Measuring the Triple Helix on the Web: Longitudinal Trends in the University-Industry-Government Relationship in Korea

Gohar Feroz Khan and Han Woo Park*
Department of Media & Communication, YeungNam University, 214-1 Dae-dong, Gyeongsan-si, Gyeongsangbuk-do 712-749, Republic of Korea. Email: gohar.feroz@gmail.com; hanpark@ynu.ac.kr

This study examines longitudinal trends in the university-industry-government (UIG) relationship on the web in the Korean context by using triple helix (TH) indicators. The study considers various Internet resources, including websites/documents, blogs, online cafes, KnowledgeIn (comparable to Yahoo! Answers), and online news sites, by employing webometric and co-word analysis techniques to ascertain longitudinal trends in the UIG relationship, which have received considerable attention in the last decade. The results indicate that the UIG relationship varied according to the government's policies and that there was some tension in the longitudinal UIG relationship. Further, websites/documents and blogs were the most reliable sources for examining the strength of and variations in the bilateral and trilateral UIG relationships on the web. In addition, web-based T(uig) values showed a stronger trilateral relationship and larger variations in the UIG relationship than *Science Citation Index*-based T(uig) values. The results suggest that various Internet resources (e.g., advanced search engines, websites/documents, blogs, and online cafes), together with TH indicators, can be used to explore the UIG relationship on the web.

Introduction

Information society theory posits that information and the flow, use, and control of information and knowledge are becoming an integral part of everyday life (Manuel, 1996). Information and communication technology have fundamentally altered the way knowledge is created and distributed within and among societies. In particular, the Internet has played a key role in facilitating the transition from postindustrial societies to knowledge-based ones (Park & Thelwall, 2006). Societies that derive their economical power from the production and diffusion of information and knowledge are referred to as knowledge-based societies or economies (Foray & Lundvall, 1996). Knowledge-based economies, like industrial ones, function based on three fundamental subdynamics, namely, wealth production, knowledge production, and normative control (Leydesdorff & Etzkowitz, 1996; Leydesdorff & Meyer, 2003). These societies have a well-established knowledge infrastructure that works as an engine for organized novelty production. It is believed that linkages and interactions among government, academia, and industry are essential elements of knowledge-based innovation (Etzkowitz & Leydesdorff, 2000)

Measuring and gauging the underlying structure and strength of the relationship among the three main components (i.e., government, academia, and industry) of knowledge-based innovation have attracted considerable attention (Leydesdorff, Dolfsma, & Panne, 2006; Leydesdorff & Fritsch, 2006; Leydesdorff & Guoping, 2001; Leydesdorff & Sun, 2009; Park, Hong, & Leydesdorff, 2005), and several models and approaches have been proposed for measuring the knowledge infrastructure (Etzkowitz & Leydesdorff, 2000; Gibbons et al., 1994). For example, the Triple Helix (TH) model is a well-known method for examining the university-industry-government (UIG) relationships (Etzkowitz, 2008; Etzkowitz & Leydesdorff). According to Park et al. (2005), "The network of university-industry-government relations can be considered as an institutional knowledge infrastructure that carries a system of operations containing science, technology, and knowledge-based innovations" (p. 6) and can be measured using indicators such as publications, patents, and communication on the Internet (Park et al., 2005).

Thus, various tools and techniques are available for measuring the institutional knowledge infrastructure. However, most of the previous studies measuring this infrastructure have focused on non-Asian (i.e., English) contexts (Lee & Jeong, 2008), and thus they have typically analyzed only

Received March 14, 2011; revised May 20, 2011; accepted May 23, 2011

*Corresponding author.

 DOI: 10.1002/asi.21595

English content when considering written communication. In addition, previous studies gauging the knowledge infrastructure or the UIG relationship have typically used well-documented and formal written communication, including patents and publications (e.g., Park & Leydesdorff, 2008; Park et al., 2005). Furthermore, most of the knowledge-based innovation indicators, including the *Science Citation Index* (SCI), are available commercially and accessible only to a limited number of subscribers. In this regard, the present study measures the strength of the UIG relationship in the Korean context by using the web. The results suggest that the web and advanced search engines, together with TH indicators (Leydesdorff & Curran, 2000), can indicate the UIG relationship in the Korean context. In particular, the study employs webometric (Bjorneborn & Ingwersen, 2004), content analysis (Holsti, 1969), and co-word analysis techniques (Callon, Courtial, & Laville, 1991) to ascertain the longitudinal trends in the UIG relationship in the last decade.

The rest of this article is organized as follows: First, we provide a brief literature review. Then we discuss the research method and present the results. Finally, we provide concluding remarks.

Literature Review

The knowledge infrastructure of postindustrial societies plays a crucial role in their economic development. A number of studies have analyzed and gauged the underlying knowledge infrastructure (Gibbons et al., 1994; Lee & Jeong, 2008; Leydesdorff et al., 2006; Leydesdorff & Fritsch, 2006; Leydesdorff & Guoping, 2001; Leydesdorff & Sun, 2009; Park et al., 2005). Several approaches have been proposed for understanding the complex nature and behavior of entities involved in knowledge creation and diffusion within knowledge-based societies (Etzkowitz & Leydesdorff, 2000; Gibbons et al., 1994). For example, Gibbons et al. differentiated the so-called "Mode 2" knowledge production and diffusion from conventional knowledge creation, which they referred to as "Mode 1," and they proposed that "Mode 2" knowledge production reflects a context-driven, problem-focused, and interdisciplinary process and is produced by interdisciplinary teams working together for a short period of time in a specific context. Criticizing Gibbons et al.'s (1994) concept, Etzkowitz and Leydesdorff proposed the TH model to explain the complex role of the knowledge infrastructure in the UIG relationship and stated that "the Triple Helix overlay provides a model at the level of social structure for the explanation of Mode 2 as a historically emerging structure for the production of scientific knowledge, and its relation to Mode 1" (p. 118).

TH Model and Indicators

According to Leydesdorff and Meyer (2003), "the Triple Helix model of university-industry-government relations tries to capture the dynamics of both communication and organization by introducing the notion of an overlay of exchange relations that feeds back on the institutional arrangements" (p. 196). The network of UIG relationships provides an institutional infrastructure that can be measured using indicators such as patents, scientometric indicators, and various Internet resources (e.g., advance search engines; Leydesdorff & Curran, 2000). Thus, the mutual information in the two or three dimensions can serve as a mechanism for measuring UIG relationships (or network data) at each point in time using entropy statistics and the mathematical theory of communication (Shannon, 1948). Using information theory, we can express the expected information content of a message outlining a certain event that occurred with a certain observed frequency distribution in terms of bits of information (Leydesdorff, 1995; Theil, 1972).

According to Shannon (1948) and Shannon and Weaver (1949), the amount of information conveyed by the occurrence of an event is inversely proportional to the probability of that event occurring, and thus if the discrepancy is considered to be a probability distribution (Σi pi), then the uncertainty associated with the distribution (H) can be defined as follows:

$$H\mathrm{i} = -\Sigma\mathrm{i}\ p\mathrm{i}\ \log 2\ (p\mathrm{i}) \tag{1}$$

Similarly, the two-dimensional distribution (Hij) can be written as

$$H\mathrm{ij} = -\Sigma\mathrm{i}\ \Sigma\mathrm{j}\ p\mathrm{ij}\ \log 2\ (p\mathrm{ij}) \tag{2}$$

where *Hij* is the sum of the uncertainty associated with the two dimensions of mutual information contained in each probability distribution, that is, the two variations overlap in their covariation and condition each other in remaining variations.

