

A recommendation-based reading list system prototype for learning and resource management

Journal of Information Science

1–16

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/01655515211006587

journals.sagepub.com/home/jis**Gobinda Chowdhury**

iSchool, University of Strathclyde, UK

Kushwanth Koya 

iSchool, Sheffield Business School, Sheffield Hallam University, UK

Mariam Bugaje

iSchool, University of Strathclyde, UK

Abstract

A reading list is a list of reading items recommended by an academic to assist students' acquisition of knowledge for a specific subject. Subsequently, the libraries of higher education institutions collect and assemble reading lists according to specific courses and offer the students the reading list service. However, the reading list is created based on localised intelligence, restricted to the academic's knowledge of their field, semantics, experience and awareness of developments. This investigation aims to present the views and comments of academics, and library staff, on an envisaged aggregated reading list service, which aggregates recommended reading items from various higher education institutions. This being the aim, we build a prototype, which aggregates reading lists from different universities and showcase it to 19 academics and library staff in various higher education institutions to capture their views, comments and any recommendations. In the process, we also showcase the feasibility of collecting and aggregating reading lists, in addition to understanding the process of reading lists creation at their respective higher education institutions. The prototype successfully showcases the creation of ranked lists of reading items, authors, topics, modules and courses. Academics and library staff indicated that aggregated lists would collectively benefit the academic community. Consequently, recommendations in the form of process implementations and technological applications are made to overcome and successfully implement the proposed aggregated reading list service. This proof-of-concept demonstrates potential benefits for the academic community and identifies further challenges to overcome in order to scale it up to the implementation stage.

Keywords

Core reading; digital information services; digital libraries; reading list; reading list service; recommender system

1. Background and proposed research

1.1. Reading lists

A reading list (RL) is a list of reading items (RIs) recommended by an academic to their class to acquire knowledge in a particular topic (module) under a specific subject area (degree programme). Without RLs, an information seeker is exposed to various factors affecting the information retrieval process, that is, semantics, ad hoc searches, experience of using a search engine and search engine efficiency [1–3]. RLs are also created for purposes of research and various academic activities [4–6]. Purposefully, this research will focus on pedagogical RLs. An RL consists of various types of reading material such as printed books, e-books, journal articles and websites.

Corresponding author:

Kushwanth Koya, iSchool, Sheffield Business School, Sheffield Hallam University, 7216d Stoddart Building, Arundel Gate, Sheffield S1 1WB, UK.

Email: k.koya@shu.ac.uk

The process of creating and sharing the RLs has now become common practice in higher education institutions (HEIs), as they are increasingly recognising the importance of student satisfaction [7]. The lists are made available to students in several ways, not limited to: (1) the academic circulating an email consisting of the list of RIs, (2) the academic posts the RL on the module's e-learning portal and (3) the HEI uses a digital information service (DIS), for example, Talis Aspire to share RLs. RLs as Stokes and Martin [8] comment offer 'a sense of direction' to the students and encourage autonomous learning [8,9]. In addition, the institution's library benefits directly with the creation of RLs as their resource purchasing models' become better informed [10,11]. The criteria for including RIs in an RL are quite varied. The characteristics suggested by various studies of a potential RI are high citation count, student comprehensibility (ease of knowledge acquisition), depth of knowledge, contemporary and availability and so on [4,12,13]. Some studies have also suggested the use of constructive alignment with a specific module's learning outcomes when creating an RL [14,15].

Despite creating RLs for students, academics are still faced with challenges such as student disengagement, incomprehensibility of the RL's relationship with the module/course and unstructured reading [12,14,16–18]. To curtail the challenges, studies suggest solutions such as organising the list based on learning outcomes, inclusion of a variety of RIs (multimedia, journal articles, books, screencasts, etc.), gathering feedback about the RLs, promoting the list during face-to-face interactions, creating annotations within the RIs and relating them to assessments [11,13,18,19].

1.2. RL services

A reading list service (RLS) is a DIS offered by an HEI to students and academics alike, which consist of RLs recommended by academics teaching a particular module within a degree programme. The RLSs make the RIs readily available, bypassing within the information retrieval architecture by placing one-click links to the RIs in the RL. Without RLSs, readers are required to search individual RIs within their library's catalogue or use existing search engines, that is, Google Scholar to retrieve the RIs electronically or physically from the library depending on the item's availability [20,21]. Physical items, that is, books are also shown along with their availability for a loan, followed by an email distributed to the students who are required to complete the module informing the availability of the RL. Most UK HEIs offer RLSs, with Talis Aspire [22] dominating the market with 105 UK HEIs implementing its service. From a research-specific angle, the service Faculty of 1000 (F1000) applies quality filters in the form of post-publication review by nominated experts to select, review and rate articles, especially in natural sciences. Ratings are based on the total sum of points achieved by an article by experts, who have an opportunity to rate an article as – Exceptional (10 points), Must read (8 points) and Recommend (6 points). The articles receive further incremental points of 3, 2 and 1, respectively, for each category [23,24]. F1000 has proved to be a significant contribution to researchers in natural sciences; however, its effectiveness in recommending RIs for students at undergraduate level is unclear, despite the inclusion of the 'tags' feature, which gives the reviewers a chance to tag an article, that is, 'good for teaching' [25].

RLSs offer several benefits such as improving the learning experience of students' by allowing them to access recommended RIs, an integrated framework to the library indicating the availability of an RI, edition checking and cloud-based access. Talis first pioneered the RLs system, that is, Aspire, by linking data in recommended RIs using open Resource Description Framework (RDF) and making the lists openly available [26,27]. Notably, it linked digital resources based on their topics/sub-topics present in them and mainly relied on open-source principles for users to populate the database [26,27]. Hence, based on an RI's ontologies, the system is able to further recommend additional RIs based on hierarchical classification, similarity and user experiences and so on [27]. It is estimated that over 95% of students are reliant on DISs such as RLS offered by their respective institutions' libraries [28,29]. Academics and library staff can appreciate the benefits such as tracking the use of RIs to understand how students use the RIs. As academics are the primary drivers of RLSs, their engagement in creating and updating the lists is critical for the RLSs to function effectively [11,30–32]. Studies have indicated that RLSs inform the libraries of titles to purchase to improve their collections and the system could be interoperated along with stock management systems for seamless library operations [33]. In addition, these services offer a variety of benefits that is automated copyright verification, systems integration with existing systems, flexible access, learning analytics and streamlining of workflows [21,22,34,35].

The RLSs also face several challenges. Standardisations in the process of creation of RLs and staff resistance due to time constraints are identified as primary challenges [31,33]. Other challenges identified are mainly surrounding student and staff engagement issues. RLs catering different types of students, performing internal studies to understand student and staff perspectives of the RLSs are some of solutions suggested by various studies to ensure engagement [19,36]. In addition, implementation of standardisation software, promotional activities and the use of annotated bibliographies to overcome the challenges were also suggested [8,11,37,38]. However, there is a much larger challenge that needs to be addressed, which is localised intelligence. As an RL is created by the recommendation of RIs by an academic belonging

to a specific HEI, the recommendations are restricted to an academic's knowledge of their field, semantics, experience and awareness of developments and so on [4]. Hence, the students may be missing the opportunity to read the best RIs for a particular course. A study further found that 18% of first year students do not use RLSs at all and 67% of students do not use more than four RIs listed in the RLSs [8]. The primary reason why students do not consult RIs was due to lengthy RIs and their contents being outdated [13]. This amplifies the localised intelligence challenge. Not only is a student's reading restricted to localised intelligence, but also they are reading less than the number of RIs recommended by their tutor. One of the solutions for this problem could be to replace the localised intelligence by just a single collective intelligence of a large number of academics in a given discipline of subject. It is obvious that an RI which has been recommended by many academics is deemed to be the best RI for a specific course. Therefore, it is possible to identify the best RIs for multiple courses by building, perhaps a national aggregated reading list (ARL) service which can collect, aggregate and display the most highly recommended RIs using the collective intelligences of various academics and scholars.

