emerald**insight**

The Electronic Library
Public libraries: roles in Big Data
Ming Zhan, Gunilla Widén,

Article information:

Users who downloaded this article also downloaded:

(2017),"Big data application framework and its feasibility analysis in library", Information Discovery and Delivery, Vol. 45 Iss 4 pp. 161-168 https://doi.org/10.1108/IDD-03-2017-0024

(2017),"Defining big data and measuring its associated trends in the field of information and library management", Library Hi Tech News, Vol. 34 Iss 9 pp. 21-24 https://doi.org/10.1108/LHTN-05-2017-0035

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Public libraries: roles in Big Data

Ming Zhan and Gunilla Widén
Department of Information Studies, Åbo Akademi, Turku, Finland

Abstract

Purpose – The purpose of this paper is to explore the roles of public libraries in the context of Big Data.

Design/methodology/approach – A mixed method approach was used and had two main data collection phases. A survey of public libraries was used to generate an overview of which professional roles connect public libraries with Big Data. Eight roles were identified, namely, educator, marketer, data organiser, data container, advocator, advisor, developer and organisation server. Semi-structured interviews with library directors and managers were then conducted to gain a deeper understanding of these roles and how they connect to the library's overall functions.

Findings – Results of the survey indicated that librarians lack a proper comprehension of and a pragmatic application of Big Data. Their opinions on the eight roles are slightly stronger than neutral. However, they do not demonstrate any strong agreement on these eight roles. In the interviews, the eight roles attained more clear support and are classified into two groups: service-oriented and system-oriented roles.

Originality/value – As an emerging research field, Big Data is not widely discussed in the library context, especially in public libraries. Therefore, this study fills a research gap between public libraries and Big Data. In addition, Big Data in public libraries could be well managed and readily approached by citizens in undertaking such roles, which entails that public libraries will eventually benefit from the Big Data era.

Keywords Big Data, Public libraries, Professional roles

Paper type Research paper

Introduction

Even though there is no accurate figure for the amount of data created daily, there is no doubt that we are living in a generation of data explosion. *Data* is defined as unprocessed information in information science (Hey, 2004), such as footfall, online browsing history and trip routes, which are automatically recorded every day. Accordingly, society is faced with, on the one hand, challenges in handling issues, such as privacy protection and data storage capabilities, and, on the other hand, opportunities that could and should be realised. As such, Big Data emerges concomitantly, which changes the way society adapts to manage and govern the data (Chen and Zhang, 2014).

According to studies by Heidorn (2011) and Gordon-Murnane (2012), libraries are now reaching a data-richness condition as well, owing to their ease of access to the internet and the worldwide availability, affordability and applicability of digital devices, the increasing number of digital resource types and the advanced technology necessary for data collecting, recording, analysing and aggregating. In other words, libraries are heading towards a situation where Big Data is continuously important due to technological developments and the condition of data-richness. Therefore, the influence stemming from Big Data is obvious in the context of libraries. As a knowledge hub, public libraries undertake the role of supporting citizens in organising their personal information. As Big Data has been demonstrated to have positive effects on pragmatic processes, such as knowledge

The Electronic Library
Vol. 36 No. 1, 2018
pp. 133-145
© Emerald Publishing Limited
0264-0473
DOI 10.1108/EL-06-2016-0134

generation (Fuchs *et al.*, 2014), user behaviour forecasting (Xiang *et al.*, 2015) and decision-making (Chen *et al.*, 2014), the ways in which Big Data is integrated into the current library system and transformed into valuable operations are relevant for public libraries in developing their services. Hence the question:

RQ1. What is the role of the public library in managing Big Data and bringing Big Data to citizens?

To answer *RQ1*, empirical research was conducted. Improvements in library services are anticipated through the investigation of the roles and uses that public libraries could employ when using Big Data. Big Data in public libraries could be well managed and, therefore, readily approached by citizens. There are two justifications for studying Big Data in public libraries. Firstly, this study is one part of a two-year government funded project, Big Cities Meet Big Data, which concentrates on the ramifications of Big Data in the public sector. Public libraries, as an element of the public sector, were chosen as an area of study for the project. Secondly, as an emerging research field, Big Data has not been widely discussed in the context of libraries, most specifically public libraries. Therefore, there is a research gap on the subject of public libraries and Big Data.

The paper is structured with a literature review of Big Data in libraries, followed by the methodology and research data. The results are presented with their implications and discussion. Conclusions are made and expectations for future studies are put forward.

