Who is who in Big Social Data?

Who is who in Big Social Data? A Bibliographic Network Analysis Study

Jari Jussila, Karan Menon, Jayesh Gupta and Hannu Kärkkäinen
Tampere University of Technology, Tampere, Finland
jari.j.jussila@tut.fi
karan.menon@tut.fi
jayesh.gupta@tut.fi
hannu.karkkainen@tut.fi

Abstract: The aim of the study is to investigate who are advancing the knowledge on Big Social Data and the related concept of Social Big Data, 'who' are these people citing and building their work on, and what are the topics and outlets where the discussion takes place. For that purpose, data was extracted from Thomson Reuters Web of Science with the search term "Big Social Data" and "Social Big Data" spanning the years from 2012 to 2016. The search resulted in 58 articles in 39 different outlets. In order to go into the depth of Big Social Data and Social Big Data, co-author bibliographic network analysis was performed on the extracted data. The co-author network analysis revealed 149 nodes (authors), and 308 edges (co-authoring relationships) between the authors. Betweenness centrality were calculated for the nodes to demonstrate who are the central authorities and their domain on the topic of Big Social Data and Social Big Data. The visualisation based on co-author network analysis provides insight into the possible clusters of authors in the topics of Big Social Data and Social Big Data. Co-citation analysis was performed for the combined network of Big Social Data and Social Big Data authors. This study was carried out using Ostinato process model for visual network analysis. The findings of the study provide insights on the leading authorities (authors) advancing the knowledge in Big Social Data. From the community of Big Social Data three authorative clusters were identified, one with authors located in Singapore and Scotland, another with authors located in Denmark, and third based in London, England. The Social Big Data communities were mainly located in Asia, with two authorative clusters, one located in Japan, and another with authors located in South-Korea and Spain. The topic modelling uncovered that the themes discussed in Big Social Data and Social Big Data communities were fairly similar, dealing with analysis of social media data in various ways. Most commonly the focus was on Twitter or Facebook data analysis. Further, the bibliometric analysis provides an indication for potential outlets (Journals and Conferences) for Big Social Data and Social Big Data themed articles, as well as, their impact on the field.

Keywords: Big Social Data, bibliography, network analysis, bibliometrics, co-author networks, co-citation analysis

1. Introduction

Big Data Analytics have become increasingly important in both the academic as well as the business communities over the past two decades (Olshannikova et al. 2017). There is already a large and vast growing body of literature on Big Data (BD) and related analytics in academic literature (see e.g. Akter & Wamba, 2016; Hilbert, 2016; Müller et al. 2016) However, the topic of Big Social Data (BSD) remains somewhat uncharted in existing literature.

In the literature BSD has been distinguished from the broader category of BD. According to Coté (2014) Big Data is any data produced as a result of the quantification of the world that may include data from sensors, multiple industrial and domestic networks as well as financial markets, whereas BSD is produced as a result of mediated communication practices of our everyday live, "*whenever we go online, use our smartphone, use an app or make a purchase*".

Social Big Data (SBD) is a related concept, which has been attributed different meanings. Bello-Orgaz et al. (2016) in their definition of SBD make the assumption that social big data comes from joining the efforts of the two previous domains: social media and big data, and go on to define SBD as "*those processes and methods that are designed to provide sensitive and relevant knowledge to any user or company from social media data sources when data sources can be characterized by their different formats and contents, their very large size, and the online or streamed generation of information*." Ishikawa (2015) in contrast, described Social Big Data science (SBD for short) as analyzing both physical real world data (heterogeneous data with implicit semantics) and social data (social media data with explicit semantics) by relating them to each other, thus not limiting SBD only to social media data. Guellil and Boukhalfa (2015) on the other hand distinguish SBD from BD based on the characteristics that are commonly attributed to social media data as described by Tang et al. (2015): "*the set of links (due to relationships between users), a nonstructural nature (due to length of messages required by* *some microblogging, the presence of spelling mistakes or other) and the lack of completeness (due to certain user requirements for data privacy)*."

Furthermore, some researchers (e.g. Mukkamala et al. 2014; Nguyen et al. 2015) use the terms BSD and SBD interchangeably (Olshannikova 2017). In majority of the articles BSD or SBD is actually either loosely defined or not defined at all.

In order to understand the emerging research stream of Big Social Data and its relationship to Social Big Data research stream, a bibliographic network analysis study was performed. The following research questions guided conducting the analysis:

1. Who are the authors that investigate Big Social Data, and who do they cite in their research?
2. How does Big Social Data and Social Big Data co-author networks relate to each other?
3. Who are the central authorities in Big Social Data research?
4. What domains can be identified from the Big Social Data studies?

2. Methodology

Bibliometric data analysis is conducted as a means to provide quantitative analysis of academic literature (Nicolaisen 2010). Bibliometrics is known as statistical analysis of written publications and citation analysis (Hajikhani 2017). As part of bibliometric study, a bibliographic network analysis was used to construct citation graph, a network or graph representation of the citations between documents.