1.3. Proposed research and perceived benefits

Currently, there is no linked system that can tell us the most highly recommended RIs in a given subject or topic based on the collective intelligence of academics teaching the concerned subject or topic across all the HEIs in a country. Talis's Aspire system proposed this idea as initially discussed; however, challenges were identified in terms of sustaining it, ownership and licencing, systems compatibility and identifying the source of recommendations [26,27]. Hence, the aim of this research is to investigate academics' and library staffs' perception of the benefits and challenges associated with an aggregated RLS system. In addressing this broad aim, we design a prototype RLS as a proof-of-concept, and conducted a qualitative study with a set of academics and librarians to investigate how they perceive the benefits of such a service, and also to investigate what could be the challenges. The following actions were therefore taken:

1. Mapping of topics and recommended RIs in academic subjects and building a searchable prototype.
2. To further understand the process of creation of RLSs at a few selected universities and staff perspectives on the working of the prototype.
3. Description of a methodological approach for the sourcing of RLSs from multiple HEIs.
4. Identification of data sources for cost data analytics of global RLSs and provide evidence of feasibility of meshing list and cost data across.

In this article, we argue that an aggregated RLS can offer practical benefits to the entire academic community, for example: (1) an academic can view RIs being recommended by academics at other HEIs; (2) students can have access to a ranked list of the best RIs recommended for a course; (3) academic libraries could build a key (must have) collection for every subject taught at their HEI and therefore improve purchasing models for their collection building based on the ranked list of RIs; (4) publishers can gather intelligence on how their published materials are being recommended, and thus make informed business decisions based on demand (the more highly recommended an item, the more demand that item possesses within academia); (5) HEIs can save monetarily by making informed deals with publishers for building a core collection of RIs for every subject and by providing access for students to the core collection and (6) offset the carbon footprint of the publishing and higher education sector [39,40]. Theoretically, the article informs the methodology involved in building an aggregated RLS system, potential challenges and solutions. The proposed research is favourably timed as UK HEIs are increasingly making RLSs compulsory for teaching, to improve student satisfaction. With the introduction of performance based funding streams for teaching and learning, that is, Teaching Excellence Framework (TEF) in the United Kingdom, HEIs are exploring several pathways to improve students' learning experience [41]. Similar benefits can be envisaged in other countries.

2. Methodology

In order to address the key research aims, we first aimed to develop a simple prototype RLS, as a proof-of-concept, and also as a demonstrator system that we could share with the participants in the qualitative part of our research so that the participants could see what an aggregated RLS would look like, and thus could address our aim within the context of an online RLS.

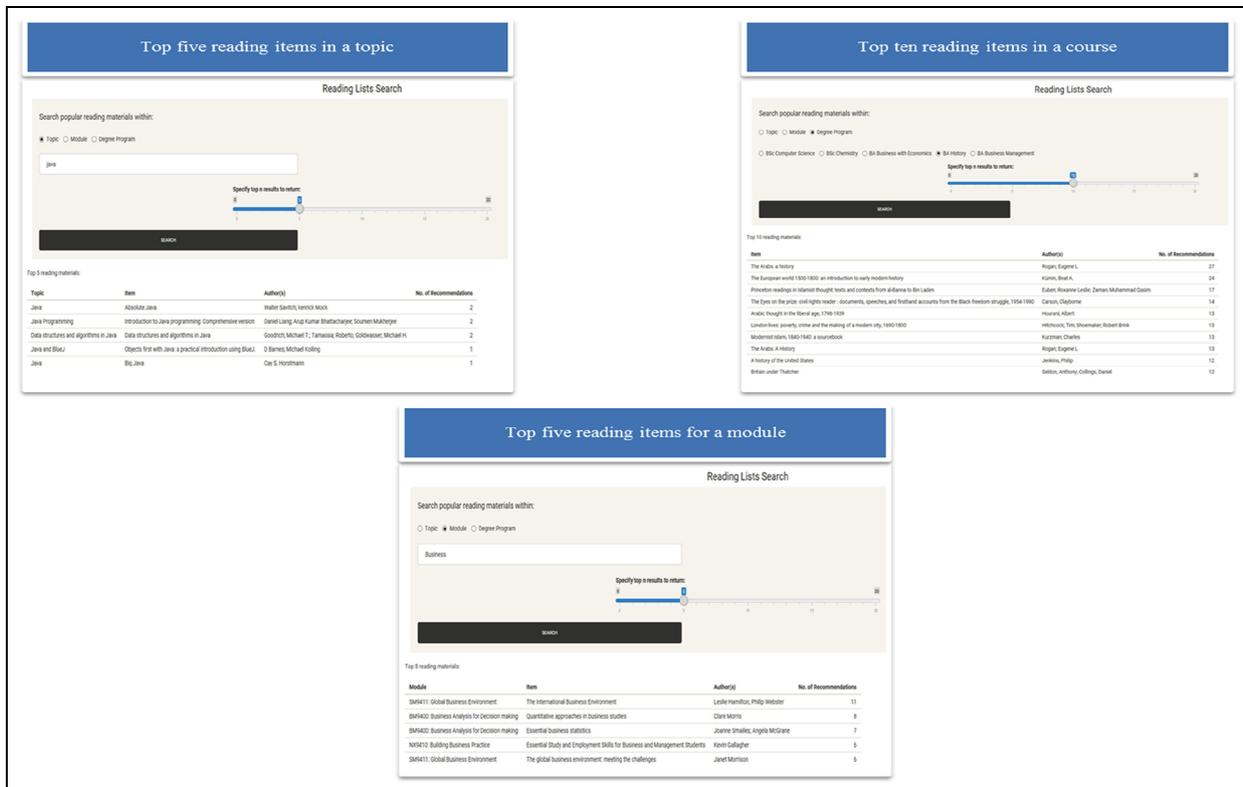


Figure 1. Collage: searching for top five recommended reading items in Java, by degree programme and module.

2.1. Initial prototype

Given that this was not the main focus of our research, the prototype building process did not follow a specific framework for guidance; however, we took inspiration from exploratory testing, based on the context of students retrieving RIs from an RLS [42]. An initial web-based prototype of the proposed system was built based on RIs from four courses being offered at the authors' university. The courses BSc Computer Science, BSc Chemistry, BA Business Management and BA Business with Economics were chosen on the basis of diversity of disciplines offered at HEIs. The RIs of all the subjects in these courses were downloaded from the authors' university's RLS website. Once downloaded, all the RIs were exported into a single MS Excel sheet. The RIs were classified in columns based on course, year, module/subject, topic, title, author, type (book, e-book, journal article, etc.) and published date. The prepared MS Excel sheet was imported to Microsoft Power BI which was used to aggregate and visualise the following questions:

1. Most recommended RIs for a specific topic, subject and course.
2. Most recommended authors for a specific topic, subject and course.

