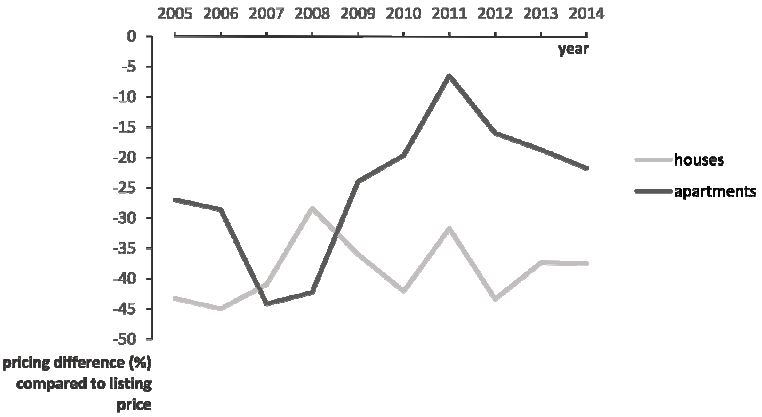

Figure 6 The difference in price (%) between listing price and realized price compared to the listing price is volatile, especially for apartments, for the time period 2005-2014.

5 DISCUSSION

5.1 About the proof-of-concepts and gaining insights in micro-dynamics of the housing market in the greater Leuven

Since the proof-of-concepts are calculated for the year 2011 and for one case study only, actual interpretations of the data and comparisons with other cities or benchmarks are not possible. The results of the proof-of-concepts did nevertheless allow to gain insights in possible analyses based on the two indicators, like the comparison with other datasets. Calculating these indicators for all municipalities in

ISBN 978-3-9504173-0-2 (CD), 978-3-9504173-1-9 (print)
Editors: Manfred SCHRENK, Vasily V. POPOVICH, Peter ZEILE, Pietro ELISEI, Clemens BEYER

Flanders or Belgium and for different years will allow to map dynamics of the housing market in time and space.

Concerning the proof-of-concept of the indicator 'speed of sale', it would be interesting to link the listings with information on the real estate agent. Does size of the real estate agency matters? Spatial analyses at neighborhood level allow insights in neighborhoods with a fast and a tame market. In collaboration with real estate companies, bench marks for a fast or slow speed of sale can be determined. We also wonder if there is an effect of the execution term of the selling deed, which is set at three months in Belgium.

The proof-of-concept for the 'listing price' indicator allows to calculate the average listing price per neighborhood, and evaluate of this average price is above or below the average listing price for the municipality, region, Flanders or Belgium. Concerning the pricing differences between the listing price and realized price, it is interesting to find out if the differences for Leuven (ranging until 45%) also appear elsewhere in Belgium. Further research at a larger scale of Belgium could verify the volatility of pricing differences for apartments. It would also be interesting to analyze the effect of the speed of sale on these pricing differences.

The goal was the development of policy-support indicators for spatial planning. The proof-of-concepts clarified that the indicators can explicit housing market dynamics in space and time at a detailed scale level. This will allow to evaluate in which neighborhoods the housing market is very dynamic and where it is dull, so where it would be opportune to stimulate or limit the housing supply. Are we planning at the right places? What if vibrant housing market dynamics appear there where we are not looking? The importance of monitoring over longer time series is related to the persistence of the observed dynamics. Housing dynamics can change rather fast, where spatial planning is characterized by a certain level of slowness. Monitoring indicators such as speed of sale and listing price allows to identify regions with a systematic higher or lower dynamic. This inspires and supports interferences on the longer run.

5.2 About distilling indicators for housing dynamics through micro-data of real estate listings

There are several reasons to untangle the spatial structures and dynamics through micro-data of real estate listings. First, distilling indicators for housing dynamics by means of micro-data is novel for Belgium. Since real estate data are not systematic published as statistical data or open data, access to these data can give insight in the spatial patterns of the housing market. We were able to elaborate on detailed real estate data per premise. This gave the means to analyze data on any geographical scale, so not to be bound by the usual statistical or administrative borders. Additional, time-series since 2005 allow the analysis of dynamics through time and through space.

