Chinese Real Estate - OSINT Analysis

August 04, 2016

Introduction

The connected world calls for a distinctively new approach to the areas of data collection, threat detection and risk assessment. The Internet has become the world's dominating data/knowledge base and represents the information-future of any business or government.

Semantic Visions (SV) is a risk assessment firm based in Prague, Czech Republic. SV runs a complex Open Source Intelligence (OSINT) System, which is in a category of its own and enables solutions to a new class of tasks. Our team has over 10 years' experience in OSINT data collection and automatic understanding of textual information across 11 languages to include Chinese. SV system analyzes and synthesizes 90% of the world's meaningful online news content in real time.

SV system enables customers to produce both operational and executive level strategic analysis of entities of interest or even entire countries. As we have successfully solved the problem of distinguishing critical signals from irrelevant noise and deriving knowledge out of billions of reports. SV system is completely independent of, and far above, Google or any other search engine. SV innovations have won multiple global technology awards.

The goal of this analysis was to prove or disprove our hypothesis that intelligence based on information coming from local sources in local language (in this case Chinese) is of higher value with respect to Chinese space-time than intelligence based on global information published in English.

1. Input Data

The period analyzed	**January 2015 – May 2016**
Total number of analyzed documents	**223.6 million**
Analyzed documents in Chinese language	**20,479,743**
Analyzed documents in English language	**203,125,685**
The number of documents relevant to the Topic: Construction & Real Estate	
Set 1: English language + China + Topic	**104,842**
Set 2: Chinese language + China + Topic	**694,718**
Set 3: Domain .cn + China + Topic	**253,176**

Chinese Public Holidays with Impact on Data Quantity

November 18 - 24, 2015	Chinese New Year (Lunar)
April 5 - 6, 2015	Ching Ming Festival
May 1,2015	Labor Day
June 22, 2015	Tuen Ng festival
October 1 - 10,2015	National Day of the People's Republic of China

2. Methodology

During the entire analyzed period, SV Data Acquisition collected over 418 million documents in 11 languages.

This report is limited to English and Chinese and all the acquired data in these two languages were semantically analyzed in SV semantic engine. The resulting semantic metadata enabled us to thoroughly analyze the chosen area of interest. In this particular case, we were looking for the topic of Construction & Real Estate, which was detected in the data as a pre-defined semantic concept..

Data Sets

Within the scope of the analysis, we were working with 40 various data sets. While all data sets which were based on the concept of Construction & Real Estate, they were combined with various other semantic features: language, source internet domain and region. The report focuses on the following three most relevant data sets:
- Set 1 - Construction & Real Estate + region CHINA + English
- Set 2 - Construction & Real Estate + region CHINA + domain .cn
- Set 3 - Construction & Real Estate + region CHINA + Chinese[1]

The three sets contain relevant quantitative data aggregated by day. Within the sets we focus on various types of sentiment. Sentiment is another type of semantic concept detected by SV engine. We apply a multitude of approaches to the sentiment analysis and each approach results in sentiment attributes, which can be combined to completely describe the sentiment of the analyzed content. SV sentiment analysis detects positive, negative or neutral opinions, tones and events in text. It is much more than just a simple matching of "good" or "bad" words from a word list.

For the purpose of detecting potential correlations, the data in this report collected by SV are extended by a time series of the TAO Real Estate Stock Index which originates at http://www.nasdaq.com/.

---

[1] Simplified Chinese, which is the character system used in the People's Republic of China

