

BRII Building the Research Information Infrastructure Project

Stakeholder Analysis

Exploratory Study into the requirements and
uses for research activity data at the
University of Oxford

Dr Cecilia Loureiro-Koechlin

BRII Project Analyst

July 2009

Acknowledgements

We are grateful to all the individuals who gave freely of their time to be interviewed for the purpose of this analysis.

A list of participants can be found in Appendix 1.

Executive summary

This study was part of the Building the Research Information Infrastructure (BRII) project, which aims at enabling efficient sharing of research activity data¹ across the University of Oxford by using semantic web technologies. The aim of the study was to examine the context in which research activities unfold at the University of Oxford. It identifies the project's stakeholders, their interests, and the kinds of settings and situations under which the research information infrastructure (RII) and its services will be used.

Data for this study were collected from interviews from a cross section of University groups, such as academic departments, administrative units and colleges. Data were categorised and analysed by using an interpretive-based content analysis method.

Findings reveal a broad range of views with respect to uses of and needs for research activity data. These views are influenced by factors such as research field, type of job (academic, non-academic) and scope (does the job involve one field or department, or more: divisional level, University level and cross disciplinary?) The findings are grouped in four main categories resulting from the data analysis process:

1. *Stakeholders*: BRII stakeholders are everyone whose work is directly or indirectly related to research activities or who are interested in them; and who have a need to access and/or display research related information. This report identifies stakeholders according to their job function: researchers, administrative staff, and strategists and disseminators. The findings suggest that the information needs for these groups vary from more specialised and narrow in scope to general and wide (e.g. across disciplines.) The study also identifies a spectrum of stakeholders according to their research culture. The two extremes are named as extreme scientists and extreme humanists. Data showed that research culture influences the way information is generated and searched for (e.g. frequency of creation and maintenance, and selection of sources.)
2. *Perspectives on research activity data*: the study uncovers a set of factors which influence the way stakeholders perceive and use research activity data. These factors are classified in *factors of influence* which are the characteristics of research activity data projected into stakeholders' work and responsibilities (e.g., trust, sensitivity and visibility of data); and *factors of use* which are the reasons for using research activity data (e.g., keeping up with the field). Additionally the study also shows how these factors are influenced by current technical systems whose functionality involves the use of research activity data.
3. *Research activities*: findings show a wide variety of kinds and shapes of research activities. These activities are influenced by the field and research culture in which they take place. Findings also suggest that the clarity in which these activities are described in data is influenced by the need for and/or existence of external funding (i.e., unfunded research activities do not appear in websites or internal systems.)

¹ Data about research in the University of Oxford

Additionally, ways for organising research activities such as collaborations or research themes are described.

4. *Current issues on data*: these are the practicalities of generating and using research activity data. Respondents reported issues such as responsibility for producing and maintaining research activity data, duplication of data and data formats. Additionally a list of research activity objects was gathered from responses. These objects correspond to information that respondents use (from internal systems or the web) or to “things” they search for, do or know about, and which are important for their work.

The findings of this study show that stakeholders can be very different from each other. Stakeholders’ characteristics and behaviour are not clear cut, but vary along more than one spectrum. Implications for the RII are therefore the need to account for *variety* across a spectrum of uses, such as different ways of searching (e.g., browsing, search box), provision of information at different depths (e.g., specialised, cross disciplinary overview, general overview), and the provision, where possible, of multiple views of the same research activity objects (e.g., projects by department, projects by researcher).

Additionally this study uncovers the characteristics of the data that will be harvested by the RII. Research activity data are heterogeneous because they represent an heterogeneous universe of research activities. Therefore sources of data that will be harvested by the RII will be in no way consistent or compatible. Also, as research activities within the University change at a fast rate, so will the data that represent them. Therefore another implication for the RII is to evolve together with the data that it holds for example by providing services to access new categories of objects.

Finally, the findings of this study provide evidence of:

- The need for an organised and easy to access store of information about research in the University of Oxford which at the same time represents its diversity.
- The need for trusted and comprehensive sources of information which are relevant for every research field in the University.
- The need to search for information about people, activities and resources from across the University at administrative and academic levels across departments and disciplines.
- The need to discover and gather information from different sources.
- The need to reuse information for different purposes.
- The need to improve research visibility by creating websites or other sources which provide overviews of research activity.
- The need for tools to access information about research from a variety of perspectives and levels of aggregation.
- The need to reduce inefficiencies.

From this list it is evident that an infrastructure with the characteristics of the RII will fit perfectly within the academic and research processes taking place in the University of Oxford.