The bibliographic network analysis was performed in two stages. In the first stage, Network Analysis Interface for Literature Studies "NAILS" (access from: http://nailsproject.net) was used to generate a network of all the documents and the citations between the documents (Knutas et al. 2015). In the second stage, Tethne bibliographic network analysis in Python (Peirson et al. 2016), tool was used to generate a co-citation network of the authors.

Data for the study was extracted from Thomson Reuters Web of Science with the topic search term "Big Social Data" and "Social Big Data" spanning the years from 2012 to 2016. The search resulted in a total of 58 articles, 37 with" Big Social Data", and 22 with "Social Big Data" search with one article appearing in both search results.

The network analysis of the 57 documents provided the basis for topic modelling of combined Big Social Data and Social Big Data documents, and also for performing a quantitative analysis of the literature, e.g. identifying the most important articles and authorities of the network that is the authors explicitly referring to their research as Big Social Data or Social Big Data research.

The co-citation network analysis revealed 149 nodes (authors being cited), and 308 edges (citations) between the authors. Betweenness centrality was calculated for the nodes to demonstrate who are the central authorities in the co-citation network and what are their domains (see e.g. Wasserman & Faust 1994). The visualisation based on co-citation network analysis provides insight into the possible clusters of authors in the topic of Big Social Data and Social Big Data. Gephi an open source software for graph and network analysis was used to create the visualisations (Bastian et al. 2009). ForceAtlas 2 (Jacomy et al. 2014) algorithm was used to layout the networks.

To overcome the main challenges and limitations of data-driven research (see e.g. Bruns 2013) this study was carried out using Ostinato process model for visual network analysis (Huhtamäki 2016; Huhtamäki et al. 2015). Ostinato process model enables several researchers to participate and collaborate in the data driven network analysis and provides guidelines on conducting the study in a way that make it both easier to follow and for other researchers to replicate.

3. Results

In this section, we first present the authors and authorities of BSD and SBD. Second, we present the domains common to BSD and SBD identified by topic modelling of abstract contents. Third, we present the most important papers using three importance measures from the citation network. Finally, we present the outlets of current BSD research.

3.1 Authors and authorities of Big Social Data

The authors of Big Social Data are illustrated by three network representations. In the first network representation, the co-authorship network of Big Social Data authors is illustrated (Figure 1). In the second network representation, the co-authorship network of Social Big Data authors is illustrated (Figure 2). In the Figure 3 a network representation of combined co-authorships of Big Social Data and Social Big Data is represented. The authorities from the combined network are then compiled into Table 1 and ordered based on network metrics (see e.g. Huhtamäki & Parviainen 2013).

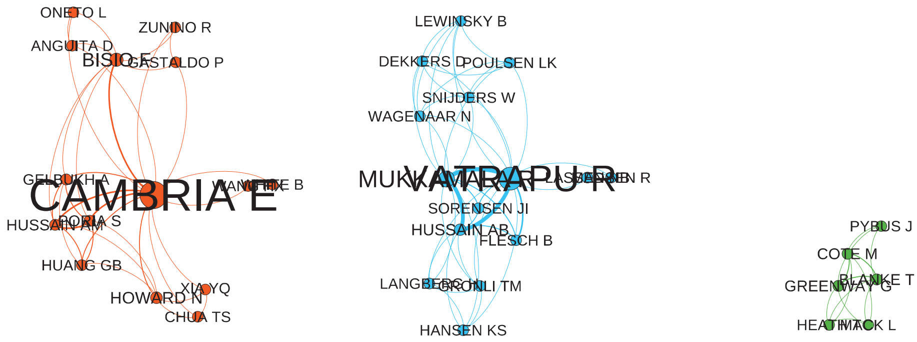


Figure 1: Co-author network of Big Social Data (see complete network at http://bit.ly/2016wosbsd).

The co-author network of BSD includes two larger clusters. Erik Cambria (CAMBRIA E) has been co-author (see e.g. Cambria et al. 2013; Cambria & Hussain 2012; Cambria et al. 2010) in most of the articles in orange cluster and Ravi Vatrapu in the light blue cluster. In addition to the two large clusters, there are several smaller less interconnected clusters.