Assume that "u" stands for "university," "i" for "industry," and "g" for "government"; then, using information theory, we can write the mutual information in two dimensions by using "T," the "transmission" between two probability distributions (e.g., the industry-government [IG] relationship), as follows:

$$\mathrm{Tig} = \mathrm{Hi} + \mathrm{Hg} - \mathrm{Hig} \tag{3}$$

Here *Hig* is zero for completely independent distributions (i.e., if the covariation is zero) and positive otherwise (Theil, 1972). Similarly, the mutual information in three dimensions can be defined (Abramson, 1963, p. 129) as follows:

$$\mathrm{Tuig} = \mathrm{Hu} + \mathrm{Hi} + \mathrm{Hg} - \mathrm{Hui} - \mathrm{Hig} - \mathrm{Hug} + \mathrm{Huig} \tag{4}$$

Note that T values for two dimensions—T(ui), T(ug), and T(ig) values—are, by definition, positive and that the trilateral relationship, that is, T(uig) values, can be negative, positive, or zero (Leydesdorff, 2003; Park & Leydesdorff, 2010) depending on the relative sizes of contributing terms. However, in the case of negative T(uig) values, the synergy is indicated by the relationship among TH components, whereas it is not in the case of positive T(uig) values. Furthermore, note from Equations 3 and 4 that bilateral terms reduce uncertainty, whereas the trilateral terms increases it, which prevail at the network level.

DOI: 10.1002/asi

TH Model and the UIG Relationship

The TH model can be used as a framework for measuring the underlying UIG relationship (Etzkowitz, 2008), and for this, various techniques have been used (e.g., Leydesdorff et al., 2006; Leydesdorff & Fritsch, 2006; Leydesdorff & Sun, 2009). According to Park et al. (2005), the knowledge infrastructure can be measured using the Internet by employing TH indicators. In addition, previous studies have employed TH indicators, together with webometric, scientometric, and co-word analysis techniques, to measure the knowledge infrastructure. Some studies have traced the UIG relationship by using the Internet. For example, Leydesdorff and Curran (2000) mapped the structural and institutional relations between universities, industries, and governments by using advanced search engine techniques, and they found that "from a theoretical point of view, it is important to note that these relations among textual units at the Internet can be considered as less codified counterparts of scientometric distributions retrieved in terms of co-word, co-authorship relations, etc." (p. 14). Park et al. (2005) combined various techniques, including co-word analysis, to map the UIG relationship on the Internet and examined the relationship by using the occurrence and co-occurrence of the words "university," "industry," and "government" and employing advanced search techniques in specific Internet domains. Similarly, Lee and Jeong (2008) used co-word techniques to analyze and map the overall evolutionary trend in Korea's national research and development efforts in robotics.

However, most of the previous studies have measured the knowledge infrastructure in non-Asian (English) contexts (Lee & Jeong, 2008), and thus they have typically analyzed only the English content of written communication such as patents and publications (Park & Leydesdorff, 2008; Park et al., 2005). Thus, the present study measures the strength of the UIG relationship in the Korean context by using the web. From a theoretical point of view, it is important to measure the TH relationships on the Internet and compare various institutional agencies. Such relationships among textual units on the Internet can be considered as less codified counterparts of scientometric distributions retrieved in terms of co-words and co-authorship relations, among others. Scientometric distributions refer to codified communication, whereas Internet relations are not controlled and based mainly on free text (Leydesdorff & Curran, 2000)

Research Methodology

Method

We employed webometric (Bjorneborn & Ingwersen, 2004), content analysis (Holsti, 1969), and co-word analysis techniques (Callon et al., 1991). According to Björneborn and Ingwersen, Webometrics is "the study of the quantitative aspects of the construction and use of information resources, structures and technologies on the web drawing on bibliometric and informetric approaches" (p. 1217). webometrics is also defined as "the study of web-based content with primarily quantitative methods for social science research goals using techniques that are not specific to one field of study" (Thelwall, 2009, p. 6). As shown in Park (2010), webometric analysis is particularly useful for understanding the dynamics of a digitalized knowledge-based economy by tracking how web objects, related terms, and hyperlinks within and across institutions are circulated in Korea's webosphere.

According to Holsti (1969), the purpose of content analysis is to objectively and systematically identify explicit characteristics of messages for the sake of inferences. Similarly, co-word analysis is an informetric technique for analyzing linkages and associations among pairs of words co-occurring in a particular domain. Content analysis, combined with the co-word technique, can help researchers to trace TH relationships from a large body of text by drawing meaningful and conceptual groupings based on the co-occurrence of words (Biddix, Park, & Wang, 2009). The strength of the co-word analysis technique lies in the fact that the object of analysis can be extended to different domains such as patents, reports, websites, and newspapers (Lee & Jeong, 2008).

Data Collection

We collected the data from Naver.com in March 2010. For this, we used the following search terms: "대학" (dae-hawg: university), "기업" (ghi-oeup: industry), and "정부" (jeong-bu: government). Although the literal translation of the word "industry" is "산업" (san-oeup) in Korean, we used "기업" (ghi-oeup) because it was deemed more suitable for the query (Park et al., 2005). Park et al. (2005) also used these search terms to provide a comparison of knowledge-based innovation system between Korea and the Netherlands. Naver started its service in 1998, and thus we collected data for the 1999-2009 period. The longitudinal data were used to track changes in the UIG relationship over this period.

More specifically, as shown in Table 1, we determined the number of hits for queries reflecting different combinations of the terms "university" (U), "industry" (I), and "government" (G) by using Boolean operators (e.g., "and" and "not"). For example, the value "21" in the cell between "U" and "1999" means that there were 21 mentions of the term "university" in 1999. During this process, we excluded documents including the terms "industry" and "government." Therefore, we retrieved only those documents containing the term "university."

We recorded the number of word occurrences in titles. In other words, we used the number of hits only for titles. Words in titles of scientific articles and webpages tend to be more meaningful and thus useful at all levels of co-word analysis (Leydesdorff, 2006; Park, 2010). Further, search engines and text analysis programs that are freely available for academic purposes can facilitate the study of dynamics of word frequency and distributions. Then, to specifically examine the UIG relationship, we collected data for each of the following five Naver categories: regular webpages, Knowledge-In (comparable to Yahoo! Answers), blogs, online cafes/the

DOI: 10.1002/asi

TABLE 1. Number of hits for triple helix components from 1999 to 2009.

Year	U	I	G	UI	UG	IG	UIG
1999	21	23	8	0	0	0	0
2000	2,210	13,836	4,712	47	10	187	1
2001	3,024	17,687	5,977	30	7	184	0
2002	4,537	28,529	5,984	80	24	170	2
2003	21,767	36,352	13,947	308	147	567	81
2004	69,717	55,637	33,825	2,249	2,399	4,631	1,088
2005	90,899	68,210	42,696	2,724	2,939	5,323	1,475
2006	233,768	138,193	91,181	9,918	9,099	16,621	4,703
2007	496,227	295,467	151,214	22,063	15,346	30,305	7,615
2008	677,336	392,342	271,605	30,568	21,062	50,311	10,255
2009	814,746	502,035	276,756	38,755	24,734	56,293	12,086

Note. U = university; I = industry; G = government; UI = university-industry; UG = university-government; IG = industry-government; UIG = university-industry-government.

bulletin board system (BBS; e.g., message and discussion boards), and news sites (news articles from media outlets).