Second, the real-estate data makes it possible to study profoundly the mechanisms in the housing market dynamics and its spatial patterns in retrospect of policy goals such as spatial efficiency within a sustainable planning.

Third, the unique data by courtesy of a private partner is a start of a collaborative planning approach between private partners, the academic world and the Flemish government. All three win from this collaboration on the long run. The private partner will have broader insights in the functioning of the real estate market by the developed indicators. The academic partner can benefit from new research on the topic of housing dynamics and elaborating indicators. The public partner can gain insights in spatial patters and dynamic to which it can steer upon as a function of desired spatial policy. Housing markets function locally, so a tailored spatial policy-framework is needed to intervene or monitor spatial development. A long lasting collaboration between the three stakeholders on the field of urban planning and on housing is also new to planning in Belgium.

5.3 About working with the database of private companies

The review and pre-processing of the company's database provided points of attention for future collaborations. The exchange of data highlighted the importance of suitable formats and standards for database export. The joint development of a protocol or model for database export could increase the efficiency of data exchange. The review and clean-up of the data revealed a number of human errors, despite the review and clean-up by the company. In addition, the metadata files with brief information on a part of the variables provided by the company appeared insufficient for the researchers to gain a full understanding of the meaning of (certain) variables and their interrelations. To handle issues with unclear variables when

calculating indicators and for the definition of benchmarks, a close collaboration with the company's employees is needed.

If the research collaboration continues, a data exchange procedure should be designed. We see two possible approaches: Extract Transform and Load (ETL) and Linked Data (LD). The first approach, Extract Transform and Load (ETL), is a database management process that includes the extraction of data from different data sources, the transformation of data for storing it in a suitable format or structure, and the loading of the transformed data in the final database. Such ETL systems are developed to integrate data from multiple applications or sources. In this case, it would be data from different real estate companies. All these 'parent' databases are typically developed and supported by different software packages, using different structures and formats.

Using ETL for the data exchange between real estate companies and the government requires the development of a database export query for each company. The initial development will need some effort in the beginning, but once operational it can be maintained with a minimal effort. Next, a transformation protocol needs to be developed, e.g. which variable from companies x, y and z refer to which variable in the final database of the government; transformation of the coordinates to a suitable format and coordination system, etc. When working with database export queries, the transformation of the data can be done by either of the parties. Finally, the transformed data is loaded into the database of the government. Given the commercial identity of real estate companies, the government could provide the technical development of the database export queries and transformation protocols.

The second approach concerns the Semantic Web and the use of Linked Data. In the semantic web, computers are capable of understanding the exact meaning of data (Berners-Lee 2000), by the linking of data and internet applications through web services (Folmer and Verdonk). Linked Data is a method for the publication of data in such a standardized and structured way that the data becomes linkable. Linked Data is considered one of the promising solutions for the integration of large volumes of data that are available online in a variety of proprietary and non-proprietary formats. The Semantic Web can help to overcome the hindrance of a fragmented supply of data (in terms of formats, publishers, etc.) by enabling the analysis, exploration and discovery of unexpected connections and relations. This way, better insights in and understanding of complex social, geographic, cultural and economic processes can be gained (Ballatore, Bertolotto et al. 2014).

5.4 About opportunities for transdisciplinary co-production in sustainable planning and housing policy

Transdisciplinary research collaboration between policy, academics and private companies is new in the context of housing policy. This study has put a first step towards involving private companies in the research collaboration of spatial policy makers and academics. Initial collaborations however appeared delicate, in terms of the building trust and balancing between commercial, academic and policy concerns.

At this stage, collaboration was limited to the consultation level. For each party involved, such collaboration appeared to offer opportunities and wins which are hardly attainable alone. This was an incentive to engage for a further exploration of future collaboration between policy and private companies. Involving stakeholders into scientific domain as advisors, informants and users moves research activities analytically and concretely into transdisciplinary co-production (Polk 2015).