Once it was clear that it was indeed possible to aggregate the RIs, a web-based searchable interface was created using Shinyapps an R studio-based application (<https://readlist.shinyapps.io/readingv2/>). Once the aggregation process was coded, Shinyapps was used to host the service. Being a searchable interface, a webpage was generated, enabling the researchers to send the prototype to interview participants. A user simply needs to search for a specific topic, that is, 'Java' and the interface will retrieve the top RIs for that specific topic based on the number of recommendations it has received, as in Figure 1. In addition, the interface also allows the user to retrieve the most highly recommended RIs by module and degree programmes. The participants in this research were asked to search for topics within the courses they taught and explore features as stated in (1) and (2).

2.2. RL collection and aggregation

A Dropbox[®] link was sent to the library staffs who have agreed to share their RLs for this study. The participants were requested to submit RLs of the courses: BSc Computer Science (UCAS Code: G400), BSc Chemistry (UCAS Code: F100), BA History (UCAS Code: V100) and BA Business Management (UCAS Code: N200). Reminders were sent at two intervals further requesting the participants to load their RLs onto the Dropbox link. A few library staff stated their inability to share RLs due to a possible copyright issue, and their University restrictions. The received RLs were downloaded from Dropbox and were exported onto MS Excel. Some of the RLs were not received in a standard format (received as citations) and hence had to be manually entered into MS Excel. This was the first indication that there is no standardised process of collecting RIs at UK HEIs. The RIs were classified in columns based on University, course, module/subject, topic, title, author, type (book, e-book, journal article, etc.) and publication date. The prepared MS Excel sheet was imported to Microsoft Power BI, where the NetworkMap plugin was used to map (1) topics, subjects and courses to RIs and (2) topics, subjects and courses to authors, along with aggregating the RIs and authors based on the number of recommendations they received from various academics who had created the RLs.

2.3. Aggregation of sample RLs

The RLs received from the libraries were exported into a single MS Excel sheet to simulate pooling of data and the sheet was transferred onto the backend R-based Shinyapps system for the web-based interface. A simple string search algorithm with additional lemmatisation was applied to retrieve the search items. As the standard of creating the RLs varied across the three HEIs, in some cases also within the HEI varied, the aggregation was performed in two methods according to the standards identified. The identified standards are as follows:

1. HEIs created their RL specifically stating which topic, subject, course, author, published date, year and week of study, and type of publication of a particular recommended RI.
2. HEIs created their RL stating subject, course, author, published date and type of publication a particular RI.

In cases where all the classifier data were available, the ARL system was able to rank the RIs based on the number of recommendations it has received for specific topics, subjects and courses. In most cases, the classifier data were not available due to the varying standards in producing the RLs. Hence, the ARL was only able to achieve partially its aggregation, either by ranking the RIs by programme or by subject.

2.4. Interviews

Semi-structured interviews were chosen for this investigation as they offered the flexibility to explore further the RL creation process at a potential participant's institution and their views on the developed prototype system [43]. Purposive sampling method was chosen and invitations were sent to both academic (within the disciplines relevant to the courses chosen to build the prototype) and library staff through professional networks, to participate in a 30-min semi-structured interview regarding the process of RLs creation, maintenance and perspectives on the developed prototype [44]. The purposive sampling criteria considered academic staff with at least a year of teaching experience in their specific discipline and library staff with experience in for digital solutions aimed at students. Overall, 14 academics and 5 senior library staff agreed to participate. The interview participants were sent another email consisting of the confirmation of the date/time of the interview and list of questions which allowed them to familiarise. Two different sets of questionnaires (Table 1) were prepared for academic and library staff so that they can contextualise, and thus provide the appropriate answers related to different aspects of RLS from the perspectives of their activities/tasks and experience.

The participants' names were coded to ensure privacy and verbal acknowledgement and permission for the recordings was sought. The interviews were recorded using a phone-based voice recorder application. The interviews were transcribed manually and content analysis was applied to analyse the data [45]. Specifically, Elo and Kyngas [46] framework for inductive content analysis was applied (see Figure 2) [46,47]. An iterative code checking process was applied at the coding stage to examine any non-emergence of new codes, which helped determine the saturation point [43].

Table 1. Interview questions for academics and library staff.

Academics	Libraries
<p>How do you create reading lists? Do you aggregate reading items based on how important they are? Would you be willing to share the reading items which you had recommended for what you are teaching? How would the proposed service be useful for various academic stakeholders?</p> <p>Would the proposed service offer students a better learning environment? Would the proposed service contribute in building a core collection for the subjects that you teach, and thus increase student satisfaction? What other technical features would you like to be implemented within this service?</p> <p>Who do you think should provide, maintain and sponsor the proposed service? Does your university have collection of reading lists for all the taught modules? Would you be willing to share the reading lists that you have created for your modules?</p> <p>NA</p>	<p>How are reading lists currently created at your University? How do you currently aggregate reading items and how easy or difficult would it be to submit (by uploading to a Dropbox or server, say) reading lists for each programme or degree course? Is it possible to include (1) student numbers and (2) download figures for the recommended reading items for each module/subject? How would the proposed service be useful for various stakeholders (librarians, academics, students and publishers)? How would the proposed service be useful for your library in providing better learning support for students? Would the proposed service contribute in building a core collection for various courses/topics, and thus improve student satisfaction, hence leading to better National Student Survey (NSS) scores? What other features would you like to see in the proposed service, in addition to those that have been suggested like search for most highly recommended authors and items in a given subject, in a specific module or topic? Who do you think should provide, maintain and sponsor the proposed service? Do you have a collection of reading lists for all the modules/topics taught at your University Library? Would you be willing to share your reading lists to contribute in building the conceptualised service? Would you be willing to upload the reading lists for three degree programmes to a Dropbox so that we can build the prototype? Would you be willing to share the costs of each RI appearing in your RIs?</p>

NA: not applicable; RI: reading item.

3. Findings

3.1. Interviews on RIs

3.1.1. Current practices in creating and using RIs in HEIs. Academics teaching the topics/subjects are generally the creators of RIs. The criteria for choosing RIs varied significantly between the academics interviewed. The following were the criteria for the recommendation of RIs:

1. Checking quality by reading inspection copies.
2. Research experience of the author.
3. Authority of the author.
4. Familiarity of authors' works.
5. Student friendly.
6. Currency.

it's a mixture of things (criteria to rank reading items). I probably know the author or familiar with the author indeed their publications and I suppose in some areas currency is very important and in some areas some textbooks from 1960s are in my reading lists as this was seminal work ... a variety to things connected with the authority of the author or publication ... or something that is new or interesting or downright controversial. – A1

Some of the items we chose in lists are not recent ones, but I feel if it is a good source ... also the I recommend items that are relevant to the module and the learning outcomes in the module descriptors ... I don't really think about what level of recommendation. – A9

The creation of RIs after choosing the RIs are found to be accomplished in two ways: either the libraries train the academics in the creation of RIs using DISs, that is, Talis Aspire or the libraries annually prompt and collect recommended