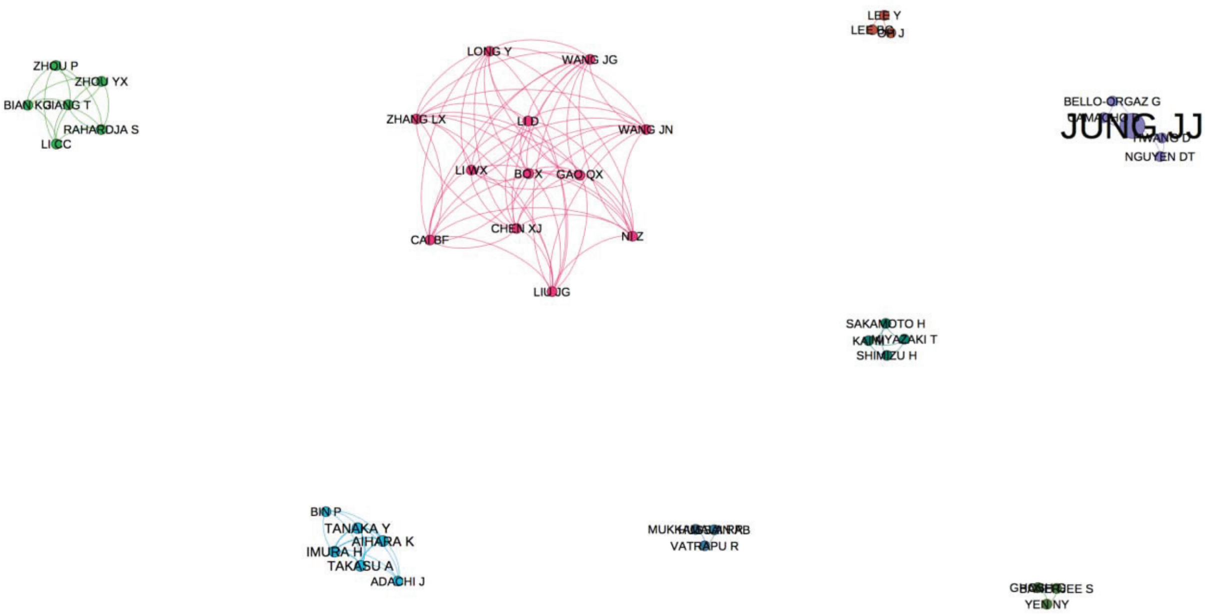


Figure 2: Co-author network of Social Big Data (see complete network at http://bit.ly/2016_wossbd)

From the co-author network of SBD only one name stands out as having a higher betweenness centrality than others. Jason J. Jung (JUNG JJ) has co-authored a paper with Bello-Orgaz and Camacho (Bello-Orgaz et al. 2016) defining Social Big Data and presenting its recent achievements and new challenges, and second paper a year earlier with Nguyen and Hwang (Nguyen et al. 2015) about detecting meaningful events from SBD. In the other clusters visible in the co-citation network of SBD there is no one author that has co-authored more papers than others. The only overlap between the SBD and BSD networks is the dark green cluster (MUKKAMALA R, VATRAPU R AND ABID AB) visible also in Big Social Data network. This is due one article (Mukkamala et al.

2014), where the authors present a modeling approach to social big data that integrate the conceptual, formal and software realms.

