Nation image and its dynamic changes in Wikipedia

Youngwhan Lee
Department of Management of Technology, Konkuk University,
Seoul, Republic of Korea, and

Heuiju Chun
Department of Statistics and Information, Dongduk Women’s University,
Seoul, Republic of Korea

Abstract
Purpose – In this volatile and increasingly fast-revolving world, it has become crucially important to monitor, measure and manage nation image and its dynamic changes in real time. However, few studies have been conducted on a model to measure the image and/or its changes. The purpose of this paper is to find an economically affordable methodology to measure nation image and its changes online in real time.

Design/methodology/approach – The study took an approach to build dynamic ontology that may reflect to change nation image in real-time. With it, the authors attempted to measure nation image in real-time.

Findings – Among many social media, the authors found that Wikipedia is particularly suitable for the purpose of measuring nation image. An ontology of nation image was built from the keywords collected from the pages directly related to the big three exporting countries in East Asia, i.e. Korea, Japan and China. The click views on the pages of the countries in two different language editions of Wikipedia, Vietnamese and Indonesian were counted.

Originality/value – The study confirms the objective: the data from a social media service, Wikipedia, may work very well as an economically affordable real-time supplement to offline nation image indices that are currently used.

Keywords Vietnam, Ontology, Indonesia, Social media, Nation brand image, Wikipedia click views

Paper type Research paper

Introduction
As many countries compete fiercely in trying to draw investments and travelers from foreign countries and to entice customers in overseas to export more products and goods than before, governments and their commissioned entities attempt to manage their nation brand image to maintain the competitiveness over competing nations and countries (Kotler and Gertner, 2002). According to Park (2010), a nation image must be managed because it plays an important role as an external clue that influences customers’ decisions at the time of purchase. Especially, several previous studies on the country-of-origin effect (also known as COE) showed that consumers often judge a product’s quality, its value and its perceived-risks by nation image and COE. Those significantly affect consumers’ purchase decisions.
(Kaynak and Kana, 2000; Tse and Gorn, 1992; Wang, 1978). Furthermore, Cho et al. (2007) found that a nation image has positive relationship with the evaluation of a product’s quality, regardless of its brand image. He also found to be true that the favorable attitudes to the country of origin show positive relationship with image evaluation.

Traditional methodology of measuring nation image has been showing some limitations because it relies heavily on person-to-person interviews. As measuring nation image has to be done on a global scale, it is prohibitively difficult and expensive. Researchers often compromise with the reality. Somewhat mediocre outputs have been produced.

To make the situation worse, nation image changes dynamically continuously. Ideally, it must be monitored repeatedly in real time. To overcome these limitations, some consulting firms maintain over tens of thousands panel pools around the world to run near real-time surveys. Maintaining tens of thousands panel pools in global scale should be very difficult and costly.

A new methodology using big data, in response to such needs, of measuring nation image online in real time is desperately sought: it must produce reports analyzing people across the world anytime needed; it must be less costly than its traditional counterpart, and, lastly but not the least, it must be applicable to different domains.

Big data analytics are frequently mentioned topics in research and practice (Buhl et al., 2013). The term “Big data” is referred to the ongoing expansion of data in terms of volume, variety, velocity and veracity (IBM, 2012). Big data analysis widens and expands the scope of traditional business intelligence (BI) which focuses on integrating and reporting structured data lied in company-internal database, by seeking and adding to extract values from semi-structured or unconstructed data originated from Webs, social media and blogs that are external to the company (Debortoli et al., 2014). Big data offer enormous opportunities and also many challenges for businesses (Buhl, 2013). To related with big data, the Mckinsey Global institute sates that the USA alone faces a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts based on their findings (Manyika et al., 2011). The possibilities of big data will continue to evolve rapidly, driven by innovation in the underlying technologies, platforms and analytic capabilities for handling data, as well as the evolution of behavior among its users as more and more individuals live digital lives (Manyika et al., 2011).

