https://doi.org/10.5392/IJoC.2017.13.1.031

Conversations about Open Data on Twitter

Seyed Mohammad Jafar Jalali
Department of Information Technology
Allameh Tabataba'i University, Tehran, Iran

Han Woo Park
Dept. of Media & Communication, Interdisciplinary Program of Digital Convergence Business, and Director of Cyber Emotions Research Institute
YeungNam University, Gyeongsan, 712-749, South Korea

ABSTRACT

Using the network analysis method, this study investigates the communication structure of Open Data on the Twitter sphere. It addresses the communication path by mapping influential activities and comparing the contents of tweets about Open Data. In the years 2015 and 2016, the NodeXL software was applied to collect tweets from the Twitter network, containing the term "opendata". The structural patterns of social media communication were analyzed through several network characteristics. The results indicate that the most common activities on the Twitter network are related to the subjects such as new applications and new technologies in Open Data. The study is the first to focus on the structural and informational pattern of Open Data based on social network analysis and content analysis. It will help researchers, activists, and policy-makers to come up with a major realization of the pattern of Open Data through Twitter.

Key words*: Open Data, Structural Pattern, Social Network Analysis, Twitter Network, Social Media.*

1. INTRODUCTION

Open Data (OD) has become a key part of public policy on digital society around the world. Some non-governmental (NGO) OD campaigns and evaluations influence aspects of national data policies that are important to citizens and to global NGOs. The purpose of current research is to understand how OD networks are structured on social media sites that allow populations to collections of connections. This study first creates discursive maps by applying a social media tool [1] to reveal the informational structures represented in the online social spaces. Twitter has been chosen because it particularly affords instant connectivity between posters and their audiences. Previous OD research has mainly examined the national data polices of individual countries [2], international comparisons [3] and international relations [4]. In sum, the majority of prior OD research has taken a normative approach. For example, they tend to examine the ways that OD polices are established and implemented while asking important questions about how those policy programs should be made and revised in various political and institutional contexts. The research materials of these studies generally include legal provisions, governmental reports, public statements, and so on. This study differs from previous studies by making three new contributions. First, because we live in a sea of posts, images, and online updates from a significant proportion of the population in a globally connected world, the data are derived from a stream of comments in Twitter. Second, personal and professional relationships among a number of actors are visualized by involving OD movements. Thus, this study captures important social interactions that move through machine-readable datasets. This study's results provide new insights and illustrations of organizational relationships between governmental agencies and influential practitioners with their shared informational sources. Last, the relationships between actors and their neighbors are uncovered using their characteristics and communication factors. The study addresses the following two questions.

Q1. Who or what are the influential actors in the OD domain? Q2. How does OD mapping take shape in Twitter?

* *Corresponding author, Email: hanpark@ynu.ac.kr*
Manuscript received Jan. 31, 2017; revised Feb. 10, 2017; accepted Feb. 17, 2017