3. Analysis

First we would like to point out and describe a few general attributes of the analyzed data: the week cycles and Chinese National Holidays, both of which cause a significant reduction in topic coverage. The Figure 1 clearly shows the repeating pattern of the five workdays with high coverage of the analyzed topic followed by two-day weekends with low coverage. In addition, Figure 2 also shows a larger gap in coverage, which corresponds to the Chinese New Year holiday. While the week cycles are identical for documents in English and Chinese data sets, the Chinese National Holidays are visible only in Chinese data sets.
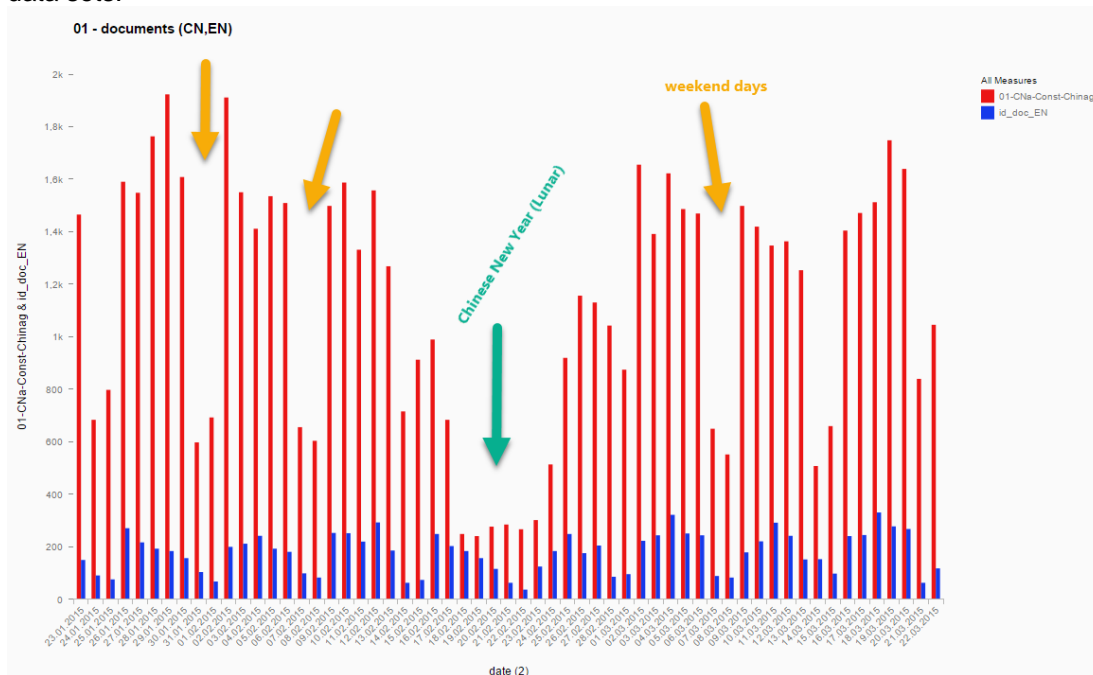
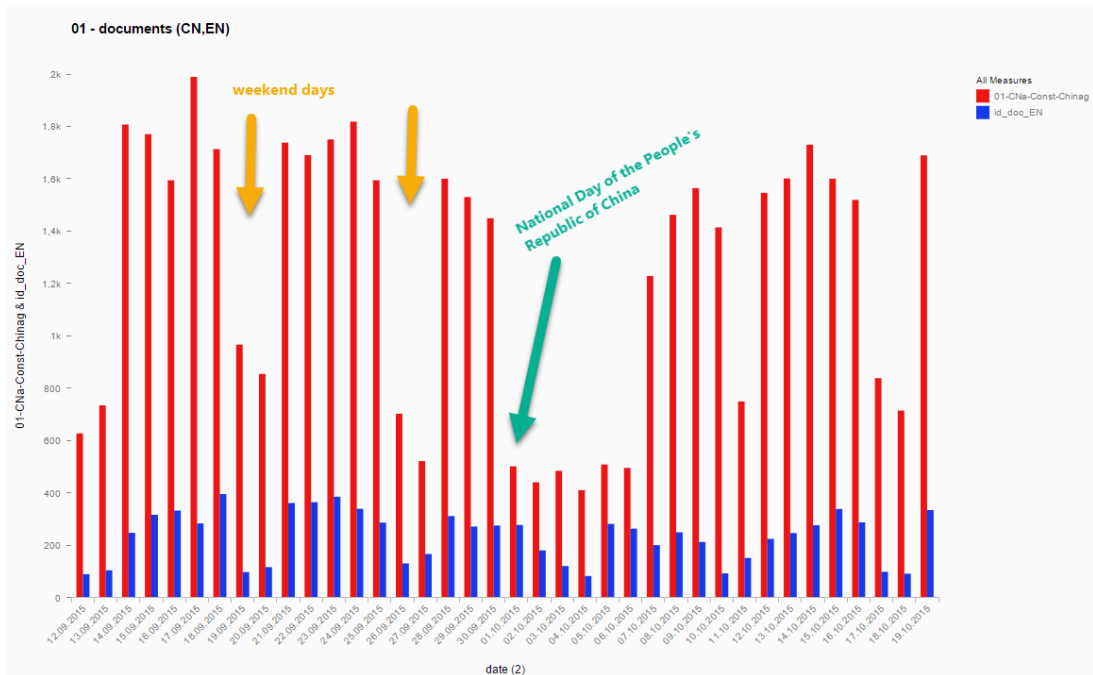Figure 1: Week cycles and Chinese New Year (Chinese: red / Set 3, English: blue / Set 1)