Analysis

We analyzed the data by using the TH indicators developed by Leydesdorff (2003) and Park and Leydesdorff (2010), who based the indicators on Shannon's information theory (Shannon, 1948; Shannon & Weaver, 1949). We also used the occurrence and co-occurrence of the terms "university," "industry," and "government" by using Boolean operators. According to Leydesdorff and Curran (2000), variations in the UIG relationship, that is, the strength and weakness of TH components, can be measured using TH indicators and advanced search engines. For example, Leydesdorff (2003) showed that mutual information on two (e.g., university-industry) and three (e.g., UIG) dimensions can be expressed using T(ui) and T(uig), respectively (Leydesdorff, 2003). Here, T(ui) and T(uig) denote the mutual information transmission in two and three dimensions, respectively. The transmission (T) is measured by "bits" or "mbits" of information. The T value can be negative, positive, or zero depending on the relative sizes of the three TH components' contributions. We calculated the TH indicators by using the relative frequency or the probability distribution discussed above (for a mathematical definition, see Leydesdorff, 2003) and T values by using a standard technique in the TH program available at http://www.leydesdorff.net/th2/index.htm.

Results

Longitudinal Trends in the UIG Relationship by Category

Figure 1 shows the longitudinal trend in T(uig) values by category. Blogs indicated the strongest trilateral relationship since 2004, reaching T (−0.400) in 2008. Webpages showed large variations in the trilateral relationship, indicating several ups and downs in the relationship. Further, news sites indicated a consistently improving trilateral relationship since 2002, as indicated by its T values (Figure 2).

Furthermore, webpages revealed large variations in the UIG relationship. Before 2002, the T(uig) values for webpages were stable, ranging from −0.350 to −0.400 mbits and indicating clear bilateral trends in the UIG relationship. However, since 2003, this relationship weakened, perhaps because of the effects of the global dot-com crash on Korea. Park et al. (2005) offered a similar conclusion. For 2004, 2005, 2006, and mid-2007, the T(uig) values were almost −0.200 mbits, indicating the weakest trilateral relationship. However, the T(uig) values decreased sharply from 2007, indicating the normalization of the UIG relationship, perhaps because of the presidential election in December 2007 and political campaigning on the web. Noteworthy is that the T values for webpages collapsed when blogs became popular in 2003–2004. In a similar vein, the T(uig) values for news sites decreased steadily since 2002. During this period, there was a substantial increase in the number of blogs and online new sites containing the terms "university," "industry," and government."

Surprisingly, compared with other categories, Knowledge-In indicated a weak UIG relationship, despite its status as the most popular service in Korea. This may be because most of the Knowledge-In documents mentioned only the term "university"; most of the Knowledge-In users are middle/high school students with many questions about universities and are thus not likely to be interested in the bilateral and trilateral UIG relationships.

The following sections discuss each category (webpages, blogs, Knowledge-In, news sites, and cafes/BBS) separately.

Webpages

The occurrences and co-occurrence of UIG terms. Figure 2 shows the occurrence and co-occurrence of the terms "university" (U), "industry" (I), and "government" (G) for the webpages category. There was an increase in the number of webpages and documents containing the terms U, I, and G since 1999, In other words, there was a steady increase in the number of the terms U, I, and G. Figure 2 shows a clear division between the top three lines (U, I, and G)

DOI: 10.1002/asi

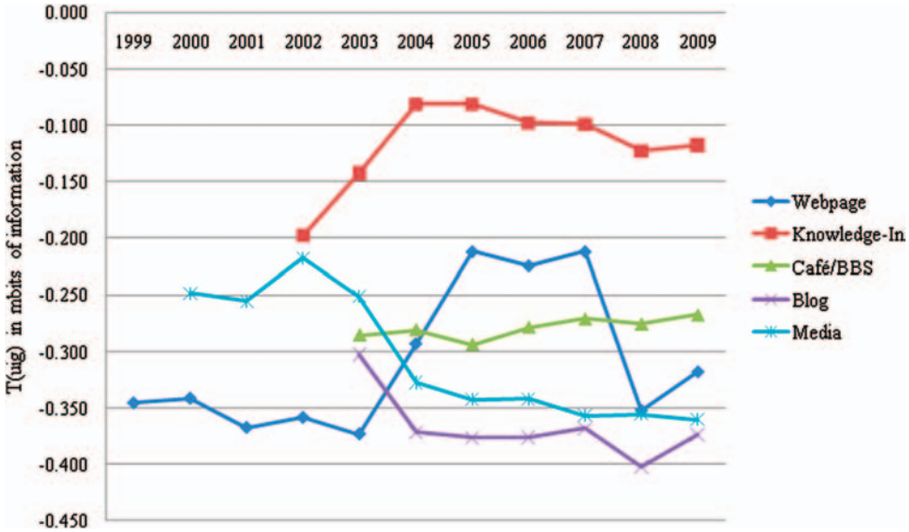

FIG. 1. Longitudinal trends in web-based T(uig) values by category. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

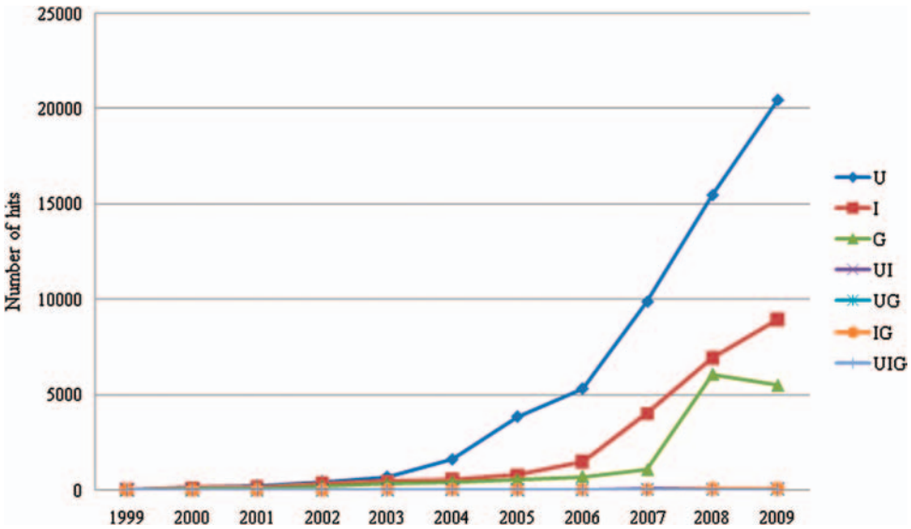

FIG. 2. Longitudinal trends in the occurrence of U, I, and G in titles of webpages. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

and the remaining lines (bilateral and trilateral relationships: university-industry [UI], university-government [UG], industry-government [IG], and UIG). Although this division is a consequence of the Boolean "and" operator, it still indicates that the terms U, I, and G occurred more frequently on webpages than the co-occurrence terms UI, UG, and IG. In addition, there was a clear division between the term U and the terms I and G, which reflects the fact that most users in Korea employ the Internet for academic purposes (Park, 2010). A large majority of Internet users are university students who produce documents covering a topic of special interest and download/upload them on the web. In addition, most of the universities in Korea have their own websites, making academic procedures available on the Internet. Thus, these factors might have limited the exposure of the terms I and G and the bilateral and trilateral UIG relationships on the web.