2. LITERATURE REVIEW

OD is defined as structured data that offers free use, publication, modification, and redistribution. It is accessible to everyone, available online without charge, and it is not hindered by copyright laws, patents, or other control mechanisms [2]. The concept of OD is not new; its historical roots date back to eighteenth century Britain's "open government" movement, which called for transparency in the proactive publication of laws and to the same movement in the US in 1966 for freedom of information [5]. There are similarities between the concepts of OD and other current "open" movements, such as open source software, open content, or the new "open government" movement. OD is a main principle of open government that potentially could produce values added to governments in various industries, such as financial services, energy, healthcare, and education. At the end of the first decade of the twenty-first century, governments are progressively expanding their citizens' access to OD. In 2009, Barack Obama was the first US President to present the ideas of transparency, participation, and collaboration of citizens in governmental activities through OD by establishing a portal website (http://www.data.gov). The idea has gained wide acceptance, and many governments that produce large volumes of data have attempted to create web portals to distribute collected governmental data. The OD concept is mentioned in the following governmental portals: http://data.gov.uk (UK), http://data.gov.in (India), and http://data.gov.gh (Ghana). The US Department of Commerce reported in 2014 on the value of governmental statistics for economic development from a new strategic perspective, stating that OD is "uniquely comprehensive, consistent, confidential, credible, relevant, and accessible" and that it optimizes the level of information by fixing failures in the market industry. Furthermore, opening data to the public via web services creates values in transparency and democratic control, self-empowerment, participation, and improved efficiency of governmental services. OD can increase individuals' active participation in their societies. The research on OD focuses on academic questions as well as governmental actions around the world. Scholars have studied OD in three categories: policymaking, policy content, and policy evaluation [3]. Regarding policymaking, the issues include public value ([6-8]), geographic coverage ([2], [6], [9-11]) and implementation ([1], [4], [7], [8]). Numerous studies have described and compared policy content regarding OD. [12] proposed a comparative framework to analyze OD in a cross-national prospective among developing and developed countries. They concluded that governments should invest in removing the barriers to OD used at the national level. Policy evaluation tends to be analyzed in an evaluation framework. These studies ([3], [5], [13-15]) have contributed to solutions to problems regarding effective assessments of OD policies. From the actors' perspectives, OD has been spotlighted by governments, activists, enterprises, international NGOs, and so on. [13] used mapping with social network analytical techniques to investigate the dynamics of risk communication among Twitter social network actors in South Korea and the US in 2013. The focus was on homeland security when North Korea conducted nuclear tests. [13] examined social networks' role in homeland security and how actors could distribute information to other influential actors. Regarding the OD of governmental issues, [1] demonstrated OD trends among national and international organizations believed to be actors in Nepal. They particularly found a strong connection among international organizations but a weak connection between national NGOs and governmental organizations in Nepal. In another study, [15] analyzed a catastrophic event in South Korea in the year 2012. They used content analysis and webometric techniques by gathering governmental data and press releases to address challenges during a catastrophic event that government intuitions faced with them. [15] compared two organizations' informational networks to identify key actors and investigated whether other agencies could create performance effectiveness objectives using social media to share information. However, the previous literature has no studies on OD using actor maps and a structural perspective.

3. MATERIAL AND METHODS

The Twitter keyword search feature on NodeXL [1], a network analysis MS Excel Add-on for data collection, analysis, and network visualization, was used. The outcomes were messages from around the world that contained the term "opendata" collected using the Twitter Search application program interface (API). Twitter's Search API protocol allows researchers to freely access tweets as far back as about 14 days. We conducted the survey in October 2015.Data for this study was collected twice: October 6, 2015, and July 27, 2016. The outcomes included Twitter accounts, messages containing the Open Data keyword, and user profile characteristics, such as the number of followers or followings, their locations, time zones, and URL (Universal Resource Locator) profiles and descriptions. To visualize and map the Open Data network, NodeXL's graph-drawing tool was used. The network analysis approach was employed to examine the nature of the connection networks in the context of Open Data Twitter engagement. Vertices or nodes (which are Twitter accounts in this study) and edges (which are following connections between nodes to represent their relationships) play the largest part in each network (graph) analytical topic. According to [11], the other important metrics in network analysis are Degree Centrality, Eigenvector Centrality, Betweenness Centrality, PageRank, Geodesic Distance, and Density. Indegree Centrality is used to determine the number of incoming ties, and Outdegree Centrality is used for total number of ties sent to when ties are directed. Betweenness Centrality is used to measure the number of shortest routes between two nodes by passing through all of the vertices. Eigenvector Centrality is used to measure the importance of a significant node in a graph. The importance of a node based on its position in the network is measured by PageRank. In a graph, Geodesic Distance is a measure based on the number of edges in the shortest path between two nodes in a network. Network Density is the proportion of highly connected direct vertexes computed as the percentage of actual connections in the entire network. The previous studies indicate that calculating Degree, Betweenness,

Eigenvector Centrality, and PageRank can be helpful to identify the influential parameters in a network ([5], [16], [17]).

We have done a depth analysis on the built network by applying the network characteristics to identify the most valuable information extracted from the network.

4. RESULTS

Network analysis is a practical way to identify a network's structure of members, such as individuals or agencies. The technique has been recently used as a useful tool to analyze connections among texts' contents. This study used it to develop an overview of an entire network. The Density was 0.00025, which was low. The average Geodesic Distance was 4.8, and the diameter between the nodes was 15. The average distance indicates a small, but loose, network. The type of network was directed. The graph's vertices were clustered using the Clauset-Newman-Moore algorithm, which tries to organize collections of densely connected vertices in separate groups or clusters. An advantage of this algorithm is its applicability to large networks with millions of vertexes [3]. The graph was laid out using the Harel-Koren Fast Multiscale layout algorithm, which is a force-directed algorithm designed to adjust all the vertexes to be about the same length and to minimize line crossings, which can create a relatively more aesthetically pleasing and readable graph. Table 1 shows the overall network metrics of OD.