There is a need for more common and collective strategies, not only in the planning but also in the management of space (Tress and Tress 2003; Dewaelheyns 2014; Dewaelheyns, Kerselaers et al. 2016). The development of new policies by governments in cooperation with private companies and citizens is considered an important way to improve the quality of government plans. Involving people in decision making allows them to learn and understand the range of problems and solutions, which can result in greater public support for the end result (Aarts and Leeuwis 2010).

The role of governments in such new governance processes shifts from being a strictly imposing authority (e.g. instrumental perspective) towards a facilitating authority, creating optimal conditions for societal projects (e.g. participation and eventually network perspective) (Aarts and Lokhorst 2012). Governments move away from organizing and managing top-down and bottom-up processes, towards anticipation and making use of self-organizing ability and initiatives of people (Aarts and Lokhorst 2012). Aarts and Leeuwis

ISBN 978-3-9504173-0-2 (CD), 978-3-9504173-1-9 (print)
Editors: Manfred SCHRENK, Vasily V. POPOVICH, Peter ZEILE, Pietro ELISEI, Clemens BEYER

(2010) explicitly stress this need for a 'network perspective' on governance, since people are active agents who interact with each other and organize themselves to get things done. This changes the way a government interacts with society to get things done. To be able to anticipate on the self-organizational ability of people, Aarts and Lokhorst (2012) suggest that governments must be alert and constantly gathering information about developments in society. This requires a pro-active attitude in contacting societal actors and opening negotiations with them (Aarts and Lokhorst 2012), which has proven fruitful in this explorative study. Such network model allows governments to better situate themselves amidst the dynamics of power and processes of self-organization.

5.5 About joined monitoring

To balance the internal and external complexity of spatial developments, quantitative analyses via monitoring systems are best combined with specific qualitative evaluations. The definition of key performance indicators and their coupling with development projects allows to check and demonstrate how planned developments contribute to an improvement of the local quality of life and to the realization of the Flemish policy ambition of spatial efficiency.

Transparency of spatial policy cannot be limited to the mere consultation of indicators. The learning-by-doing processes and transdisciplinary research collaborations demand more attention for mutual communication. This requires new instruments that allow citizens, companies, organizations, etc. to share information, accompanied by clear agreements on how to handle the shared information. The sharing of information by publishing spatial indicators in new formats that are freely accessible at the Internet (such as Linked Data) increases the utility of available information outside the own ICT-environment. This is not only the case for governments, but also for private companies as indicated by our research. This allows to discover correlations between data of diverse sources, inviting new insights in spatial issues of companies, organizations and citizens.

6 CONCLUSIONS

Data on the supply and demand side of the housing market is not always (readily) available. We explored possible strategies to develop two policy-support indicators on local housing market dynamics for the buyers' market. These indicators were "speed of sale" and "listing price".

We focused on two methods, being the technical and quantitative approach of web data extraction systems, and the participatory qualitative approach of research collaboration and co-production. The web data extraction system proved a valuable method for exploring new ideas and concepts for new indicators. This is especially the case when data is not readily available. Yet, when moving to an operational phase, web data extraction systems are not the best suited method to collect large amounts of data at high temporal frequencies. In addition, legal issues are to be respected: all evaluated portal sites prohibited the storage of the information published on their website in databases of third parties, being exactly the goal of web data extraction systems.

The qualitative approach of research collaboration and co-production was explored via interviews with two companies publishing real estate listings online. The use of explorative interviews appeared a good entry to dialogue with stakeholders. It became clear that both companies shared our goal and responsibility. These are two preconditions for further collaboration, regardless of the intensity of collaboration (consultancy, collaboration and self-mobilization). The interviews also allowed to gain preliminary insights in the housing market which proved valuable along the process. The development of new policies by governments in cooperation with private companies and citizens is considered an important way to improve the quality of government plans. Points of attention for the exchange of data between a private company and research comprise the joint development of a protocol or model for database export and close collaboration with the company in the form of a feedback panel (conceptually and technically).

