



## Library Hi Tech

Asynchronous social search as a single point of access to information

Marco Buijs, Marco Spruit,

### Article information:

To cite this document:

Marco Buijs, Marco Spruit, (2017) "Asynchronous social search as a single point of access to information", Library Hi Tech, Vol. 35 Issue: 4, pp.656-671, <https://doi.org/10.1108/LHT-01-2017-0007>

Permanent link to this document:

<https://doi.org/10.1108/LHT-01-2017-0007>

Downloaded on: 08 January 2018, At: 05:20 (PT)

References: this document contains references to 40 other documents.

To copy this document: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)

The fulltext of this document has been downloaded 36 times since 2017\*

### Users who downloaded this article also downloaded:

(2017), "Human behavior analysis for library and information science", Library Hi Tech, Vol. 35 Iss 4 pp. 442-444 <<https://doi.org/10.1108/LHT-10-2017-0213>>

(2017), "Library usage mining in the context of alternative costs: The case of the Municipal Library of Prague", Library Hi Tech, Vol. 35 Iss 4 pp. 565-583 <<https://doi.org/10.1108/LHT-01-2017-0019>>

Access to this document was granted through an Emerald subscription provided by emerald-srm:213934 []

### 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](http://www.emeraldinsight.com/authors) for more information.

### About Emerald [www.emeraldinsight.com](http://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.

LHT  
35,4

656

Received 7 January 2017  
Revised 29 June 2017  
Accepted 29 July 2017

# Asynchronous social search as a single point of access to information

Marco Buijs and Marco Spruit

*Department of Information and Computing Sciences, Utrecht University,  
Utrecht, The Netherlands*

## Abstract

**Purpose** – The purpose of this paper is to present asynchronous social search as a novel and intuitive approach to search for information in which people collaborate to find the information they are looking for.

**Design/methodology/approach** – A prototype was built to test the feasibility in a business environment. A case study was performed at an organisation with over 1,000 employees to evaluate the quality of asynchronous social search as a single point of access to information.

**Findings** – Based on the results, the authors conclude that asynchronous social search has great potential as a single point of access to organisational information. Key strengths include that the implementation requires no integration with the existing information technology infrastructure of organisations and participants were very satisfied with the results provided by the prototype.

**Originality/value** – This work demonstrates that asynchronous social search indeed provides a very good starting point for a single point of access to information, as integration with existing software systems is not necessary, and due to the lightweighness of the approach it also performs really well which, in turn, stimulates the technology's acceptance by its end-users.

**Keywords** Information retrieval, Information science, Enterprise search, Asynchronous social search, Bookmarklet prototype, Single point of access, Single case study

**Paper type** Research paper

## 1. Introduction

Sharing information and knowledge in a business environment has become an important topic for organisations. From a business point of view, many of their information resources should only be available internally. Reasons to keep access restricted to people from the organisation are often related to privacy and competitive advantages. By using intranets, the barriers to share knowledge have been lowered. However, direct access to the resources is only available to a small part of the organisation. In many cases, only specific business units or workgroups have access to information that would also be valuable to many others in the organisation. In such an environment information is stored in so called data silos. One major problem that leads to inaccessibility of resources is information overload. The information infrastructure is often complicated to understand and the amount of knowledge available is way too high. The right information at the right time becomes as good as invisible as a needle in a haystack (Offsey, 1997). There is a need for a solution which does not require extensive change to the information systems of the organisation. According to Accenture, one of the first steps that should be taken is to provide the people of the organisation with a single point of access (Nanterme and Daugherty, 2014), for which we have already proposed a working solution (Buijs and Spruit, 2015).

Based on work from Evans and Chi (2008) and Golovchinsky *et al.* (2009), we define asynchronous social search as “Information seeking supported by a network of people where collaboration takes place in a nonconcurrent way”. Important concepts in asynchronous social search are user-generated content and user feedback (e.g. Wijaya *et al.*, 2008). In this paper, we describe a prototype that is based on asynchronous social



---

search and we evaluate its feasibility with a case study. The main research question during this case study was:

*RQ1.* Can asynchronous social search function as a single point of access to information within an organisation?

## 2. Related work

According to literature, there is indeed a need for a single point of access to information within organisations (Nanterme and Daugherty, 2014). Whereas a conventional business intelligence approach can provide robust access to an organisation's historical body of knowledge (e.g. Vleugel *et al.*, 2010), there is also a need in organisations for Enterprise Search software, that is more flexible, easy to use and easy to deploy. In total, 60-80 per cent of queries posed on organisational intranets are navigational queries, in which the user wants to access a known resource of which he does not know how to reach it directly or he thinks the fastest way to access it is via the Enterprise Search engine (Mukherjee and Mao, 2004; Broder, 2002).

The concept of a single point of access also emerged from a governmental perspective, referring to the integration of public services from a citizen's point of view. It suggests that its "customers" may request any public service through a single point of access using the communication channel of their choice (Tambouris and Wimmer, 2008). This perspective has also been referred to as a strategic perspective of electronic democracy (Watson and Mundy, 2001).

A recent systems perspective onto the single point of access approach is realised in a particular data context service called Ground, which is a decoupled architecture of multiple applications and backend services that together serves as a single point of access (Hellerstein *et al.*, 2017). The Dédalo middleware has similar functionalities (Inacio *et al.*, 2016). The strategic objective of a single point of access can furthermore be found in the health policy domain, where it implies a coordination of healthcare services design (Methley, 2016), facilitated by an enterprise data warehouse environment (Smith *et al.*, 2016; Spruit and Sacu, 2015).

A truly personal approach to information retrieval on the WWW has been taken by Delicious. On Delicious people can create an account, add bookmarks to it and retrieve those bookmarks later on based on tags that can be assigned to bookmarks. They can also befriend people and search in the bookmarks of their friends. Several studies were performed on whether such an approach could improve web search (Heymann *et al.*, 2008; Yanbe *et al.*, 2007; Noll and Meinel, 2007). Heymann *et al.* (2008) discovered that a significant part of bookmarks are tagged with terms that also occur in the title, content or metadata of the bookmarked source. This suggests that it is unlikely that using bookmarks in web search would lead to much better results than a full text search. Another problem with bookmarking they discovered is coverage of the web. Only a small portion of the web is bookmarked compared to the portion of the web that is indexed by large search engines. However, they also conclude that pages that are bookmarked are interesting pages and 25 per cent of these pages have not been indexed by search engines. Furthermore, they conclude that bookmarked webpages are disproportionally common in search results compared to their coverage.