Table 1. Network metrics of Open Data

Network Metric	Value of the year 2015	Value of the year 2016
Vertices	6934	9917
Unique edges	11999	16629
Edges with duplicates	3672	5695
Total edges	15671	22324
Reciprocated vertex pair ratio	0.0468	0.03080
Reciprocated edge ratio	0.0894	0.05977
Connected components	673	826
Single-vertex connected components	349	431
Maximum vertices in a connected component	5502	8177
Maximum edges in a connected component	14068	20430
Maximum geodesic distance (Diameter)	15	13
Average geodesic distance	4.841	4.294011
Graph density	0.00025	0.00017

To map the Open Data network, the structure of the analysis was reconciled to the six structures of Twitter communications about the nature of conversations, including divided, unified, fragmented, clustered, inward and outward hub, and spoke structures [1]. Because of the Harel-Koren Fast Multiscale layout algorithm network graph, the Open Data Twitter network structure was compatible with the "community clusters" structure. Community clusters is an approach widely used to diversify angles on a subject based on its relevance to different audiences, which reveals diversity of opinion and perspective on a topic. These structural conversations appear as multiple centers of activity creating a collection of medium-sized groups. The Twitter map of Open Data is illustrated in Fig. 1 and 2 for 2015 and 2016. The following sections report the results of the analysis of the most important parts of the Open Data network and compares them.

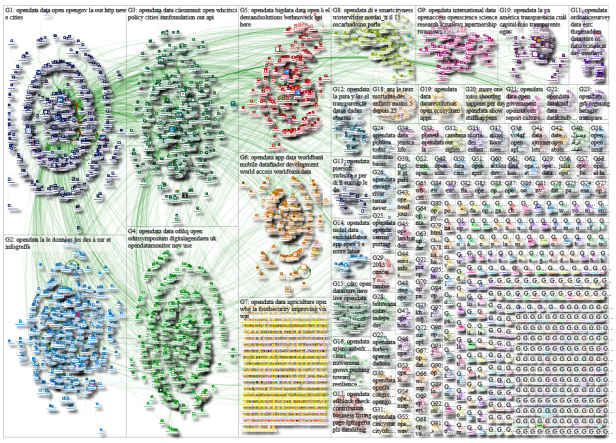

Fig. 1. Open Data structure in 2015

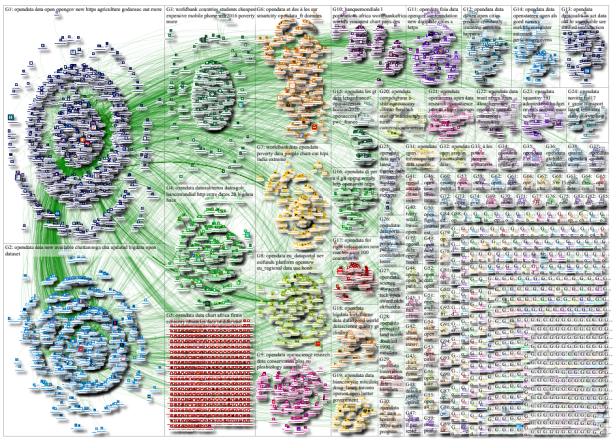

Fig. 2. Open Data structure in 2016

The ranking of the top 10 important Twitter accounts was determined by Indegree, Betweenness, Eigenvector Centrality, and PageRank metrics, which also determined the users' popularity. The data comprised 6,934 Twitter accounts (network vertices) in 2015 and 9,917 Twitter accounts in 2016 that contained Open Data in tweets in the requested range or that were replied to or mentioned in those tweets. There were 11,999 unique network edges in 2015 and 16,629 unique network edges in 2016. The number of duplicated tweets was 3672 in 2015 and it was 5695 in 2016. Users with high Betweenness were recognized as moderators with many connections to others. Tables 2 and 3 show the top 10 Twitter