Figure 2: Week cycles and National Day of the People's Republic China (Chinese: red / Set 3, English: blue / Set 1)

Our analysis focuses on periods of non-standard coverage, and we explore the reasons and events that could affect the numbers of documents in these periods.

The largest coverage in Chinese language occurred on March 30, 2015. For this reason, we have selected the period around this date and subjected it to detailed analysis, and examined it in the context of the whole year 2015.
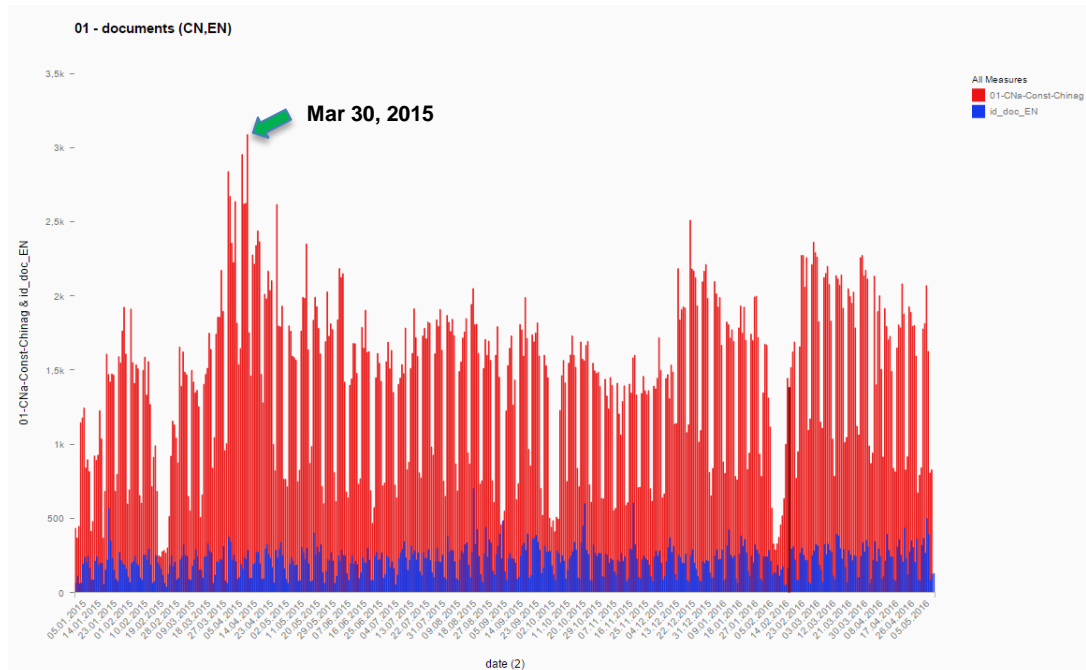

