

Publication Output on the Topical Area of "Energy" and Real Estate (Education)

Bob Martens



Outline

- 1. Introduction
- 2. Data sources
- 3. Text mining
- 4. ERES Education Seminar 2017 case

1. Introduction

- Disciplines evolve over time
- Associations like ERES need to track this evolution
 - support and adjust to emerging themes
- How can we identify those changes?
 - How substantial are they?
- Which topics are rising? Which topics are declining?

2. Data sources

- Records of the ERES conferences 2012-2017 as stored in the "conf-vienna" database
 - Submission environment for the annual conference
 - Representation of the scholarly discussion
- Statistics: Titles and abstracts of all accepted papers to these annual conferences
 - 1,780 papers - 452,074 words - 15,399 different words
 - Every word is repeated 29.47 times in average
 - Most frequent word “the” is found 32,000 times

MySQL-view

TABELLEN & VIEWS

-  confaccount
-  confaccountbup
-  confauthor
-  confexcursions
-  confexcurspriorities
-  confhotel
-  conflocation
-  confpapers
-  confparticip
-  confparticipview
-  confpotchair
-  confpotdisc
-  confreviewerprefs
-  confreviewers
-  confreviews
-  confslot
-  confslotloc
-  conftheme

Suche: PAPERNUMBER =

PAPERNUMBER	TITLE	THEME	SLOT	NR	EINGEL
1	Valuing Visual Accessibility of Sce...	L	31010	4	2016-
2	Decision support system for prop...	NULL	NULL	NULL	NULL
3	Land valuation in infrastructure pr...	NULL	NULL	NULL	2017-
4	Refitting Vacancy for the Creative I...	L	29030	1	2017-
5	A stress testing scenario analysis f...	C	29050	1	NULL
6	Specifying cost-influential factors...	L	29050	1	NULL
7	Placing the theory of 'affordances'...	NULL	NULL	NULL	NULL
8	Active innovation resistance - the...	K	30010	2	NULL
9	Measuring the Institutional Eleme...	NULL	NULL	NULL	2017-
10	Assessing the Housing Attributes...	NULL	NULL	NULL	NULL
11	Is Distance Matters to the Elderly?	J	NULL	NULL	NULL
12	Homeownership: Boon or Bane? Both!	NULL	NULL	NULL	2016-
13	Benchmarking of workplace perfor...	O	30070	1	NULL
14	The impact of changing user cost...	NULL	NULL	NULL	NULL
15	Developing a new green office buil...	L	29050	3	NULL
16	Legal aspects of the circular buildi...	NULL	NULL	NULL	NULL
17	Urban and peri-urban residential r...	I	29050	2	NULL

3. Text mining

- An analytical method based on **text** rather than **numbers**
- The method is dependent on language
- It needs to take into account – to some extent – the **complexity of languages**:
 - Text structure (sections, paragraphs), sentence structures, grammar
 - Synonyms (two words mean the same), polysemes (one word with different meanings)
 - Importance of context, irony, satire

Text mining – basic concepts

- Analysis of text in a (large) number of documents
- Text mining requires simplifying assumptions and cannot extract meaning perfectly from text
- Text mining incorporates linguistics, statistics, mathematics & computer science
- **Bag of words model**: Within each document the text is split into words. Every word is treated the same irrespective of its position in the document. Relationships between words are ignored (at this stage).

Text mining – basic concepts

- Not every word is really relevant:
 - Some words occur very often, but do not carry much information. E.g., “the”, “a”, “of”, “and”, “in”. They are **stopwords** and excluded from the analysis.
 - Some words occur in one document but most likely not in any other. Often these are numbers. E.g., “0.3993” will most likely not occur in another document. Such words are also excluded.
 - Some words are different as strings, but mean the same. Differences due to grammar. E.g., “originate”, “originates”, “originating”. Such terms are changed to their common word stem → stemming

Data sources: „cleaning“

rank	term	count	rank	term	count
1	real	3,682	9	data	1,155
2	estate	3,485	10	value	1,138
3	market	2,649	11	development	1,132
4	housing	2,184	12	investment	1,090
5	property	1,851	13	price	1,054
6	paper	1,610	14	model	977
7	study	1,356	15	analysis	934
8	research	1,317			