accounts of each indicator ranked by the metrics in the years 2015 and 2016, respectively. In rank order, the first to tenth ranked influential users in 2015 were devoted to bobbycaudill, cliffordious, odihq, worldbankdata, sunfoundation, worldbank, dyntra_org, okfn, opendatasoft, and etalab. Furthermore, the top 10 influentials on opendata in the Twitter network in 2016 were bobbycaudill, cliffordious, worldbank, worldbankdata, eu_dataportal, odihq, owenboswarva, datosgob, opendatacon, and open_govern. In 2015, the account with the most followers was bobbycaudill (Bobby Caudill), who is a marketing and technology consultant targeting global and startup companies. The account ranking second was cliffordious (Clifford McDowell), founder of Doorda.com and advisor to OD communities. The third ranked account was odihq (Open Data Institute), an organization in London that addresses problems and challenges of Internet data issues. The next ranked accounts were, in rank order from fourth to tenth, worldbankdata (World Bank Data), sunfoundation (Sunlight Foundation Company), worldbank (World Bank), dyntra_org (Dyntra organization), okfn (Open Knowledge Intl), opendatasoft (OpenDataSoft), and etalab (Etalab).

Table 2. Top significant users in 2015

Twitter User	Betweenness Centrality	Eigenvector Centrality	PageRank
bobbycaudill	17229445.35529	0.022	104.94
cliffordious	7273350.8382	0.011	66.88
odihq	2404534.366	0.005	39.05
worldbankdata	2394157.085	0.004	51.32
sunfoundation	1457670.371	0.003	27.56
worldbank	1377414.476675	0.003	50.01
dyntra_org	1357173.177111	0.000	52.38
okfn	1254514.256603	0.000	9.89
opendatasoft	1246005.093504	0.000	17.11
etalab	1045445.35237	0.002	19.99

In 2016, the first and second rankings were the same as in 2015. The next ranked accounts were worldbank (The World Bank), worldbankdata (World Bank Data), eu_dataportal (European Data Portal). They are regarded as international organizations harvesting metadata from public sector information available on public data portals across Europe. There are owenboswarva (UK-based data analytics firm working mostly in information rights), datosgob (Spanish national web portal that has organized and managed the Public Sector Information Catalog in Spain since 2009), opendatacon (annual International Open Data Conference), and open_govern (Spain-based company that focuses on OD transparency in Spain).

Table 3. Top significant users in 2016

Twitter User	Betweenness Centrality	Eigenvector Centrality	PageRank
bobbycaudill	41028283.111	0.019	243.47
cliffordious	13022129.076	0.007647	110.32
worldbank	10486979.118	0.002	225.15
worldbankdata	4708438.948	0.002	68.86
eu_dataportal	3279029.687	0.002	31.57
odihq	2386035.498	0.002	31.50
owenboswarva	1801399.329	0.002	29.35
datosgob	1792525.091	0.001	29.28
opendatacon	1745681.782	0.002	23.57
open_govern	1721730.646	0.001	9.02

Table 4 provides the top 10 words in terms of frequency (except for mentions and re-tweets) in which the top three terms were opendata (9,017), data (1,986), and open (1,053) in 2015 and opendata (12,339), data (3,217) and open (1,474) in 2016 with respect to frequency in the entire network. Table 5 shows the top 10 hashtags by frequency. The most common hashtags in the network in 2015 were #opendata (9,276), #bigdata (388), and #cfasummit (310), although other hashtags were frequently mentioned in message content. The most common hashtags in 2016 were #opendata (12,734), #bigdata (846), and #opengov (598).

Table 4. Top 10 words by frequency

Rank	Word (Year 2015)	Frequency	Word (Year 2016)	Frequency
1	opendata	9017	Words in Sentiment List#1: Positive	5896
2	data	1986	Words in Sentiment List#2: Negative	1898
3	open	1053	Words in Sentiment List#3: Angry/Violent	9
4	la	940	Non-categorized Words	212228
5	le	509	Total Words	220022
6	des	427	opendata	12339
7	app	414	data	3217
8	new	410	open	1474
9	out	393	https	1010
10	bigdata	383	new	977

Table. 5. Top 10 hashtags by frequency

Rank	Hashtags in tweets of the year 2015	Frequency	Hashtags in tweets of the year 2016	Frequency
1	opendata	9276	opendata	12734
2	bigdata	388	bigdata	846
3	cfasummit	310	opengov	598
4	data	268	cha	376
5	opengov	237	data	343
6	ogov	164	smartcity	259
7	transparencia	159	datosabiertos	239
8	odmsymposium	149	openscience	236
9	api	143	dataviz	200
10	whcitsci	142	africa	190

Because the Twitter website applies a 140-character limit, most of the URLs were identified by switching from the short version URL to the full URL, which identified the top five URLs of opendata tweets shown in Table 6.