Figure 3: The number of analyzed documents / per day (Chinese: red / Set 3, English: blue / Set 1)
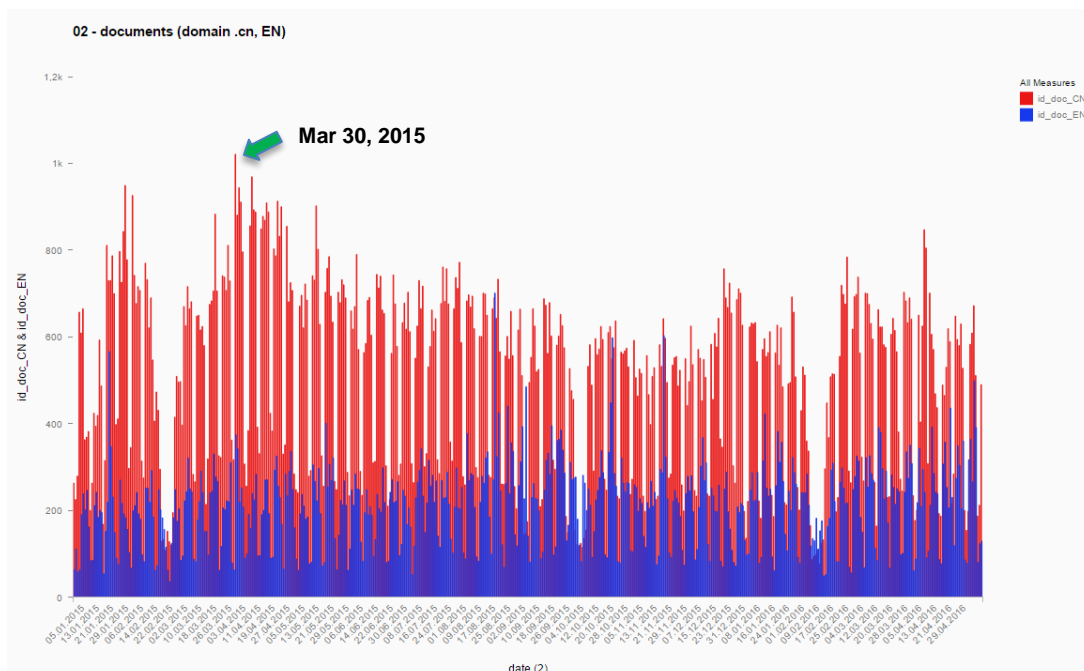

Figure 4: The number of analyzed documents / per day (Chinese: red / Set 2, English: blue / Set 1)

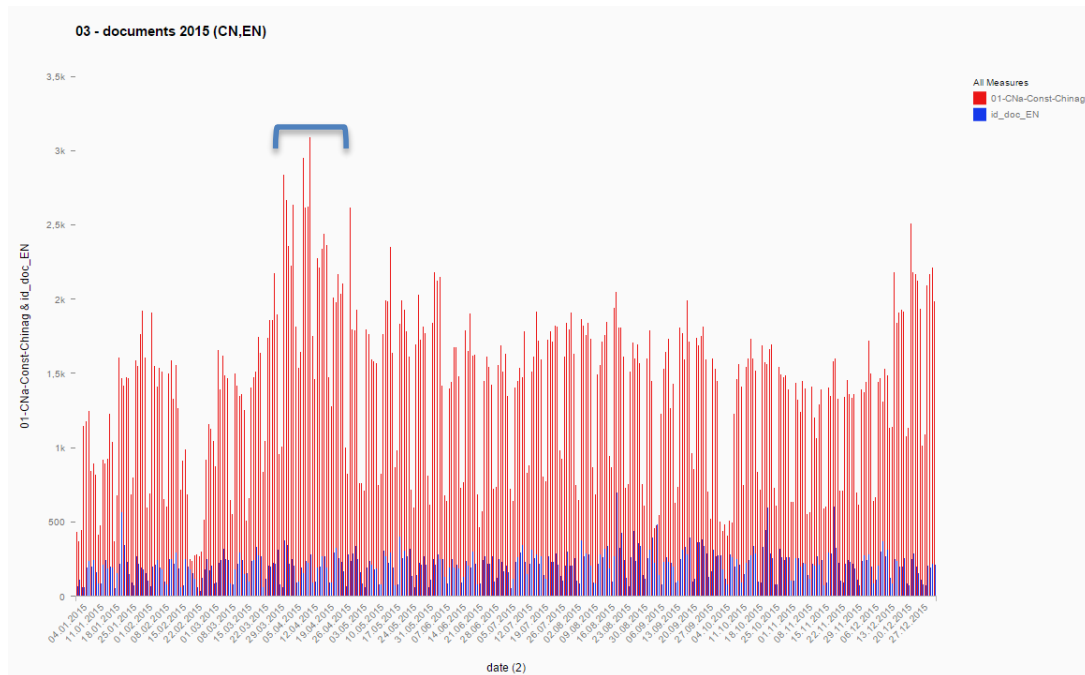In the next step, we zoomed down and worked with selected dates of the given year.