This paper presents an attempt to measure nation image analyzing utterances for nations using unstructured data. The researchers made an attempt to measure image from the data that can be obtained in social media and compare them to image collected by traditional offline surveys. Among many social media service providers, the data from Wikipedia well satisfy the constraints for the purpose. Unfortunately, nation image from Wikipedia is partial only limited to its awareness. Even so, the data include surprisingly rich ontological information about nation image in both breadth and depth that offer different aspects to the ones from traditional methodology. The researchers believe that the image driven from Wikipedia can be used as an online real-time supplement to offline expensive traditional models measuring nation image.

1. Literature review
Image of a nation may be a determining factor for travelers to select the countries to travel. Good nation image works as an absolute and necessary element in enticing international capital investments and/or attracting best possible workforce from overseas. Therefore, many governments around the world try to manage their nation image as good as they can be (Park, 2010).
Nation image is defined to be an individual item or a compound of the nation name, emblems, designs, quality, policy, awareness, etc., which differentiate a nation from other nations in the areas of foreign policy, culture, exports, immigration and investments, tourism and so on. Although many have attempted to develop a model to measure nation image and analyze the assets belong to them, neither a standard nor a consensus has been reached.

In general, despite no agreed standard, the index proposed by Anholt (2005) is the first of analytical ranking of the world’s nation brand. Anholt’s NBI (also known as Alholt-Gfk Roper Nation Brand Index) is a result of a survey that is run on a worldwide panel of 28,000 consumers in 35 countries on the perception of nations in cultural, political, commercial and human assets, investment potential and tourist appeal. According to Anholt, nation brand is the sum of people’s perceptions of a country across the six areas of nation’s competence:

Every country has its brand strengths and weaknesses, so there is a different “winner” for each point of hexagons. The overall “Top Nation Brand” is the one with the highest marks across all points of the hexagon (Anholt, 2008).

The Anholt NBI measures the power and appeal of nation’s brand image and obtains the information how consumers around the world see the character and personality of the brand. Also, Country Brand Index (CBI) is maintained by a global brand consulting company, Future Brand (2008). This model surveys a qualitative analysis, expert opinions and statistical data to compute the national rankings. The survey uses “hierarchical decision model” to measure the steps of awareness, familiarity, associations, preference, consideration, decisions/visitations and advocacy.

Samsung Economic Research Institute (SERI) and Presidential Council of Nation Brand (PCNB) in Korea jointly developed Nation Brand Double Octagon (NBDO) measure substance and image: Substance is measured using 125 statistical data offered by many institutes and governments in the world such as IMD, WEF, World Bank, UNESCO, etc.; image is computed from a survey from 13,500 opinion leaders in 26 countries (Lee, 2012).

Although offering researchers, marketers and policymakers rich insights, Anholt NBI, Future Brand’s CBI and SERI-PCNB’s NBDO models all share similar limitations:

- Collecting and maintaining a global pool of consumers is prohibitively costly. Although it can be done and is being done, customers of the companies, which are countries and cities, may have difficult time to justify the expenses. An affordable way to analyze a nation brand is needed.
- The traditional models are for offline and hardly provide real-time analysis. In the rapidly changing global economy, a true sense of real-time analysis for nation brand is desperately needed.
- Perhaps the most critical issue among the above limitations so far, the models are designed for off-line interviews and surveys and not well suited for on-line analysis. A model for on-line analysis for the data that can be obtained in the internet must be developed.

A methodology for nation brand analysis to overcome the limitations listed above must meet the following constraints:

- The resulting reports must depict the nation brand as good as Anholt NBI and Future Brand CBI models would. That is, the nation image must give as insightful, useful and important information as traditional nation brand image did, to researchers, marketers and policy-makers.
2. Analysis of country image in Wikipedia

To develop a model to analyze country image, the characteristics of data obtained from various social media are discussed in the Subsection 3.1. From this, the reason why Wikipedia is determined to be suitable for the study is discussed.