Table of Contents

Acknowledgements.....	1
Executive summary.....	2
1. Introduction.....	5
1.1. Background	5
1.2. Project brief	5
1.3. Aims of this study	5
2. Methodology.....	6
3. Findings	8
3.1. Stakeholders	9
3.2. Perspectives on research activity data	14
3.3. An overview of research activities	21
3.4. Current issues on data	28
4. Conclusions.....	30
4.1. Insights and outcomes of the analysis	31
4.2. Implications for development.....	32
4.3. User needs - summary	33
5. Appendices.....	34
5.1. Appendix 1: List of interviewees	34
5.2. Appendix 2: List of data contributors.....	36
6. References	37

1. Introduction

1.1. *Background*

The day to day activities of doing research impinge on many groups and individuals within and across higher education institutions. All those involved need access to research activity data (i.e., data about research) and all create such data. However, most current initiatives to collect research information within Oxford have their own aims and business processes at the forefront, and, where a particular business need is pressing, making the data collected available for other purposes is of secondary consideration or altogether out-of-scope.

1.2. *Project brief*

The Building the Research Information Infrastructure (BRII) project² will enable efficient sharing of research activity data. It will produce a scalable infrastructure to capture data of this type from a disparate range of sources. Research activity data describe the people, resources, backgrounds, ideas and outcomes related to research activities of all kinds at the University of Oxford. From this infrastructure BRII will create exemplar services that disseminate and re-use that data using a lightweight solution based on semantic web technologies.

The BRII project is funded under the Joint Information Systems Committee (JISC) Institutional Innovation Programme. The project began on October 2008 and will end in March 2010. The expected outputs of BRII are:

- Collection of ontologies and taxonomies
- University Blue Pages of research activity information
- Example use of API e.g. integrating with a themed website such as Cancer
- Oxford Research Information Infrastructure

1.3. *Aims of this study*

This study was an effort to understand the context in which research activities unfold at Oxford. In doing so it outlined the kind of settings and situations under which the Research Information Infrastructure (RII) and its services will be used. This study identified the project's stakeholders. It assessed their interests and the ways those interests could affect or be affected by the RII. Perspectives were also collected on current needs and uses of research activity data at different levels and across different fields within the University. All this information is used in this report to define profiles of the people and activities, patterns of behaviour and beliefs, work settings and research cultures which determine the kinds of data needed from the infrastructure and the way in which that data will need to be delivered. The content of this report will inform the design of use cases and more detailed development documentation in the future.

² More information about the BRII project can be found in the project's website <http://brii.ouls.ox.ac.uk> and weblog <http://brii-oxford.blogspot.com>

2. Methodology

This study was guided by an interpretive approach to information systems. An interpretive stand addresses the world from the point of view of the people being studied (Hammersley, 1992). It believes that the world is a product of the mind (Williams and May, 1996) and the concern is on the “meanings that people give to their environment, not the environment itself” (May, 1997). In this study, an interpretive approach allowed access to first-hand knowledge and understanding of stakeholders’ interpretations (or beliefs) of research activity data (what are they? do they use/need them?) as well as the effects of the BRII project on their work (why is it relevant to their work?). Perceptions about the RII were elicited from stakeholders involved directly or indirectly with research activities in the University of Oxford. Their perceptions are crucial to the success of the BRII project as the RII will be embedded within practices across diverse information systems dealing with research activity data. In doing so the RII will play an important role within Oxford research networks facilitating research and administrative interactions.

As very little was known with respect to uses of and needs for research activity data in the University of Oxford, an exploratory design was used. Focusing on involving as many areas as possible rather than analysing in depth people’s perceptions, this study tried to uncover patterns of beliefs and behaviour across the University’s different academic and administrative units.

A comprehensive strategy was devised to interview individuals, who were representative of all dependencies of the University, including:

- At least one department from each academic division,
- Three offices with the University Administration and Services
- 2 University Colleges
- 2 other administrative areas

Note: A complete list of interviewees is provided in appendix 1.

Additionally, staff representing a variety of roles were chosen for interviews, ranging from researchers, who are directly involved with research activities, to administrative staff who support researchers; and from departmental and divisional officers to University level administrators who need to oversee research activities across different areas. With this strategy we made sure that we were collecting information which accounted as much as possible for the heterogeneity of Oxford’s administrative and research structures and cultures.

Names of interviewees were collected through known contacts (e.g., users of ORA³) or staff who had been previously interviewed. This was done to ensure that only relevant people were asked for interviews. Initial contacts were done by emails. Known contacts were always used as references. This helped with getting trust and credibility from potential interviewees. Emails contained a brief of the project and the purposes of the interviews. Contacts were asked for a 30 minute face-to-face interview. If that was not possible, they were offered the

³ Oxford University Research Archive <http://ora.ouls.ox.ac.uk>

option to complete an online survey created at [SurveyMonkey](https://www.surveymonkey.com)⁴. Most of the contacts that responded to our first email preferred to be interviewed in person. Only five people completed the online survey. Two of them were interviewed afterwards as well.