Figure 5: Mar 26. – Apr 30, 2015 Increase in the number of documents in Chinese but not in English
(Chinese: red / Set 3, English: blue / Set 1)

At the end of March and during April 2015, we see the increase of the number of documents with positive sentiment.
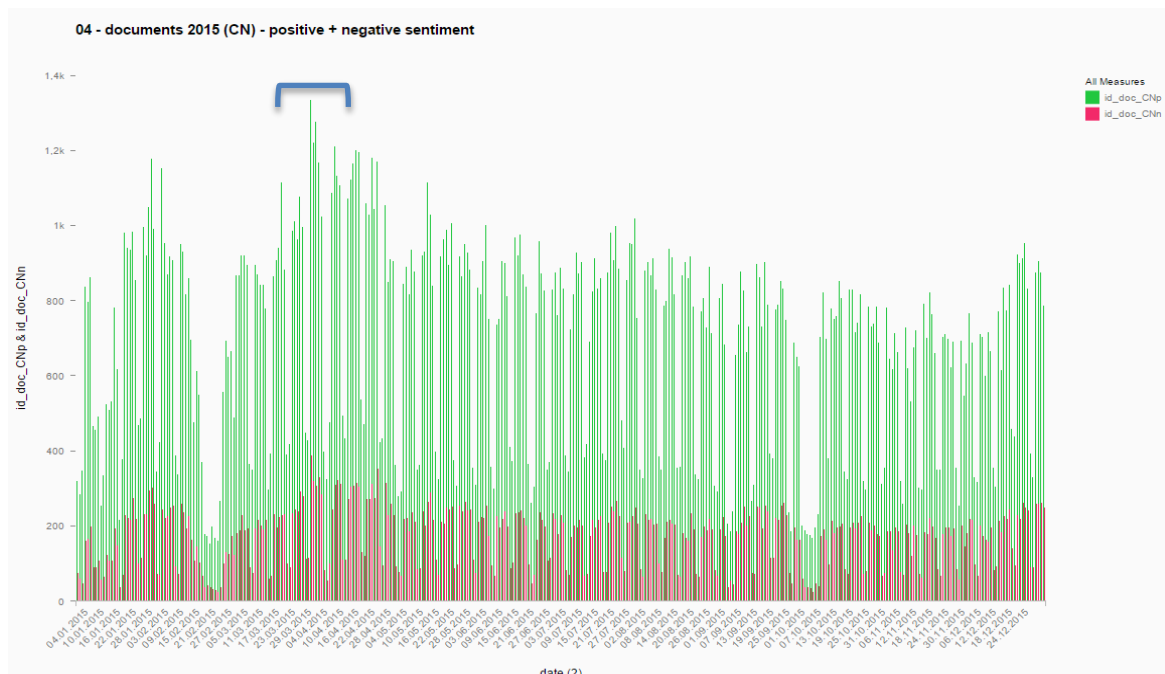

Figure 6: The number of documents with positive (green) and negative (red) sentiment during 2015 in Chinese / Set 3.

The number of documents with negative sentiment in this period does not fluctuate significantly and their numbers are at a similar level throughout 2015. We see that the volume of positive news coverage mostly creates the peak.
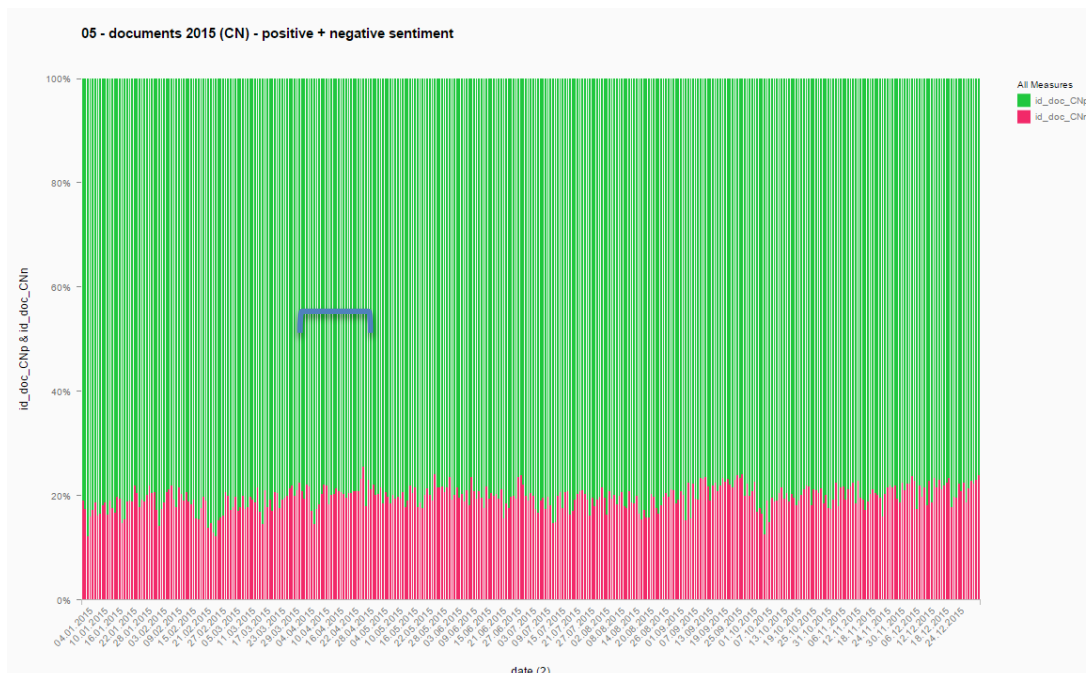