Bookmarking on Delicious is a form of collaborative tagging. Golder and Huberman (2006) performed research in this field of study and they define collaborative tagging as "the process by which many users add metadata in the form of keywords to shared content". During their research they observed that people use a great variety of tags, but also consensus is reached in such a way that stable patterns emerge in tag proportions with respect to tagged resources. They also identify the main reason behind tagging, which is personal use. They conclude that

the stable patterns in tagging can be used to organise and describe how web resources relate to each other.

Morris, in combination with several other researchers, performed studies with respect to the social aspect of search (Morris and Horvitz, 2007; Morris, 2008; Hecht *et al.*, 2012; Morris, 2013). Morris and Horvitz created a prototype which supported both synchronous and asynchronous collaborative web search. One feature of their prototype was particularly found useful: insight in the browsing history of their own and their collaborators queries. They conclude that although there is a desire for active, small-group collaborative search, existing technologies do not adequately support this (Morris and Horvitz, 2007). In 2008, Morris wrote a paper in which he explicitly concludes that new interaction techniques and new interfaces are necessary to support collaborative search. In addition, he wrote another paper in 2013 in which he concludes that there is great potential for technological innovation in the field of collaborative search (Morris, 2013). In 2012, Hecht, Teevan, Morris and Liebling published an article on another search engine prototype named SearchBuddies. SearchBuddies is integrated into Facebook which provides a social environment to operate in. This enabled them to make use of alternative rich feedback mechanisms such as likes. They observed that such a socially embedded search engine enables people to get answers to questions that could not be answered by search engines before (Hecht *et al.*, 2012).

Models have been created to understand the process of social search and to get insight in the anatomy of social search engines (Evans and Chi, 2008, 2010; Horowitz and Kamvar, 2010). In this context, social search is defined by Evans and Chi (2008) as “information seeking and sense-making habits that make use of a range of possible social interactions: including searches that utilize social and expertise networks or that may be done in shared social workspaces. This notion certainly encompasses collaborative co-located search, as well as remote and asynchronous collaborative and collective search”.

Golovchinsky *et al.* (2009) identified different sorts of systems for collaboration in online information seeking. They proposed a model which can be used to classify social search support tools. In this model, different kinds of collaboration are distinguished and four dimensions are provided: intent, depth of mediation, concurrency and location. Intent identifies into which extent collaboration is explicit or implicit. Depth of mediation refers to the level on which people collaborate, where level refers to, for example, the user interface or the underlying algorithms. In the concurrency dimension, it is indicated whether a search method supports synchronous or asynchronous collaboration. Location identifies whether collaboration is co-located or distributed, i.e., are people physically in the same room or not. Evans *et al.* (2010) conducted research in which they compared an explicit, user interface level, synchronous, distributed search method to traditional search methods. They observed that collaboration can help people with information seeking tasks although this type of support tool for collaboration did not perform better than traditional search methods. Such collaboration should therefore be used in combination with existing search methods they concluded.

Just like Morris, Bilenko and White found that browsing history from multiple web users can be useful input to search methods. They took things a step further by not only looking at browsing history, but to user activity in general (Bilenko and White, 2008). They state that user activity as input for a search method can lead to improved precision and recall. Matthijs and Radlinski (2011) address the need of personalised search results and use browsing history to create a user profile, which in turn is used to rerank the top 50 search results returned by existing search engines. They conclude that their reranking outperforms the Google's default ranking.

In 2013, Gerlitz and Helmond describe a phenomenon they refer to as the Like economy. They describe what metrics were used to determine the value of webpages through the years. In the early days of the World Wide Web, the number of hits, webpage visits,

was used for this. This is what the authors refer to as the Hit economy. In the late 1990s, Google introduced the hyperlink as a metric, which led to great success. This was a first step in including relational value and social validation in search engine algorithms. Gerlitz and Helmond refer to this as the Link economy. A shift from the informational web to the Social web led to the introduction of more user-focused web metrics to determine the value of webpages (e.g. Pachidi *et al.*, 2014). A key feature in the social web is the use of social buttons. Social buttons allow users to like, recommend, share and bookmark web resources. They enable new ways of information exchange and value determination. This is what Gerlitz and Helmond call the Like economy. In short, the emergence of the Like economy puts the users in charge by enabling them to provide relevance feedback on web resources.

### 3. Research design

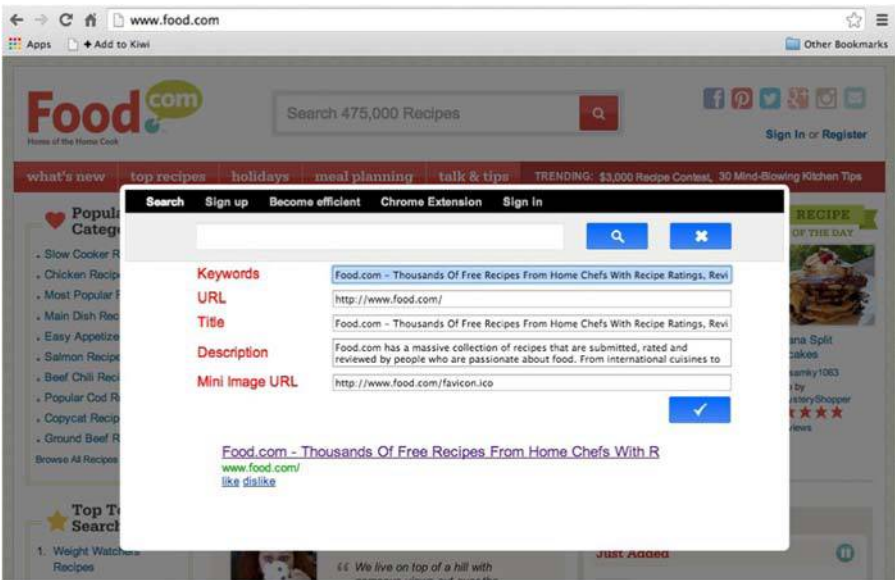
To be able to validate the hypothesis that asynchronous social search provides better quality in search results than traditional search methods a prototype was built. The main goal of the prototype was to function as a single point of access to information within an organisation. To make the prototype easy to access, it was accessible using a web browser. Its main functions were the ability to add links to information resources either manually or automatically and retrieve those resources later on based on queries posed by the users. There are three ways in which results can be added to the prototype. The first is manually, by filling in a URL, title, description and keywords. Figure 1 provides a screenshot of what this way looks like in practice. The second way is by adding a bookmarklet to your favourites in your web browser. When a user has the bookmarklet in his favourites list in his web browser and he visits a website, he can click on the bookmarklet. This results in a popup of the search engine with a form shown to add a result to the search engine. In this form, the URL, title, description and keywords are already filled in based on the page that the user is currently visiting. This second way of adding webpages to the search engine is less time consuming than the first. An example is shown in Figure 2. The third way to add search results to the search engine is by installing an extension for the Chrome web browser. By installing this extension, all the websites that are visited by the user are added to the search engine automatically. When a result is added using the extension, all content of the page that is being added is indexed in case that the page makes use of the HTTP protocol and not of the HTTPS protocol. To guarantee a decent corpus size, the API of bookmarking website Delicious was used to enrich the corpus with resources tagged publicly on Delicious.