15 most frequently used words in the corpus:

→ after stopword removal

rank	term	count	rank	term	count
1	real	3,682	9	develop	2,117
2	market	3,539	10	study	1,883
3	estate	3,524	11	building	1,661
4	house	2,990	12	paper	1,644
5	use	2,458	13	model	1,560
6	value	2,318	14	research	1,425
7	property	2,239	15	differ	1,400
8	price	2,194			

→ after stemming

Data sources: „ranking“

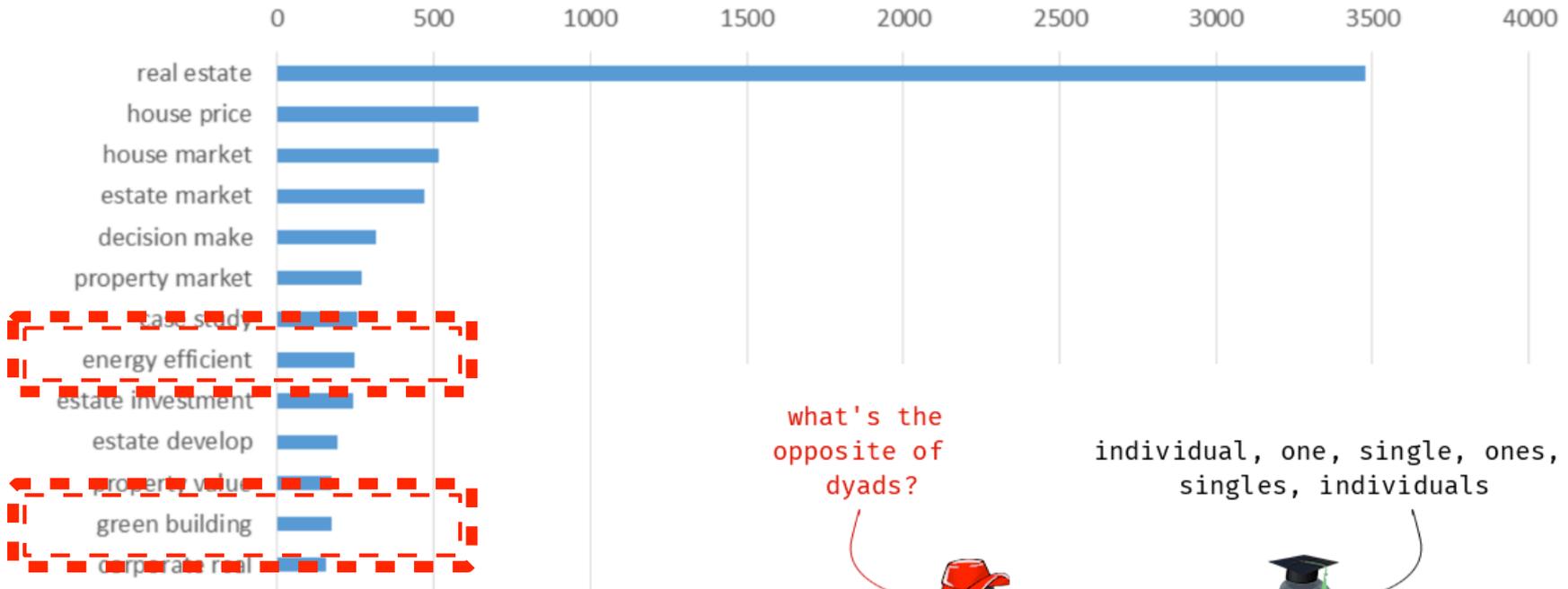
term	rank					
	2012	2013	2014	2015	2016	2017
real	2	1	1	1	1	2
market	1	2	3	3	4	3
estate	3	3	2	2	2	4
house	4	4	9	4	3	1
use	7	5	4	7	6	5
value	6	9	6	5	7	6
property	8	7	5	8	5	9
price	5	6	7	6	8	11
develop	9	10	10	9	9	7
study	11	11	8	11	10	8
building	12	8	22	14	11	10
paper	10	12	11	12	14	15
model	13	14	12	10	22	14
research	18	15	18	13	13	12
differ	14	13	13	15	18	18