Figure 7: The number of documents with positive (green) and negative (red) sentiment during 2015 in Chinese / Set 3. (normalized stacked chart)

The identified "peak" in March 2015 can be attributed to the new policy of the Chinese Government in the Real Estate market. This new policy announced China's Central Bank together with selected Chinese Government Departments on 27 March 2015. The objective of this new policy was to reduce the cost of mortgages, to provide incentive for buying new properties by more people, and to support the real estate market.

In our analysis we also expected that this new Chinese Government policy in the Real Estate industry also reflected on stock markets. In order to verify this hypothesis, were selected the stock index Guggenheim China Real Estate ETF TAO and looked for correlations.

Stock index TAO:

- Guggenheim China Real Estate ETF (TAO) – stock index focused on the Chinese Real Estate market, which includes more than 50 companies from this industry
- 5 companies with the biggest effect on this index:

Ticker	Name	Shares / Par Value	Market Value	Weighting
823	LINK REIT	111,136	$797,519	5.96%
4	WHARF HOLDINGS LTD/THE	121,306	$775,863	5.79%
83	SINO LAND CO LTD	443,474	$758,285	5.66%
16	SUN HUNG KAI PROPERTIES LTD	55,195	$735,938	5.50%
1113	CHEUNG KONG PROPERTY HOLDINGS LTD	101,500	$681,252	5.09%

In the first quarter of 2015, the value of TAO stock index did not change significantly. During March, it decreased and then in April it begins to rise. The decline of the index value may be associated with the expectations of the new policy of the Chinese Government, and the subsequent rise began after the release of the policy.
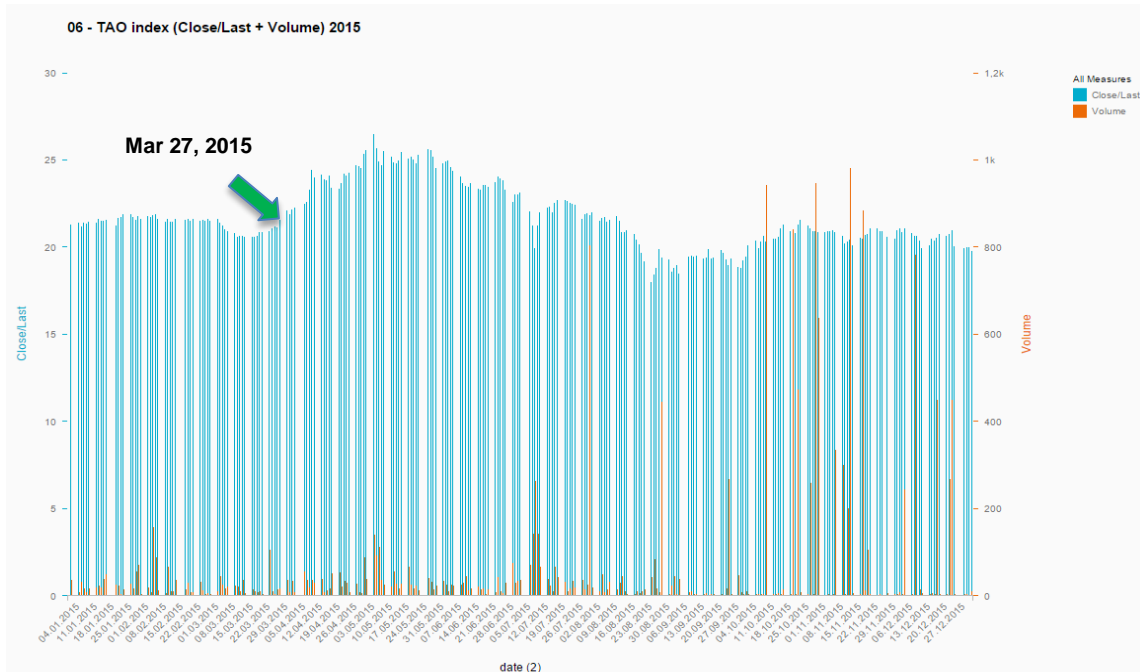
TAO reached the maximum on May 4, 2015.