Literature review
Big Data and the library
Libraries are in the Big Data era, as indicated by the availability of various digital devices; diverse origins of data; the enhanced ability to collect, store and handle data; and the predisposition to use data for decision-making (Affelt, 2015; Gordon-Murnane, 2012). As such, the model of the library has begun to evolve from Library 2.0 (emphasising user participation) and Library 3.0 (facilitating the management of user-generated content) to Library 4.0 "where not only inference and research are available, but the system will analyse information by itself and discuss findings with users" (Noh, 2015, pp. 791-792). After reviewing the relevant literature, Noh further notes that in Library 4.0, the volume of data and services to be managed by future libraries will be massive. Therefore, future libraries can be called *massive data libraries*, where Big Data plays the main role. Furthermore, libraries will not only be influenced by the advent of Big Data but also fuel the development of Big Data (Wittmann and Reinhalter, 2014). Therefore, it is rational to investigate public libraries in the context of Big Data.

What is Big Data?
Currently, there is no consensus on a definition for Big Data in librarianship. Hoy (2014, pp. 321-322) referred to Big Data as "the idea that computers can gather trillions of pieces of information about billions of different things and find useful patterns in that information" after reviewing definitions generated from previous studies. Federer (2016, p. 36) defined Big Data as four Vs: volume, the scale of data; velocity, the speed at which data are created; variety, the type of data; and veracity, the reliability and integrity of data. Aiming to outline the essential features of Big Data, De Mauro *et al.* (2016) reviewed 1,437 conference papers and journal articles and concluded that technology, information, method and impact are four features of Big Data. They defined Big Data as an "information asset characterised by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value" (p. 131). Although different aspects have been emphasized

in defining Big Data, one universal point is that Big Data has increased technological transformation in the library, and such a transformation requires librarians and information professionals to take on new roles (Affelt, 2015; Gordon-Murnane, 2012; Hoy, 2014; Wittmann and Reinhalter, 2014).

The influence of Big Data on library roles

According to Gordon-Murnane (2012), Hoy (2014) and Wittmann and Reinhalter (2014), librarians should take on more data-specific roles as Big Data enhances the data services within libraries. Libraries are needed to organise data, provide access to internal and external data sets, authorise copyright on property issues and facilitate the processes for reusing data and training users. To fulfil these functions, the skills of librarians should be updated. For instance, indexing and abstracting skills could be incorporated with Big Data technologies to locate valuable resources within a larger data resource. Reference interviewing skills are also important to understand what a customer needs and, thus, help to provide the best Big Data solution (Affelt, 2015; Hoy, 2014). Affelt (2015) also highlights three roles for information professionals working with Big Data: curator who determines where and how to obtain it; data cleanser who removes erroneous and duplicated data; and data archive manager who builds and maintains data warehouses. Eventually, the role of libraries is expected to evolve and encompass new roles to be performed by librarians and information professionals in the context of Big Data. According to each of the studies done by Hoy (2014), Huwe (2014) and Teets and Goldner (2013), libraries are well positioned to work with Big Data. Teets and Goldner (2013) stated that libraries should take on the function of sharing vast amounts of library collection data to benefit a larger audience and establish systems to forecast user patterns by collecting Big Data. For example, a PhD student from a small European institution may be concentrating on a similar project to that of an American professor working in a major research organisation; Big Data may help recognise that similarity and help them combine resources.

Big Data generates new roles, responsibilities and challenges for libraries. Shen and Varvel (2013) emphasized, that although the promise of Big Data might exist, challenges must be met by academic libraries. Hence, effective data management services are required. To develop practical ideas for data management services in academic libraries, a case study of Johns Hopkins University (JHU) was conducted. A service-model framework was established in JHU's data management services, which includes three aspects: environmental responsiveness, socio–technical readiness and marketing and collaborations. The primary factors for success which were discussed were how to ensure services based on Big Data. One way is adoption, which refers to how much the service is used and how much data are used and reused. Another factor is acceptance, which means how the service is varied as well as how it is appreciated in general. Overall, it was concluded that various factors, such as an organisation's financial situation, staffing abilities and organisational cultures, should all be considered when providing data management services based on Big Data.

All these studies indicate that Big Data and libraries could be readily combined with each other. Nevertheless, Big Data is mainly discussed in the library field from a general point of view. Or, to put it another way, only a few studies have examined Big Data in specific kinds of libraries, such as public libraries. This motivates the current study to outline some roles that public libraries could and should undertake in the context of Big Data to better serve citizens, communities and organisations.

Data collection and methodology

The purpose of the present study is to explore the roles of public libraries in the context of Big Data. A mixed method approach was used and had two main data collection phases. A survey of public libraries was used to generate an overview of which professional roles connect public libraries with Big Data. Then, semi-structured interviews with library directors and managers were conducted to gain a deeper understanding of these roles and how they connect to the library's overall duties.