Interviews were designed to be unstructured and open ended, to change and adapt to each individual's particular area of expertise. They intended to generate insights into the practices and ideas that constitute the management and performance of research in Oxford. This allowed for some flexibility in data collection, allowing issues to emerge and then be examined at a later stage.

Interviews always started with the interviewer giving a summary of the project, and by explaining the interviewee his/her role in the development of the RII. In most cases interviewees responded to the introduction by commenting on issues they found relevant. Some of them were supportive of the RII whereas others highlighted complexities and problems which could hinder the BRII project. When the interviewees did not comment after the introduction, the interviewer started by asking questions about the respondents' jobs and responsibilities. After that, questions were asked in the following three areas of enquiry:

- Creation and management of Research Activity data.
- Use of Research Activity data (perhaps from other sources.)
- Issues and future uses of Research Activity data.

Note: the online survey contained multiple-choice questions covering the same areas mentioned above.

The focus of BRII is on collecting freely available research activity data. However questions about restricted data were also asked. These questions involved the use of financial and project management information; and issues regarding the visibility of this information on the web. Responses helped to define the boundaries as to what kind of data BRII will be allowed to collect and display publicly.

As the interviews were unstructured, questions varied according to the respondents' roles and areas of expertise; sometimes covering only one or two of the areas mentioned above. Participants talked about their daily routine and the practicalities of using research activity data. They were also asked to explain the reasons for using or manipulating data the way they did. In addition respondents reflected on issues, problems and differences with other areas that they were aware of. At the end of interviews respondents were asked whether they were willing to contribute their research activity data to the project. The full list of potential contributors is given in appendix 2.

Interviews took place within a period of three months. However additional interviews are being scheduled with staff from other areas of the University. These interviews will be used to enrich and validate the findings of this report. So far interviews have lasted between 20 minutes and 1:45hr each. All interviews were recorded with a digital recorder with the permission of respondents. Each interview was transcribed into Ms Word documents⁵.

⁴ www.surveymonkey.com

⁵ Audio and Word files containing all interviews are kept in the BRII project records.

Transcriptions were not done verbatim due to time constraints. However they captured the ideas covered in every interview. All transcription documents were entered in a qualitative data analysis tool which helped with the coding of texts.

A content analysis (Priest et al., 2002) method was used to analyse the data. This method involves the coding of texts under categories. These categories emerge from the reading of texts as interpreted by the researcher and reflect the purposes of the research. Coding of data is a process whereby segments of texts are grouped according to commonalities in their meanings. From the texts grouped under each category, concepts and propositions defining those categories are developed at high levels of abstraction. This means that they may stand for similar pieces of data, and are useful to explain general patterns of behaviour and beliefs in areas not covered by the interviews. This kind of process is called interpretive generalisation (Walsham, 1995). Data coding is an iterative process in which categories are further refined after each round. The more iterations of data coding, the more refined the categories and their concepts. Unfortunately due to time constraints of this study, only one round of data categorisation was possible. This one iteration is sufficient to grasp stakeholders' perspectives and start designing a pilot for the RII. Further work on user needs will be necessary if the pilot is extended and enhanced beyond the official term of the project.

The following are the top-level categories that emerged from this study's interviews:

1. Stakeholders: this category contains information which defines the kinds of stakeholders involved with the BRIL project. These are not proper names but profiles and roles which are distinctive to Oxford's research community.
2. Perspectives on research activity data: this category is the core of this study. It reveals the views, positive or negative, that stakeholders have on the project. They range from ways and reasons for doing certain things in the research arena to the principles that guide their actions.
3. Research activities: these data describe the structures of research activities as well as the cultures that inspire them.
4. Current issues on data: refers to the characteristics and qualities of research activity data, as they are currently used.

A detailed discussion of each category is given in section 3 of this report. Discussion on category 1 will identify BRIL's stakeholders which is one of the aims of this report. Discussion in category 2 deals with the second aim which is to explore the applicability to and impact of the project on the stakeholders. Category 3 complements the first two categories as it describes the contexts under which stakeholders work and under which the RII will be used. Category 4 describes the practicalities of generating and using research activity data.

3. Findings

This section describes the findings after data collection and analysis. It presents in a narrative and organised way all issues raised during the interviews. Main insights and outcomes of this analysis will be highlighted in section 4 of this report as well as some development implications.

Overall, the data analysis carried out on the information collected from interviews with stakeholders across the University of Oxford reveals a broad range of views with respect to uses of and needs for research activity data. These views are influenced by factors such as research field, type of job (academic, non-academic) and scope (does the job involve one field or department, or more: divisional level, University level and cross disciplinary⁶?) Where possible, distinctions will be highlighted between the perspectives influenced by these factors.