The rankings of the terms are surprisingly stable over the years

Text mining – word pairs

- Shortcoming of the "bag of words" approach is that it does not store the relation between terms
- From the fact that “**real**”, “**market**”, and “**estate**” are the most frequent terms (text corpus) **we cannot conclude** that their combination – “**real estate market**” – is a key topic
- Deviation from the bag of words model and identify word pairs → **dyads** – for example “**energy efficiency**”
- Calculating the **frequencies of dyads**
→ Leads to the representation of ERES core areas

Overall frequency of dyads



what's the
 opposite of
 dyads?



individual, one, single, ones,
 singles, individuals



Dyads: Volatility of ranking

- The secondary aspects “energy efficient” and “green building” are much more volatile in their position over the annual conferences (2012-2017)
- “Energy efficient” ranked 20th in 2012, 4th in 2016 and 30th in 2017
- The dyad “green building” ranked 5th in 2012, 38th in 2015 and dropped out of the top 40 in 2017

4. ERES-EDU Case

- Data mining towards "Energy" AND "Real Estate"
- Extraction of data with all the records (2012-2017), in which "energy", "real" and "estate" appear at the same time
- Sorted by a score, which is the product of the respective number of three terms
- 68 papers – "energy" appears 272 times

=====

CONF:ERES2012, PAPERNUMBER: 120, SCORE: 578

TITLE: ARE "GREEN FACTORS" ABLE TO DRIVE THE REAL ESTATE VALUE UP?

ABSTRACT: In the era of ecology awareness sustainable development plays a key role and is not different. It is said that in EU buildings are responsible for 40% of total greenhouse gas emissions, really crucial for climate change. Real estate sector in Poland is aware of the importance of green buildings. Nevertheless green building market in Poland suffers lack of proper interest. The value of the house which reflects environmental costs and benefits. However real estate market despite widely propagated financial and environmental benefits for green buildings.

AUTHOR(S): Sliczna, Malgorzata G., Poland (malgorzatasli@hotmail.com)

=====

CONF:ERES2015, PAPERNUMBER: 152, SCORE: 195

TITLE: Do Smart Grids Increase Real Estate Market Values?

ABSTRACT: Purpose –Buildings energy efficiency is generally considered in terms of energy efficiency goals. It is commonly agreed that the greater the building energy efficiency, the more attractive retrofitting was set in place and simultaneously photovoltaic power plants (PV) are attractive for both institutional and small private investors. Nonetheless Governments can substantially change their energy consumption patterns. In this scenario overall real estate market values, if the energy consumption pattern of the household does not properly performed. Aim of the paper is to investigate whether Smart Grids can give Smart grids give de facto producers and consumers, the opportunity to be active in the energy market. We provide a model based on the real option theory to determine the value of the investment.

Score	Title
578	Are "Green Factors" Able to Drive the Real Estate Value Up?
195	Do Smart Grids Increase Real Estate Market Values?
180	Using spatial econometric models for valuing buildings energy performance
175	The Energy Efficiency of Corporate Real Estate Assets: The Role of Professional Management for Corporate Environmental Performance
144	Corporate Real Estate Incorporating Sustainability
140	Sustainability in existing Real Estate: successful only true creation of threefold mutual business gains
117	The Energy Efficiency of Corporate Real Estate Assets: The Role of Energy Management for Corporate Environmental Performance
112	Does the Energy Label (EL) matter in real estate market? A stated preferences analysis in Barcelona
112	Hunt for the green value
108	The Effect of Energy Performance Certificates on the Value of Buildings
100	Do 'Green' Residential Properties Exist in Wroclaw's County?

Word cloud





Acknowledgements

I would like to express my sincere gratitude to **Gunther Maier**, for maintaining "confvienna" as well as the data mining for this contribution!