An online survey was conducted during the first stage. The aim was to collect opinions from librarians so as to pinpoint their preferences regarding the different roles that could be undertaken by public libraries in the context of Big Data. As the combination of public libraries and Big Data is in its infancy, no mature questionnaires could be used or referred to, so a questionnaire was designed. To include suitable content in the questionnaire, three librarians from a university library were interviewed. University librarians were selected for this step of acquiring a more general view because university libraries collaborate with researchers who have often already worked with the challenges stemming from Big Data. Therefore, it was assumed that these experiences might provide additional interpretations of the concept and enrich the discovery of library roles. In the end, seven roles were identified, namely, educator, marketer, data organiser, data container, advocator, advisor and developer. As public libraries provide services not only to individuals but also to organisations, interpreted by Stejskal and Hajek (2015) as having roles regarding the providing of services for organisations as well as individuals, the additional role of organisation server was included in the survey. Thus, eight roles were identified and their findings are supported by numerous studies (Affelt, 2015; Federer, 2016; Gordon-Murnane, 2012; Hoy, 2014; Teets and Goldner, 2013; Wittmann and Reinhalter, 2014). On basis of these eight roles, a questionnaire was designed. This questionnaire was sent to professionals for reviewing before being officially sent out to ensure its efficacy and reliability. Table I includes an explanation of the eight roles identified.

The questionnaire was delivered online. It was sent out on 27 January 2016. Respondents had until 10 March 2016 to complete the survey. Results were recorded automatically in a Web-based format. The URL of the survey was sent to the reference e-mail address of public libraries in Finland (available at the website www.libraries.fi). Eventually, 49 responses were successfully attained. One was duplicated and thus deleted. A total of 48 cases were analysed using SPSS 17.0.

Semi-structured interviews were conducted with library managers and directors to gain a deeper comprehension of the feasibility and applicability of public libraries with respect to the eight roles. The interviewees were selected from big city libraries in Finland. The rationale was to choose city libraries that are pioneers in Finnish librarianship and have branch libraries in small towns in their region with densely populated suburbs. Therefore,

Role	Explanation
Marketer	Making the Big Data concept and the use of Big Data known to citizens
Educator	Helping citizens understand what Big Data is
Data organiser	Cleansing and maintaining sets of Big Data
Data container	Storing Big Data
Advocator	Supporting individuals in using Big Data
Adviser	Provide advice to solve personal issues from a Big Data point of view
Developer	Using Big Data to develop current and new services
Organisation server	Serving other organisations from a Big Data point of view

Table I.
Explanation of the eight roles

the opinions of the leaders of such libraries should have a strong influence on forming ideas about Big Data and libraries, and represent a broad range of views. Eleven interviews were carried out between 26 October and 24 November November 2016. The average time for each interview was 30 min. All interviews were conducted in English and were recorded and transcribed manually. During the interview, the interviewees' daily practices in relation to data, their attitudes towards the eight roles identified in the survey and their opinions on Big Data were discussed. Content analysis was used to analyse the interview transcripts.

Results
Results of the survey
The survey is composed of three parts: demographic information, the perception of Big Data and the librarians' attitudes towards the eight roles. There were twice as many females as males, and most people were aged between 26 and 40 years. More than half of the respondents had a master's degree, and nearly 30 per cent of the representatives had a bachelor's degree, together accounting for over 80 per cent of the respondents. More than 81 per cent of the respondents had worked in a public library for at least three years. The results also show that 46 per cent of the respondents work with "library loans/document delivery" and that 60 per cent of them perform more than one kind of job. The responses regarding degrees, length of employment and broad working areas imply that these librarians are sufficiently professional to provide their opinions on the topic. The perception of Big Data is designed to determine whether current librarians have practical experience with Big Data, thus making it possible to interpret how well they understand this concept. As shown in Figures 1 and 2, most of the respondent librarians had heard of Big Data as a concept. Nevertheless, more than half of them did not have hands-on experience of dealing with Big Data. Meanwhile, it is worth highlighting that the number of librarians with experience of Big Data is not much lower than that of those who have not.