3.1. Stakeholders

As the University of Oxford is a research-intensive institution, most of its activities are directly or indirectly related to research. University staff members are involved with research activities in different ways. Due to the size and heterogeneity of the University and its staff, there is a need to organise stakeholders in groups to better facilitate understanding and capturing of their characteristics. This report will use two typologies to classify stakeholders which are considered relevant to the study. The typologies are a) Function and b) Research Culture. Stakeholders' profiles highlighting their responsibilities and perspectives are described within each typology as follows.

Stakeholders according to their function

- a. **Researchers** are those who define and undertake research activities. They define, carry out and produce relevant research outputs and outcomes. Researchers possess knowledge and expertise as key elements of success.

Researchers work at different levels of responsibility. Principal Investigators lead research activities. Collaborators, associates or fellows work alongside principal investigators. Research assistants, postdoctorals and doctoral students also provide support while acquiring relevant experience.

Research activity data are all about researchers, what they do, and what they produce. Researchers produce most research activity data, from project funding applications and project overviews, to their own research profiles and publications. Related administrative, and project management data are generated from their activities as well. However, the task of managing this information (maintaining, uploading into internal systems and websites) is usually done by administrative staff.

Information from interviews suggests that as researchers are specialised in a particular area of expertise, their information needs are bounded by such an area. However, current areas of research are not necessarily defined by traditional (discipline-based) fields. As the need for originality and innovation by cross disciplinary research is on the increase, researchers need to be aware of activities located at the edges of their fields and/or in the intersection with other areas. This compels researchers to be more proactive at extending their networks of contacts and information beyond their discipline.

Researchers are interested in the academic aspects of research activity data but also in the ones who own and produce that information. Preferred sources for information about

⁶ The terms cross disciplinary, multidisciplinary and interdisciplinary are interchangeable in this report

completed research are academic publications (e.g., electronic journals and databases.) Preferred sources for information about research in progress are conferences, and project and sponsors websites. Additionally some respondents stated that they access sources via a number of departmental or University services (databases, spreadsheets) which contain information on current research.

- b. **Administrative staff** deals with the management component of research activities at different levels. Depending on the size of their areas of responsibility they can oversee activities from small to large sized groups whose scope also ranges from departmental to divisional or international levels.

They usually help researchers with the administration, control and finances of their projects. Variety in this group is defined by the type of support and expertise they provide to research activities throughout their lifecycle. Support ranges from preparation of funding applications, personnel recruitment to monitoring and tracking of progress, as well as financial control. Typical administrative roles are: departmental administrators, research administrators and financial officers. A role which sometimes overlaps functions with the ones of researchers is the research facilitator. In some cases research facilitators need to have specific knowledge or expertise in particular areas of research to be able to provide support to researchers. Tasks which overlap research and administrative functions are for example, the writing of reports at departmental level which assess their research weaknesses and design strategies to tackle them.

Administrative staff generate administrative and project management information to ensure the success of the research effort. The nature of these data is longitudinal as it has information taken at different points in time which help track activities. This information includes allocation of finances for research activities, distribution of costs and overheads, statistics of progress and completion. Administrators maintain this information with that being produced by the researchers (academic, scientific information.) This is the case for instance of websites which have to be updated periodically. Although staff members help with the management of technical information they do not have to be experts in it, but they have to be able to acquire it from local or external sources.

Interview data show that a key driver for administrative staff is the achievement of financial benefits in the delivery of projects leading to further increase of financial resources. This means they strive to keep existing projects within budget and on time so that researchers can demonstrate a good degree of trustworthiness to funders. This driver is complemented by a mentality of acquiring and maintaining an excellent reputation.

- c. **Strategists and disseminators** are those individuals who work at University level and who are not directly involved with research activities. However they need to keep well informed so they can develop strategies for future research priorities and disseminate research information across the University, nationally or internationally. Roles within this category include head of departments and divisions, research and strategy committees, vice chancellors and the Press Office.

Similarly to administrators, this group is not necessarily composed of academic experts. However unlike researchers and administrative staff, strategists and disseminators do not produce research activity data but are heavy consumers of them. They are continuously fed with information which they have to interpret, adapt, reformat or publicise in different outlets. The information they use comes in the form of snapshots, accounting for activities which are taking place at one point in time. Products of their work could be classified as reports or summaries written in non-specialised language and whose scope covers all University's research disciplines. They possess a high level perspective of research activity at the University.

The above groups and their scope are represented in the following diagram. Figure 1 conveys the idea that the more specialised and detailed the information needed, the narrower the search space becomes for an individual. The more specialised an individual is the more specialised the information they need. The diagram also shows the idea that the higher the level of management is the more general information is required and vice versa.