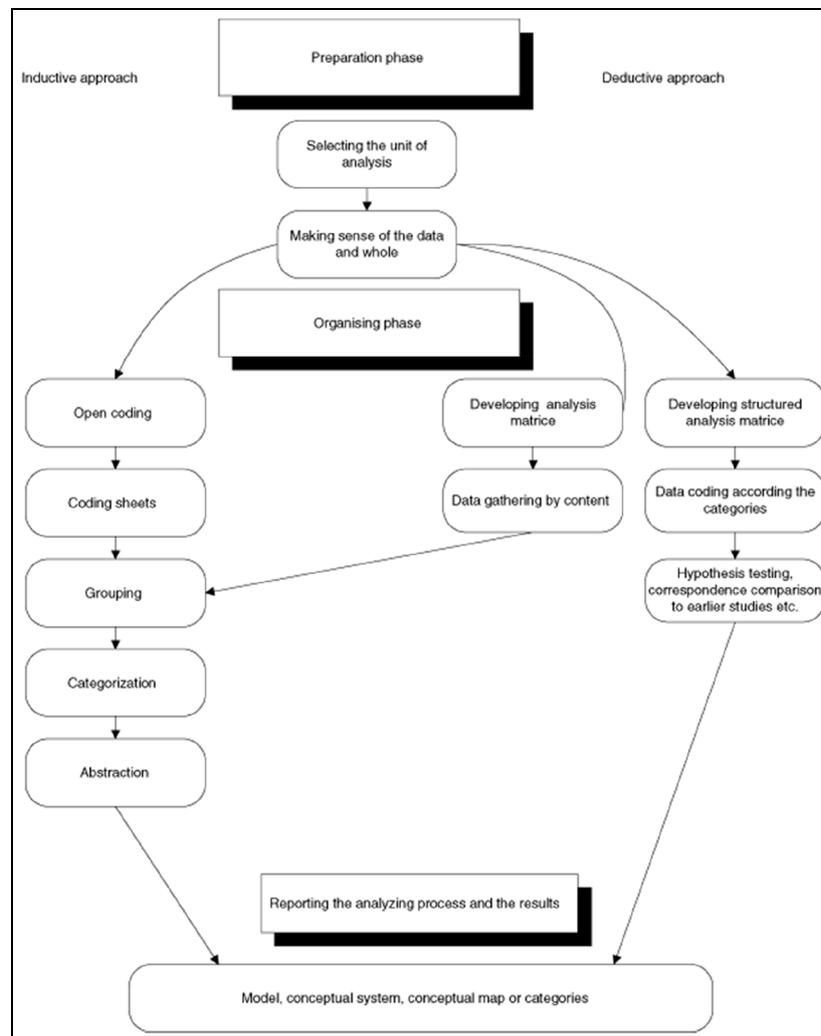


Figure 2. Elo and Kyngas [46] framework for content analysis.

RIs from the academics to create RLs using DISs. Figure 3 indicates the process currently followed at the participants' institutions in the creation of RLs. Both the academics and the library staff agreed that the students found RLSs useful because of its interactive nature, although a list of RIs is always available in their respective VLEs:

Academics produce reading lists as part of their course handbook ... we are using Talis Aspire reading list service ... we have about 1100 lists for this year ... academics send in the documents and we make it available for them. – L1

Every year we receive a message from the library asking for reading lists or updates to existing list ... we then go through the process and the library makes them available ... students find reading lists helpful, although a list of reading material is available in the module descriptors ... reading lists service is interactive. – A1

In some cases where the libraries annually prompt the academics, the libraries usually send an MS Excel template for the academics to complete; however, it was said that only a few academics include all the metadata (topic, week of delivery, type of item, etc.). Although it does not hinder the process of creating RLs, it may cause problems when aggregating data. This is another indication regarding non-standardised procedures in creating RLs and any guidance through University policy could assist both staff and students.

3.1.2. RLs: need for aggregated RLs. Both the groups, academics and library staff responded positively about the idea of ARL as it will be useful for both academic staff and students. However, the library staff were not quite sure about the

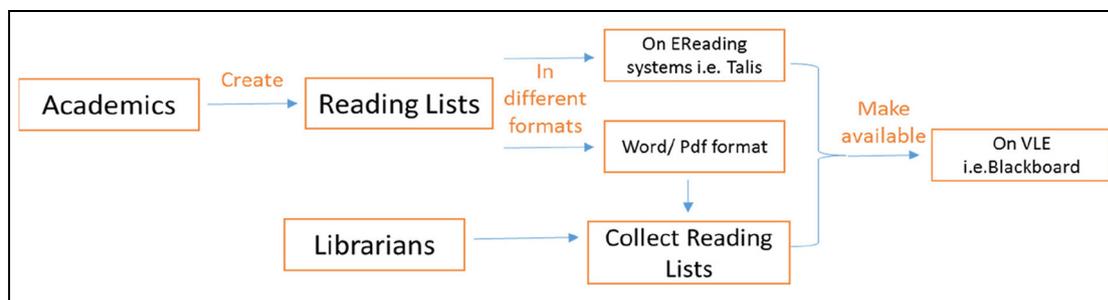


Figure 3. Current process of RLs creation at UK HEIs.

exact process which academics follow to recommend RIs for their courses; they suggested it would be better to talk to the academics, which we did and found that academics do not follow a standard procedure in recommending an RI as mentioned in the previous section and noted the usefulness of the ARL system. While in general both the academics and librarians felt that, an ARL service would benefit students, and suggested that a survey or focus group study with students might be useful to understand potential student benefits. Some library staff and academics pointed out that publishers will benefit the most from such a service, as it offers them intelligence on the performance of their publications, and hence, they should be brought on board for discussions and implementation of such a service. A recent report suggested the implementation of topic-wise RLs, which the proposed ARL does, as students felt overwhelmed by RLs [18]. Hence, the ARL service can have better student acceptance rate:

the whole idea is good ... academics do need help in making decisions and they don't have the time ... so often they decide to do that as a fast activity ... I think this service can help academics make informed decisions. – A7

I can see a number of ways in which this could be helpful ... I don't know how academics go about producing their reading list. I suspect they stick to what they know ... or use for their research ... this service can help academics consult other reading lists or the top reading lists to see the signpost (standard) of what is being recommended ... from the library perspective it can provide value in helping us keep our collections relevant ... reading lists are the primary source of information for us to know what to buy ... this service adds value (to buy highly popular items in good numbers rather than having to buy a bunch of items which are hardly used). – L3

When asked specifically about the potential benefits of ARL and data analytics service for academic libraries, all the respondents commented its potential usefulness in building an informed collection, and subsequently helping to build a core collection of 'must have' RIs for every discipline. In addition, the respondents generally thought the ARL can provide the academics a standard to compare against while recommending RIs for their RLs. Furthermore, the libraries can make better purchasing decisions, eventually leading to the students reading the best material, in addition to making good use of the item. However, there was a concern that building such a 'must have' RL for arts and humanities subjects may be difficult because usually the RLs for these subjects are often long and disparate:

the whole idea is good ... academics do need help in making decisions and they don't have the time ... so often they decide to do that as a fast activity ... I think this service can help academics make informed decisions. – A7

this kind of information is very useful to have (in purchasing recommended reading items considering the number of students attached to it; better collection purchase model for the library). – L4

I can see a number of ways in which this could be helpful ... I don't know how academics go about producing their reading list. I suspect they stick to what they know ... or use for their research ... this service can help academics consult other reading lists or the top reading lists to see the signpost (standard) of what is being recommended ... from the library perspective it can provide value in helping us keep our collections relevant ... reading lists are the primary source of information for us to know what to buy ... this service adds value (to buy highly popular items in good numbers rather than having to buy a bunch of items which are hardly used). – L3

The participants were further requested to use the initial prototype and suggest additional features that they would like to see implemented. The participants suggested that personalisation of the prototype can benefit different stakeholders in ways which met their objectives: for example, for academics and students to identify the most highly recommended authors and RIs on a given course, module of topic; for libraries to identify the core (or must have collection) as well as optional list for collection development/management and for publishers to show the performance of a specific title in their portfolio, their market competitors and so on. Some also agreed that download data, if available, might be useful.