In addition, the perceptions of the respondents were implicitly evaluated. The characteristics of Big Data (volume, variety, velocity and value) (Huang *et al.*, 2015) were expressed, respectively. The frequency (from 0 for *never* to 6 for *always*) of dealing with data that have such characteristics was measured. The more often they work with such data, the deeper their perception of Big Data is expected to be. In the end, the overall mean value was 2.84, which indicates that the frequency of librarians working with Big Data occurs just a few times a month. The means of the items concerning volume, velocity and value are far lower than that of the item concerning variety. The unbalanced working frequency regarding different data characteristics was easily understood based on the responses of the librarians.
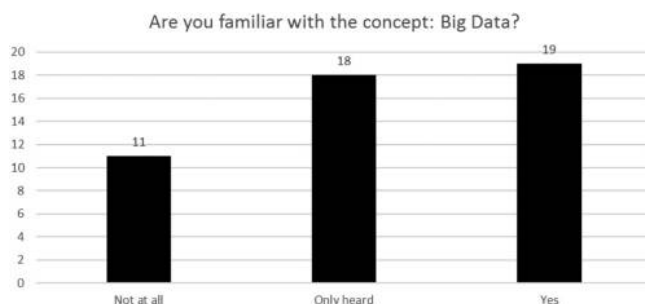

Figure 1.
Responses about familiarity with the Big Data concept

The overall mean values of each role are all higher than four, which is the scale of "either agree or disagree". It can be concluded that librarians do not disagree completely, especially regarding roles, such as marketer, educator, advocator, advisor and developer; thus, they tend to somewhat agree with adopting these roles. The role as organisation server received the lowest score, almost four. This could imply that the librarians do not have a clear opinion on this role. Regarding individual items, Item 17 ("The public library should establish data warehouses to store and preserve data generated from users or from projects incorporating other public sectors") was given the lowest grade, a mean value slightly lower than four, which means the respondents slightly disagree with this statement. Item 20 ("The public library should give users tools or instructions rather than the actual result when users have trouble in managing personal information") received the highest mean value in Table II, indicating that most librarians positively agree with this statement. Checked individually, the results are visualised in Figure 3.

As presented in Figure 3, 42 per cent of the respondents agree with the proposed roles. Half of the respondents are in the middle, from neutral to slightly agreeing. Among them, eight people had a mean value higher than 4.5, which means they show a slight tendency to

Figure 2.
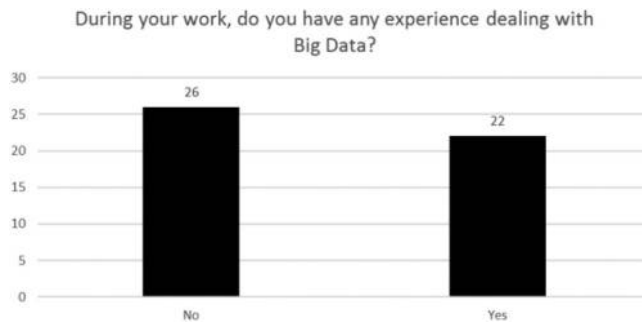Responses about experiences dealing with Big Data

Roles	Item	Mean value	Overall mean value
Marketer	I12	5.27	5.06
	I13	4.85	
Educator	I14	4.83	5.22
	I15	5.60	
Data organiser	I16	4.54	4.62
	I18	4.69	
Data container	I16	4.54	4.64
	I17	*3.81*	
	I19	5.58	
Advocator	I20	*5.79*	5.33
	I21	4.86	
Advisor	I22	5.38	4.96
	I23	4.54	
Developer	I24	5.00	4.86
	I25	4.71	
Organisation server	I26	4.10	4.10
	I27	4.10	

Table II.
Result of the librarians' opinions on different library roles

agree. Thus, the number of respondents agreeing is 28, nearly 60 per cent of all respondents. The lowest individual score was 2.8, which implies that this librarian disagrees with all these proposed roles to some extent, and the highest score was 6.8, which is very close to the maximum scale of the survey: 7.

At the end of the survey, an open question was asked to see whether additional roles could be added. Nevertheless, no role was put forward and the librarians expressed their opinions on Big Data in these terms instead: "the concept of Big Data itself is rather vague"; and "not sure what Big Data is".

Such expressions echo the result of the librarians' perception on Big Data and signify an insufficient comprehension of the benefits and challenges of Big Data for current librarianship. In conclusion, the result of the survey indicates that librarians lack a proper comprehension of and a pragmatic application of Big Data. Their opinions on the eight roles are slightly stronger than neutral. However, they do not demonstrate any strong agreement on these eight roles.

Interview results
The interview data were analysed using content analysis to systematically categorise the findings. The focus is on identifying the awareness and perceived possibilities of Big Data in public libraries and on the possible roles that librarians in public libraries can undertake in this context.

Big Data is a new phenomenon in Finnish public libraries. Almost every interviewee mentioned that they were unfamiliar with Big Data because it is new and has not been widely discussed in their libraries:

> I think we are just at the beginning of Big Data; and

> [...] in our library, there are 200 people working and I would say that many of them even don't know what Big Data is.