Longitudinal trends in the UIG relationship. Figure 3 shows the longitudinal trend in bilateral and trilateral UIG relationships for webpages (expressed in terms of T values). The bilateral T values for U and I were the highest, indicating the important role played by webpages in the UI relationship. Noteworthy is the conflicting behavior of the UI and UG bilateral relationships. The T(ui) values started to increase sharply in 2005 and peaked in 2007 (0.600), whereas the T(ug) values started to decrease in 2005 and bottomed out in 2007 (0.100). This provides evidence of some tension in

DOI: 10.1002/asi

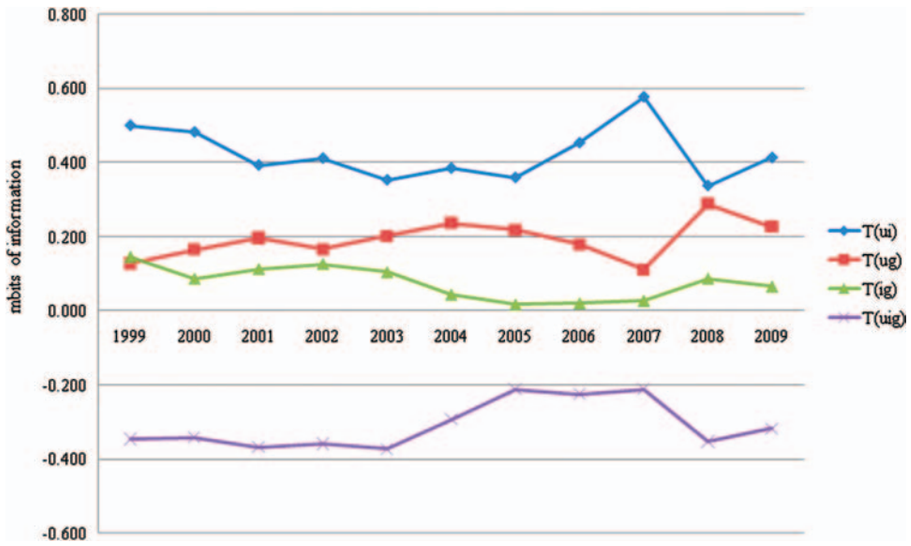

FIG. 3. Longitudinal trends in bilateral and trilateral UIG relationships for webpages. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

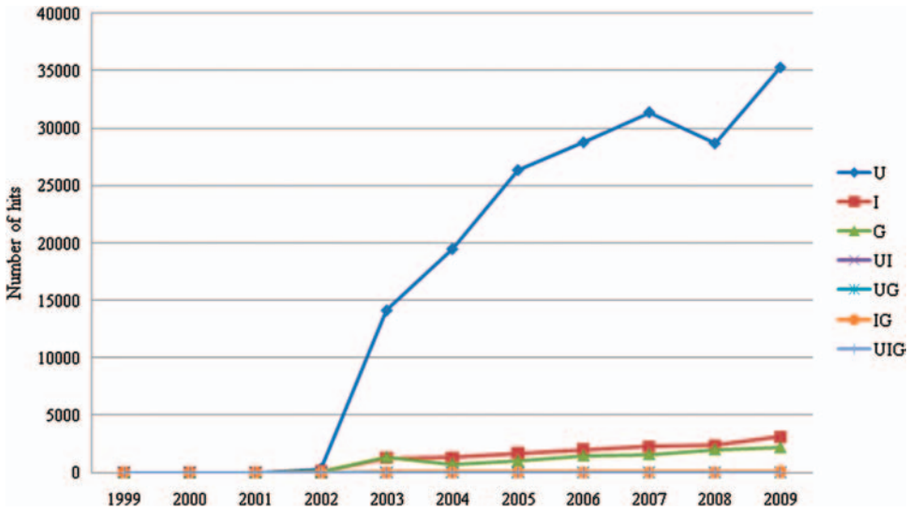

FIG. 4. Longitudinal trends in the occurrence of U, I, and G in titles of Knowledge–In documents. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

the longitudinal UIG relationship in Korea. For example, between 1999 and 2009, the strengthening of the bilateral UI relationship was always accompanied by the weakening of the bilateral UG and IG relationships and vice versa, which may be because of the lack of coordination among the three TH components. This is also indicated by the lack of a strong trilateral relationship, as indicated by T(uig) values, from 2005 to 2007.

Further, the IG relationship showed the lowest T values for all the years. The IG relationship was weakest between 2003 and 2007, which is likely because of President Roh's preference for the UI relationship over the IG/UIG relationships (Park & Leydesdorff, 2010; Shapiro, So, & Park, 2010). The UIG relationship, which showed the lowest degree of mutual information sharing during the Roh administration, provides support for this view. Further, the T(uig) values started to decrease from 2007, when President Lee took office.

Knowledge-In

The occurrences and co-occurrence of UIG terms. Figure 4 shows the longitudinal trend in the occurrence of the terms U, I, and G in titles of documents obtained from Naver's Knowledge-In service. Only the term U increased noticeably since Naver started the Knowledge-In service in October 2002. Noteworthy is that whereas U increased dramatically, reaching 35,285 hits, the other terms remained

DOI: 10.1002/asi

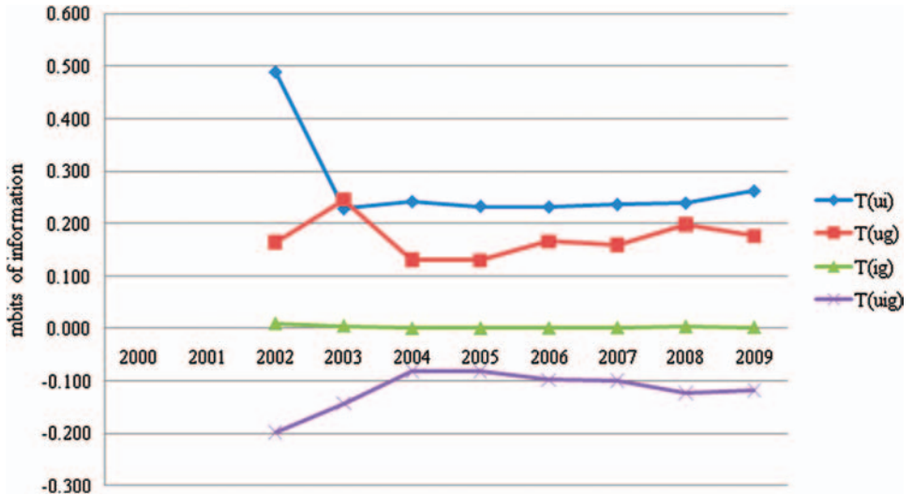

FIG. 5. Longitudinal trends in bilateral and trilateral UIG relationships for Knowledge-In. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

below 5,000. In particular, the co-occurrence of the terms U, I, and G did not increase as much as that of individual terms. Most of the dramatic increases in the occurrence of U, I, and G in titles of Knowledge-In documents occurred from 2002 onward, demonstrating an upward trend. Within 4 months of the service, a total of 400,000 Knowledge-In documents were uploaded, and there were 6.5 million daily visitors. The occurrence of U increased sharply and continued at a high rate, whereas the occurrence of I and G remained low. As mentioned earlier, this may be because most of the Knowledge-In users are students whose main interest is in universities. In addition, even before the advent of Web 2.0, the convenience of searching, retrieving, and producing information by using Knowledge-In may have contributed to the explosion of questions and answers. In addition, since Naver started offering Knowledge-In free of charge in 2002, its collection of web documents has expanded exponentially. Thus, it is likely that the hits for web content containing the term U were influenced by the Knowledge-In service.

Longitudinal trends in the UIG relationship. Although Knowledge-In users' interest is rooted in university-related topics, it is possible to observe and track longitudinal trends and variations in the UIG relationship by using T values. Figure 5 shows the longitudinal trend in the bilateral and trilateral UIG relationships for Knowledge-In. As indicated by the T(ui) values, the bilateral UI relationship was stronger in mid-2002, when the Knowledge-In service was introduced. However, this relationship weakened dramatically since mid-2002 because of the dot-com crisis. The other bilateral and trilateral relationships also weakened during this period. The UIG relationship never normalized or reached the initial 2002 level. On the other hand, the T(ig) and T(uig) relationships were almost nonexistent. However, the higher T(ui) values suggest that Knowledge-In users are somewhat interested in the UI relationship. This may be because these users are likely to perceive this relationship as a source of financial support (e.g., finding information on employment opportunities related to their college major) for their higher education. In addition, the government has been implementing policies to improve this relationship, which is supported by the slight improvement in the UIG relationship and the bilateral UI relationship in 2007, when President Lee was in office.