We compared our design with the following existing, somewhat similar solutions in a competitor analysis: Google, Bing, Facebook, Knowledge Graph, Wolfram Alpha, Delicious, DuckDuckGo, Sindice, and FAROO. From this qualitative analysis, we conclude that our design has at least four competitive advantages. First, our prototype design has significantly better discovery capabilities based on the combination of discovery methods that are used. Second, we combine a unique set of ranking methods to enable democratic ranking, of results and improved ranking quality in general. Third, we give users the

The screenshot shows a web interface for adding links manually. At the top, there is a navigation bar with links: Search, Sign up, Become efficient, Chrome Extension, and Sign in. Below this is a search bar with a magnifying glass icon and a close icon. The main form has five input fields, each with a red label on the left: 'Keywords' (with a placeholder example: 'e.g., fruit, meat, dinner, best vegetarian recipes, top 10 herbs'), 'URL' (with a placeholder example: 'e.g., food.com'), 'Title', 'Description', and 'Mini image URL'. A blue checkmark button is located at the bottom right of the form.

**Figure 1.**  
Screenshot of how a  
link can be added  
manually

**Figure 2.**  
Example of how a  
link can be added  
using the bookmarklet



explicit choice between personalised and non-personalised search. Fourth, we provide a user interface for content that is presented in unreadable format for humans such as JSON, to make content in the deep web accessible to users.

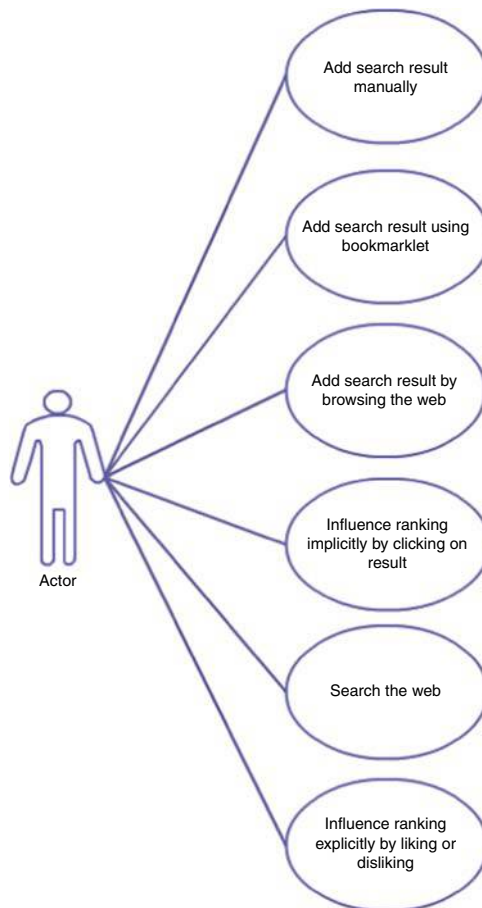
#### 4. Case study design

A ranking of search results given a query  $Q$  is based on credits. Credits are assigned to search results regarding queries in two ways. First, a user can like and dislike a search result given a query. Second, when a user clicks on a search result, credits are assigned to the result with respect to the posed query.

Users can sign up for an account in the prototype. Because of privacy issues, the Chrome extension differentiates between the HTTP and HTTPS protocols. When the HTTP protocol is used when a user accesses a web resource, the resource is added to the search engine publicly, meaning that everyone can retrieve it. When the HTTPS protocol is used, however, the web resource is only stored when the user is signed in on the search engine and will only be accessible to the user himself.

To provide an overview of the system in a systematic and generally accepted format, a use case diagram and a components diagram were created. Use case diagrams are used to model the interaction that people should be able to have with a system, whereas component diagrams are used to show the software architecture and the flow of messages within that architecture, also showing the interfaces that are available in the architecture (D'souza and Wills, 1998). Figure 3 shows a use case diagram stating all the actions that a user should be able to do with respect to the prototype. Figure 4 gives an overview of the software architecture using a components diagram. Both diagrams make use of the Unified Modeling Language notation.

An embedded, closed, single-case holistic study as described by Yin (2009) was performed with the proposed search method. The object of study was the proposed social search method. With the case study we evaluated the capabilities of asynchronous social search methods to function as a single point of access within an organisation.