This, on one hand, explains the low perception of Big Data reflected by the findings of the survey owing to the infancy of Big Data in libraries. On the other hand, it can indicate that all the library directors are mainly unfamiliar with Big Data. Although not knowing the exact definition of Big Data, they readily related Big Data to a "huge amount of data" or "data generated on social media", which falls under the definition of Big Data to some extent. Therefore, they hold positive opinions towards Big Data being applied in public libraries:
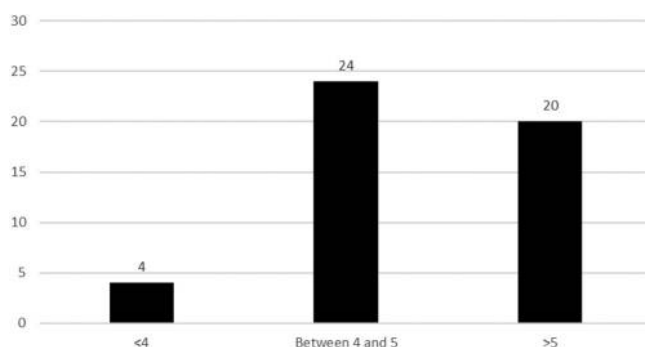

Figure 3.
The overall mean value of the proposed roles

I am very positive about it. I think using Big Data will give better tools to serve citizens, to show politicians what we are doing.

There are lot of things where we can dig into when we learn Big Data.

Well, there are risks as we told. But it could be a huge possibility to us to use that (Big Data).

140

The interviewees consider Big Data to be an effective approach for understanding the requirements of library patrons. Furthermore, tools could be developed with Big Data to support decision-making processes. In addition, Big Data boosts data reusability. As discussed, data reusability means reusing the data not only inside the library but also other public data; for instance, the use of parking places around the library area – people might not go to the library very often if parking places around the library are hard to find. It suggests that libraries could develop services, such as drive-in collection points. This example demonstrates the significance of comprehending the needs of the whole of society. As suggested by half of the library managers and directors, it is salient for libraries to understand their society and their city before they start to provide services that use Big Data.

However, their optimistic attitude towards Big Data is predicated on three preconditions. In their opinions, the three preconditions should be met; otherwise, it would be challenging for libraries to utilise Big Data. The conditions are as follows:

- sufficient finances to improve the current infrastructure and technically harness Big Data;
- sufficient personnel to exploit the values of Big Data; and
- authorisation and legislation regarding the protection of individual privacy.

As individual records are stored in the library system, obtaining the permission of citizens to use their personal data is a challenge for libraries wishing to use Big Data. Furthermore, laws need be passed to provide public libraries with the legal right to explore personal data.

Opinions about the roles examined in the survey. Owing to the lack of hands-on experience in handling Big Data, library directors and managers discussed their understandings of the eight roles relating to the projects that they have done. Even though not every role was realised in each library, these roles are agreed as covering the main scope of the responsibilities of a Finnish public library. However, libraries should consider their main responsibility when deciding which roles should be undertaken. Every interviewee mentioned that the role of storing data should be undertaken by the National Library of Finland or outsourced to commercial systems so that all the other public libraries could easily access the data. One responsibility of the National Library of Finland is to provide data services to all libraries (public or academic) in Finland. Therefore, it would be wise to select the National Library of Finland as the only storage centre for library Big Data. In addition, four directors and managers repute the idea that such a data storage role could be accomplished by library system vendors. Currently, the library system in northern Finland is provided and maintained by private companies. Millions of pieces of information and data are recorded on the system server. Therefore, it may be effective to outsource such a role to library system providers as they have more experience at storing data.

The data organiser role is accepted by library directors and managers. As public libraries in each region might have their own development strategy, it might lead to various requirements for the data. For example, geographic data are currently analysed for mapping out user needs in different locations in several cities. Therefore, public libraries could extract

different data from the data storage centre and establish their own data sets for further utilisation. They also need to act as a data organiser after establishing the data set. However, small scale libraries might not participate in this role in that they are led by the main library in their region.

In the interviews, the role of educator, marketer, adviser and advocator is reflected by each library's involvement in social media and the internet. For instance, three library managers mentioned that courses concerning social media are arranged in primary schools and in the homes of elderly people. During the course, the benefits and applications of social media are introduced. Additionally, citizens could ask for specific advice to manage their information needs with the help of social media. According to three directors, a Facebook home page, posters and face-to-face communication are used to market new developments in their libraries. Thus, the interviewees saw the possibility and significance of public libraries performing these roles with the advent of Big Data. However, the limited time and resources of librarians make it difficult for libraries to undertake all these four roles together. Phrases such as "no money" and "short of staff" were constantly mentioned during the interviews. Therefore, a combined role was put forward by two managers: that of facilitator:

[. . .] the way we could cover educator or marketer, advocator, and adviser is if the libraries work as a facilitator [. . .] so the library again is the medium between the actual experts of Big Data [. . .] and provide the possibilities, the means to meet people, and help them meet the actual experts.