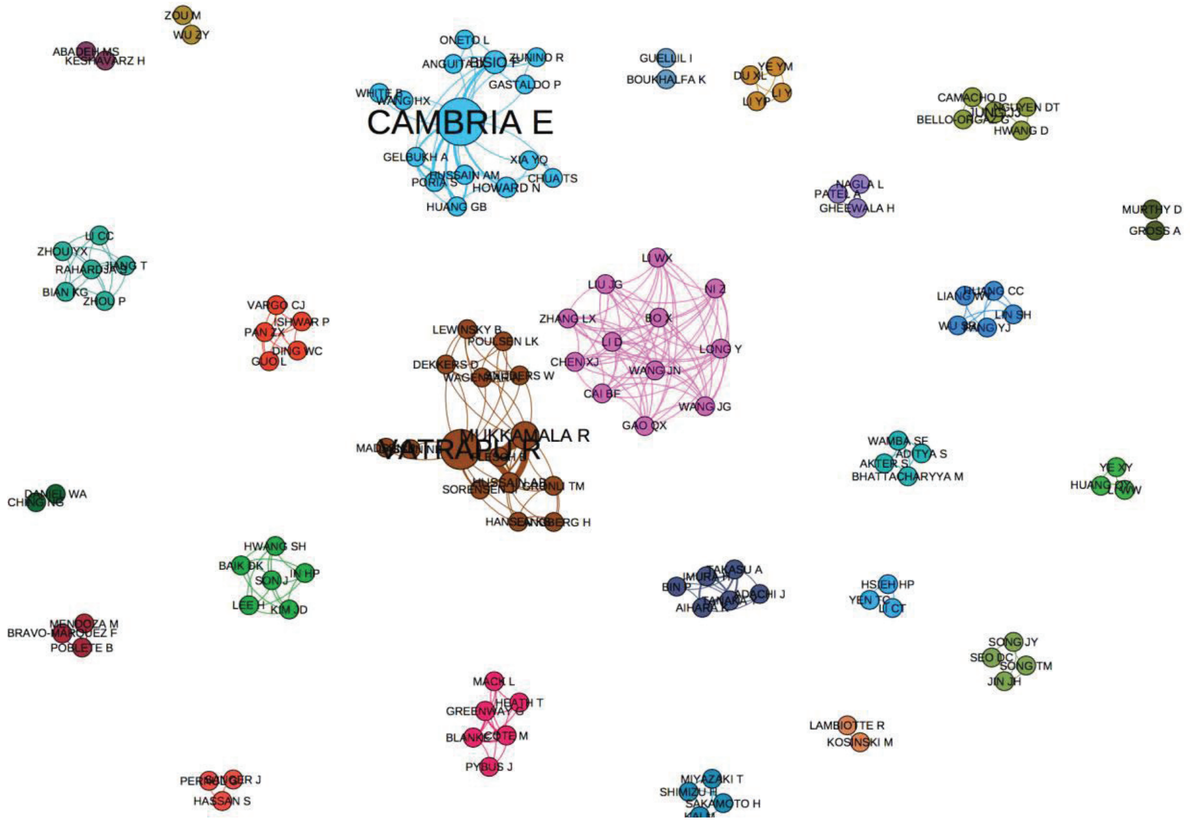


Figure 3: Combined co-author network of Big Social Data and Social Big Data (see complete network at http://bit.ly/2016_wosbsd-sbd)

In Figure 3 the BSD and SBD co-author networks are combined. This makes it possible to evaluate the authorship and authorities of the research stream in a comparable way. The complete co-citation network shows several clusters of 2-6 authors that work together, but are not connected to the larger clusters. Two of the largest clusters, light blue and brown, are BSD communities and the third large cluster, purple, represents SBD cluster. Table 1 presents the authors with the highest betweenness centrality in the combined co-author network.

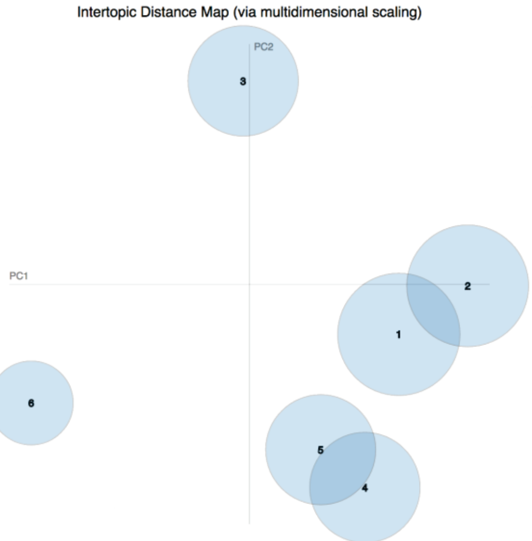
Table 1: The authors with the highest betweenness centrality in the co-author network.

Author	Institution	Betweenness centrality	PageRank
CAMBRIA E	NANYANG TECHNOL UNIV, SINGAPORE	55,80	0.018
VATRAPU R	WESTERDALS OSLO SCH ARTS COMMUN & TECHNOL, NORWAY	41,30	0.015
MUKKAMALA R	IT UNIV COPENHAGEN, DENMARK	17,30	0.013
HUSSAIN AB	COPENHAGEN BUSINESS SCH, DENMARK	2,30	0.008
HUSSAIN AM	UNIV STIRLING, SCOTLAND	0,83	0.008
PORIA S	UNIV STIRLING, SCOTLAND	0,83	0.008
COTE M	KINGS COLL LONDON, ENGLAND	0,67	0.008
GREENWAY G	KINGS COLL LONDON, ENGLAND	0,67	0.008
BLANKE T	KINGS COLL LONDON, ENGLAND	0,67	0.008
BISIO F	UNIV GENOA, ITALY	0,33	0.011
JUNG JJ	CHUNG ANG UNIV, SOUTH KOREA	0,17	0.011
HOWARD N	UNIV OXFORD, ENGLAND	0,13	0.008
TAKASU A	NATL INST INFORMAT, JAPAN	0,02	0.007
AIHARA K	NATL INST INFORMAT, JAPAN	0,02	0.007
HUANG GB	NANYANG TECHNOL UNIV, SINGAPORE	0,02	0.007
GELBUKH A	INST POLITECN NACL, MEXICO	0,02	0.007
IMURA H	HOKKAIDO UNIV, JAPAN	0,02	0.007
TANAKA Y	HOKKAIDO UNIV, JAPAN	0,02	0.007