Blogs

The occurrences and co-occurrence of UIG terms. Figure 6a shows the longitudinal trend in the occurrences and co-occurrence of U, I, and G in blog titles. Noteworthy is that the occurrence of U and I increased at almost the same rate since the blog service started in 2003. The occurrence of U increased from 254 hits in 2003 to 64,208 in 2009, and the occurrence of I increased from 173 to 64,209. The occurrence of G also increased from 2003 to 2008, the last year of the Roh administration, but it leveled off in 2009. Compared with Knowledge-In (Figure 4), blogs (Figure 6a) facilitated more discussions about I and G because of their professional focus. However, the co-occurrence of U, I, and G was relatively low.

Longitudinal trends in the UIG relationship. Figure 6b shows the longitudinal trend in bilateral and trilateral UIG relationships for blogs. Noteworthy is that blogs showed the strongest trilateral relationship. The trilateral relationship remained steady throughout the 2003-2009 period, and the T(uig) value reached -0.400 mbits in 2008. Further, among bilateral relationships, the UI relationship was the strongest. The UI relationship peaked in 2003 in terms of the T(ui) value, and since then, the relationship weakened dramatically and stayed weak from 2004 to 2007. The relationship weakened again in 2008 but returned to previous levels in 2009. In addition, the T(ug) and T(ig) values remained lower than the T(ui) values, and there were no noticeable variations.

DOI: 10.1002/asi

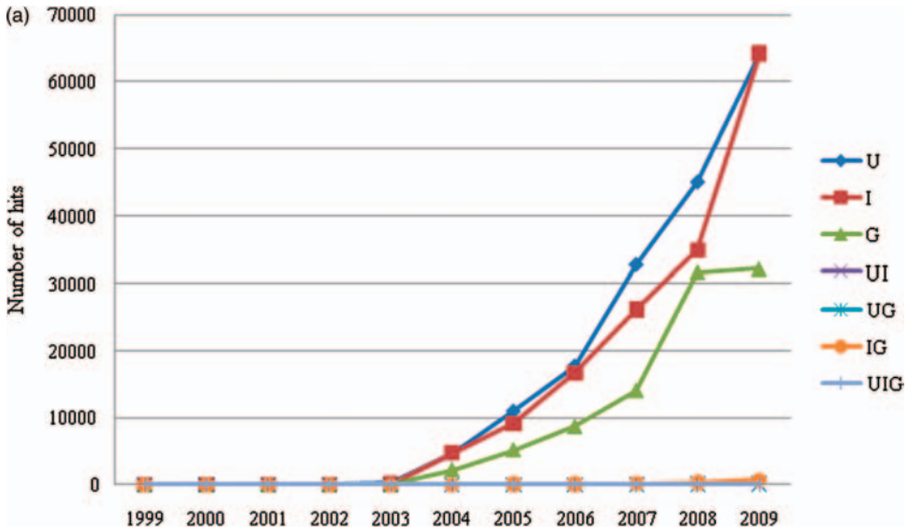

FIG. 6a. Longitudinal trends in the occurrence of U, I, and G in blog titles. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

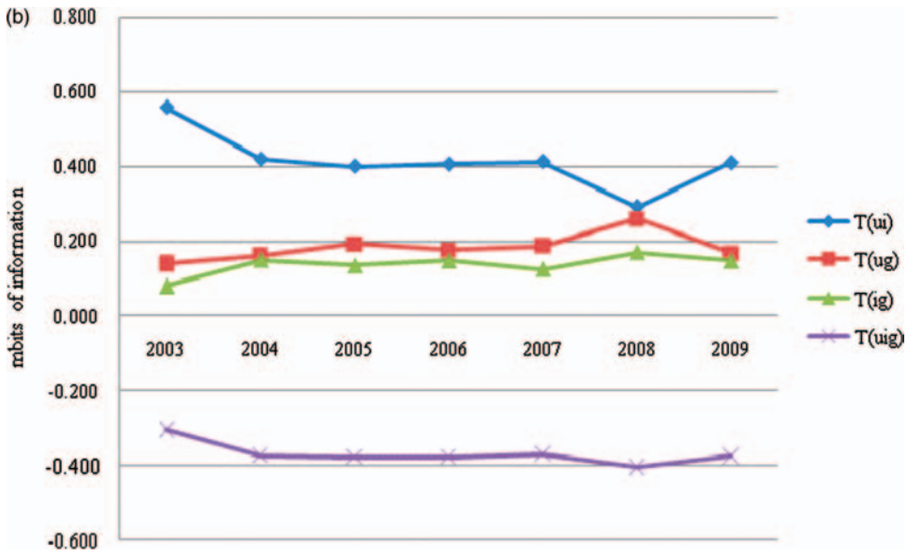

FIG. 6b. Longitudinal trends in bilateral and trilateral UIG relationships for blogs. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

Noteworthy is the conflicting behavior of the T(ui) and T(ug) relationships. An increase (decrease) in T(ui) values was accompanied by a decrease (increase) in T(ug) values.

Cafes/BBS

The occurrences and co-occurrence of UIG terms. Figure 7a shows the longitudinal trend in the occurrences and co-occurrence of the terms U, I, and G in titles of cafe/BBS documents. The results are generally consistent with those for the other categories. There were 1,549 hits for U in 2003, when Naver started its cafe/BBS service, and 664,941 hits in 2009. There were 952 hits for I in 2003 and 353,196 hits in 2009. Noteworthy is that there were nearly 70,000 hits for IG between 2008 and 2009 and approximately 50,000 hits for UI, which were the highest numbers of hits across the categories for the co-occurrence of IG and UI, which may be because of the professional nature of cafes and the BBS and users' interest in the country's affairs and business.

Longitudinal trends in the UIG relationship. Figure 7b shows the longitudinal trend in bilateral and trilateral UIG relationships for cafes/BBS in terms of T values. Cafes/BBS provided the highest T(ui) values, which showed an upward trend. The bilateral UI relationship peaked in 2007. The T(ui) and T (ug) values diverged beginning in 2004. An

DOI: 10.1002/asi

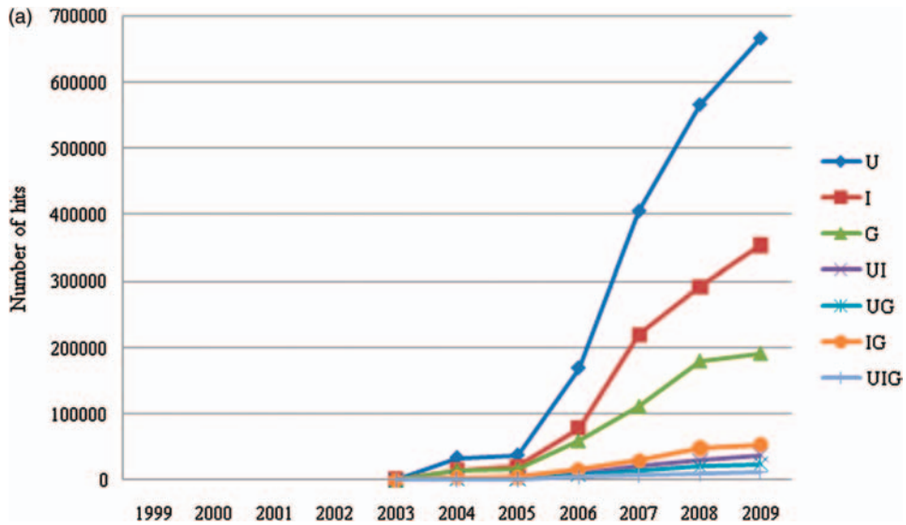

FIG. 7a. Longitudinal trends in the occurrence of U, I, and G in titles of cafeBBS documents. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