Regarding the roles of developer and organisation, there are not many experiences that can be used to reflect on Big Data. Nevertheless, library directors and managers are positive about these two roles. As mentioned, many technologies or new tools could be applied in the library field as they are of potentially great value for developing services. However, a lack of knowledge and experienced professionals is an obstacle. In addition, the interviewees are looking forward to collaborating with other organisations. They hope that certain public sector organisations and private companies would consider the public library a trustworthy place to which they might turn, as suggested by the following quote: "our library system should be able to give our data for private companies".

Thus, the public library could be expected to offer some library data sets to other organisations. These organisations could then utilise the data sets according to their own strategy to attain a specific result. The aim of such a process is to broaden the audience for library services from citizens to organisations. However, as all public libraries now lack staff, few librarians would have the time to make plans to put everything into practice.

Discussion

Big Data, as a term, is no longer totally new in libraries, which is confirmed by the fact that more than 75 per cent of the respondents in the survey had heard of Big Data, not to mention that nearly 46 per cent of them had had practical experience of handling Big Data. However, hearing about Big Data does not mean understanding it. This situation could be reflected by the unbalanced working frequency in managing data with Big Data characteristics, and this was apparent in the answers to the open question. As they do not usually handle data in great volume and rapidly increasing speed, and rarely create value from such amounts of data, this could indicate that they might have a limited understanding of Big Data. This, to some extent, could be attributed to the statement by library directors and managers that Big Data is a new phenomenon. However, according to the responses from the survey and statements in the interviews, the respondents' unfamiliarity of Big Data tends to be on the conceptual level. That is, they are not familiar with the definition of Big Data. Nevertheless,

they are capable of relating Big Data to relevant items; therefore, inferences on how it could be used can be made owing to the librarians' reflections on similar practices.

The natural match between public libraries and Big Data
On the basis of the positive attitude of the library managers and directors and the results of the open question in the questionnaire, it can be concluded that there is a natural match between public libraries and Big Data. Respondents made statements, such as the following:

Libraries should be at the front and centre in developing the use of Big Data [. . .]

Libraries have a big and important role in dealing with Big Data [. . .]

The library sits on Big Data so the subject is unavoidable [. . .]

All the statements and the feedback obtained during the interviews imply that using Big Data would be a natural process for public libraries. This viewpoint echoes previous studies (Hoy, 2014; Huwe, 2014; Noh, 2015; Wittmann and Reinhalter, 2014). A public library is not only a data generator but also a data container. Numerous data are recorded within a public library. As was discussed during the interviews, millions of collection data, library economic statistics, user information and borrowing histories are normally stored in a public library. In addition, digitalisation in public libraries is promoted by the Finnish Government, boosting the volume, variety and velocity of digital data recorded within them. Therefore, managing Big Data is not something a library plans to do; it could naturally be associated with the increase in data. However, there is no guarantee that this natural match will make it easier to manipulate Big Data in a public library. The challenges of Big Data management in other areas could also exist in librarianship. No appropriate designs for handling vast amounts of data and a lack of proper resources for analysing data are two major challenges (Kaisler *et al.*, 2013), as mentioned in responses to the survey and during the interviews.

Unbalanced resource allocation could also be an issue because small libraries would have more difficulty in analysing and using Big Data, which implies a long process before Big Data is used at a national level. The need to understand society is also required before Big Data can be applied more effectively and specifically. However, that brings challenges in knowing current society. Therefore, the natural match between Big Data and public libraries is naturally accompanied by big challenges.

Service-oriented roles
The analysis regarding library roles in the survey indicates that not all respondents completely agree with the proposed roles, partly owing to their poor understanding of Big Data. However, after considering the interview results, it is concluded that librarians tend to agree with service-oriented roles, which are educator, marketer, advocator, adviser, developer and organisation server. These roles were put forward to mainly help library users understand and utilise Big Data or benefit from Big Data. The realisation of service-oriented roles could help answer one part of the research question: What roles should the public library undertake to bring Big Data to citizens? According to the survey results, librarians slightly agree with having the role of educator, marketer, advocator and adviser, and they see the opportunity for using Big Data for the benefits of users. In their answers to the survey's open question, words such as "guidance" or "help individuals" were mentioned several times, which implies that the responsibility of libraries to help users in the generation of Big Data is shared by librarians as well. Moreover, the opinions put forward during the interviews highlight the inclination to use these four roles. All the library

directors and managers consider these four roles indispensable for Big Data. In addition, they hold strong expectations about developing services from Big Data and communicating with other public sectors and private organisations through Big Data. To better interpret these roles, three perspectives were generated for further discussion.