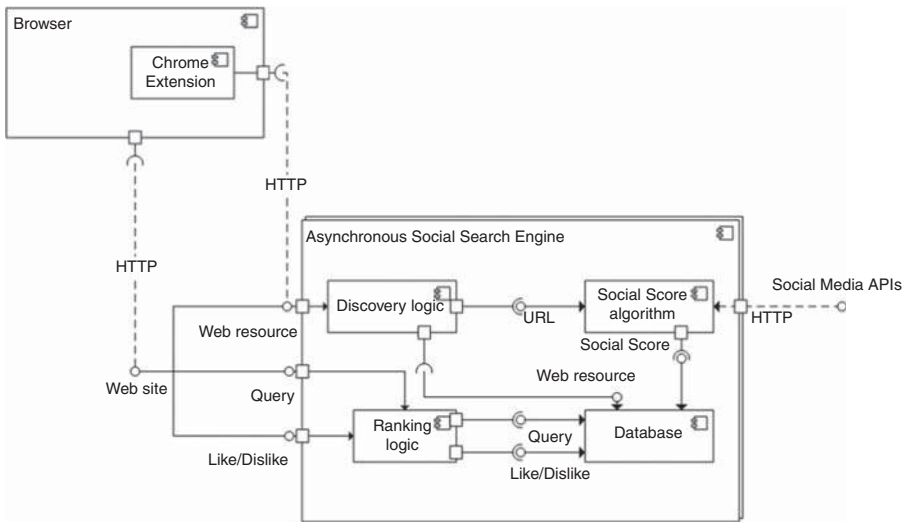


**Figure 3.**  
Use case diagram of  
the asynchronous  
social search engine  
prototype

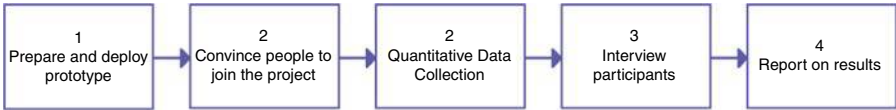
The search engine was hosted internally on the intranet such that knowledge became only available within the organisation. Employees used the Search Engine for five weeks and in that period more and more resources became available through the search engine. At the end of the five weeks semi-structured interviews were conducted to get information about how they experienced using the Search Engine. To be able to answer the main research question in this case study, people from the organisation were interviewed that used the search method. Furthermore, quantitative data were assessed. This process has been illustrated in a workflow diagram in Figure 5.

Another important question during case study design was whether we were measuring what we wanted to measure. We guaranteed construct validity by making use of both qualitative and quantitative data. In the case study, the cause was the introduction of the new search method in the organisation. We measured the effects of this introduction using both qualitative and quantitative data. This way we tried to guarantee internal validity. External validity based on a single case study is not really possible. However, if the results of the case study would indicate that the social search method is seen as added value to the organisation, this would be consistent with literature found that states that there is a need for a single point of access to organisational information within organisations.

**Figure 4.**  
Components diagram  
of the asynchronous  
social search engine  
prototype



**Figure 5.**  
The steps followed in  
the process of  
performing the case  
study illustrated in a  
workflow diagram



Furthermore, there is consistent literature that emphasises the need for data lakes instead of data silos (Nanterme and Daugherty, 2014). Four criteria were used for the selection of an organisation to perform the case study at:

- (1) the employees make use of web resources in their daily work, both internally hosted as externally hosted;
- (2) the organisation should have an information infrastructure with multiple data sources;
- (3) the organisation has more than 100 employees, full filling over 100 full-time equivalents; and
- (4) the organisation must be, at least partially, physically located in the Netherlands since it would be impractical to perform the research further away from the researchers and university given the time and budget limitations of this research.

During this case study, the Technology Acceptance Model 2 (TAM2) was used to determine the usefulness of the prototype and the potential of social search to function as a single point of access within organisations (Venkatesh and Davis, 2000). TAM2 is mainly based on TAM (Davis, 1989). TAM2 was chosen over TAM because it explicitly defines external factors that influence the perceived usefulness. Perceived usefulness was used as a surrogate for usefulness. Another option would have been to use the unified theory of acceptance and use of technology (UTAUT) (Venkatesh *et al.*, 2003). We chose to use TAM2 over UTAUT because UTAUT is more complex with 41 independent variables and 8 dependent variables (Bagozzi, 2007). UTAUT has also been criticised for being less parsimonious than TAM2



(Van Raaij and Schepers, 2008). Therefore, TAM2 is more practical to use for this case study in which only a small number of people participated and no conclusions could be drawn solely based on quantitative data.

Asynchronous  
social search

663

## 5. Case study results

An organisation was selected that is active in the human resource management (e-HRM) and automation of wage and salary administration business with approximately 1,000 employees. The organisation is mainly based in the Netherlands and possesses one of the largest development teams in the Netherlands. The schedule shown in Table I was followed during the case study.

Only a preselected list of domains was indexed by the Chrome extension that was deemed to be relevant for the daily work. This list of domains that were automatically indexed using the Chrome extension was composed based on input from participants that actually installed the Chrome extension. The bookmarklet for adding results and the personal account feature were, unfortunately, not properly introduced to the participants and therefore not all participants were aware of the fact that these features existed. In total, 22 domains were indexed by the Chrome extension, of which some cannot be shown for reasons of confidentiality. A few that are not confidential are linkedin.com, microsoft.com, and stackoverflow.com.

The case study ran for 25 workdays from the 10th of February 2014 onwards. In total, 16 employees participated in the case study. All 16 installed the Chrome Extension and used the search method for a period of four to five weeks. Participation was voluntary, which is a relevant factor according to the TAM2 model. According to TAM2, subjective norm should not have any influence on intention to use if usage of a system is on voluntary basis (Venkatesh and Davis, 2000). In the interviews conducted in the fifth week of the case study, all other factors that can influence the perceived usefulness of a system according to TAM2 were asked about implicitly and sometimes explicitly. From the 16 participants, 15 were interviewed. The last one was on holidays. Table II gives an overview of the number of people per function in the participants group. In total, 22 unique visitors were identified and

Activity	Time period
Prepare and deploy prototype	3-7 February 2014
Convince people to join the project	10-14 February 2014
Quantitative data collection	10 February-14 March 2014
Interviews with participants	3-14 March 2014
Report on results	17-21 March 2014

**Table I.**  
Time schedule for  
the case study

Function	No. of participants
Software engineer	5
Support consultant	2
Development manager	2
User experience designer	2
Product owner	1
Test engineer	1
Application specialist	1
Business intelligence specialist	1
Chief information officer	1

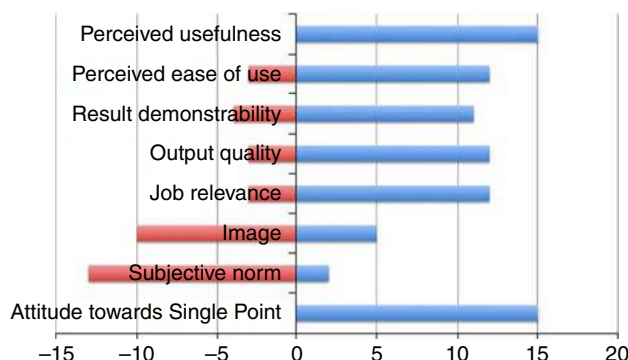
**Table II.**  
Number of people per  
function in the  
participants group of  
the case study

a total of 232 visits to the website were recorded with a total of 641 page views. Figure 6 gives an overview of the direct influencing factors from the TAM2 model on perceived usefulness and the actual influence the factors had on the perceived usefulness based on the interviews held with participants.