The academics in particular wanted a feature that can bring forward the recent up-to-date RIs for particular topics, in addition to having direct access to the digitised RI:

The service will be highly useful for academics ... but once the system is established it creates a bubble for everyone ... what I mean is we will tend to recommend the same items repeatedly restricting access to latest items ... – A7

if you had links to the actual books as well ... that would help ... because then you actually want to see the description of the book and then make the recommendation ... and probably link to social media (to share the usefulness of a reading item). – A6

When asked about who should maintain such a service, there was a general agreement that an ARL and data analytics service should be managed and provided by a central agency. Some suggested that a business model has to be developed for this, but all agreed that given the clear benefits to all the stakeholders as mentioned earlier and that this will not require a huge amount of work once the system and the workflow was designed, implementation and delivery of such a service should not be difficult. With regard to openly sharing their RLs, academics were willing to share their lists; however, library staff expressed concerns about copyright issues:

I would imagine it to be fairly centralised ... it's the kind of thing that somebody like Jisc would be able to do themselves. – L3

I think this needs to be, obviously centrally provided ... in some way by an organisation that has the kind of profile and visibility to get a buy-in from the board (users) ... Jisc seems to be an obvious choice in that space. – L5

I don't mind sharing the lists which I recommend ... many academics share theirs through personal webpages ... which are again on their faculty profiles (website) ... – A2

The lists are perhaps governed by the university's copyrights ... might not be possible to share them openly ... it's a grey area ... – L3

4. Discussion and recommendations

Academics and library staff believe that an aggregated and ranked RL for every topic and every subject would benefit everyone. They also believe that introduction of such a service will be useful and popular among academics, librarians, students and publishers. Although students and publishers would benefit from the ARL service, their views were not recorded in this feasibility study, and therefore, future studies must gather relevant information about their needs and potential benefits.

As indicated by Jones [31] and Cross [33], standardisation in the creation of RLs is a major issue. This research identifies two types of standardisation issues:

1. The process of recommendation of an RI.
2. The process of attaching all the necessary metadata of an RI during recommendation.

The process of recommending an RI should be a liberty that academics must hold as they know which material will assist their students. However, for the ARL to work, it is necessary to maintain a specific standard right from the collection of RLs to developing a model for core collection. The findings indicated that it was indeed possible to aggregate large sets of RLs resulting in ranking of RIs according to course, subject and topic. However, standardisation held the key in achieving the desired result. Extrapolating from the findings, there are five key challenges to overcome in order to achieve a fully fledged ARL: (1) collection of data; (2) preparation or standardisation of data in scenarios where collecting standardised data is not possible; (3) coding and aggregation of RIs belonging to various courses; (4) presentation of the aggregated data sets to the appropriate communities within academia and (5) developing a model for core collections.

4.1. Collection of data

One of the key challenges of the aggregation process is the collection/harvesting of RLs from multiple UK HEIs. The interviews revealed willingness of library staff and academic staff to share/upload their RLs into a central system or server. As updating RLs is an annual process, both academics and library staff feel that it is not a demanding job, given the variety of benefits the ARL service can offer. However, wider discussions with the HEI library community and some advocacy may be needed in order to ensure that every library submits/uploads the RLs. It may be noted that 100% compliance in terms of uploading RLs may not be achieved in the first instance, but the submission rate will grow as the community sees the benefits of the new service. Also, the ARL system can be automated for regular updates in the RIs, which is especially useful for disciplines where knowledge evolves at a faster rate [48].

4.2. Preparation or standardisation of data

In scenarios where collecting standardised data is not possible, that is, either following a certain citation style or filling in the data into a certain template, some form of pre-processing/cleaning may be required prior to the data being coded and aggregated. A possible solution may be to advise academics to use a prescribed format or template. In the sample data set gathered from universities, the template or the citation style was not uniform. Hence, we had to convert the entire data set into a standard template. However, it was also observed that the academics did not fill in all the metadata required to describe the data sets. Therefore, it is equally important to add all the metadata [49]. A possible template if the RLS are prepared in MS Excel will be Table 2, which stores all the metadata of a specific RI, that is, course, year of study, module/subject within the course, topic with the subject, author, type of item and published year.

In scenarios where filling data into a template is not possible, academics should recommend the RI in a standard citation style, that is, Harvard, APA, Chicago and so on. As academics are used to following standardised procedures, that is, submitting journal articles with a required style of referencing, it is essential to convince academics to follow a particular standard, given the benefits the ARL service can offer. In addition, applications have been developed within the web semantics, text mining and natural language fields to alleviate this issue and create standardised data and necessary taxonomies [50–52]. This study has demonstrated a set of potential benefits of an aggregated RL system for libraries, academics, students and also publishers by identifying recommended RIs in specific disciplines, courses, modules and sub-topics. This is based on converting localised intelligence into collective intelligence of academics. Further research is required to address some of challenges associated with standardisation, syntax, semantics and further coding of RLS.

4.3. Data coding and aggregation

Coding and aggregation of RLS from subjects can be standardised using any nationalised discipline or course coding system. For example, in the United Kingdom, the University & College Admissions Service (UCAS) codes can be used. The libraries submitting the RLS may be advised to use the standard UCAS code for each course rather than use the course title or description. However, it may be difficult to standardise specific modules/subjects since they are localised. Hence, for module/subject level aggregation of RLS, a manual or semantic verification coding may be required. Alternatively, module/subject titles may be automatically parsed to create keywords and the recommended RIs may be linked to the keywords, and subsequently aggregated under the module/subject code. This will assist in producing two sets of ARLs: (1) one for each subject using UCAS subject code and (2) another for keyword level aggregation showing recommended RLS for each keyword/topics under a module/subject. However, in order to produce best results, the keyword lists need to be controlled using a subject dictionary or thesaurus, and this may need some research and development activities or web semantics, text mining and natural language applications can be implemented [50–52]. Several technical architectures have been developed recently in terms of offering aggregation services and the service needs to further articulate a suitable structure for the ARL offering [53,54].

4.4. Data presentation

Different sections within the academic community may use the ARLs for different purposes. For example:

1. An academic or a student may like to find out the most highly recommended RIs for a subject or a topic.
2. A librarian may like to generate a list of the most highly recommended RIs in a given topic, subject and course to build a core (must have) collection.
3. A publisher may like to find out how a product published by them has been recommended – for which subject, how many times, what are the competing titles and so on.