The institutions of the authors in Table 1 provide some further clues about the BSD and SBD communities. Authors with the highest betweenness centrality belong to the BSD community. It can be observed that there is a large community of BSD authors (e.g. Cambria E, Hussain AM, Poria S, Howard N, Bisio F, Huang GB, Gelbukh A) from Singapore and Scotland (light blue cluster in Figure 3), see e.g. Oneto et al. 2016; Cambria et al. 2016; Poria et al. 2015) and another Denmark and Norway centric BSD community (brown cluster in Figure 3, e.g. Vatrapu R, Mukkamala R, Hussain, AB). It must be noted, however, the Ravi Vatrapu has two affiliations, primary affiliation with Copenhagen Business School in Denmark and another with Westerdals Oslo School of Arts, Communication and Technology in Norway. Also the current affiliation of Raghava Rao Mukkamala is Copenhagen Business School in Denmark. Ordering the authors by betweenness centrality also highlights one additional central BSD community based in London (red cluster in Figure 3), with authors Coté, Greenway and Blanke (e.g. Blanke et al. 2014) having a high betweenness centrality. Based on the betweenness centrality also two SBD communities can be identified, one SBD community is represented by green cluster top right corner in Figure 3 with authors from South-Korea and Spain (e.g. Jung JJ, Nguyen DT, Bello-Orgaz G) and another by dark blue cluster in Figure 3 with authors from Japan (e.g. Takasu, A, Aihara K, Imura H, Tanaka Y).

3.2 Topic modelling of Big Social Data studies

Topic modeling technique has been applied to analyze the abstracts contents. The technique is a type of statistical model for discovering the abstract "topics" that occur in a collection of documents in order to explore hidden semantic structures in a text body (Blei 2012). Application of the “Latent Dirichlet allocation” introduced by Sievert and Shirley (2014) was utilized in order to perform the topic generation of the analyzed abstracts. Figure 4 is an illustration of the popular distant topics/themes related to Big Social Data and Social Big Data. (The interactive visualization for the topical abstract analysis is available for download from https://github.com/jjussila/ECSM-2017/tree/master/results/topicmodelvis)



Top-10 Most Salient Terms

1	2	3	4	5	6
data	social	data	social	social	results
social	data	language	media	data	noise
analysis	set	natural	paper	user	users
big	analysis	learning	study	big	sentiment
mining	analytics	polarity	data	services	subjectivity
research	big	processing	analysis	network	opinions
networks	based	methods	crises	platform	twitter
information	approach	information	organization	suicide	study
media	formal	based	user	service	green
large	model	analysis	facebook	framework	nyc

Figure 4: Topic modeling of Big Social Data Studies. (Interactive visualization available for viewing from https://htmlpreview.github.io/?https://github.com/jjussila/ECSM-2017/blob/master/results/topicmodelvis/index.html)

Based on the topic modeling of abstracts with "NAILS" two overlapping clusters can be identified. Cluster 1 can be labeled as "Social Set Analysis" and cluster 2 as "Big Social Data". The second pair of overlapping clusters, include cluster 5 that can be labeled "Social Big Data", which overlaps with cluster 4 "Social Media Analysis". These two pairs of topic clusters correspond also to Figure 1 (Big Social Data co-author network) and Figure 2 (Social Big Data co-author network). Interestingly, these two pairs of clusters author on closely related research topics, but independently of each other. Cluster 3 and cluster 6 are more farther apart. Cluster 3 can be labeled as "Natural Language Processing" and cluster 6 can be labeled as "Opinion Mining". In the Top-30 most salient terms also individual social media services were visible: facebook was among the most salient in "Big Social Data" cluster and "Social Media Analysis" cluster, whereas Twitter was among the most salient terms in "Natural Language Processing" cluster.

3.3 Most important articles

The most important papers are identified below using three importance measures: 1) in-degree in the citation network, 2) citation count provided by Web of Science (only for papers included in the dataset), and 3) PageRank score in the citation network. The top 15 highest scoring papers are identified using these measures separately. The results are then combined and duplicates are removed. Results are sorted by in-degree, and ties are first broken by citation count and then by the PageRank.

Table 2: The most important articles

	Article	In-degree	Citation count	Page Rank
1	Mukkamala et al. 2014 ***Fuzzy-Set Based Sentiment Analysis of Big Social Data***	4	5	0.0007182
2	Poria et al. 2014 ***EmoSenticSpace: A novel framework for affective common-sense reasoning***	2	31	0.0006597
3	Cambria et al. 2014 ***Guest Editorial: Big Social Data Analysis***	2	20	0.0006562
4	Bravo-Marquez et al. 2014 ***Meta-level sentiment models for big social data analysis***	1	21	0.0006469
5	Bello-Orgaz et al. 2016 ***Social big data: Recent achievements and new challenges***	1	19	0.0006436
6	Poria et al. 2015 ***Sentiment Data Flow Analysis by Means of Dynamic Linguistic Patterns***	1	17	0.0006320
7	Nguyen et al. 2015 ***Time-Frequency Social Data Analytics for Understanding Social Big Data***	1	4	0.0006353
8	Mukkamala et al. 2014 ***Towards a Set Theoretical Approach to Big Data Analytics***	1	1	0.0006320
9	Mukkamala et al. 2015 ***Social Set Analysis of Corporate Social Media Crises on Facebook***	1	1	0.0006320
10	Poria et al. 2016 ***Fusing audio, visual and textual clues for sentiment analysis from multimodal content***	0	15	0.0009830
11	Lambiotte & Kosinski 2014 ***Tracking the Digital Footprints of Personality***	0	5	0.0006264
12	Halavais 2015 ***Bigger sociological imaginations: framing big social data theory and methods***	0	3	0.0006264
13	Wu & Zhou 2014 ***An incremental community detection method for social tagging systems using locality-sensitive hashing***	0	3	0.0006264
14	Aihara et al. 2014 ***Crowdsourced Mobile Sensing for Smarter City Life***	0	2	0.0006264
15	Flesch et al. 2016 ***Social Set Visualizer (SoSeVi) II: Interactive Computational Set Analysis of Big Social Data***	0	2	0.0006264