Being able to explore micro data obtained through the research collaboration is an advantage. First, access is provided to an original dataset of all listings for the whole of the country, including historical data of a 10 years' range. This prevents extensive and technically complex searches to be carried out via web scraping of the Internet. Second, a dialogue with the private partner can provide interesting and first hand insights regarding the functioning of the housing market. Third, data on the level of single listings give a range of potential opportunities to analyze spatial patters and dynamics as from regional to local levels as to

neighborhoods. As our case study shows, differences in pricing and vacancy rate can be found throughout the neighborhoods and the urban and suburban municipalities.

The proof-of-concept indicators on 'speed of sale' and 'listing price' could be further tested on fixed geographical areas such as administrative divisions, policy based boundaries (e.g. urban areas, zoning plans), or (segmented) housing markets. Based upon time series, dynamics throughout history could be made visual in space. In the future, we intend to broaden up the indicators to the whole of Flanders or Belgium and benchmark them in time. A further differentiation is possible due to the different submarkets (buyers' market, rental market, development of newly build homes). This will enable us to benchmark the results of the greater Leuven with other urban metropolitan zones.

To conclude, our study proves that smart governance can gain from smart technologies, but should not lose sight of active collaborations.

7 ACKNOWLEDGEMENTS

The Flemish Policy Support Centre on Space (2012-2015) funded this research. The views expressed in this paper are not necessarily those of the Flemish Government. Detailed real estate data were kindly provided by Zimmo.

8 REFERENCES

Aarts, N. and C. Leeuwis. "Participation and Power: Reflections on the Role of Government in Land Use Planning and Rural Development." The Journal of Agricultural Education and Extension 16(2): 131-145. 2010.

Aarts, N. and A. M. Lokhorst. The Role of Government in Environmental Land Use Planning: Towards an Integral Perspective, 2012.

Albrechts, L. "Reframing strategic spatial planning by using a coproduction perspective." Planning Theory 12(1): 46-63. 2013.

Antrop, M. "Landscape Change and the Urbanization Process in Europe." Landscape and Urban Planning 67(1-4): 9-26. 2004.

Atkinson-Palombo, C. "New housing construction in Phoenix: Evidence of "new suburbanism"?" Cities 27(2): 77-86. 2010.

Ballatore, A., M. Bertolotto, et al. "Linking geographic vocabularies through WordNet." Annals of GIS 20(2): 73-84. 2014.

Barker, K. Review of Housing Supply: Delivering Stability - Securing our Future Housing Needs. London, HM Treasury: 14. 2004.

Bauer, F. and M. Kaltenböck. Linked Open Data: The Essentials. A Quick Start Guide for Decision Makers. edition mono/monochrom. Vienna, Austria, 2012.

Bengs, C., K. Schmidt-Thmoé, et al. ESPON 1.1.2. Final Report: Urban - rural relations in Europe. ESPON 1.1.2. C. Bengs and K. Schmidt-Thomé. Helsinki, Centre for Urban and Regional Studies, Helsinki University of Technology: 482. 2006.

Berke, P. R. "Does Sustainable Development Offer a New Direction for Planning? Challenges for the Twenty-First Century." Journal of Planning Literature 17(1): 21-36. 2002.

Berners-Lee, T. Weaving the Web: the original design and ultimate destiny of the World Wide Web by its inventor. HarperBusiness. San Francisco, 2000.

Bervoets, W., M. van de Weijer, et al. "Towards a sustainable transformation of the detached houses in peri-urban Flanders, Belgium." Journal of Urbanism: International Research on Placemaking and Urban Sustainability 8(3): 302-330. 2015.

Brandt, P., A. Ernst, et al. "A review of transdisciplinary research in sustainability science." Ecological Economics 92: 1-15. 2013.

Clauretie, T. and P. Thistle. "The Effect of Time-on-Market and Location on Search Costs and Anchoring: The Case of Single-Family Properties." The Journal of Real Estate Finance and Economics 35(2): 181-196. 2007.

Creswell, J. W. Research Design; Qualitative, Quantitative and Mixed method Approaches. 3rd edition. Sage Publications. Thousand Oaks CA, 2008.

Darbellay, F. "Rethinking inter- and transdisciplinarity: Undisciplined knowledge and the emergence of a new thought style." Futures 65: 163-174. 2015.