Table 6. Top 5 URLs in the entire network by rank

Rank	URL in 2015	Frequency	URL in 2016	Frequency
1	http://www.dyntra.org/las-capitales-de-america-a-examen-en-transparencia-publica/	131	http://blogs.worldbank.org/opendata/can-we-quantify-learning-globally-measure-progress-sdg-4?CID=ECR_TT_worldbank_EN_EXT	200
2	http://blogs.worldbank.org/opendata/accessing-development-data-dynamic-mobile-world?CID=ECR_TT_worldbank_EN_EXT	102	http://blogs.worldbank.org/opendata/chart-worlds-youngest-populations-are-africa	189
3	http://blogs.worldbank.org/opendata/accessing-development-data-dynamic-mobile-world	98	http://blogs.worldbank.org/opendata/where-are-cheapest-and-most-expensive-countries-own-mobile-phone?CID=ECR_TT_worldbank_EN_EXT	137
4	http://blogs.worldbank.org/opendata/fr/entre-1990-et-2015-le-taux-de-mortalite-des-enfants-de-moins-de-cinq-ans-diminue-de-plus-de-moitie-l?cid=EXT_Twitterbanquemondiale_P_EXT	70	http://blogs.worldbank.org/opendata/chart-obstacles-firms-face-middle-east-and-north-africa	105
5	https://snlg.ht/KcBKB	61	http://www.goodnewsfinland.com/feature/open-data-creates-life-simplifying-apps/	94

The URLs of the Open Data Twitter network in 2015 were analyzed to obtain the top six URLs, which were related to Dyntra, The World Bank, and snlg. These URLs were websites for transparency of all the capital cities of the Europe continent and accessing development data in a dynamic mobile world. The seventh URL is an article on the theodi website that provides information about UK innovations across sectors and regions. The article explains the advantages of OD in social, environmental, and economic contexts for creating innovative products and services to fill gaps in markets that generate income. The eighth URL relates to the French blog on the Etalab mission prepared regarding an event related to OD in France. The ninth URL links to the PBS website (a domain name) regarding the power of OD with respect to shootings in the US. The tenth URL is Stephen Larrick's article on Sunlight Foundation's website in which he shares his insights about OD policies in cities. Regarding 2016, analysis of the URLs of Open Data found that the top four, the sixth, and the tenth ranked URLs were devoted to articles of The World Bank in the context of OD. The fifth ranked URL was an article on the goodnewsfinland.com website about OD as an application and software in the realm of ICT (Information and communication Technology) explaining the creation of friendly applications by using OD. The seventh ranked URL was an article on the European Commission website (europa.eu) declaring the amount of money invested in research and innovation for OD authorized by the European Commission was EUR 8.5 billion. The eighth ranked URL links to an "I Quant NY" website article that reveals the role of OD in New York City's budget error. The ninth URL is a call for proposals for "Africa Open Data Network Research Coordination" on the website of Open Data for Government (od4d.com), which was built to demonstrate the effects of OD on sustainable development challenges. Table 7 shows the top 10 domains of Open Data in the Twitter network in 2015 and 2016. It shows that Worldbank.org, Twitter.com, and Google.com were the most frequent among the domains in 2015. Furthermore, Worldbank.org, Twitter.com, and paper.li were the most frequent domains in 2016. Table 7 reveals that OD activists were partial to using the Worldbank website and governmental domains, such as gov.uk, as well as social media websites (namely, Twitter.com) and the search engine, google.com, as primary sources of information. Importantly, a comparison between the top domains in tweets of Open Data in 2015 and 2016 found some distinctive differences. By sharing OD, some domains, such as godan.info, provided information services on agriculture and nutrition, and thereby rose into the top 10 domains in 2016. These results suggest the importance gained from commercial use of OD in 2016 and reveal that OD had moved into other industries.