The search and display interface should cater for the specific needs of each of these categories of users. Depending on the user, different types of output may be generated ranging from a ranked list of RIs along with number of recommendations, to a visual interface showing RIs for a given subject or topic along with the number of recommendations and so on.

4.5. Developing a model for core (must have) collections

The outcome of this study is to demonstrate the possibility for the HEIs (and FEs) to identify a list of RIs that may form the core collection for every subject. Once such a list is generated (and this list may vary every year depending on the changes in the recommendation patterns of academics for their subjects each year), it may be used to build a common collection of for every subject nationally. This core collection for every subject will not only benefit students and academics by providing 24 × 7 access to key RIs in a given subject, it will also facilitate generation of a variety of data and

Table 2. A sample MS Excel template for the preparation of reading lists.

Programme	Year	Module	Week	Topic	Item	Author	Type	Year
BSc Computer Science	I	KF 4011 Systems Analysis	I	UML	Object-oriented Systems Analysis and Design using UML	Simon Bennett, Steve McRobb, Ray Farmer	Book	2010
BSc Computer Science	I	KF 4011 Systems Analysis	I	UML	Systems Analysis and Design: An Object-oriented Approach with UML	Alan Dennis, Barbara Wixom, David Tegarden	Book	2015
BSc Computer Science	I	KF 4011 Systems Analysis	I	UML	Systems Analysis and Design: An Object-oriented Approach with UML	Alan Dennis, Barbara Wixom, David Tegarden	Article	2016
BA Business with Economics	I	NX9410 Building Business Practice	I	Study skills	The Study Skills Handbook	Stella Cottrell	Book	2013
BA Business with Economics	I	NX9410 Building Business Practice	I	Study skills	Essential Study and Employability Skills for Business and Management Students	Kevin Gallagher	Book	2016
BA Business with Economics	I	SM9411 Global Business Environment	I	Global business environment	The International Business Environment	Leslie Hamilton, Philip Webster	Book	2015
BA Business with Economics	I	SM9411 Global Business Environment	I	Global business environment	Business Environment in a Global Context	Andrew Harrison	Book	2014
BA Business with Economics	I	SM9411 Global Business Environment	I	Global business environment	The Global Business Environment: Meeting the Challenges	Janet Morrison	Book	2011

UML: unified modelling language.

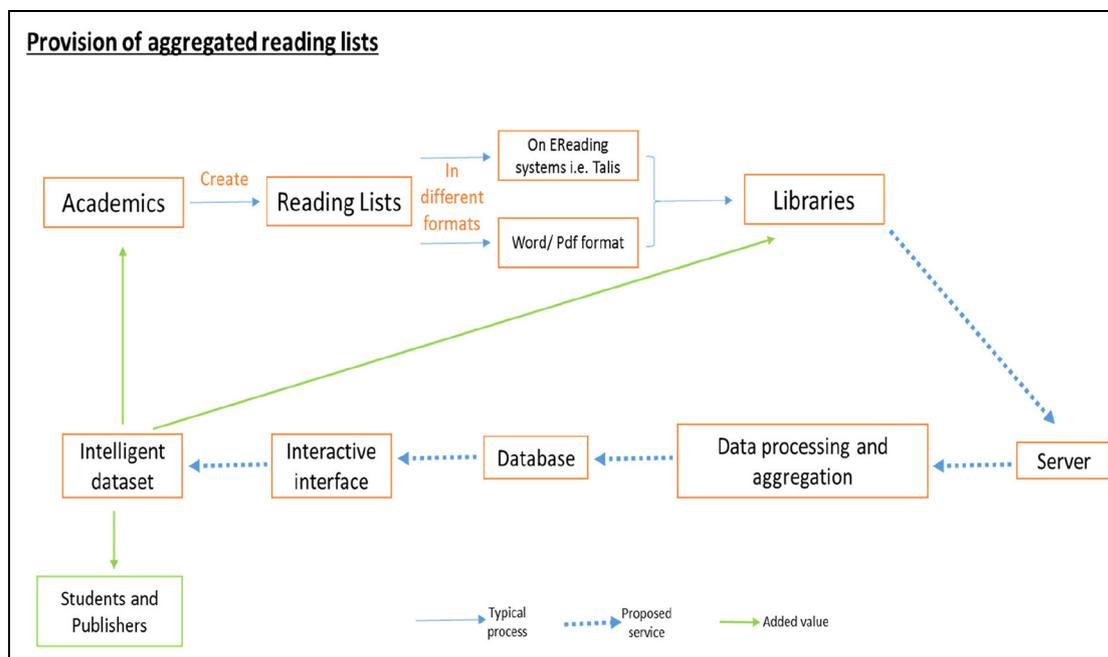


Figure 4. Process involved in the creation of an aggregated reading list service.

intelligence based on access and use of these RLs. This intelligence will be useful for the HEIs and academics in making useful decisions for improvements/intervention in students' learning, and publishers to understand the reading patterns of students that may help improve their business processes. In addition, the envisaged collection will have the usual economic and environmental benefits that are achieved by replacing printed with digital content [39,40]. In addition, building a core collection promotes better learning options being offered to students, hence can possibly lead to better student satisfaction scores. With the introduction of the performance-based funding aspects in teaching, that is, the TEF, UK HEIs would want to offer better learning facilities to obtain higher funding [41]. Once the libraries identify the core collection, they could design better purchasing models for collection building, leading to better financial management.

4.6. Proposed process to achieve the aggregated RL model

This study has shown that it is indeed possible to build an ARL service where RLs are created in templates consisting of complete metadata or in cases where a uniform citation style was used. Furthermore, a standardisation process can be implemented to convert differently created RLs into a single form.

Considering the challenges identified and envisaging solutions to overcome the challenges, Figure 4 indicates recommended processes which can achieve the ARL service. Currently, RLs are created by academics teaching a particular module/subject once a year when prompted by their libraries. The lists are either made available in their respective VLEs by the library staff or uploaded to their RLS, that is, Talis Aspire by the academic if they had received training. In scenarios where the academics did not receive training or are not responsible for their upload, the library staffs' ensure their upload. The formats of the RLs currently vary significantly across the HEIs. Aggregation of the RLs is a straightforward process; however, without a specific format, it is very challenging to aggregate all the RLs. Hence, the academics or the libraries need to upload/submit RLs in a standard format. Once the RLs are uploaded into the server, depending on the previous step, the RLs can either be curated or sent directly into the database where the aggregation process takes place. Through a simple user interface with built-in personalisation facilities, different users can retrieve a ranked list of items for every course, subject/module and topic. In addition, as the ARL only collects recommended RLs, but not the names of HEIs or individual academics, there is no risk of violating intellectual property issues, thus ensuring reassurance to both the HEIs and the academics.