De Decker, P., M. Ryckewaert, et al. Ruimte voor wonen. Trends en uitdagingen (Space for living. Trends and challenges). Garant Publishers. Antwerpen, 2010.

De Roo, G. and G. Porter, Eds. Fuzzy planning: the role of actors in a fuzzy governance environment. Ashgate Publishing Limited. Hampshire, Burlington, 2007.

Dewaelheyns, V. The garden complex in strategic perspective. The case of Flanders. PhD, KU Leuven. 2014.

Dewaelheyns, V., E. Kerselaers, et al. "A toolbox for garden governance." Land Use Policy 51: 191-205. 2016.

Dewaelheyns, V., E. Vanempten, et al. "The Fragmentation Bias in Valuing and Qualifying Open Space." Journal of Urban Design: 1-20. 2014.

Engelen, G., C. Lavalle, et al. The MOLAND modelling framework for urban and regional land-use dynamics. Modelling Land-Use Change. Progress and Applications. . E. Koomen, J. Stillwell, A. Bakema and H. J. Scholten. Springer: 297-320. The Netherlands, 2007.

European Commission. (2003). "MOLAND: Monitoring Land Use/Cover Dynamics - Towards sustainable Urban and Regional Development: Analysis of Fragmentation of Urban Landscapes." Retrieved 2003, from http://moland.jrc.it/technical_tools/frag/moland_fra.htm.

European Environment Agency. Urban Sprawl in Europe - The Ignored Challenge. European Environmental Agency and European Commission Directorate General Joint Research Centre. Copenhagen, European Environmental Agency. 2006.

European Environment Agency. Land use. Copenhagen, European Environment Agency,. 2013.

Ferrara, E. and R. Baumgartner (2011). Design of automatically adaptable web wrappers. Proceedings of the 3rd International Conference on Agents and Artificial Intelligence.