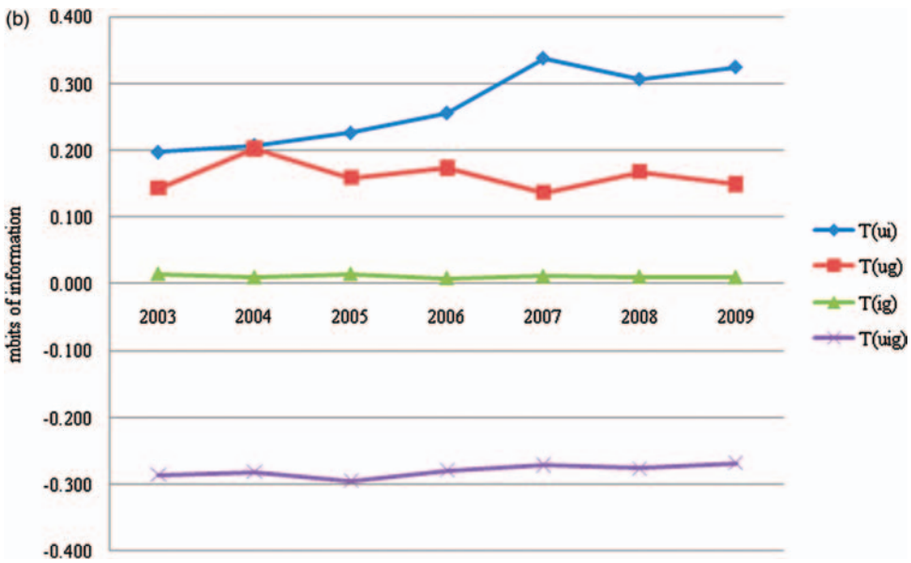

FIG. 7b. Longitudinal trends in bilateral and trilateral UIG relationships for cafesBBS. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

improvement in the UI relationship weakened the UG relationship and vice versa. On the other hand, the IG relationship was almost nonexistent. Finally, cafes/BBS showed a strong trilateral relationship, but here there were no large variations in the relationship.

News Sites

The occurrences and co-occurrence of UIG terms. Noteworthy is that the titles of documents from online news sites, unlike those of documents from other categories, provided the highest number of hits for I, followed by G (Figure 8a). This is consistent with the media's keen interest in business and politics. The occurrence of the term I led that of the other terms and showed a steady upward trend since 1999 (with minor ups and downs). The occurrence of the term G increased steadily since 1999, rising sharply in 2007 and then declining slightly in 2008. In contrast to the other categories, news sites provided the lowest number of hits for the term U, which may be because of the media's lack of interest in academic matters. Finally, the terms U, I, and G did not receive much attention from online news sites.

Longitudinal trends in the UIG relationship. Consistent with the results in previous sections, news sites showed the strongest bilateral IG relationship in terms of the T value (Figure 8b). Noteworthy is that an improvement in the

DOI: 10.1002/asi

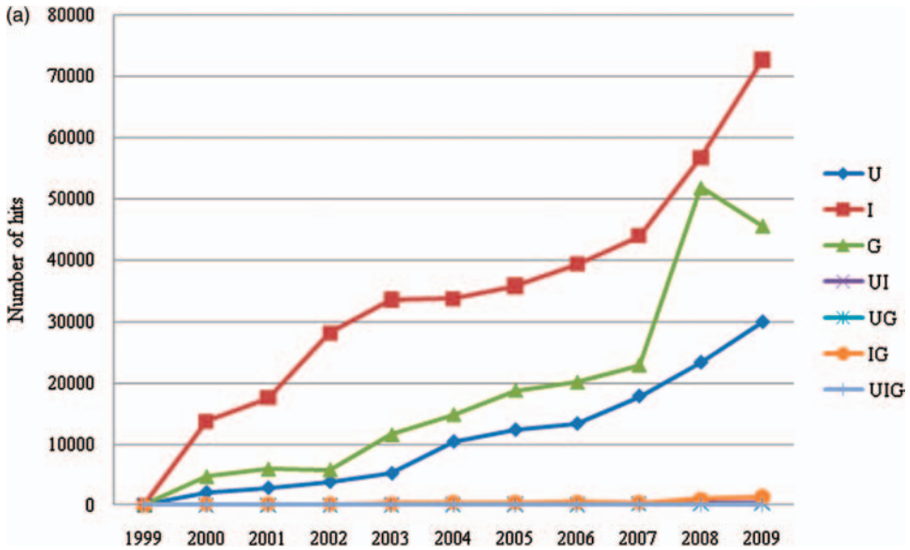

FIG. 8a. Longitudinal trends in the occurrence of U, I, and G in titles of documents on online news sites. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

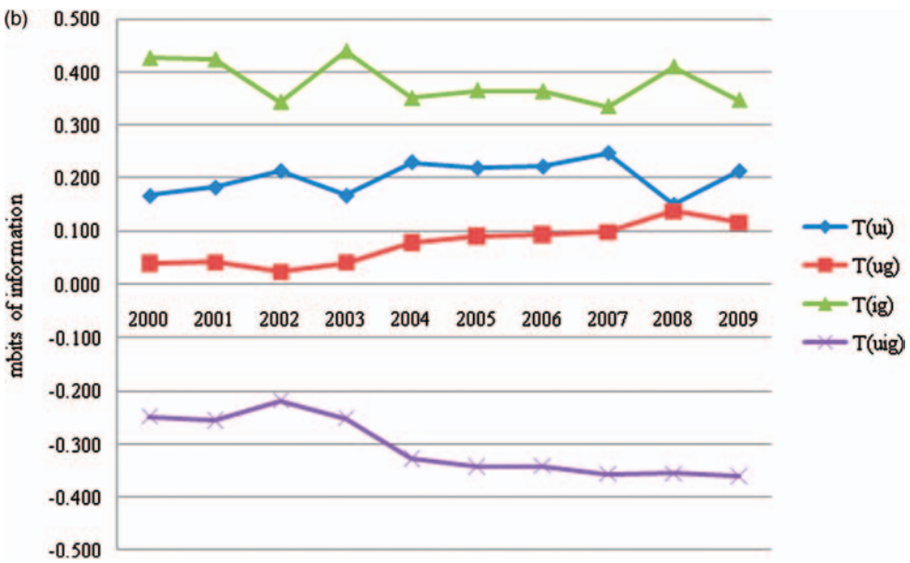

FIG. 8b. Longitudinal trends in bilateral and trilateral UIG relationships for news sites. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

bilateral IG relationship was accompanied by a decline in the bilateral UI relationship and vice versa. Further, the trilateral relationship strengthened since 2002 onward, peaking in 2009. Similarly, the UG relationship was at its weakest (almost nonexistent) in 2002 but showed some improvement in subsequent years. However, this relationship was weaker than the UI and IG relationships.

A Comparison Between SCI-Based T(uig) and Web-Based T(uig)

We compared web-based T(uig) values with SCI-based T(uig) values, which were calculated using SCI publications for the same period for each TH component (Figure 9). Using SCI-based T values to examine bilateral and trilateral relations among TH components is well known (Leydesdorff & Fritsch, 2006; Park et al., 2005; Park & Leydesdorff, 2010). It is clear from Figure 9 that web-based T(uig) values accounted for a larger portion of the variation in the UIG relationship than SCI-based T(uig) values, which, to some extent, remained steady throughout the sample period. This striking difference may be because Internet resources are more diverse than SCI-based indicators, which are strictly codified and available commercially only to a restricted number of users.