Wikipedia offers two different kinds of data: page edits and click views. In the Subsection 3.2, the researchers discuss why and how the page edits used in pursuit of developing a model satisfying constraints for this study. The page edits were used to construct an ontology model. The discussion includes how it was built from the edits, from Wikipedia English edition, directly related to countries in focus. The countries are chosen for the study are the big three exporting countries in East Asia, South Korea, China and Japan.

Having the set of ontology built from Wikipedia English edition, it was translated and applied to categorize the page views in Wikipedia Vietnamese and Indonesian editions. As the page views indicate how people perceive about focused countries, the researchers attempted to have a close look into the page views of the big three exporters and compare them against each other. From this, the researchers conclude that they show the awareness of nations among the users of a Wikipedia language group. Although the data sets themselves from Wikipedia are somewhat limited, because word selections are neutral, they give insightful information, and it is worthy for nations to monitor and manage. The researchers propose to call them the awareness of a nation image.

2.1 Social media services and the characteristics of their data

To compute a real-time online nation brand image or something similar, the characteristics of social network services in a global scale are investigated. They are Facebook, Twitter, YouTube and Wikipedia.

The data from Facebook are not easy to draw because it allows a third party business application to access to only the one that obtains a user's permission. As the research team cannot even get the data, Facebook is ruled out. As for YouTube, it is difficult to analyze video clips with comments. YouTube for analysis is also ruled out.

Twitter data are readily available, as tweets can be easily collected and analyzed, although difficulty exists as a large portion of the data uses cants and slangs in addition to acronyms and emoticons. The most serious problem of all is that the data show users' turbulent emotional cornering that are heavily inconsistent. Such cornering and inconsistency may be useful when the emotional changes of a group are the concern. Therefore, Twitter for the research is also ruled out.

As for Wikipedia, the data are consisted of two kinds: the number of edits and the number of clicks:

(1) The number of edits is the concerns of experts which effectively close to experts’ opinions in Delphi method; and

(2) Editing in Wikipedia is tightly controlled to ensure the quality of article content.
Wikipedia strongly encourages editors to take neutral point of view:

Editing from a neutral point of view (NPOV) means representing fairly, proportionately, and, as far as possible, without bias, all of the significant views that have been published by reliable sources on a topic. All Wikipedia articles and other encyclopedic content must be written from a neutral point of view. NPOV is a fundamental principle of Wikipedia and of other Wikimedia projects. This policy is nonnegotiable and all editors and articles must follow it (http://en.wikipedia.org/wiki/Wikipedia:NPOV).

For this reason, Wikipedia articles present words showing neither emotional affinity nor loyalty, which are the attributes that general nation brand indices depend on. Concerning this, the researchers decided to limit the study only to measure what is dictated by the data of Wikipedia. This limitation turned out to be a blessing for the study because the fact that the data set is more objective and consistent than other data sets from many other social media, such as Twitter, added significant and exciting values to the final output of the study.

While the number of edits shows experts’ opinions, the number of clicks shows the concerns of average users.

Also, while the number of edits may be arbitrary and controlled, the number of clicks is spontaneous and voluntary. In measuring the dynamic shifts of average people’s concerns (or awareness) to a nation, the researchers find the number of clicks ideal.

2.2 Nation image Wikipedia ontology

Traditionally, the equity types of country brand proposed and accepted by many are composed of awareness, performance, emotional affinity, and loyalty. Awareness represents familiarity to a country, while performance measures qualitative evaluation. Emotional affinity means a user’s personal liking or sympathy to a country. Finally, loyalty comprises visiting loyalty and purchasing loyalty, which means that a person wants to visit a country and to buy a country’s product, respectively.

To make sure that the model constructed from this study is sound and complete, models used in other approaches were inspected, compared and incorporated. In particular, the categories used in the models of future brand CBI, Anholt’s NBI (Nation Brand Index), city brand index and SERI-PCNB’s NBDO were carefully examined.