Educator and marketer: from the perspective of enlarging the audience for Big Data
Public libraries should take on the role of educator – and, thus, help citizens understand what Big Data is – and marketer, to make the concept of Big Data and the use of Big Data familiar to citizens. With the help of these two roles, more citizens would become acquainted with Big Data. As pinpointed by Hoy (2014, p. 324), "librarians will need to help their patrons understand what Big Data can and cannot do, and how it can best be used to achieve their research goals". Even though Hoy's research mainly concentrated on academic universities, the essence of libraries acting as an educator is shared by public libraries. Thus, public libraries ought to work as an educator thinking of Big Data as an emerging science, and teach people how to utilise it for their own good. Public libraries should also act as professional marketers when promoting the opportunities and benefits of Big Data as it has great potential for being developed into new services within the library. Similar ideas are also mentioned in other studies (Gordon-Murnane, 2012; Wittmann and Reinhalter, 2014).

Adviser, advocator and organisation server: from the perspective of the user
For public libraries, users are individuals, communities and other organisations. To smooth the utilisation of Big Data from the user's point of view, three roles are proposed that serve different users. Advocator and adviser are aimed at individuals and communities, and organisation server is intended to serve other organisations. As mentioned by Hoy (2014), libraries have no exemption from being active as a technology adopter. Two ideas are put forward for involving Big Data in the library: guiding users to use potential databases and helping researchers in data management, such as data sharing and archiving. Such guidance and help could also be given to other organisations; thus, these two ideas suitably support libraries acting as an advocator, adviser and organisation server.

Developer: from the perspective of service improvement
Public libraries should assume the role of developers who discover useful information from Big Data and transform it into services. As Xiang *et al.* (2015) demonstrated Big Data techniques provided new insights for hotels wishing to realise the requirements of guests. A public library can also develop some services to further realise a patron's needs. For instance, public libraries can cooperate with health-care centres to promote knowledge concerning disease prevention by using Big Data analytics. In other words, to function as a service developer, public libraries can not only pay attention to Big Data within themselves but also consider resources outside the library because the amount of open data has increased. Methods for creating services in the current data-driven generation could be a major role for public libraries to undertake.

It should be outlined that many resources are required to develop service-oriented roles. A lack of money, professionals or time would be obstacles to functioning in these roles. Therefore, the combined role of a facilitator was put forward to help the library realise these roles in an effective way. Being a facilitator integrates the functions of the four roles of educator, marketer, adviser and advocator, and considers libraries as a platform where external professionals can be introduced to citizens. Compared with these four roles, the idea of a facilitator is more consistent with the current situation in public libraries. As the

available resources are too limited to accomplish all four roles, it would be wise to let other people help libraries accomplish this task.

System-oriented roles
Compared with service-oriented roles, the respondents tended to hold conservative opinions on system-oriented roles, which are data organiser (cleansing and maintaining big data sets) and data container (storing Big Data). They relate to the perspective of sustainable development, which emphasises the task of storing and archiving data for further use. System-oriented roles answer the other part of the research question: What roles should public libraries undertake to manage Big Data? These two roles are accepted by library directors and managers. However, the roles of being data organiser and data container are related to the responsibility and the size of a public library. It is suggested that the National Library of Finland should act as the only place for storing data for Finnish libraries and the main library of each region should organise data for branch libraries. Therefore, the respondents in small libraries tended not to agree on these two roles; however, those from the main library considered them necessary roles. Thus, the attitudes towards those roles are not clear in the survey, suggesting that the scale and responsibility of a public library could decide whether these two roles are undertaken. In addition, no matter which role a public library undertakes, legislation issues, such as copyright and the right to use personal data, should be solved in advance; otherwise, the willingness of libraries to use Big Data in practice will be diminished.

Conclusion
The aim of this study was to explore roles that should be undertaken by public libraries in the context of Big Data. According to the study, there is a natural match between Big Data and public libraries even though there are challenges. The proposed eight roles of educator, marketer, advocator, adviser, developer, organisation server, data container and data organiser are largely accepted by the librarians in the survey. These eight roles can be classified into two groups: service- and system-oriented roles. For service-oriented roles, the library's resources are significant because the more sufficient resources a library has the easier it is to operate in service-oriented roles. If resources are insufficient, public libraries can act as a facilitator and employ external resources to provide services regarding Big Data. As for system-oriented roles, they are not necessary for every library, especially for small libraries or branch libraries, because a library's responsibility and scale make the difference in system-oriented roles.