DOI: 10.1002/asi

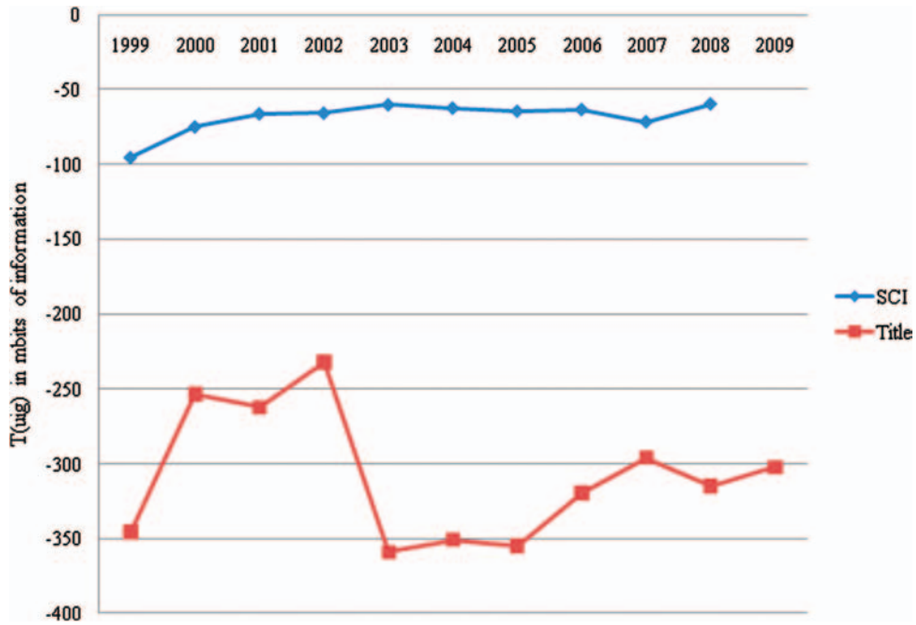

FIG. 9. A Comparison between web-based T(uig) and SCI-based T(uig) values. [Color figure can be viewed in the online issues, which is available at wileyonlinelibrary.com.]

Discussion

This study gauges the longitudinal UIG relationship by using the web in the Korean context. The results provide evidence of some tension in the longitudinal UIG relationship in Korea. For example, an improvement in one bilateral relationship was always accompanied by a decline in another bilateral relationship and vice versa. This tension was mostly apparent in the bilateral UI and UG relationships, which may be because of political and economic constraints influencing these parties. In addition, the UIG relationship varied according to the government's policies.

Webpages showed the highest T(ui) value (0.600) for the bilateral UI relationship. Further, webpages accounted for the largest portion of the variation in the trilateral relationship. On the other hand, blogs showed the lowest T(uig) value (−0.400) for the trilateral relationship. In other words, blogs can be a good indicator of the trilateral relationship. These results suggest that webpages and blogs are the most reliable Internet sources for examining the strength of and variations in the bilateral and trilateral UIG relationships and that webpages and blogs are important indicators of the bilateral and trilateral UIG relationships on the web

Further, except for news sites, all the categories showed the highest T value for the UI relationship; news sites provided the highest T value for the IG relationship. In addition, except for news sites, all the categories provided the highest number of hits for U, followed by I and G, in that order; news sites provided the highest T value for IG, followed by UI and UG, in that order. These results suggest that, unlike the IG relationship (which is typical for a political economy, not a knowledge-based economy), the UI relationship can be considered a proxy for the integration of knowledge-based innovation systems. Furthermore, this suggests that U and I have the most visible relationship on the web, followed by UG and IG, in that order, and that, unlike other Internet sources, online news sites are interested mainly in the bilateral IG relationship, which may be because of the media's keen interest in business and politics, which is consistent with the media's fundamental motivation.

The highest numbers of hits were for IG and UI in titles of documents from cafes and the BBS, which may be because of the professional nature of cafes and the BBS and users' keen interest in the country's affairs and business. Compared with SCI publications, web-based resources (i.e., blogs, online news sites, Knowledge-In, webpages, and cafes/BBS) showed a stronger trilateral relationship. In addition, web-based T(uig) values accounted for a larger portion of the variation in the UIG relationship than SCI-based T(uig) values, which, to some extent, remained steady throughout the sample period. These results indicate that Internet resources (e.g., advanced search engines and web documents), together with TH indicators, can be used to examine the UIG relationship (Leydesdorff, 2003; Leydesdorff & Curran, 2000; Park et al., 2005).

Finally, most of the curves since 2002/2003 were very flat. In other words, the measurements in these cases were reliable in reproducing similar values. For example, when the current state is the best predictor of the next state, the system can be said to satisfy the Markov property. Although we were not able to strictly test the Markov process by conducting a multivariate time series analysis, this study's data on most Internet resources (e.g., news sites, blogs, and Knowledge-In)

DOI: 10.1002/asi

are likely to provide reliable results by satisfying the Markov property (for a good example of the Markov property test for the TH relationship on the Internet, see Leydesdorff, 2006, pp. 246–249).

Contributions and Limitations

This study contributes to the literature by using the web to measure longitudinal trends in the UIG relationship. This study has two major contributions. First related to the contents of the web (e.g., blogs and news sites) and second related to the techniques of searching and categorizing these contents (e.g., search engines).

First, the study measures the UIG relationship by using the web (Leydesdorff, 2003; Park et al., 2005) and empirically validates the existence of the relationship. For this, this study proposes the use of a number of Internet resources, including blogs, online news sites, Knowledge-In, webpages, and cafes/BBS, for measuring the UIG relationship on the web. This research further strengthened the view that the Internet can help people in making decisions about politics and education, for instance, and can act as a platform for social justice by giving access to information that foster social mobility (Introna & Nissenbaum, 2000). For example, Lim and Park (in press) used web data to examine the visibility of Korea's National Assembly members across various web-based platforms, including blogs, images, and news. Park and Thelwall (2008) used the web to investigate the relationship between linkage patterns and social structures in Korea and showed that the Internet can be used to ascertain Internet-based sociopolitical phenomena. Similarly, Park et al. (2010) employed content analysis, together with semantic analysis, to analyze the profiles of Korean politicians on a social networking site and showed that online political communication (e.g., comments and posts) can be used to illustrate the meaning and collective sentiment of comments.

Second, this study demonstrates how one can take advantage of advanced search engine options to demonstrate a phenomenon residing over the Internet. The difficulty of categorizing TH components (UIG) from formal scientific publications discourages researchers to employ TH indicators. However, it will be relatively easy to examine the UIG relationship using search engine data (Park et al., 2005; Park, in press). For example, Leydesdorff and Curran (2000) mapped the structural and institutional UIG relationships by using the Internet, particularly advance search engines, and suggested that such relationships can be measured and compared using the Internet. These kinds of Internet data are particularly appropriate for the study of innovation-related issues including national TH system, as shown in the current article. The current research teaches us that the data collected over the Internet using search engines are particularly appropriate for gauging an Internet-born phenomenon at a particular moment in time and comparing the results over the time axis dynamically and conveniently (Leydesdorff & Curran, 2000). Although search engines do not cover the web completely (Lawrence & Giles, 1998), they still represent a useful tool for accessing and gauging Internet-related phenomena such as sociopolitical activity (Park, in press). Thus, the study provides a better understanding of how a diverse, open environment such as the web can be used to gauge and examine longitudinal trends in the UIG relationship and offers an alternative to strictly codified and commercially available indicators such as SCI publications. Thus, the web can be used to provide supporting evidence (Thelwall, 2002). Using similar methods employed in this study may empower and give voice to the researchers and institutes (particularly in poor countries) who, traditionally, have been weaker and lacking resources (e.g., access to SCI publications) to pursue research objectives. Further, a new hybrid indicator combining SCI with Internet data could be developed to compliment traditional citation-based indexes.