Table 7. Top 10 domains in entire network

Rank	Domain in 2015 Tweeter Network	Frequency	Domain in 2016 Tweeter Network	Frequency
1	worldbank.org	637	worldbank.org	1814
2	twitter.com	617	twitter.com	1604
3	google.com	288	paper.li	561
4	theodi.org	185	google.com	388

5	gouv.fr	159	chattlibrary.org	376
6	gov.uk	140	europa.eu	274
7	ac.uk	139	theodi.org	156
8	dyntra.org	131	gov.uk	137
9	co.uk	116	godan.info	132
10	paper.li	114	ow.ly	131

5. DISCUSSION AND CONCLUSIONS

This study examined the networked structure of Open Data by focusing on the Twitter social network data. It helps the authors determine the overall shape of the Open Data community. Network metrics were employed. Degree Centrality, Betweenness Centrality, Eigenvector Centrality, PageRank, and Clustering Coefficient were used to identify the ten most significant actors. In sum, most of the Twitter accounts were related to public organizations and persons that actively sent tweets about their ideas and activities in the context of OD. The results suggest that most of the tweets were related to concepts concerning OD, new applications, and new technologies, such as big data.The analysis of the URLs in the tweets found that the World Bank's website made a remarkable contribution to the Open Data discussion in Twitter in 2015 and 2016 based on the high frequency of connections to its webpages with OD content. Future research should include more quantitative and qualitative analysis, such as interviewing with the Open Data experts to provide a better understanding of the diffusion of the Open Data wave in worldwide level. Besides, applying a wider time-span namely 10 years (we applied a two-year time-period) is strongly suggested to understand better the diffusion shape of Open Data. This research is also valuable for scholars who study the social network websites such as Twitter and is also worthwhile for the governments, studying the statue of Open Data in a worldwide leve.

ACKNOWLEDGEMENTS

The authors have no conflict of interest to declare. Special thanks go to Dr. Marc Smith for his valuable assistance on data collection and sharing via NodeXL Gallery. The authors thank Hyo-Chan Park for his assistance on editing the manuscript.

REFERENCES

[1] R. Sandoval-Almazan and J. R. Gil-Garcia, "Toward an integrative assessment of open government: Proposing conceptual lenses and practical components," Journal of Organizational Computing and Electronic Commerce, vol. 26, no. 1-2, 2016, pp. 170-192.

[2] K. Jung and H. W. Park, "A semantic (TRIZ) network analysis of South Korea's "Open Public Data" policy," Government Information Quarterly, vol. 32, no. 3, 2015, pp. 353-358.

[3] A. Zuiderwijk and M. Janssen, "Open Data policies, their implementation and impact: A framework for comparison," Government Information Quarterly, vol. 31, no. 1, 2014, pp. 17-29.

[4] B. N. Madsen, S. Brier, K. E. L. Johansson, B. Hjørland, H. E. Thomsen, and H. S. Sørensen, "The Landscape Of Philosophy Of Science," Term Bases and Linguistic Linked Open Data, 2016, p. 212.

[5] J. Weng, E. P. Lim, J. Jiang, and Q. He, "Twitterrank: finding topic-sensitive influential Twitterers," In Proceedings of the third ACM international conference on Web search and data mining, ACM, 2010, pp. 261-270.

[6] T. M. Harrison, S. Guerrero, G. B. Burke, M. Cook, A. Cresswell, N. Helbig, and T. Pardo, "Open government and e-government: Democratic challenges from a public value perspective," Information Polity, vol. 17, no. 2, 2012, pp. 83-97.

[7] G. F. Khan, H. Y. Yoon, J. Y. Kim, and H. W. Park, "From e-government to social government: Twitter use by Korea's central government," Online Information Review, vol. 38, no. 1, 2014, pp. 95-113.

[8] M. Kassen, "A promising phenomenon of Open Data: A case study of the Chicago Open Data project," Government Information Quarterly, vol. 30, no. 4, 2013, pp. 508-513.

[9] H. Eskelinen, L. Frank, and T. Hirvonen, "Does strategy matter? A comparison of broadband rollout policies in Finland and Sweden," Telecommunications Policy, vol. 32, no. 6, 2008, pp. 412-421.

[10] S. Hunnius, B. Krieger, and T. Schuppan, "Providing, guarding, shielding: Open government data in Spain and Germany," In 2014 EGPA Annual Conference, Speyer, Germany, 10-12 Sep. 2014.

[11] R. Matheus, M. M. Ribeiro, and J. C. Vaz, "Brazil Towards Government 2.0: Strategies for Adopting Open Government Data in National and Subnational Governments," In Case Studies in e-Government 2.0, Springer International Publishing, 2015, pp. 121-138.