4.7. Further road map

The current prototype was able to highlight the possibility of aggregation in four courses belonging to different academic disciplines. The next step would be to expand the application of the prototype in other disciplines to identify the various

Search | Browse

Reading List Service

java

Type	Title	Author	Recommendations	Year Published	Publisher
Book	Data Structures and Algorithm Analysis in Java, Thi	Clifford A. Shaffer	1 time(s)	2011	Dover Publications Inc.
Book	Practical Database Programming with Java	Ying Bai	1 time(s)	2012	John Wiley & Sons Inc
Book	Iron-Clad Java	Jim Manico, August Detlefsen	1 time(s)	2014	McGraw-Hill Education - Europe
Book	Concurrent and real-time programming in java: threads, RTSJ and RMI	Badr Benmammar	1 time(s)	2018	Wiley-ISTE
Book	Java: How to Program (early objects)	Paul J. Deitel, Sherif G. Aly, Saleh Al-Hazbi, Harvey M. Deitel	1 time(s)	2015	Prentice Hall
Book	Java Web Services: Up and Running: A Quick, Practical, and Thorough Introduction	Martin Kalin	1 time(s)		
Book	JAVA Gently : programming principles explained	Judy Bishop	1 time(s)	1998	Addison-Wesley

Figure 5. Second version of the prototype.

challenges that RLs in other disciplines may contain. In addition, engaging HEI libraries to obtain their advocacy should be the next target for this project. Not only will the libraries be the primary users of the proposed ARL, but also it streamlines the data collection process. The HEI libraries will play a key role in obtaining RIs from respective academics in prescribed formats or templates. In parallel, running roadshows/workshops to create awareness of the benefits of the ARL service and solicit HEI libraries to participate in the uploading RLs and encourage academics to follow a prescribed format in creating RLs will assist in the seamless aggregation of RLs.

Finally, building a self-sustainable business model to run the ARL service will be the final challenge to overcome. However, once the workflow process – creation of RLs using a standard citation style or format, uploading of data to a server and the process of coding data from the uploaded RLs – is standardised, the implementation and overall management of the ARL service will be less resource intensive. Moreover, once fully implemented, the benefits will easily surpass the costs of implementation and management of the service, with the creation of a better learning environment for every student in every discipline at every university in the country.

The key benefits of the ARL are aimed at core collection building and perhaps better financial savings for the HEIs. The HEIs can benefit from creating better learning environments for students, generate collective intelligence from the UK academic community which can also be used by the publishers and authors. Students can benefit immensely from this collective intelligence. HEIs can benefit financially and by creating better student experience, leading to better TEF scores. Academics can also learn from the collective intelligence available within their community and adjust their recommendations as needed.

4.7.1. Second version of the prototype. Based on the feedback from academics and library staff, the authors have built a second version of the prototype (<http://readlist.eu.meteorapp.com>). This version consists of a searchable interface to retrieve the ARLs of 19 disciplines and further sub-disciplines. In addition, the user can browse through the ARLs in specific disciplines (Figure 5).

4.7.2. Limitations and future work. The sample in this study does not consist of students and publishers, who are essential stakeholders of the ARL proposed system, whose views can inform the further development. Future studies need to sample academic staff, library staff, students and publishers to acquire comprehensive views on the system. In addition, system developers need to be consulted to study the feasibility of the ARL for large-scale operations. Future research can also investigate the HEIs' openness in sharing their RLs, another factor which can further enrich the system and at a localised level build bridges between disciplines [55]. Taking forward, it is also important to build advocacy within academia as such a service does not exist, hence showcasing the service and scaling up offers a good chance of it being implemented as a standard service offered to UK Universities. Also, extensive research is needed for standardising and formatting RLs so that the required bibliographic and semantic data can be automatically extracted, coded and analysed in order to build various data analytics models.

5. Conclusion

This study indicates a feasibility for building an ARL service, which can be centrally hosted to provide UK HEIs a core list of RLs for any given topic, subject and course, by showcasing:

1. The possibility to map topics and recommended RLs in various academic subjects and build a searchable prototype.
2. A need for such a service in academia.
3. A methodological approach to build the service taking into account the current processes in creating RLs, academic's and library staffs views on the proposed service.
4. Envisaged benefits from the proposed service.

The study demonstrates that RLs can indeed be mapped and also ranked based on recommendations to their respective topics, subjects and courses at UK HEIs. Interviews with academics and library staff revealed that there is indeed a significant amount of interest in such a service as academics would want to learn from what their counterparts in other HEIs are recommending and the libraries' best interests lay in being able to create core-RLs based on collective intelligence and hence build better purchasing models. This in turn helps the HEIs offer a better learning environment for students where every student will have access to all the core RLs in their respective disciplines and subject, possibly leading to better TEF performance. The study is able to showcase the various challenges which need to be overcome to develop the proposed service and offers a road map on interventions which can standardise multiple RLs from UK HEIs, and thus help the service provider aggregate them to identify core RLs. To summarise, the study demonstrates the key benefits of an RLS from the perspectives of academic and library staff, and overall, there has been a positive response to an aggregated RLS. In addition, the study demonstrates several benefits for an aggregated RLS for the students, for academics, and also for library services for collection development and services for students. However, the study also identifies some key challenges that need to be addressed and resolved in order to build and implement an aggregated RLS on a national scale.

Acknowledgements

The authors acknowledge the time and effort offered by the academics and senior library staff who had agreed to participate in this study.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by the JISC Innovation Grant, awarded to the authors in 2016.