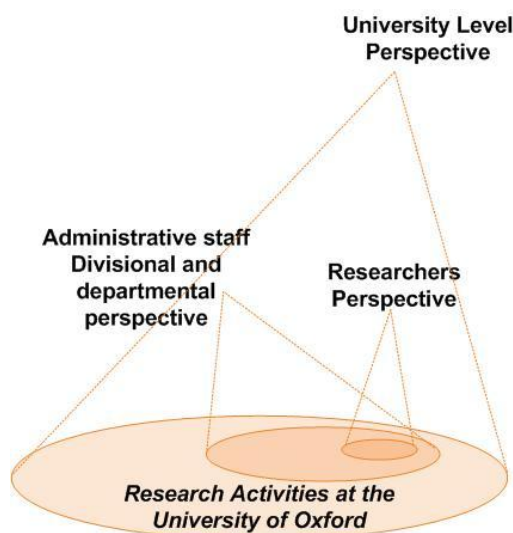


Figure 1 Scope of research activity according to functions.

Stakeholders according to research culture

Within this typology two extremes of a spectrum can be identified. The extremes are stakeholders with orientation to Science or Humanities. Between these two extremes many other stakeholders can be identified whose orientation is not often clear cut. This typology was designed based on researchers' culture but also involves the administrative staff who support them.

- a. **Extreme scientists** are individuals who normally work in large groups⁷ with other scientists. Their research agenda is often externally driven. The nature of their research activities requires them to acquire sophisticated research resources and to collaborate with people within or outside the University. They continuously

⁷ One academic from the Medical Sciences division stated that he had co-authored his last paper with 250 people.

monitor opportunities for funding to support long term efforts. Because of the pressure to gain international recognition and excel, they work in a highly competitive environment.

Scientific research is often empirical in nature, based on hypotheses which are tested in laboratories or in fieldwork. Their research outputs are diverse. They can include for instance: new devices, software, patents, drugs and treatments. In producing such outputs it is essential to develop and maintain strong links with industry. Scientific outputs are ephemeral; their value today might not be the same as tomorrow.

As they need to maintain larger groups and a wide range of resources their funding needs can easily increase with time, making the visibility of their efforts a key element of success.

Administrative staff members working in scientific-driven research areas are usually required to be proactive at providing their support for researchers. For example, there will be research facilitators responsible for seeking funding opportunities and potential collaborators.

- b. **Extreme humanists** on the other hand are people who can easily work alone or in small groups. Their quest for funding is not crucial to their activity therefore funding can be intermittent, if not seen as unnecessary (they feel they are already paid to do their work.)

For some, humanistic research does not have a consistent nature, comprising disparate disciplines such as history, linguistics and philosophy. However some general research patterns can be identified. For example humanistic research approaches tend to be hermeneutic (Boland, 1985) in nature adding to existing theories by interpreting and understanding relevant literature (as opposed to empirical research.) Humanists tend to work in their offices. They locate themselves in quiet places where they can work with no interruptions. Work involves reading and writing and there are few, if any, experiments or fieldwork.

Interviewees reported that the main outputs of humanists come in the forms of ideas and are disseminated via books or journal articles. In this sense the research is more perennial and its value often lasts longer in time.

Support from administrative staff is less necessary in this area as there are fewer officially defined, externally funded activities and as research activities are simpler (i.e. it is simpler to organise writing sessions between researchers than to build and acquire equipment for a laboratory). Administrative staff who oversee and support research activities can be shared between departments, and usually manage grants, recruitment and resources.

As said before interviewees possessed a mixture of skills, interests and goals which combined both of the above orientations. For instance some researchers develop collaborative work in small groups and generate joint research papers; other activities would involve only theoretical developments from research data from previous

experiments or fieldwork. Figure 2 illustrates the spectrum with the mentioned features.

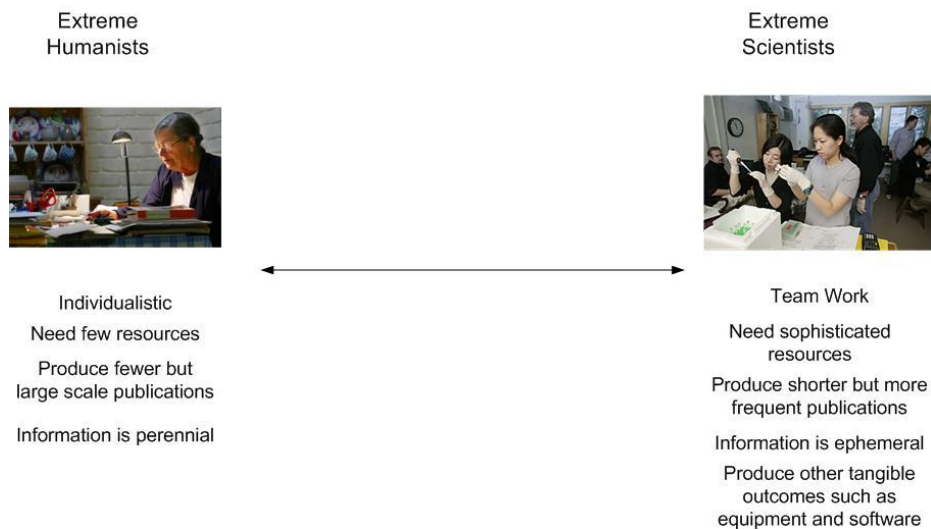


Figure 2 Nature of Research Activities according to Research Cultures

A final consideration in the distinction of stakeholder groups, particularly academic related, is the size of the area where they belong to. The larger the group the more competitive the department becomes, the bigger the organisational challenges and the more resources that need to be available. Larger departments for instance have already implemented and use information systems (processes and in some instances software) to support their work.