5.1 Perceived ease of use

Most people found the interface clear, simplistic, easy to understand and all except for one participant was happy with the performance of the system. For example, the product owner said about the interface: “It is clear, light and simple. It doesn’t take any effort to add results”. People were able to access to prototype rapidly because they had stored the link in their bookmarks, used it as their default search method from the omnibox or set it up as their default page for opening new browser pages and page tabs. There were, however, people that did not use Chrome as their default browser and this group of people had to take more effort to access and use the prototype because it was only properly functioning in the Chrome browser. There were also some people who would have liked some more explaining text in the interface to get them started. The majority, however, had a clear understanding of the way the interface was meant to be used. There were also two persons who argued that it is not user friendly that you first have to visit a page before it is indexed by the search engine. They, and a few others, would have liked to have an additional sort of crawler in place that would index pages connected to visited pages automatically. There were also a few participants who said that adding results manually should have been easier.

Most of the people were able to retrieve the documents they added either manually or implicitly by visiting pages using the Chrome extension without any trouble. People also indicated that it was rather easy to modify the ranking of results. There were a few people who found that it was too easy for people to manipulate the rankings on their own. They did not like the fact that likes could be assigned an indefinite number of times and not only once for one person. People also indicated that there should be a quick way to pose the same query to other search engines such as Google. A button should be provided which would instantly pose the same query to Google. Or, even better, Google results should be integrated into the interface in such a way that they come below the results from the prototype as it works right now. Several also indicated that it could be useful to show how a score for a



**Figure 6.**  
Direct influencing  
factors on perceived  
usefulness in the  
TAM2 model

**Note:** The red horizontal bars indicate the number of people that experienced the factor as having a negative influence on perceived usefulness whereas the blue bars indicate the number of people that experienced the factor as having a positive influence on perceived usefulness

resource is composed. Did it get 65 likes because one person liked it up on its own? Or did 40 different people click on the link and 25 people explicitly assigned a like to the resource? One of the participants that came up with this suggestion also noted that not too many changes should be made to system as is, because things would only become more complex.

### 5.2 Result demonstrability

Most people knew both how results could be added and how they were then ranked during querying. They correctly indicated that ranking is based mainly on likes and results can be added manually and using the Chrome extension. As software engineer 5 phrased it: "What I see is that the number of likes is key in ranking results".

One participant also found out about the bookmarklet, although this was not the intention during the case study. It turned out, however, that this participant was very enthusiastic about the bookmarklet. It was not always clear to people how likes could be assigned to search results. Most participants knew you could add them manually. Some indicated that likes were also generated implicitly based on page visits or clicks. Notice that the number of visits did not actually play a role in ranking although this was presumed by some participants. An interesting note is that only one person mentioned the fact that keywords are also used when ranking results. The others probably took this for granted, because they all knew they could pose queries to the system.

### 5.3 Output quality

Participants were asked whether they were satisfied with the results they got from the system, what percentage of queries lead to success and whether they could find back recourses they added to the system both implicitly and explicitly. On average, approximately one in three queries returned good results. Nearly all people were satisfied with their success ratio on queries. The participating application specialist stated: "One in four, and this ratio increased. I was satisfied with one in four. You know that it is something that is being built up and more information is becoming available all the time".

Some people indicated that there was a big difference in success ratio between intranet and internet search. Many indicated that the ratio differed with what you were looking for. The majority claimed that the tool was much more suitable for intranet search than for internet search. As software engineer 4 indicated: "(The success ratio) depends on what you are looking for. Documentation like sources and intranet sources (high success ratio), but no technical sources like Stackoverflow (low success ratio)".

Quite some people indicated that where search tools of internal systems failed, the prototype was a good alternative. Such internal systems included a Wiki, Sharepoint and Team Foundation Server. Not everyone tried to retrieve pages they added, but most of the people who added pages could find them back easily within the prototype. In general, we can conclude that people were rather satisfied with the output quality of the system.

### 5.4 Job relevance

Virtually all employees indicated that the information available through the prototype was relevant for their daily work. As the test engineer indicated:

You get ideas by using the system about resources you might not have known about before.

Mainly, the internally hosted resources on the intranet were found useful by most participants. Search is a task that all participants had to do regularly and mainly for the intranet they found that the support of the prototype was good. For external sources, such as Stackoverflow, however, the support was often not good because not all resources were indexed by the system. Several people also indicated that they should have asked to index a

few more domains with the Chrome extension automatically that were relevant to their work. People were also asked whether the tool made them do their work more efficient and effective. Most people thought it would, although there were also quite a few who said that some improvements needed to be made before this would actually be the case. Suggestions came in to support search on group level and individual level, taking only a small group of people into account, e.g., your own department. A side note to these suggestions was that tunnel vision should be prevented and you should not lose all access to information outside your own group. That would ruin the inspiring aspect about the prototype. People also asked whether sources hosted on disk could also be indexed and other document types such as PDF format. Another person indicated that it would be useful to work with tags to be able to distinguish information topics. One participant stated that the current prototype could only search in unstructured data and that it would be very useful to him if also structured data sources could be queried in natural language. He compared this feature to Microsoft enquiry.

