Urban Planning and Smart Cities: When is ‘Smart’ really Smart

Danny Vandenbroucke (SADL)
Outline

• When is ‘Smart’ really Smart?
• Spatial planning and the housing market
• Web technologies to enrich SDI’s
• Conclusions and ongoing work
Outline

• When is ‘Smart’ really Smart?

• Spatial planning and the housing market

• Web technologies to enrich SDI’s

• Conclusions and ongoing work
When is ‘smart’ really smart

A smart city is an urban development vision to integrate multiple information and communication technology (ICT) and Internet of Things (IoT) solutions in a secure fashion to manage a city's assets – the city's assets include, but are not limited to, local departments' information systems, schools, libraries, ...

- Is it only about ICT and IoT?
- Is it also not about generating intelligent information from a bulk of data … ? About smart governance … ?
Smart cities

What makes a smart city smart?

- Open Data
- Smart Energy
- IoT
- Smart Retail
- Smart Health
- Smart Mobility
- Smart Home
- Education
When is ‘smart’ really smart

- We need more than (open) data …
  - We need intelligence, information and insight …

How do children look to traffic situations?

How do car drivers look for parking space?
When is ‘smart’ really smart

• OpenTransportNet.eu hub
  o Monitoring traffic and safety in the city of Turnhout using ANPR technology
  o WebGLayer API available
Outline

• When is ‘Smart’ really Smart?

• Spatial planning and the housing market

• Web technologies to enrich SDI’s

• Conclusions and ongoing work
Micro-dynamics of housing market?
Transdisciplinary research collaboration

- **Objective**: better understand the housing market in view of managing the urban sprawl (densification instead of spreading)

- **Method**: collaborative planning, fuzzy planning, adaptive co-management of space, …

- **Involved partners**: university researchers on geosciences and –technology, the spatial planning policy department of the Flemish government, and private businesses active on the real estate market
Housing market

• Housing is one of the main drivers of spatial development (EEA, 2013)
  ○ Together with employment and mobility

• Spatial and urban planning should play greater attention to price signals

• Information on the housing market is collected and managed by the private sector
  ○ Independent from any SDI development
Indicators

- **Speed of sale**: duration that houses are listed for sale on the market (‘time-on-market’)
  - Proxy = time that a listing is published online

- **Listing price**: price of the real estate listing, this differs with the realized price

➤ Challenge: how to collect the data and enrich the base data from the existing SDI?
Outline

• When is ‘Smart’ really Smart?

• Spatial planning and the housing market

• Web technologies to enrich SDI’s

• Conclusions and ongoing work
Isabelle Loris, based on data from Statistics Belgium (beSTAT.be) and census data (Census2011.be).
Immo (housing) websites
Web scraping

- **Web Data Extraction System**: extract and collect 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)

- **Import.io allows**
  - Extraction in different formats (XLS, HTML, JSON and CSV)
  - You have to ‘teach’ the tool based on at least 5 web pages
    - Extraction and **crawling** supported by the use of XPath
Micro-data real estate listings

• Scraping small(er) independent real estate agents websites across Belgium

• Method was used to collect additional information
  o Lot size, Energy Performance Certificate (EPC), value, building year

• Data cleaning using OpenRefine
  o Selection of relevant and sufficiently documented records
  o Transformation of data to usable formats (listing price, housing numbers, dates, geographic coordinates, etc.)
Indicator calculation

- ‘Speed of sale’ = difference in number of days between the first and last date of publication on the portal website
  - Variables “Pub1 start” and “Pub1 stop”

- ‘Listing price’ = variable ‘initial asking price’
  - Compared with
    - the average selling price tracked by Belgian censuses
    - the average realized price for the entire dataset (2005-2014)
    - Pricing difference (%) =
      \[ \frac{\text{realized price} - \text{listing prize}}{\text{listing prize}} \times 100 \]
Results of the testing

1. Web data extraction
   - webdata → database

2. Exploration
   - web data
   - and testing
   - prototype indicators
   - technical & juridical barriers
   - suboptimal data transfer
   - collaboration is a better option

• Two major issues
  o Web scraping takes much time
    (44h/website)
  o Grey legal zone!

“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
Transdisciplinary approach

• Extending the method: collaboration with one company
  o 10 year database of listings
  o More than 110,000 listings on the Web
Micro-data real estate listings
Speed of sale (# of days for sale)
- × not yet sold
- ● 2 weeks
- ○ 1 month
- ● 3 months
- ● 6 months
- △ from 6 months up to 4 years
- ◼️ urban area

### Speed of sale

<table>
<thead>
<tr>
<th>Speed</th>
<th>[%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>unknown yet</td>
<td>23</td>
</tr>
<tr>
<td>&gt; 6 months</td>
<td>18</td>
</tr>
<tr>
<td>6 months</td>
<td>11</td>
</tr>
<tr>
<td>3 months</td>
<td>19</td>
</tr>
<tr>
<td>1 month</td>
<td>9</td>
</tr>
<tr>
<td>2 weeks</td>
<td>19</td>
</tr>
</tbody>
</table>
Outline

• When is ‘Smart’ really Smart?

• Spatial planning and the housing market

• Web technologies to enrich SDI’s

• Conclusions and ongoing work
Conclusions and further work

• Web scraping methods are working fine but …
  o Still slow – that can be solved
  o Legal considerations

• Complement with a collaborative approach
  o Although the company uses web scraping themselves (😊)

• Combine it with other techniques
  o Design an automated ETL procedure
  o Apply Linked Data techniques

Geospatial data on the web (OGC & W3C)
• **Way forward**
  
  o map dynamics of the housing market in time and space
    • bench marks for a fast or slow speed of sale?
  o evaluate in which neighborhoods the housing market is very dynamic and where it is dull

• The importance of monitoring over longer time series: housing dynamics can change rather fast, where spatial planning is characterized by a certain level of slowness.

  “*Where would it 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?*”
1. Web data extraction

2. Exploration web data and testing prototype indicators

3. Transdisciplinary collaboration
   - users' agreement between research partner and one company
   - technical & juridical barriers
   - suboptimal data transfer
   - collaboration is a better option

4. Future co-production

Smart governance for smart cities

far future self-mobilisation
The Flemish Policy Support Centre on Space (2012-2015) funded this work. The views expressed in this presentation are not necessarily those of the Flemish Government.

Detailed real estate data were kindly provided by Zimmo.

Credits to SADL staff:

Valerie Dewaelheyns
Thérèse Steenberghne
Anuja Dangol
Thank you!