Connecting the two typologies

Data from interviews suggest that researches and support staff (as presented in the first typology) are very much bounded by the research cultures which surround them. Strategists and disseminators (at divisional and University level) are not necessarily influenced by such cultures. However, their needs for information require them to access resources which range between the two extremes. For example, to produce a *country profile*, information needs to be collected on all areas of the University where research relevant to that country is done. This involves collaborations, literature reviews and data collection done in or about a particular country. Information collected could range from development of new drugs to analysis of economic crises as well as linguistic and language studies.

Role of University Colleges in research

Traditionally colleges have been places for teaching. Little if any research is done there. Two interviewed academics, both college fellows, concurred that this varies depending on the research culture of the researcher. Scientists undertake their research work in their faculties, where they have access to laboratory or other resources. They think of colleges as places where they go to teach. Humanists tend to research in their college offices, especially when they need only to read and write. However, little acknowledgement of the research activities of fellows appears on colleges' websites.

One researcher explained that colleges are now being rewarded financially for having research active fellows⁸. He said he did not understand why, because research would be carried out elsewhere.

3.2. Perspectives on research activity data

This section of the report discusses views of research activity data as seen by the interviewees. A definition of this type of data is not provided here. Instead the focus is on the nature of the processes surrounding research activity data and the issues that arise from their use. This approach provides insights into stakeholders' interests and on the implications that the use of these data has on their work.

The data analysis uncovered categories corresponding to factors of influence and use. These two will be presented separately. Influence factors are the characteristics of research activity data which are projected into the work and responsibilities of the stakeholders. Practical reasons why research activity data are used the way they are form the factors that govern their use. Additionally perceptions are also influenced by current technical systems whose functionality involves the use of research activity data. A list of those systems is provided at the end of this section. Some of these systems are perceived to have overlaps with the RII. Both factors, influence and use, and perceptions of related systems are interdependent and together constitute the preconceptions or mental images stakeholders currently have of research activity data.

Factors of influence

- a. **Trusted research activity data:** this issue was raised by researchers and not mentioned by other types of stakeholders. Academics are critical of the information they use, as it has to be valid and reliable within their field of expertise. Trusted sources of information include peer reviewed journal articles and electronic bibliographic databases. Also most researchers are building or have established networks of contacts that they consider to be reliable sources of information.
- b. **Sensitivity of research activity data:** sensitive information is that kind of information which, if made public, could harm the reputation of researchers and institution and hinder the execution of research activities. In general information about people and their publications is not seen as confidential. However there are some areas of sensitivity within the actual research activity realm. Data from interviews show four potential areas of conflict where data would be preferred not to be made visible. The first area is finances and project management information. Administrative staff particularly, regard this kind of information as internal and used for administrative purposes rather than for display. The second is controversial topics of research. Researchers are concerned about publication of information on these topics because of their sensitivity. An example of this is research on animals. The third area of conflict is where confidentiality agreements exist, particularly with

⁸ This is via a system called Joint RAM or JRAM developed by a working group (JRAWG), which includes college and University representatives.

industrial sponsors. The final area is unfinished research. Making public incomplete research would harm the reputation of the researcher or cause other difficulties.

- c. **Visibility of research activity data** is addressed by interviewees from two points of view. The first is the traditional view by which researchers make their work accessible to others through academic sources such as peer reviewed publications, conferences, and research networks which may or may not be provided online. These are regarded as highly relevant and sometimes as the only necessary media by researchers. Researchers make their work visible in this way to register their work and to certify themselves as the owners of that research. The second view is publicising the research on the open Web via channels other than those traditional sources. There is a general agreement that exposing research in this way is extremely beneficial for researchers who are looking for external funding. This has been highlighted in the interviews by administrative staff. They believe that making profiles and project overviews freely available in this way plays an important role in the success of subsequent individual applications for funds. This is enhanced when the researchers' academic unit has a highly respected reputation.

Although that, in most cases, publications can be made available online, online databases and journals are not considered as online sources per-se but as efficient ways for getting a copy of the actual papers.

Personal and organisational reputations are also important drivers of visibility, particularly in competitive fields.

Administrative staff does not perceive visibility of research on the web as crucial where research does not depend on external funding.