As shown in Figure 1, the researchers measured the number of edits in Wikipedia English edition. The most frequently edited pages directly related with the focused countries from the top are collected. Four hundred keywords for each country and 1,200 keywords altogether were collected. The keywords then were mapped into four categories: politics, society, economy and culture.

Future brand CBI was partly used because it was not directly comparable to all other models.

Each category was then divided into subcategories. For example, the politics category is divided into president, military, government, administration and foreign relation. The society category is divided into geography, education, history, language, religion and people. The economy category is divided into industry, welfare, finance, infrastructure and science and technology. Finally, culture is divided into art, music, film, sport, cuisine and entertainment. Considering the subcategories, the edited pages (i.e. keywords) were divided into 22 subcategories which became the building blocks of an ontology set named Nation image Wikipedia (NI-Wiki) ontology.

Then, the number of page views of each article was collected. As stated, the number of edited articles represents experts’ interests about a nation, while the number of page views
represents general interests in it. The articles are classified according to NI-Wiki ontology. NI-Wiki ontology maps the Wikipedia articles into both the four main categories and the specialized classification mentioned above. The sum of page views of each category shows awareness of the category.

2.3 Comparisons of South Korea, China and Japan

2.3.1 Nation image Wikipedia ontology to the page views of Wikipedia English edition. The page click views during a year period from June 2012 to May 2013 are counted. Table I shows the page views of each country. Interestingly and somewhat surprisingly, the page views of Japan are the lowest. The views about China show as twice as those about Japan.

As for the background information, the total numbers of pages related to the three nations in the English Edition of Wikipedia are as in the Table II. It shows South Korea has the lowest number of pages. The pages about China are as three times as the ones about South Korea, whereas the ones about Japan is as twice as them.

As shown in Figure 2, NI-Wiki ontology extracted from English edition, three countries have different page views according to the four main categories. Among the main categories, South Korea shows politics 37 per cent, economy 1 per cent, society 33 per cent and culture 29 per cent; China shows politics 57 per cent, economy 1 per cent, society 36 per cent and culture 6 per cent. Finally, Japan shows that politics 14 per cent, economy

<table>
<thead>
<tr>
<th>Country</th>
<th>Politics</th>
<th>Economy</th>
<th>Society</th>
<th>Culture</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Korea</td>
<td>100.8</td>
<td>2.8</td>
<td>81.1</td>
<td>92.0</td>
<td>276.8</td>
</tr>
<tr>
<td>China</td>
<td>240.4</td>
<td>6.4</td>
<td>150.4</td>
<td>25.7</td>
<td>422.9</td>
</tr>
<tr>
<td>Japan</td>
<td>19.0</td>
<td>2.4</td>
<td>148.7</td>
<td>29.1</td>
<td>199.1</td>
</tr>
</tbody>
</table>
2 per cent, society 62 per cent and culture 22 per cent. All of the three countries show the lowest number of page views in economy. The highest page views are all different in each country. In the case of South Korea and China, politics has the highest, while in the case of Japan, society has the highest. Also, it is worthy of attention that in the case of South Korea, politics, society and culture have similar page views:

Focusing on the culture category, Figure 2 shows that heavy click views for South Korea compared to the views for Japan and for China. The researchers were able to find several reasons. Some of them are well-known to many: for example, Gangnam Style, a Korean pop music, was a world-wide sensation for the period the study was taken.

Some are not well-known and even surprising to many: for example, Running Man, a weekly entertainment series from a South Korean TV station broadcasted in Korean, gets enormous click views week after week as shown in Figure 3. In fact, it is ranked 602nd in traffic among all pages written in English Edition of Wikipedia. Running Man is not the top entertainment program series in Korea, but it is on the top of the many other competing programs in its kind in the world. This may be unknown to most of the people around the

Table II.
The pages in the English edition of Wikipedia (as of May 31, 2013)

<table>
<thead>
<tr>
<th></th>
<th>South Korea</th>
<th>China</th>
<th>Japan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total pages</td>
<td>43,899</td>
<td>132,566</td>
<td>107,832</td>
</tr>
</tbody>
</table>