Finally, this study has some limitations. First, it was conducted in a diverse, open environment (i.e., the web), and the data, collected using Internet search engines, may reflect some political and technical constraints (Introna & Nissenbaum, 2000). However, the openness and diversity of the web are its important strengths, and thus they need to be employed constructively. Second, different search engines may yield different results even if they use the same search criteria (e.g., this study's search criteria) because their logic and algorithms are likely to be different. Thus, the generalizability of the findings may be limited. However, the reliability of data collected from the Internet may not depend on search engines. For example, Park (in press) suggested that search engines do not need to be exhaustive, reliable, or objective, but that they should be used for collecting useful data in a limited amount of time. Further, Elgesem (2008) argued that "search engines are objective in the sense that these engines try to be consistent with their own stated policies" (p. 239). Finally, we assumed that the bilateral and trilateral relationships (measured in terms of T values) are only positive. However, there may be cases in which such relationships are negative, and thus future research should ascertain the existence of negative relationships on the web by using T values.

Conclusion

This study examines the longitudinal trends in the UIG relationship on the web in the Korean context by using various Internet resources, including webpages, blogs, online cafes, Knowledge-In, and online news sites. The results indicate that the UIG relationship varied according to the government's policies and that there was some tension in the longitudinal UIG relationship. Further, websites and blogs were the most reliable sources for examining the strength of and variations in the bilateral and trilateral UIG relationships. These results suggest that webpages and blogs are important indicators of the bilateral and trilateral UIG relationships on the web. In addition, web-based T(uig) values showed a stronger trilateral relationship and larger variations in the UIG relationship than SCI-based T(uig) values. Finally, the results suggest that various Internet resources (e.g., advanced search engines),

DOI: 10.1002/asi

together with TH indicators, can be used to explore the UIG relationship.

Acknowledgments

We are grateful to Ji-Young Kim for her assistance in data collection and acknowledge partial support from the Social Science Korea (SSK) and World Class University (WCU) Programs (National Research Foundation of Korea; NRF-2010-330-B00232 for SSK and 515-82-06574 for WCU).

References

Abramson, N. (1963). Information theory and coding. New York: McGraw-Hill.

Biddix, J.P., Park, H.W., & Wang, T. (2009). Co-word analysis of open-end answers from 719 Chinese Internet users: An alternative content analysis method for qualitative 720 research. East Asian Journal of Humanities Research, 16, 415–447.

Bjorneborn, L., & Ingwersen, P. (2004). Toward a basic framework for webometrics. Journal of the American Society for Information Science and Technology 55(14), 1216–1227.

Callon, M., Courtial, J., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. Scientometrics, 22(1), 155–205.

Elgesem, D. (2008). Search engines and the public use of reason. Ethics and Information Technology, 10, 233–242.

Etzkowitz, H. (2008). Triple helix innovation: Industry, university, and government in action. London: Routledge.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From national systems and "mode 2" to a triple helix of university-industry-government relations. Research Policy, 29(2), 109–123.

Foray, D., & Lundvall, B.-A. (1996). The knowledge-based economy: From the economics of knowledge to the learning economy. In D. Foray & B.-A. Lundvall (Eds.), Employment and growth in the knowledge-based economy (pp. 11–32). Paris: OECD.

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). The new production of knowledge, the dynamics of science and research in contemporary societies. London: Sage

Holsti, O.R. (1969). Content analysis for the social sciences and humanities. Reading, MA: Addison-Wesley.

Introna, L.D., & Nissenbaum, H. (2000). Shaping the web: Why the politics of search engines matters. Information Society, 16(3), 169–185.

Lee, B., & Jeong, Y.-I. (2008). Mapping Korea's national R&D domain of robot technology by using the co-word analysis. Scientometrics, 77(1), 3–19.

Leydesdorff, L. (1995). The challenge of scientometrics: The development, measurement, and self-organization of scientific communications. Leiden, the Netherlands: DSWO Press.

Leydesdorff, L. (2003). The mutual information of university-industry-government relations: An indicator of the triple helix dynamics. Scientometrics, 58(2), 445–467.

Leydesdorff, L. (2006). The knowledge-based economy: Modeled, measured, simulated. Boca Raton, FL: Universal Publishers.

Leydesdorff, L., & Curran, M. (2000). Mapping university-industry-government relations on the Internet: The construction of indicators for a knowledge-based economy. Cybermetrics, 4(1), Paper 2.

Leydesdorff, L., Dolfsma, W., & Van der Panne, G. (2006). Measuring the knowledge base of an economy in terms of triple-helix relations among technology, organization, and territory. Research Policy, 35(2), 181–199.

Leydesdorff, L., & Etzkowitz, H. (1996). Emergence of a triple helix of university-industry-government relations. Science and Public Policy 23, 279–286.

Leydesdorff, L., & Fritsch, M. (2006). Measuring the knowledge base of regional innovation systems in Germany in terms of a triple helix dynamics. Research Policy, 35(10), 1538–1553.

Leydesdorff, L., & Guoping, Z. (2001). University-industry-government relations in China: An emergent national system of innovations. Industry and Higher Education, 15(3), 179–182.

Leydesdorff, L., & Meyer, M. (2003). The triple helix of university-industry-government relations. Scientometrics, 58(2), 191–203.

Leydesdorff, L., & Sun, Y. (2009). National and international dimensions of the triple helix in Japan: University, industry, and government versus international coauthorship relations. Journal of the American Society for Information Science and Technology, 60(4), 778–788.

Lim, Y.S., & Park, H.W. (in press). How do congressional members appear on the web?: Tracking the web visibility of South Korean politicians. Government Information Quarterly

Manuel, C. (1996). The rise of the network society. Oxford: Blackwell.

Park, H.W. (2010). Mapping the e-science landscape in South Korea using the webometrics method. Journal of Computer-Mediated Communication, 15, 211–229.

Park, H.W. (in press). How do social scientists use link data from search engines to understand Internet-based political and electoral communication? Quality & Quantity. doi:10.1007/s11135-010-9421-xOnline First™

Park, H.W., Hong, H.D., & Leydesdorff, L. (2005). A comparison of the knowledge-based innovation systems in the economies of South Korea and the Netherlands using triple helix indicators. Scientometrics, 65(1), 3–27.

Park, H.W., & Leydesdorff, L. (2008). Korean journals in the Science Citation Index: What do they reveal about the intellectual structure of S&T in Korea? Scientometrics, 75(3), 439–462.

Park, H.W., & Leydesdorff, L. (2010). Longitudinal trends in networks of university-industry-government relations in South Korea: The role of programmatic incentives. Research Policy, 39(5), 640–649.

Park, H.W., & Thelwall, M. (2006). Web-science communication in the age of globalization. New Media & Society, 8(4), 629–650.

Park, H.W., & Thelwall, M. (2008). Link analysis: Hyperlink patterns and social structure on politicians' web sites in South Korea. Quality & Quantity, 42(5), 687-697. doi:10.1007/s11135-007-9109-z

Shannon, C.E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27, 379–423 and 623–356.

Shannon, C.E., & Weaver, W. (1949). The mathematical theory of communication. Urbana, IL: University of Illinois Press.

Shapiro, M.A., So, M., & Park, H.W. (2010). Quantifying the national innovation system: Inter-regional collaboration networks in South Korea. Technology Analysis & Strategic Management, 22(7), 845–857.

Theil, H. (1972). Statistical decomposition analysis. Amsterdam: North-Holland.

Thelwall, M. (2002). Research dissemination and invocation on the web. Online Information Review, 26(6), 413–420.

Thelwall, M. (2009). Introduction to webometrics: Quantitative web research for the social sciences. San Rafael, CA: Morgan & Claypool Publishers.