ORCID iD

Kushwanth Koya  <https://orcid.org/0000-0002-7718-1116>

References

- [1] Malizia A, Olsen KA, Turchi T et al. An ant-colony based approach for real-time implicit collaborative information seeking. *Inform Process Manag* 2017; 53(3): 608–623.
- [2] Mukherjee P and Jansen BJ. Conversing and searching: the causal relationship between social media and web search. *Internet Res* 2017; 27(5): 1209–1226.
- [3] Zhai C and Lafferty J. A study of smoothing methods for language models applied to information retrieval. *ACM T Inform Syst* 2004; 22(2): 179–214.
- [4] Raamkumar AS, Foo S and Pang N. Using author-specified keywords in building an initial reading list of research papers in scientific paper retrieval and recommender systems. *Inform Process Manag* 2017; 53(3): 577–594.
- [5] Jardine JG. *Automatically generating reading lists*. Computer Laboratory, University of Cambridge, 2014, <https://tinyurl.com/yy4b62hs>
- [6] Qayyum MA. Capturing the online academic reading process. *Inform Process Manag* 2008; 44(2): 581–595.
- [7] Boyle L and Mitchell D. Dusting off the reading list: is it fit for purpose? *J Perspect Appl Acad Pract* 2016; 4(1): 82–86.
- [8] Stokes P and Martin L. Reading lists: a study of tutor and student perceptions, expectations and realities. *Stud High Educ* 2008; 33(2): 113–125.
- [9] Rowley M, Hartley J and Larkin D. Learning from experience: the expectations and experiences of first-year undergraduate psychology students. *J Furth High Educ* 2008; 32(4): 399–413.
- [10] Carter S and Ostendorf D. Processes and strategies for collaboratively purchasing electronic resources. *Collab Librarian* 2017; 9(1): 10.
- [11] Atkinson S, Conway N, Taylorson B et al. Reading lists – pitching it right. *SCONUL Focus* 2010; 50: 75–79.
- [12] Croft D. Embedding constructive alignment of reading lists in course design. *J Libr Inf Sci* 2020; 52(1): 67–74.
- [13] Brewerton G. Implications of student and lecturer qualitative views on reading lists: a case study at Loughborough University, UK. *New Rev Acad Librarian* 2014; 20(1): 78–90.
- [14] Croft D. *What is a reading list for? A guide for module leaders on aligning reading with learning outcomes*. Oxford Brookes University, 2018, <https://tinyurl.com/y25reql6>
- [15] Biggs JB. *Teaching for quality learning at university: what the student does*. London: McGraw Hill Education, 2011.
- [16] Römhild J. ‘How does this book relate to me?’ Personal engagement and critical enquiry in the literary studies classroom. *High Educ Res Dev* 2019; 38(1): 51–62.
- [17] Douglas K, Barnett T, Poletti A et al. Building reading resilience: re-thinking reading for the literary studies classroom. *High Educ Res Dev* 2016; 35(2): 254–266.
- [18] Siddall G and Rose H. Reading lists – time for a reality check? An investigation into the use of reading lists as a pedagogical tool to support the development of information skills amongst Foundation Degree students. *Libr Inf Res* 2014; 38(118): 52–73.
- [19] McGuinn K, Stone G, Sharman A et al. Student reading lists: evaluating the student experience at the University of Huddersfield. *Electron Libr* 2017; 35: 322–332.
- [20] Walters WH. Google Scholar coverage of a multidisciplinary field. *Inform Process Manag* 2007; 43(4): 1121–1132.
- [21] Zhang Y. Developing a holistic model for digital library evaluation. *J Am Soc Inf Sci Tec* 2010; 61(1): 88–110.
- [22] Talis Aspire. Why do universities choose Talis Aspire? 2019, <https://talis.com/talis-aspire/>
- [23] Li X and Thelwall M. F1000, Mendeley and traditional bibliometric indicators. In: *Proceedings of the 17th international conference on science and technology indicators*, Montreal, Canada, 5 September 2012, vol. 2, pp. 451–551. Montréal, QC, Canada: Science-Metrix and OST.
- [24] Mohammadi E and Thelwall M. Assessing non-standard article impact using F1000 labels. *Scientometrics* 2013; 97(2): 383–395.
- [25] Bornmann L and Haunschild R. Which people use which scientific papers? An evaluation of data from F1000 and Mendeley. *J Informetr* 2015; 9(3): 477–487.
- [26] Heath T, Singer R, Shabir N et al. Assembling and applying an education graph based on learning resources in universities. In: *Proceedings of the linked learning (LiLe)*, 17 April 2012, <https://tinyurl.com/rk82qah>
- [27] Shabir N and Clarke C. Using linked data as a basis for a learning resource recommendation system, <https://tinyurl.com/sr8kd59>
- [28] Beetham H, Newman T and Knight S. *Digital experience insights survey 2018: findings from the pilot of teaching staff in UK further and higher education*. JISC, <https://tinyurl.com/y4cpoavr>
- [29] Beetham H, Newman T and Knight S. *Digital experience insights survey 2018: findings from Australian and New Zealand university students*. JISC, <https://tinyurl.com/y5vnyqkn>
- [30] Morgan R. The online reading list project using Talis List at the University of Glamorgan. *Program* 2007; 41: 262–275.
- [31] Jones H. Reading lists in Cambridge: a standard system? <https://tinyurl.com/y3lcuqb>
- [32] O’Neill M and Musto L. Faculty perceptions of Loughborough’s Online Reading List System (LORLS) at Dublin Business School (DBS). *New Rev Acad Librarian* 2017; 23(4): 353–376.
- [33] Cross R. Implementing a resource list management system in an academic library. *Electron Libr* 2015; 33: 210–223.
- [34] Bamidis PD, Konstantinidis ST, Bratsas C et al. Federating learning management systems for medical education: a persuasive technologies perspective. In: *Proceedings of the 2011 24th international symposium on computer-based medical systems (CBMS)*, Bristol, 27–30 June 2011, pp. 1–6. New York: IEEE.

- [35] Pereira CK, Campos F, Ströele V et al. BROAD-RSI – educational recommender system using social networks interactions and linked data. *J Internet Serv Appl* 2018; 9(1): 7.
- [36] Piscioneri M and Hlavac J. The minimalist reading model: rethinking reading lists in arts and education subjects. *Arts Humanit High Educ* 2013; 12(4): 424–445.
- [37] Derven C. ANLTC/SWETS Research Fund evaluating services and specifications for reading list systems, 2011, <https://tinyurl.com/y3usyrbby>
- [38] Wusteman J, Eden B and Masson A. VRE library services. *Library Hi Tech*, 12 June 2009, <https://tinyurl.com/y5v5u77g>
- [39] Chowdhury G. Carbon footprint of the knowledge sector: what's the future? *J Doc* 2010; 66: 934–946.
- [40] Chowdhury G. Building environmentally sustainable information services: a green is research agenda. *J Am Soc Inf Sci Tec* 2012; 63(4): 633–647.
- [41] Gillard JW. An initial analysis and reflection of the metrics used in the Teaching Excellence Framework in the UK. *Perspectives* 2018; 22(2): 49–57.
- [42] Kaner C, Bach J and Pettichord B. *Lessons learned in software testing*. New York: John Wiley & Sons, 2008.
- [43] Braun V and Clarke V. *Successful qualitative research: a practical guide for beginners*. London: SAGE, 2013.
- [44] Wellington J. *Educational research: contemporary issues and practical approaches*. London: Bloomsbury Publishing, 2015.
- [45] Huberman M and Miles MB. *The qualitative researcher's companion*. Thousand Oaks, CA: SAGE, 2002.
- [46] Elo S and Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008; 62(1): 107–115.
- [47] Irvine KN, Warber SL, Devine-Wright P et al. Understanding urban green space as a health resource: a qualitative comparison of visit motivation and derived effects among park users in Sheffield, UK. *Int J Env Res Pub He* 2013; 10(1): 417–442.
- [48] Ekstrand MD, Kannan P, Stemper JA et al. Automatically building research reading lists. In: *Proceedings of the 4th ACM conference on recommender systems*, Barcelona, 26–30 September 2010, pp. 159–166. New York: ACM.
- [49] Beall J. Metadata and data quality problems in the digital library. *J Digit Inform* 2006; 6(3): 20.
- [50] Ko A and Gillani S. A research review and taxonomy development for decision support and business analytics using semantic text mining. *Int J Inf Tech Decis* 2020; 19(1): 97–126.
- [51] Nadkarni PM, Ohno-Machado L and Chapman WW. Natural language processing: an introduction. *J Am Med Inform Assoc* 2011; 18(5): 544–551.
- [52] Müller HM, Kenny EE and Sternberg PW. Textpresso: an ontology-based information retrieval and extraction system for biological literature. *PLoS Biol* 2004; 2(11): e309.
- [53] Cunha D, Neves P and Sousa P. PaaS manager: a platform-as-a-service aggregation framework. *Comput Sci Inf Syst* 2014; 11(4): 1209–1228.
- [54] Serbanescu V, Pop F, Cristea V et al. Architecture of distributed data aggregation service. In: *Proceedings of the 2014 IEEE 28th international conference on advanced information networking and applications*, Victoria, BC, Canada, 13–16 May 2014, pp. 727–734. New York: IEEE.
- [55] Bird KS and Pitman L. How diverse is your reading list? Exploring issues of representation and decolonisation in the UK. *High Educ* 2020; 79(5): 903–920.