[12] R. P. Nugroho, A. Zuiderwijk, M. Janssen, and M. de Jong, "A comparison of national Open Data policies: Lessons learned," Transforming Government: People, Process and Policy, vol. 9, no. 3, 2015, pp. 286-308.

[13] T. Nam, and T. A. Pardo, "The changing face of a city government: A case study of Philly311," Government Information Quarterly, vol. 31, 2014, pp. S1-S9.

[14] G. Puron-Cid, "Factors for a successful adoption of budgetary transparency innovations: A questionnaire report of an open government initiative in Mexico," Government Information Quarterly, vol. 31, 2014, pp. S49-S62.

[15] T. Stamati, T. Papadopoulos, and D. Anagnostopoulos, "Social media for openness and accountability in the public sector: Cases in the Greek context," Government Information Quarterly, vol. 32, no. 1, 2015, pp. 12-29.

[16] W. W. Xu, I. H. Chiu, Y. Chen, and T. Mukherjee, "Twitter hashtags for health: applying network and content analyses to understand the health knowledge sharing in a Twitter-based community of practice," Quality & Quantity, vol. 49, no. 4, 2015, pp. 1361-1380.

[17] X. V. Meza and H. W. Park, "Organic Products in Mexico and South Korea on Twitter," Journal of Business Ethics, vol. 135, 2016, pp. 587-603.

Seyed Mohammad Jafar Jalali

He is a research fellow at Allameh Tabataba'i University in Information Systems. He graduated in Information Technology Engineering and he received his Masters of Science in Management of Advanced Information Systems in 2016 at Allameh Tabataba'i University, Iran, under the supervision of Dr. I. Raeesi Vanani. His thesis topic focuses on the study of the Business Analytics Emerging Trends using Text Mining approaches. His research interest is focused on Machine Learning, Data mining, Business Analytics, Social Network Analysis and Text analytics. He has published journal and conference papers in recent years in international communities.

Han Woo Park

He is a Full Professor in the Dept. of Media & Communication, Interdisciplinary Program of Digital Convergence Business, and Director of Cyber Emotions Research Institute at YeungNam University, South Korea. His research focuses on the use, impact, and role of open and big data in extending academic, governmental, and business networks in scientific, technical, and innovative activities. He was a pioneer in network science of open and big data in the early 2000s (often called Webometrics, Hyperlink Network Analysis, Link Analysis, etc.) when he used to work for Royal Netherlands Academy of Arts & Sciences (KNAW in Dutch), and was Principal Investigator of the World Class University (WCU) research project (2009~2011) funded by South Korean government. Professor Park has published more than 100 articles in Web of Science Journals (often called Social Science Citation Index). Several publications were included in top 10 list of downloads and citations. He has been co-awarded the best paper in EPI-SCImago in 2016 and included in the list of core-candidates of the Derek de Solla Price Memorial Medal in 2017. He co-authored more than 10 books by contributing a chapter on various topics on social media and virtual methods. He is currently the president (since 2013) of the World Association for Triple Helix and Future Strategy Studies (WATEF). Since the establishment of open data law in Korea, he has actively participated in the Strategic Committee of Public Data chaired by Prime Minister and international events including 'Big Data and the 2030 Agenda for Sustainable Development' workshop organized by UN-ESCAP. He has founded a prestigious conference on big data, government 3.0, and triple network in Asia, called DISC (Data, Innovation, Social Network, & Convergence) that is only conference acknowledged by the International Networks for Social Network Analysis (INSNA). Professor Park was a local organizer of the Internet Research15 conference of the Association of Internet Researcher (AoIR) that held in Asia for the first time. He also sat on the consulting boards of European Union project on open data (http://www.data4policy.eu) in collaboration with Oxford Internet Institute where he used to be a visiting scholar. He has been invited to give a keynote speech in a number of workshops and conferences including the International Communication Association (ICA), International Conference on e-Democracy and Open Government (CeDEM), and NetGlow. Further, he is strongly affiliated with several prestigious journals on open government and big data such as Big Data & Society, Technology Forecasting & Social Change, and Scientometrics. Further, he is a principal investigator of OpenData500.com on Korean side (http://www.opendata500.com/kr). Additionally, Professor Park was recently appointed as MICE (Meeting, Incentive tour, Convention, Exhibition) Ambassador for Daegu, Korea.