### 5.5 *Image*

Most participants indicated that they did not know what their colleagues were thinking about the system or even knew that their colleagues did not know about the system. Software engineer 3, for example, said: “Didn’t hear my colleagues about it and my manager doesn’t know about it”. There were some participants that talked about the prototype with each other, all in a positive way. As the CIO stated: “Didn’t have the time to share with my colleagues, one colleague found it a nice idea”.

In general, we conclude that image had a negative influence on perceived usefulness, mainly because not much was known about it in the organisation outside the participants group. Using the system would therefore not enhance one’s status within the organisation.

### 5.6 *Subjective norm*

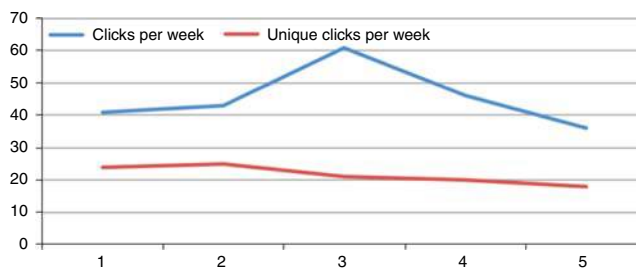
Participants were using the system for at least four weeks, but many of them did not use the system more than a few times a week. In general, we can say that people did not have much experience with the system and therefore, subjective norm does have an influence on perceived usefulness in our case study. The vast majority of participants did not have the feeling that their supervisors and colleagues were expecting them to make use of the system. Therefore, subjective norm had a direct negative influence on perceived usefulness. Furthermore, according to the TAM2 model, subjective norm has a positive correlation with image. Therefore, in our case study, subjective norm has a negative influence on image and thereby also an additional indirect negative influence on perceived usefulness.

### 5.7 *Quantitative data*

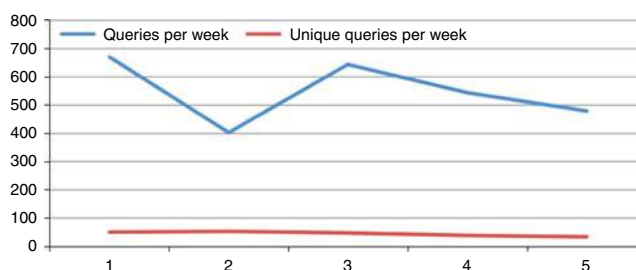
Quantitative data were used to validate the qualitative data. The number of clicks, queries and the size of the index was recorded over time. Figure 7 gives an overview of the number of clicks recorded per week. It also shows the number of unique clicks per week. The number of unique clicks is defined as the number of queries in which one or more clicks were registered during a visit. The total number of clicks shows a bit of a random pattern, although you see a clear decline in the last two weeks. The number of unique clicks shows a more stable pattern, also decreasing slightly over time. Figure 8 gives an overview of the number of recorded queries per week, where each additional keystroke is recorded as a different query because of the instant search feature. Also the number of unique queries per week is shown. Here, the number of unique queries is defined as the number of unique queries posed to the system in one visit. In this case, the number of unique queries also shows a more stable, slightly decreasing pattern. Last, Figure 9 gives an overview of the

total index size at the start of every week. Notice that the index size at the start of the experiment was 40, because there was already one participant who started using the system one day before the start of the measuring period for the case study. During the third week of the experiment there was a decline in index growth. This can be explained by the fact that at the end of the second week, an update of the Chrome Extension was released. This had to do with the fact that additional domains were requested to be indexed. Unexpectedly, this led to Chrome automatically disabling the extension until the user explicitly gave permission that this additional domain could also be indexed. Therefore, many people had the extension disabled during the third week.

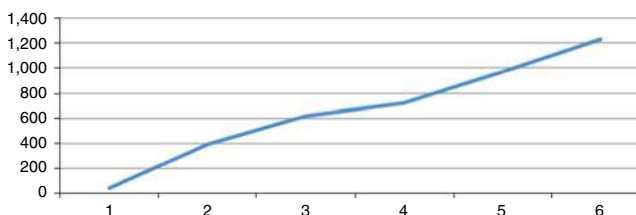
The growth function of the index seems to be rather linear, but there is noticeable decrease in growth already indicating logarithmic growth. This can be explained as follows. The majority of relevant resources have not yet been indexed by the system. Because slightly fewer resources are being indexed towards the end of the case study we can infer that people are also revisiting already indexed resources resulting in a decrease of index growth. This holds only by the assumption that participants did not change their frequency in browse behaviour using Chrome. In total, 3,034 attempts were made to add a resource to the prototype, either using the Chrome extension or manually. Based on the interviews, we assume that over 99 per cent of attempts to add resources were made without any user effort, so using the Chrome extension. This means that approximately 3,000 pages were



**Figure 7.**  
Number of recorded  
clicks per week



**Figure 8.**  
Number of recorded  
queries per week,  
where every  
additional keystroke  
during query  
typing is considered  
to be an additional  
query due to the  
instant search feature



**Figure 9.**  
Size of the index  
measured at the start  
of every week

visited by the 16 participants within the 22 domains that were allowed to be automatically indexed. Because the final index contained 1,238 unique URLs, on average, every page in the index was visited over 2.4 times by the participants. Although we did not measure the number of attempted additions over time, we expect that this factor increased over time since, at the start of the case study, the index was very small and the chance to visit new resources using the Chrome was bigger than near the end of the experiment.

The quantitative data indicate a slight decrease of use of the system over time. That could be explained as that people did not like the system. However, this would be inconsistent with qualitative data acquired via the interviews in which most people stated that the system could be of great help. Another explanation would be that people forgot about the system because they were too busy and the prototype was not an integral part of their daily work. This is also what we found during the interviews with participants.