Figure 2.
NIWI-English: South Korea, China and Japan

Notes: (a) South Korea; (b) China; (c) Japan

Figure 3.
Running Man click views

Running Man (TV Program) click views
world. Perhaps, even the people producing the TV series may not know what is happening in Wikipedia, therefore in the world. This kind of phenomena is something that cannot be easily pinpointed because the popularity is NOT as big a global syndrome as seen in Gangnam Style. It is moderate and yet significant enough to look into because it is enough to change the dynamics of a nation image. This kind of popularity that changes in a nation image cannot easily be spotted.

Figure 4 shows that the sub-categories of the culture category comparing the NI-Wiki-English of the three countries reveal detailed information. South Korea gets more click views in the music sub-category than the other two, whereas China gets more in film and cuisine.

2.3.2 NI-Wiki ontology to the Vietnamese and Indonesian edition. NI-Wiki ontology was translated and applied to Vietnamese and Indonesian Editions to obtain NI-Wiki Vietnamese and Indonesian for the three countries. The researchers attempted to show how the countries are perceived differently by the people in the language groups.

Table III shows the Vietnamese Edition. The people using Vietnamese language clicked more on China than the other two countries. Among the three nations, South Korea shows that the nation is very strong in its culture category.

The click views for the period from June 2012 to May 2013 show the significant differences ($p < 0.0001$). This is visualized in Figure 5.

To visualize the differences of awareness, the researchers used multiple corresponding analysis. Figure 6 shows the result of the corresponding analysis in Wikipedia Vietnamese Edition. As shown in the Figure 6, the three countries get different number of clicks for different subcategories to the users of the Edition: Korea gets more clicks, therefore more awareness from the users, on music, finance and entertainment than the other two; Japan on military, infrastructure and cuisine; China on language, geography, people, foreign relations and so on.

Table IV shows the numbers of click views in Wikipedia Indonesian Edition of the big three countries. It shows the three countries show significant differences ($p < 0.0001$). The people using Indonesian language click more on Japan.

As for South Korea, society and culture categories show large portions of click views at 52.8 and 40.4 per cent, respectively. On another hand, China gets a large portion of click views in the society category.

<table>
<thead>
<tr>
<th></th>
<th>Politics</th>
<th>Economy</th>
<th>Society</th>
<th>Culture</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Korea</td>
<td>62748.1</td>
<td>93201.2</td>
<td>1310148</td>
<td>1852017</td>
<td>3318384</td>
</tr>
<tr>
<td>China</td>
<td>426225</td>
<td>66352.3</td>
<td>435523</td>
<td>325632</td>
<td>5173532</td>
</tr>
<tr>
<td>Japan</td>
<td>118294</td>
<td>54678</td>
<td>1385068</td>
<td>479222</td>
<td>2037262</td>
</tr>
<tr>
<td>Total</td>
<td>607267</td>
<td>214231</td>
<td>705809</td>
<td>2656871</td>
<td>10529178</td>
</tr>
</tbody>
</table>

Notes: Unit in thousand, %; Chi-square = 2.74e6; df = 6, $p < 0.0001$
views to society at 82.6 per cent, followed by culture, politics and economy at 13.9, 3.0 and 0.5 per cent, respectively. As for Japan, it shows society at 78.9 per cent, culture at 19.3 per cent and politics at 1.7 per cent. In summary, clicks on South Korea show balanced clicks between culture and society but China and Japan show.

Figure 5.
Number of click views of the big threes in Wikipedia Vietnamese edition

Notes: (a) Japan; (b) China; (c) South Korea

Figure 6.
Corresponding analysis of the click views in Wikipedia Vietnamese edition
Figure 7 visualizes the number of click views in Wikipedia Indonesian Edition. As shown in it, South Korea is more balanced between society and culture categories than other two countries.