Figure 8: Stock index TAO in 2015

When we compare the value of TAO index to the number of documents, we can be observe the increase of both values in the following time sequence.

During the course of the entire analyzed period (2014-2016), TAO index maximum was in May 2015, and this growth began just after the release of the new policy of the Chinese Central Bank and the Chinese Government in the mortgage and real estate market.
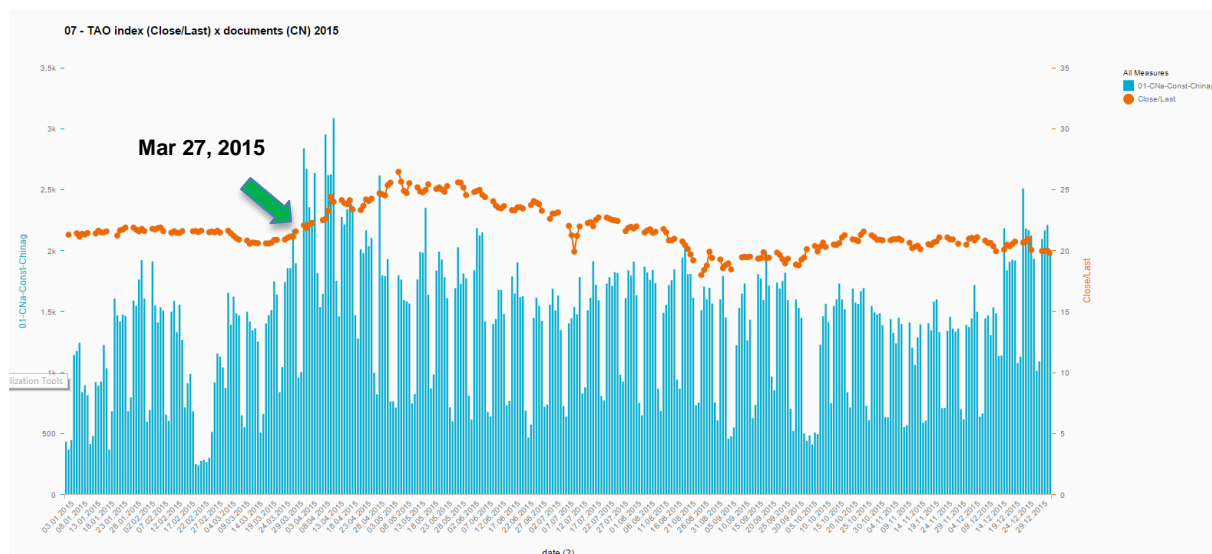