- d. **Difficulties at accessing data:** some respondents, all of them administrative, reported two types of difficulties when wanting to access research activity data. The first is accessing some information resources such as when needing to access RAE⁹ statistics to inform strategic decision making, or to obtain information on successful and failed grant applications to be used for strategic purposes and to improve future applications. Respondents stated that this information was not available to them for policy reasons. The second type of difficulty is that of connecting data at different aggregation levels. One particular example given was extracting the names of Oxford researchers from country level data or a bibliography on a particular country.

In general academic staff are confident when seeking research activity information in their specialist discipline. They are experts in the subject and therefore are familiar with the terminology, and have acquired knowledge of efficient and effective search techniques and a network of contacts which they use consistently.

⁹ The *Research Assessment Exercise (RAE)* assesses the quality of research in universities and colleges in the UK

More difficulties are found among administrative staff where tasks involve searching and using a combination of data from multiple sources, and which is written in specialised language.

- e. **Interference:** only one academic expressed concern over the possible interference that an infrastructure like the RII could cause in their work. This academic explained that researchers organise and maintain their research activity information in the way they prefer and when it is convenient for them. External demands will put only more pressure on their work. This respondent also added that an initiative like BRII should be driven by a bottom-up approach, i.e., by the academics. However from the administrative point of view, an initiative like BRII would allow re-use of data (e.g., at creating and updating websites) facilitating administrators work by not having to request that information from the academics many times.
- f. **Structuring-categorising-organising data:** this category reflects on the actual organisational and research structures of the University and on how these structures are reflected in the data.

It was noted during the course of the analysis that research topics or subjects evolve quicker than the structure of the University. Therefore there can be a mismatch between current descriptions of academic disciplines used by the active research community and the University's naming of research areas. This creates confusion among people who are not familiar with specific research topics or who are new to the research arena (i.e. potential students.) Drivers for change among research topics are progress and new discoveries in research and competition all of which compel researchers to adapt and be more original at a faster rate. Research topics are more easily adaptable than organisational structures. New topics are created and others cease to exist. They split or merge with others. Some divisions deal with these rapid changes by creating research themes¹⁰ which group researchers according to their interests.

When asked, researchers and administrative staff classified their research activities under their research topic(s), leaving the department or unit, where they are taking place, as a secondary element in their description. In websites, academic publications are more likely to be listed under their authors who in turn will have a link to their department.

Additionally a relevant way of classifying research is by the name(s) of the researchers who are responsible. One interviewee stated that in some circles and circumstances researchers refer to research activities under the name of the PI and not listed as a named research group¹¹.

Research topics are also used to classify researchers, in forms of research interests or research keywords. Some researchers dislike any form of classification. Other academics would like to be able to choose their own keywords. When keywords are

¹⁰ E.g., Cardiovascular Science theme www.cardioscience.ox.ac.uk

¹¹ E.g., Department of Biochemistry, research groups
www.bioch.ox.ac.uk/aspsite/index.asp?sectionid=rgroups

created by researchers they are field specific and specialised, and as no coordination is made across disciplines fake intersections and uncertainties can be created.

Some administrators and strategists stated that there is a need for a unique, centralised and cohesive set of keywords for the whole University. In this way there will not be ambiguities, but one system to learn and use. Perceptions are that having a University scoped tool such as the RII will help to create such a set of keywords. Contrary to this, one researcher questioned the effort of trying to systematize information of such diverse kinds of research.

There is clearly a conflict between perceptions of academics and administrators in this area. Whereas academics chose keywords which demonstrate their expertise in specific fields, administrators and strategists would like to have an easy to search and easy to understand set of keywords.

- g. **Time issues** refer to issues regarding the relevance of research activity data across time. The speed at which research fields change influences the way academics and support staff search for information. Researchers from the scientific extreme of the spectrum trust publications to find information about current research. As it takes only a few months for a paper in their area to be published and its content is thought to be relevant in the immediate future after publication, researchers are in constant search for information to keep up to date with the field. However, after a while papers lose their validity as more research is reported. Humanists work at a slower pace and, the validity of their outcomes tends to be longer. Even though papers in the humanities might take years before publication, most of them will remain useful as research material.

Strategists and disseminators need time to collect information and adapt it to different audiences. They believe that as soon as they finish their work the information that they have produced becomes outdated (e.g., a list of experts in one field.)

h. Audience

This category represents the kinds of people that usually consult research related websites. The list below is a combination of actual and potential readers. Actual readers are those who stakeholders are sure read their websites. Potential readers are those who stakeholders expect to read their websites but who cannot be fully confirmed as such.

- Researchers (within or outside Oxford) looking for information in their field of expertise
- Researchers (within or outside Oxford) looking for information which is not in their field of expertise
- Non-academics: administrative staff, strategists, disseminators
- Novices: potential students, early career researchers
- Funding bodies

Academics and administrators take into account the above list to design the content and format of their websites. For example academics would have an easy to read or

non-scientific section if they would like their site to be visited by potential students. They would also have a scientific-language version of their website for other researchers and funders.