Ferrara, E., P. De Meo, et al. "Web data extraction, applications and techniques: A survey." Knowledge-Based Systems 70: 301-323. 2014.
Filippova, O. and S. Fu. "Time-On-Market and House Prices in Auckland, New Zealand." Pacific Rim Property Research Journal 17(1): 70-91. 2011.
Flemish Government. The Green Paper Spatial Policy Plan: Flanders in 2050 – Human scale in a metropolis. Spatial Development Department Flanders. Brussels: 88. 2012.
Folmer, E. and Y. Verdonk. Linked data in beeld. Platform Linked Data Nederland Nederland.
Gallent, N. "The future of housing and homes." Land Use Policy 26, Supplement 1: S93-S102. 2009.
Guarte, J. M. and E. B. Barrios. "Estimation Under Purposive Sampling." Communications in Statistics - Simulation and Computation 35(2): 277-284. 2006.
Healey, P. Collaborative Planning: Shaping Places in Fragmented Societies. UBC Press, 1997.
Healey, P. "Collaborative planning in a stakeholder society." Town Planning Review 69(1): 1-1. 1998.
Healey, P. "The Treatment of Space and Place in the New Strategic Spatial Planning in Europe." International Journal of Urban and Regional Research 28(1): 45-67. 2004.
Healey, P. Urban Complexity and Spatial Strategies: Towards a Relational Planning for Our Times. Routledge Taylor & Francis Group. London, 2007.
Helgers, R. and E. Buyst. Woningprijzen: een regional woningmarktmodel. (Housing prices: a regional housing market model). Leuven, STeunpunt Wonen: 51. 2014.
Hirsch Hadorn, G., D. Bradley, et al. "Implications of transdisciplinarity for sustainability research." Ecological Economics 60(1): 119-128. 2006.
Johnson, K. H., J. D. Benefield, et al. "The probability of sale for residential real estate." Journal of housing research 16(2): 131-142. 2007.
Kasanko, M., J. I. Barredo, et al. "Are European cities becoming dispersed? A comparative analysis of 15 European urban areas." Landscape and Urban Planning 77(1-2): 111-130. 2006.
Katz, J. S. and B. R. Martin. "What is research collaboration?" Research Policy 26(1): 1-18. 1997.
Kerselaers, E., E. Rogge, et al. "Changing land use in the countryside: Stakeholders' perception of the ongoing rural planning processes in Flanders." Land Use Policy 32(0): 197-206. 2013.
Khusro, S., F. Jabeen, et al. "Linked Open Data: Towards the Realization of Semantic Web-A Review." Indian Journal of Science and Technology ISSN 7(6): 745-764. 2014.
Koontz, T. M. "The farmer, the planner, and the local citizen in the dell: how collaborative groups plan for farmland preservation." Landscape and urban planning 66(1): 19-34. 2003.
Laender, A. H. F., B. A. Ribeiro-Neto, et al. "A brief survey of Web data extraction tools." SIGMOD Record 31(2): 84-93. 2002.
Lang, D., A. Wiek, et al. "Transdisciplinary research in sustainability science: practice, principles, and challenges." Sustainability Science 7(1): 25-43. 2012.
Leeuwis, C. "Reconceptualizing Participation for Sustainable Rural Development: Towards a Negotiation Approach." Development and Change 31(5): 931-959. 2000.
Maxwell, J. Designing a qualitative study. Handbook of applied social research methods. L. Bickman and D. J. Rog. Sage: 69-100. Thousand Oaks, CA, 1997.
Messely, L. On regions and their actors. An analysis of the role of actors and policy in region-specific rural development processes in Flanders. PhD, Universiteit Gent. 2014.
Messely, L., E. Rogge, et al. "Using the rural web in dialogue with regional stakeholders." Journal of Rural Studies 32(0): 400-410. 2013.
Miller, N. and M. Sklarz. "Pricing Strategies and Residential Property Selling Prices." Journal of Real Estate Research 2(1): 31-40. 1987.
Oktay, D. "Human Sustainable Urbanism: In Pursuit of Ecological and Social-Cultural Sustainability." Procedia - Social and Behavioral Sciences 36: 16-27. 2012.
Olsson, P., C. Folke, et al. "Adaptive Comanagement for Building Resilience in Social–Ecological Systems." Environmental Management 34(1): 75-90. 2004.
Patton, M. Q. Qualitative Research & Evaluation Methods. SAGE Publications, 2002.
Polk, M. "Transdisciplinary co-production: Designing and testing a transdisciplinary research framework for societal problem solving." Futures 65: 110-122. 2015.
Pretty, J. N. "Alternative Systems of Inquiry for a Sustainable Agriculture." IDS Bulletin 25(2): 37-49. 1994.
Rogge, E., J. Dessein, et al. "Stakeholders perception of attitudes towards major landscape changes held by the public: The case of greenhouse clusters in Flanders." Land Use Policy 28(1): 334-342. 2011.
Sarawagi, S. "Information Extraction." Found. Trends databases 1(3): 261-377. 2008.
Ståhle, A. "More green space in a denser city: Critical relations between user experience and urban form. ." Urban Design International 15 (1): 47-67. 2010.
Strauss, A. and J. Corbin. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. . Sage Publications. Thousand Oaks, 1998.
Teddlie, C. and F. Yu. "Mixed Methods Sampling: A Typology With Examples." Journal of Mixed Methods Research 1(1): 77-100. 2007.
Tress, B. and G. Tress. "Scenario Visualisation for Participatory Landscape Planning - a Study From Denmark." Landscape and Urban Planning 64(3): 161-178. 2003.
Tress, B., G. Tress, et al. "Integrative studies on rural landscapes: policy expectations and research practice." Landscape and Urban Planning 70(1-2): 177-191. 2005.
Verbeek, T., K. Boussauw, et al. "Presence and trends of linear sprawl: Explaining ribbon development in the north of Belgium." Landscape and urban planning 128(0): 48-59. 2014.
Verborgh, R. and M. De Wilde. Using OpenRefine. Packt Publishing, 2013.
Zscheischler, J. and S. Rogga. "Transdisciplinarity in land use science – A review of concepts, empirical findings and current practices." Futures 65: 28-44. 2015.