Figure 9: Stock index TAO x number of documents in Chinese language (Set 3)
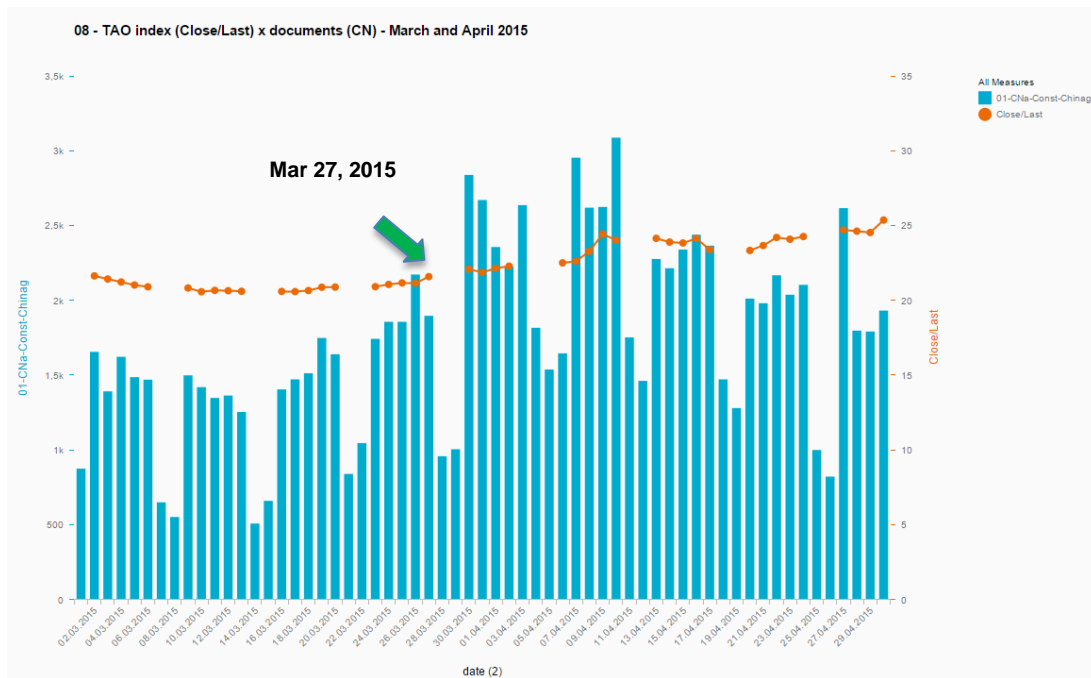
Figure 10: Stock index TAO x number of documents in Chinese language (Set 3) - March and April 2015)

4. Summary

At the end of March 2015, the Chinese Government and Bank of China jointly released a policy change in the Real Estate Industry. While the impacts of this policy change on the news coverage in Chinese Sources were significant, the response in global English language sources was marginal.

From business prospective, an even more important fact is that the information coming in Chinese language from Chinese local sources correlated with the Chinese Real Estate Index (TAO), while there was no correlation with the global English news coverage.

Brief description of the detected Chinese mortgage policy change:
- The main change was a reduction of minimum deposit when buying a second property from 60-70% down to 40%
- The seller is exempt from tax if the property is owned for more than two years (previously it was 5 years)

Impacts of the detected Chinese mortgage policy change:
- It was followed by a significant increase of news coverage of Real Estate in local Chinese sources.
- The whole increase consisted almost exclusively of positive news coverage in local Chinese sources.
- There was practically no increase of negative sentiment news coverage in local Chinese source;
- There was a correlating increase of the Real Estate stock index "TAO"
- In global English language sources, however, there was no discernable increase in news coverage of Real Estate in China (although information about the policy change did make it to the news).

Conclusion:
- Local information has an essential role in business related intelligence and business decision-making processes.

Semantic Visions delivers actionable business intelligence based on local sources.

5. Annex: Selected example articles from Chinese sources (in Chinese language)

Note: For ease of reference, the article titles were translated into English.

- How warm wind coming down the property market policy

  政策暖风吹 楼市未来怎么走

- Bank: two suites mortgage interest rates is expected to return a reference

  银行：二套房按揭利率有望回归基准

- Chinese real estate market overjoyed new decree on the economy is expected to boost short-term effect

  中国房地产新政令市场喜出望外，对经济料有短期助推效果

- Property Rates New Deal will repeat 2009's?

  楼市新政出台 房价会否重演 2009 年的"疯狂"？

- Mortgage Deal or pro-hand housing transactions more active to improve demand release

  房贷新政或促二手房交易更活跃 改善性需求释放

- Two suites down payment loans fell more than 4 percent of 2-year second-hand housing exempt from sales tax

  二套房贷款首付比降至 4 成 2 年以上二手房免征营业税

- Property-than-expected new regulations loosening efforts to help release of rigid demand

  房产新规松绑力度超预期 有助释放刚性需求

- New Deal "after another" is intended to promote the improvement of housing demand release

  楼市新政"迭出"意在促改善性住房需求释放

- Ministries jointly push the New Deal to the inventory of the property market is the main theme of this year

  两部委联手推新政 去库存是今年楼市主旋律