### 5.8 Discussion

We believe that some essential changes need to be made to the prototype to make it really useful to people in supporting their daily work as a value-added SaaS solution (e.g. Abdat *et al.*, 2011). All participants we interviewed indicated that there was definitely a need for a single point of access through which all information related to the organisation could be accessed. All of them saw the prototype as a very good starting point for such a single point of access. Many of the participants had additional feature requests for and remarks about the system, the most important ones being:

- Browser independency: quite some participants made use of multiple browsers. The system should be made available in all browsers to make it easier to access the system.
- The like system should be improved to prevent malicious use: the ranking system with the likes should be made less prone to malicious use. Now people can assign an indefinite number of likes to a resource for a certain query. There should be a limit in the influence that one person can have on the ranking of a document. It could also be useful for people to see what a score of a page is composed of.
- Trending topics: it would be useful to know what colleagues are looking at a lot, also in the short term. Therefore, a list of trending topics and recently added pages that get many views could be very useful to employees. So next to the ranking based on all likes, there should also be a ranking based on recent likes.
- Manually adding results should be made easier: it takes too much effort to add results manually and the process is not always straightforward. Introducing the bookmarklet would be a great solution to this problem. Looking back at the case study, it was a mistake not to provide the participants with the bookmarklet.
- Dynamically identify relevant domains: the Chrome extension made use of a predefined set of domain names that were indexed such that irrelevant domain with respect to work the organisations were not indexed automatically. This list should be easier to modify or even be changed dynamically based on artificial intelligence. Now people had come to us to ask for additional domains to be added to the list, we had to perform an update of the extension and then people had to accept the new rights required due to the update. In short, an inefficient process.
- Internet vs intranet search: the prototype indexed both internal and external sources. Most participants stated that the prototype worked best for intranet search and worse for internet search. To keep the goal of the system clear, it could be smart to exclude external internet resources. This way, the goal and scope of the system would be clearer, resulting in enterprise search.

- Personalisation and groups: some people described the system as a sort of shared bookmarks. They would like to also have an overview only their own bookmarks. They would also like to have sort of intermediate level between organisation wide ranking and personal ranking. Employees should be able to join groups, like their own department or their own function. Such groups of people should then be able to have their own set of resources. This could also involve tagging of resources with group names.
- Fallback search methods: it can be frustrating to search the prototype, mainly when searching on the web because not many resources are indexed compared to the size of the web. Therefore, there should be a fallback search method or several fallback search methods such that if the number of results provided by the system is limited, the user can, without much effort, pose the query to another search method or this should even be done automatically. This could be achieved with a button for each search method, or more sophisticated by appending for example Google Search results to the list of results. Another option would be to use an automatic redirect of the browser when there are no results found.
- Additional information: an info page explaining how the system works, what the goal of the system is and how it works would be useful according to several participants.

Through this exploratory case study we have been able to test and evaluate in the field our current design for this conceptually new solution. With the collected feedback and feature requests, we now better understand the extent of the problem-solution alignment, generating the input for the next iteration of the prototype in a follow-up case study. Another part of our longer-term vision is the aim to design a meta-algorithmic model as a deterministic, methodological “recipe” for performing effective organisational and internet searches by domain experts such as librarians and information professionals using asynchronous social search, based on the requests and remarks of the case study participants as outlined above (Spruit and Jagesar, 2016). An implementation of such a model could subsequently provide lightweight IT support for operationalise efficient information access through a single point of access.

## 6. Conclusions

Asynchronous social search is a very good starting point for a single point of access to organisational information. That is what can be concluded based on the performed case study. This new way of discovering and indexing webpages has as a key advantage that integration with existing software systems is not necessary. Therefore, the search method can be implemented relatively easily within any organisation, and faster and cheaper than existing Enterprise Search Engines. Consistent with current research, we found that there is indeed a need for a single point of access within organisations. The exploratory case study has provided many relevant insights on how to further improve our approach to realise a collaborative single point of access through a lightweight asynchronous social search bookmarklet.

## References

- Abdat, N., Spruit, M. and Bos, M. (2011), “Software as a service and the pricing strategy for vendors”, in Strader, T. (Ed.), *Digital Product Management, Technology and Practice: Interdisciplinary Perspectives, Advances in E-Business Research (AEBR) Book Series*, IGI Global, pp. 154-192.
- Bagozzi, R.P. (2007), “The legacy of the technology acceptance model and a proposal for a paradigm shift”, *Journal of the Association for Information Systems*, Vol. 8 No. 4, available at: <http://aisel.aisnet.org/jais/vol8/iss4/12/>

- Bilenko, M. and White, R.W. (2008), "Mining the search trails of surfing crowds: identifying relevant websites from user activity", *Proceedings of the 17th International Conference on World Wide Web, ACM*, pp. 51-60.
- Broder, A. (2002), "A taxonomy of web search", *ACM SIGIR forum*, Vol. 36, pp. 3-10.
- Buijs, M. and Spruit, M. (2015), "Determining the Relative Importance of Webpages Based on Social Signals Using the Social Score and the Potential Role of the Social Score in an Asynchronous Social Search Engine", in Fred, A., Dietz, J., Aveiro, D., Liu, K. and Filipe, J. (Eds), *Knowledge Discovery, Knowledge Engineering and Knowledge Management – 6th International Joint Conference, (IC3K 2014)*, ScitePress, Rome, 21-24 October 2014, pp. 118-131.
- Davis, F.D. (1989), "Perceived usefulness, perceived ease of use, and user acceptance of information technology", *MIS Quarterly*, Vol. 13 No. 3, pp. 319-340.
- D'souza, D. and Wills, A.C. (1998), *Objects, Components, and Frameworks with UML: The Catalysis(sm) Approach*, Addison-Wesley Professional Reading.
- Evans, B.M. and Chi, E.H. (2008), "Towards a model of understanding social search", *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work ACM*, pp. 485-494.
- Evans, B.M. and Chi, E.H. (2010), "An elaborated model of social search", *Information Processing and Management*, Vol. 46 No. 6, pp. 656-678.
- Evans, B.M., Kairam, S. and Piroli, P. (2010), "Do your friends make you smarter?: an analysis of social strategies in online information seeking", *Information Processing & Management*, Vol. 46 No. 6, pp. 679-692.
- Gerlitz, C. and Helmond, A. (2013), "The like economy: social buttons and the data-intensive web", *New Media & Society*, Vol. 15 No. 8.
- Golder, S.A. and Huberman, B.A. (2006), "Usage patterns of collaborative tagging systems", *Journal of Information Science*, Vol. 32 No. 2, pp. 198-208.
- Golovchinsky, G., Pickens, J. and Back, M. (2009), "A taxonomy of collaboration in online information seeking", arXiv preprint arXiv: 0908.0704.
- Hecht, B., Teevan, J., Morris, M.R. and Liebling, D.J. (2012), "Searchbuddies: bringing search engines into the conversation", *ICWSM*, Vol. 12, pp. 138-145.
- Hellerstein, J., Sreekanti, V., Gonzalez, J., Dalton, J., Dey, A., Nag, S. and Donsky, M. (2017), "Ground: a data context service", *Proceedings of the Conference on Innovative Data Systems Research (CIDR)*, CIDR 17, Chaminade, CA, 8-11 January.
- Heymann, P., Koutrika, G. and Garcia-Molina, H. (2008), "Can social bookmarking improve web search?", *Proceedings of the International Conference on Web Search and Web Data mining, ACM*, pp. 195-206.
- Horowitz, D. and Kamvar, S.D. (2010), "The anatomy of a large-scale social search engine", *Proceedings of the 19th International Conference on World Wide Web, ACM*, pp. 431-440.
- Inacio, E., Siqueira, F. and Dantas, M. (2016), "Dédalo: a middleware with a single point of access for heterogeneous distributed computing environments", 2016 IEEE Symposium on Computers and Communication (ISCC), IEEE, pp. 964-969.
- Matthijs, N. and Radlinski, F. (2011), "Personalizing web search using long term browsing history", *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, ACM*, pp. 25-34.
- Morris, M.R. (2008), "A survey of collaborative web search practices", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM*, pp. 1657-1660.
- Morris, M.R. (2013), "Collaborative search revisited", *Proceedings of the 2013 Conference on Computer Supported Cooperative Work, ACM*, pp. 1181-1192.
- Morris, M.R. and Horvitz, E. (2007), "SearchTogether: an interface for collaborative web search", *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology, ACM*, pp. 3-12.
- Mukherjee, R. and Mao, J. (2004), "Enterprise search: Tough stuff", *Queue*, Vol. 2 No. 2, pp. 36-46.