3.4 Outlets of Big Social Data research

We investigated only academic outlets, namely journal articles and conference articles, that discuss BSD and SBD research. Based on the Web of Science search, 25 journal articles were discovered illustrated in Table 1, and 32 conference articles illustrated in Table 3.

Table 3: Distribution of journal articles over time by publication outlet

	Number of articles per year		
Journal	2014	2015	2016
Cartography and Geographic Information Science			1
Econtent	1		
Expert Systems with Applications			1*
Future Generation Computer Systems			1
IEEE Access			1
Information Communication & Society		1	

Information Fusion			2
International Journal of Distributed Sensor Networks		1	
Journal of Adolescent Health			1
Journal of Environmental Management		1	
Journal of Information Science			1
Journal of Organizational and End User Computing			1
Journalism & Mass Communication Quarterly			1
Knowledge-based systems	3		
Neural Networks	3		
Neurocomputing			1
Proceedings of the IEEE		1	
Scientia Iranica		1	
The IEEE Computational Intelligence Magazine		1	2
Total	7	6	12

The journal publication outlets were mostly computer science focused, but also included journals in the fields of health, environmental management and mass communication. We found no BSD or SBD articles published in Journal outlets before the year 2014. One of the articles (Seo et al. 2017) was published in the year 2017, however accepted and made available in the year 2016, so we chose to include it in the study. The conference publication outlets are illustrated in Table 4.

Table 4: Distribution of conference articles over time by publication outlet.

	Number of articles per year		
Conference	2014	2015	2016
ACM International Conference on Multimedia		1	
Conference on IT in Business, Industry and Government	1		
Conference on Swarm Intelligence and Evolutionary Computation			1
International Conference on Advances in Social Networks Analysis and Mining			1
International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare			1
IEEE International Congress on Big Data	1	1	1
IEEE International Conference on Big Data	1	2	
IEEE International Conference on Big Data Computing Service and Applications			1
IEEE International Conference on Communications			1
IEEE International Conference on Information Reuse and Integration		1	
IEEE International Conference on Programming and Systems		1	
Industrial Conference on Data Mining			1
International Conference on Advances in Social Networks Analysis and Mining			1
International Conference on Big Data and Smart Computing	1		
International Conference on Computational Aspects of Social Networks			
International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare			1
International Conference on Distributed, Ambient and Pervasive Interactions			1
International Conference on Service Science		1	
International Conference on Service-Oriented Computing and Applications	1		
International Conference on Telecommunications			1
International Enterprise Distributed Object Computing Conference		1	
International Enterprise Distributed Object Computing Workshop	2		1
International Joint Conference on Neural Networks		1	
Proceedings of the 2015 ACM International Conference on Computer-Supported Cooperative Work and Social Computing		1	1
Symposium on Intelligent Distributed Computing		2	
Wireless and Optical Communication Conference		1	
Total	7	13	12

The BSD and SBD articles were also published between the years 2014-2016, with exception of one article, which was published in the International Conference on Computational Aspects of Social Networks (CASoN) in the year 2012 (Lee et al. 2012).

4. Discussion and Conclusions

The bibliographic network analysis focusing on co-author and co-citation networks of BSD and SBD was able to uncover clusters of authors that have co-authored on BSD and SBD. Based on the co-authoring and co-citation behaviour two large clusters were identified from the BSD communities, one based in Denmark and another with central authors in Singapore and Scotland. A third, smaller, but central BSD community was formed around authors located in London, England. There was only one large cluster of authors co-authoring in SBD communities (purple cluster in Figure 3), yet two smaller SBD clusters had higher betweenness centrality. The big picture from the analysed WOS bibliometric data does indicate that both BSD and SBD research is fragmented into relatively small communities around the world, with most active BSD research done in Europe and SBD research in Asia. Further, there is little overlap in co-authorship between researchers that label their research as BSD and the researchers that label their research as SBD, even though the topic modelling results point out to highly similar research topics.