To visualize the differences of awareness, the researchers once again used corresponding analysis. Figure 8 shows the result of the analysis in Wikipedia Indonesian Edition. As shown in the Figure, the three countries get different number of clicks for different subcategories to the users of the Edition: Korea gets more clicks, therefore more awareness from the users, on entertainment, film, sport, and geography than the other two; Japan on music and military; China on science and technology, language, people and so on.

3. Conclusion and future studies
In this study, the researchers attempted to find a way to measure nation image online real time, whereas the traditional way heavily depend on offline face-to-face surveys.

Considering data from many different social media, the ones from Wikipedia is particularly well-fit for nation image. The data from Wikipedia are somewhat limited because Wikipedia encourages its editors to take neutral positions on issues. Even though the selected words are neutral, data from Wikipedia offer insightful information, which is worthy to give an attention. The researchers attempted to find a way to measure an online real-time nation image that may supplement traditional nation image.

Nation image from Wikipedia is partial only limited to click views. Even if partial, it includes surprisingly rich ontological information in both breadth and depth that help to understand how people in a language group perceive about a country. It gives as insightful, useful and important information as traditional nation image would offer.

Taking a close look into the click views of the big three exporters and comparing them to each other, the researchers found that they offer interesting facts in the Wikipedia editions. The corresponding analysis revealed that the users of Vietnamese edition and the users of Indonesian edition have difference in their perception toward the big three exporting countries chosen in the study as the number of clicks are somewhat different.

<table>
<thead>
<tr>
<th></th>
<th>Politics</th>
<th>Economy</th>
<th>Society</th>
<th>Culture</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Korea</td>
<td>64,023(3.1)</td>
<td>74,996(3.7)</td>
<td>1,080,683(52.8)</td>
<td>828,044(40.4)</td>
<td>2,047,746(100)</td>
</tr>
<tr>
<td>China</td>
<td>48,181(3.0)</td>
<td>7,424(0.5)</td>
<td>1,323,658(82.6)</td>
<td>223,037(13.9)</td>
<td>1,602,300(100)</td>
</tr>
<tr>
<td>Japan</td>
<td>71,355(1.7)</td>
<td>3,417(0.08)</td>
<td>3,289,958(78.9)</td>
<td>803,007(19.3)</td>
<td>4,167,737(100)</td>
</tr>
<tr>
<td>Total</td>
<td>183,559(2.4)</td>
<td>85,837(1.1)</td>
<td>5,694,299(72.8)</td>
<td>1,854,088(23.7)</td>
<td>7,817,783(100)</td>
</tr>
</tbody>
</table>

Table IV. Number of click views of the big threes in Wikipedia Indonesian edition

Notes: (a) Japan; (b) China; (c) South Korea

Figure 7. Number of click views in Wikipedia Indonesian edition
Yet, looking close to the keywords (subcategories), categories are somewhat similar. Korea gets clicks on entertainment and film. Those are keywords related with Korean pop culture that has become very popular recently among the young people in the both language groups. On another hand, Japan gets clicks on military and China on people and language.

The findings gave the researchers enough to assert that measuring the number of clicks in Wikipedia editions is an affordable candidate to monitor real-time awareness of a nation. However, it must be noted that they offer neither performance and emotional affinity nor loyalty, each of which is considered as an important part of equity to nation brand.

Clearly, data obtained from Wikipedia are advantageous over traditional methodologies because it may be repeatedly run in real time and does produce consistent and dynamic results. For this, it may well qualify to be used supplement to traditional nation brand image.

As for the future study, it would be ideal to find a way to get other kinds of brand equity that are missing from Wikipedia which are performance, emotional affinity and loyalty. The researchers would like to continue to search ways to find those other kinds of brand equity.

As for the limitation of this particular study, we, the researchers, admit that we could not develop an algorithm for automated ontology construction. In another research, however, we were successful in developing such an algorithm. The work was reported in another paper (Lee, 2016).

References


**Corresponding author**
Heuiju Chun can be contacted at: hjchun@dongduk.ac.kr

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com