- Nanterme, P. and Daugherty, P. (2014), "From digitally disrupted to digital disrupter", technical report, Technology Labs, Accenture.
- Noll, M.G. and Meinel, C. (2007), "Web search personalization via social bookmarking and tagging", *The Semantic Web*, Springer, pp. 367-380.
- Offsey, S. (1997), "Knowledge management: linking people to knowledge for bottom line results", *Journal of Knowledge Management*, Vol. 1 No. 2, pp. 113-122.
- Pachidi, S., Spruit, M. and van der Weerd, I. (2014), "Understanding users' behavior with software operation data mining", *Computers in Human Behavior*, Vol. 30, pp. 583-594.
- Smith, M., Roos, L., Burchill, C., Turner, K., Towns, D., Hong, S., Greg, F., Lisa, L.M., Marni, B., Mahmoud, A., Ruth-Ann, S., Patrick, N.J., Boris, S., Adrian, L., Sarah, G. and Ginter, J. (2016), "Health services data: managing the data warehouse: 25 years of experience at the manitoba centre for health policy", *Data and Measures in Health Services Research*, Springer, Boston, MA, pp. 1-26.
- Spruit, M. and Jagesar, R. (2016), "Power to the People! Meta-algorithmic modelling in applied data science", in Fred, A., Dietz, J., Aveiro, D., Liu, K., Bernardino, J. and Filipe, J. (Eds), *Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, KDIR, ScitePress, Porto, 11-13 November, pp. 400-406.
- Spruit, M. and Sacu, C. (2015), "DWCMM: the data warehouse capability maturity model", *Journal of Universal Computer Science*, Vol. 21 No. 11, pp. 1508-1534.
- Tambouris, E. and Wimmer, M. (2008), "Online one-stop government: a single point of access to public services", in Anttiroiko, A. (Ed.), *Electronic Government: Concepts, Methodologies, Tools, and Applications*, IGI Global, Hershey, PA, pp. 2805-2829.
- Van Raaij, E.M. and Schepers, J.J. (2008), "The acceptance and use of a virtual learning environment in china", *Computers & Education*, Vol. 50 No. 3, pp. 838-852.
- Venkatesh, V. and Davis, F.D. (2000), "A theoretical extension of the technology acceptance model: four longitudinal field studies", *Management Science*, Vol. 46 No. 2, pp. 186-204.
- Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003), "User acceptance of information technology: toward a unified view", *MIS Quarterly*, Vol. 27 No. 3, pp. 425-478.
- Vleugel, A., Spruit, M. and van Daal, A. (2010), "Historical data analysis through data mining from an outsourcing perspective: the three-phases method", *International Journal of Business Intelligence Research*, Vol. 1 No. 3, pp. 42-65.
- Watson, R.T. and Mundy, B. (2001), "A strategic perspective of electronic democracy", *Communications of the ACM*, Vol. 44 No. 1, pp. 27-30.
- Wijaya, S., Spruit, M. and Scheper, W. (2008), "Webstrategy formulation: benefiting from web 2.0 concepts to deliver business values", in Lytras, M., Damiani, E. and Ordóñez de Pablos, P. (Eds), *Web 2.0: The Business Model*, Springer, pp. 103-132.
- Yanbe, Y., Jatowt, A., Nakamura, S. and Tanaka, K. (2007), "Can social bookmarking enhance search in the web?", *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital libraries*, ACM, pp. 107-116.
- Yin, R.K. (2009), *Case Study Research: Design and Methods*, Vol. 5, Sage.

### Further reading

- Methley, A.M., Chew-Graham, C.A., Cheraghi-Sohi, S. and Campbell, S.M. (2017), "A qualitative study of patient and professional perspectives of healthcare services for multiple sclerosis: implications for service development and policy", *Health & Social Care in the Community*, Vol. 25 No. 3, pp. 848-857.

### Corresponding author

Marco Spruit can be contacted at: m.r.spruit@uu.nl

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)

**This article has been cited by:**

1. ChenMu-Yen, Mu-Yen Chen, LughoferEdwin David, Edwin David Lughofer, YenNeil Y., Neil Y. Yen, ChenChia-Chen, Chia-Chen Chen. 2017. Human behavior analysis for library and information science. *Library Hi Tech* 35:4, 442-444. [[Citation](#)] [[Full Text](#)] [[PDF](#)]