For the authors of BSD and SBD research this study indicate potential new collaboration opportunities both thematically and from the research network perspective, also possible routes in bridging the gap between BSD an SBD research. For researchers and practitioners interested in gaining state-of-the-art knowledge on BSD and SBD this research highlights the most authorative authors, as well as, the most important publications based on the importance measures.

However, this study has several limitations. First of all, the search was limited to only one database, Web of Science, and only articles that matched the topic search string “Big Social Data” or “Social Big Data” were included. Using topic search strings “social data” (804), “social media data” (557), and “big data” (13995) could be included in further research. A systematic literature review study (e.g. Hajikhani 2017) between these three different concepts and Big Social Data could reveal to what extent the domains discussed and the authorities referred to are similar and different. Also one interesting avenue for further research would be to investigate bibliographic coupling and co-citation networks of BSD and SBD. The methodology described in this paper could be also used as basis for creating pre-understanding on any research topic. Further, it might serve as means in discovering research gaps in existing literature.

References

Aihara, K., Imura, H., Takasu, A., Tanaka, Y. and Adachi, J., 2014, November. Crowdsourced mobile sensing for smarter city life. In Service-Oriented Computing and Applications (SOCA), 2014 IEEE 7th International Conference on (pp. 334-337). IEEE.

Akter, S. and Wamba, S.F., 2016. Big data analytics in E-commerce: a systematic review and agenda for future research. Electronic Markets, 26(2), pp.173-194.

Bastian, M., Heymann, S. and Jacomy, M., 2009. Gephi: an open source software for exploring and manipulating networks. ICWSM, 8, pp.361-362.

Bello-Orgaz, G., Jung, J.J. and Camacho, D., 2016. Social big data: Recent achievements and new challenges. Information Fusion, 28, pp.45-59.

Blanke, T., Greenway, G., Pybus, J. and Cote, M., 2014, October. Mining mobile youth cultures. In Big Data (Big Data), 2014 IEEE International Conference on (pp. 14-17). IEEE.

Blei, D.M., 2012. Probabilistic topic models. Communications of the ACM, 55(4), pp.77-84.

Bravo-Marquez, F., Mendoza, M. and Poblete, B., 2014. Meta-level sentiment models for big social data analysis. Knowledge-Based Systems, 69, pp.86-99.

Bruns, A., 2013. Faster than the speed of print: Reconciling ‘big data’social media analysis and academic scholarship. First Monday, 18(10).

Cambria, E., Speer, R., Havasi, C. and Hussain, A., 2010, November. SenticNet: A Publicly Available Semantic Resource for Opinion Mining. In AAAI fall symposium: commonsense knowledge (Vol. 10, No. 0).

Cambria, E. and Hussain, A., 2012. Sentic computing: Techniques, tools, and applications (Vol. 2). Springer Science & Business Media.

Cambria, E., Schuller, B., Xia, Y. and Havasi, C., 2013. New avenues in opinion mining and sentiment analysis. IEEE Intelligent Systems, 28(2), pp.15-21.

Cambria, E., Wang, H. and White, B., 2014. Guest editorial: Big social data analysis. Knowledge-Based Systems, (69), pp.1-2.

Cambria, E., Howard, N., Xia, Y. and Chua, T.S., 2016. Computational intelligence for big social data analysis [guest editorial]. IEEE Computational Intelligence Magazine, 11(3), pp.8-9.

Coté, M., 2014. Data motility: The materiality of big social data. Cultural Studies Review, 20(1), p.121.

Flesch, B. and Vatrapu, R., 2016, September. Social Set Visualizer (SoSeVi) II: Interactive Computational Set Analysis of Big Social Data. In Enterprise Distributed Object Computing Workshop (EDOCW), 2016 IEEE 20th International (pp. 1-4). IEEE.