This study focuses solely on public libraries in Finland. In future studies, public libraries in other countries should be examined. Furthermore, this study does not concentrate on the importance of each role. A study examining role-ranking based on significance could enrich the content of the present study and help attain more systematic findings.

References

Affelt, A. (2015), *The Accidental Data Scientist: Big Data Applications and Opportunities for Librarians and Information Professionals*, Information Today, Medford, NJ.

Chen, C.L.P. and Zhang, C.Y. (2014), "Data-intensive applications, challenges, techniques and technologies: a survey on Big Data", *Information Sciences*, Vol. 275, pp. 314-347.

Chen, M., Mao, S. and Liu, Y. (2014), "Big Data: a survey", *Mobile Networks and Applications*, Vol. 19 No. 2, pp. 171-209.

De Mauro, A., Greco, M. and Grimaldi, M. (2016), "A formal definition of Big Data based on its essential features", *Library Review*, Vol. 65 No. 3, pp. 122-135.

Federer, L. (2016), "Research data management in the age of Big Data: roles and opportunities for librarians", *Information Services and Use*, Vol. 36 Nos 1/2, pp. 35-43.

Fuchs, M., Pken, H.W. and Lexhagen, M. (2014), "Big Data analytics for knowledge generation in tourism destinations: a case from Sweden", *Journal of Destination Marketing and Management*, Vol. 3 No. 4, pp. 198-209.

Gordon-Murnane, L. (2012), "Big Data: a big opportunity for librarians", *Online*, Vol. 36 No. 5, pp. 30-34.

Heidorn, P.B. (2011), "The emerging role of libraries in data curation and e-science", *Journal of Library Administration*, Vol. 51 Nos 7/8, pp. 662-672.

Hey, J. (2004), "The data, information, knowledge, wisdom chain: the metaphorical link", *Intergovernmental Oceanographic Commission*, available at: www.dataschemata.com/uploads/7/4/8/7/7487334/dikwchain.pdf (accessed 4 March 2016).

Hoy, M.B. (2014), "Big Data: an introduction for librarians", *Medical Reference Services Quarterly*, Vol. 33 No. 3, pp. 320-326.

Huang, T., LiangLan, L., Fanga, X., An, P., Min, J. and Wang, F. (2015), "Promises and challenges of Big Data computing in health sciences", *Special Issue on Computation, Business, and Health Science*, Vol. 2 No. 1, pp. 2-11.

Huwe, T.K. (2014), "Big Data and the library: a natural fit", *Computers in Libraries*, Vol. 34 No. 2, pp. 17-18.

Kaisler, S., Armour, F., Espinosa, J.A. and Al, E. (2013), "Big Data: issues and challenges moving forward", in Croll, P.R. (Ed.), *46th Hawaii International Conference on System Science (HICSS)*, *IEEE*, pp. 995-1004.

Noh, Y. (2015), "Imagining library 4.0: creating a model for future libraries", *The Journal of Academic Librarianship*, Vol. 41 No. 6, pp. 786-797.

Shen, Y. and Varvel, V.E. (2013), "Developing data management services at the Johns Hopkins University", *The Journal of Academic Librarianship*, Vol. 39 No. 6, pp. 552-557.

Stejskal, J. and Hajek, P. (2015), "Effectiveness of digital library services as a basis for decision-making in public organizations", *Library and Information Science Research*, Vol. 37 No. 4, pp. 346-352.

Teets, M. and Goldner, M. (2013), "Libraries' role in curating and exposing Big Data", *Future Internet*, Vol. 5 No. 3, pp. 429-438.

Wittmann, R.J. and Reinhalter, L. (2014), "The library: Big Data's boomtown", *The Serials Librarian*, Vol. 67 No. 4, pp. 363-372.

Xiang, Z., Schwartz, Z., Gerdes, J.H., Jr. and Uysal, M. (2015), "What can Big Data and text analytics tell us about hotel guest experience and satisfaction?", *International Journal of Hospitality Management*, Vol. 44, pp. 120-130.

About the authors

Ming Zhan is a Doctoral Candidate of the Department of Information Studies at Åbo Akademi, who is committed to exploring the possibilities of applying Big Data to public libraries. Ming Zhan is the corresponding author and can be contacted at: mzhan@abo.fi

Gunilla Widén is a Professor of Information Studies at the Åbo Akademi University. Her research interests are knowledge management and information behaviour. She has lead several large research projects financed by Academy of Finland, and currently leads a project about the impact of information literacy in the digital workplace (2016-2020). She has published widely in her areas of expertise and been appointed expert in several evaluation committees.