Guellil, I. and Boukhalfa, K., 2015, April. Social big data mining: a survey focused on opinion mining and sentiments analysis. In Programming and Systems (ISPS), 2015 12th International Symposium on (pp. 1-10). IEEE.
Knutas, A., Hajikhani, A., Salminen, J., Ikonen, J. and Porras, J., 2015, June. Cloud-based bibliometric analysis service for systematic mapping studies. In Proceedings of the 16th International Conference on Computer Systems and Technologies (pp. 184-191). ACM.
Kumar, S., 2015. Co-authorship networks: a review of the literature. Aslib Journal of Information Management, 67(1), pp.55-73.
Hajikhani, A., 2017, January. Emergence and dissemination of ecosystem concept in innovation studies: A systematic literature review study. In Proceedings of the 50th Hawaii International Conference on System Sciences.
Halavais, A., 2015. Bigger sociological imaginations: Framing big social data theory and methods. Information, Communication & Society, 18(5), pp.583-594.
Hilbert, M., 2016. Big data for development: A review of promises and challenges. Development Policy Review, 34(1), pp.135-174.
Huhtamäki, J. and Parviainen, O., 2013. Verkostoanalyysi sosiaalisen median tutkimuksessa. Otteita verkosta-Verkon ja sosiaalisen median tutkimusmenetelmät. Vastapaino, Tampere.
Huhtamäki, J., Russell, M.G., Rubens, N. and Still, K., 2015. Ostinato: The exploration-automation cycle of user-centric, process-automated data-driven visual network analytics. In Transparency in Social Media (pp. 197-222). Springer International Publishing.
Huhtamäki, J 2016, Ostinato Process Model for Visual Network Analytics: Experiments in Innovation Ecosystems. Tampere University of Technology. Publication, vol. 1425, Tampere University of Technology.
Ishikawa, H., 2015. Social big data mining. CRC Press.
Jacomy, M., Venturini, T., Heymann, S. and Bastian, M., 2014. ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. PloS one, 9(6), p.e98679.
Lambiotte, R. and Kosinski, M., 2014. Tracking the digital footprints of personality. Proceedings of the IEEE, 102(12), pp.1934-1939.
Lee, Y., Oh, J. and Lee, B.G., 2012, November. Logical push framework for real-time SNS processing. In Computational Aspects of Social Networks (CASoN), 2012 Fourth International Conference on (pp. 47-51). IEEE.
Mukkamala, R.R., Hussain, A. and Vatrapu, R., 2014, September. Fuzzy-set based sentiment analysis of big social data. In Enterprise Distributed Object Computing Conference (EDOC), 2014 IEEE 18th International (pp. 71-80). IEEE.
Mukkamala, R.R., Hussain, A. and Vatrapu, R., 2014, June. Towards a set theoretical approach to big data analytics. In Big Data (BigData Congress), 2014 IEEE International Congress on (pp. 629-636). IEEE.
Mukkamala, R.R., Sørensen, J.I., Hussain, A. and Vatrapu, R., 2015, September. Social set analysis of corporate social media crises on facebook. In Enterprise Distributed Object Computing Conference (EDOC), 2015 IEEE 19th International (pp. 112-121). IEEE.
Müller, O., Junglas, I., vom Brocke, J. and Debortoli, S., 2016. Utilizing big data analytics for information systems research: challenges, promises and guidelines. European Journal of Information Systems, 25(4), pp.289-302.
Nicolaisen, J., 2010. Bibliometrics and citation analysis: From the science citation index to cybermetrics.
Journal of the American Society for Information Science and Technology 61(1), pp.205–207.
Nguyen, D.T., Hwang, D. and Jung, J.J., 2015. Time-frequency social data analytics for understanding social big data. In Intelligent distributed computing VIII (pp. 223-228). Springer International Publishing.
Olshannikova, E., Olsson, T., Huhtamäki, J. and Kärkkäinen, H., 2017. Conceptualizing Big Social Data. Journal of Big Data, 4(1), p.3.
Oneto, L., Bisio, F., Cambria, E. and Anguita, D., 2016. Statistical learning theory and ELM for big social data analysis. IEEE Computational Intelligence Magazine, 11(3), pp.45-55.
Peirson, B. R. Erick., et al. 2016. Tethne v0.7. http://diging.github.io/tethne/
Poria, S., Gelbukh, A., Cambria, E., Hussain, A. and Huang, G.B., 2014. EmoSenticSpace: A novel framework for affective common-sense reasoning. Knowledge-Based Systems, 69, pp.108-123.
Poria, S., Cambria, E., Gelbukh, A., Bisio, F. and Hussain, A., 2015. Sentiment data flow analysis by means of dynamic linguistic patterns. IEEE Computational Intelligence Magazine, 10(4), pp.26-36.
Poria, S., Cambria, E., Howard, N., Huang, G.B. and Hussain, A., 2016. Fusing audio, visual and textual clues for sentiment analysis from multimodal content. Neurocomputing, 174, pp.50-59.
Seo, Y.D., Kim, Y.G., Lee, E. and Baik, D.K., 2017. Personalized recommender system based on friendship strength in social network services. Expert Systems with Applications, 69, pp.135-148.
Sievert, C. and Shirley, K.E., 2014, June. LDAvis: A method for visualizing and interpreting topics. In Proceedings of the workshop on interactive language learning, visualization, and interfaces (pp. 63-70).
Tang, J., Chang, Y. and Liu, H., 2014. Mining social media with social theories: a survey. ACM SIGKDD Explorations Newsletter, 15(2), pp.20-29.
Wasserman, S. and Faust, K., 1994. Social network analysis: Methods and applications (Vol. 8). Cambridge university press.
Wu, Z. and Zou, M., 2014. An incremental community detection method for social tagging systems using locality-sensitive hashing. Neural Networks, 58, pp.14-28.