As BRIL aims at disseminating and providing an easy search facility for the same kind of information but at a greater scale, the above list represents the potential users of the RIL.

Factors of use

Interviewees search for and use research activity data for different reasons. The following were the most highlighted.

- a. **Keeping up with the field** by acquiring knowledge which can be useful for one's own research. Most academics develop their own personal ways of doing this. For example by subscribing to mailing lists or browsing colleagues' websites on a regular basis. Academics also keep updated by means which does not involve research activity data, for example by attending important conferences attended by relevant people in their field.
- b. **Controlling and keeping track** of grants is one of the priorities for administrators, particularly where their departments are funding dependant. Funders demand compliance with strict requirements. Although the responsibility of the research activities relies on the researchers, they need support with the administrative aspects of the activities. Some of these requirements are progress reporting, milestones, budgets.

Another area for control is the maintenance of information in websites. Again most of their content is produced by the researchers, but administrators have to make sure they are up to date, particularly when a further application for funding is submitted or when a project milestone is reached.

Some tasks require centralised control, a given example being the writing of departmental quarterly reports for sponsors. One member of staff is assigned to collect information from all research activities across the department and then assemble it into a coherent report.

- c. **Finding someone or something:** In particular circumstances researchers, administrators, strategists and disseminators need to find details about specific people or activities. Searches could involve an initial exploration of a variety of sources if little is known about the subject. Reason for searching ranged from finding experts, to finding collaborators associated with known people and finding information about resources within the University (e.g., their location or the expertise needed to build and maintain a particular resource), as well as finding information about important grants and discoveries.
- d. **Disseminating and publicising** research activities involve similar tasks as explained in "controlling and keeping track" (letter b in factors of use). Usually the aim is to produce documents which provide an overview of research in certain fields or

places, or to produce narratives focused on a particular research activity but which are aimed at non-expert audiences. For disseminators research activity data are the raw data they need to start their work. They sometimes need to contact the original sources in person to verify and obtain more information.

- e. **Building strategies:** information on current and past research activities is essential for building new strategies. Strategists need information which can help them identify strengths and weaknesses in their area. For example, they would like to know with which sponsor they are more successful and with which they are not. Such information can be found in previous applications, RAE (REF) reports and sponsors websites. They analyse that information and produce reports which are used to support new strategies.

In addition, data from interviews provided insights into how respondents search for research activity information. Three patterns were found which correspond to the *function* typology (see section 3.1.) Because academics are interested in publications they focus on academic journals and bibliographic databases. Administrators are more interested in information on grants. Strategists and disseminators are interested in general information. The following list contains some of the most mentioned sources. Note that this list is not exhaustive but gives a good idea of commonly used sources of data.

Source	Academic	Administrative	Strategist/ disseminator
Google	x	x	x
Websites	x	x	x
Pubmed	x		
Web of Science	x		
Bibliographic databases	x		
Citation indexing	x		
RSS Feeds	x	x	
Notifications, automatic emails	x	x	
RDInfo (RDFunding),		x	
ResearchResearch		x	
Research Services reports		x	
Databases provided by funding bodies		x	
Through other people, in meetings, word of mouth	x	x	x
Via the press			x

Relevant Systems

During the interviews names of several University information systems (i.e. computer applications) were mentioned in relation to the use of research activity data, or to the management of research. Not all respondents use or have used these systems but in some cases are users of their outcomes (e.g., reports). The systems are as follows:

<i>Pinfox: (PI Information Oxford) management of research activity data such as financial and grant information and researchers profiles</i>	<i>Nuffield Department of Clinical Medicine (NDM)</i>
---	--

NDM staff reported a degree of satisfaction with the system. One respondent stated that “everybody trusts Pinfox, they know it works.” The same respondent added that every

department should have a similar system to collect research related information because very little could be done to keep track and disseminate activities on the web if there are no data to work on.

Pinfox was also mentioned by other interviewees who do not work in NDM: Social Sciences, Cardiovascular Medicine and International strategy. They all concurred in that a system providing similar functions would be useful and their departments should have a similar one.

DIS a system to manage staff, teaching, allocation of resources and research information ***Social Sciences division***

The Social Sciences division is currently designing a divisional information system which it will offer to all its departments. Within the research activity element of that system, members of the division would like to have similar functions as the ones offered by Pinfox.

A meeting was held between members of BRIL and members of the Social Sciences division to discuss overlaps and dependencies between the two systems. It was agreed that the two systems would be complementary and that BRIL could use information provided by the Social Sciences division. The Division can in turn use this together with additional and enhanced data supplied by BRIL to provide other services (e.g. accessing lists of publications by author or topic from ORA and BRIL.)

Oracle: finances ***Central University***
Resolve: costing
OpenDoor: Human Resources

These systems are used by administrative staff to control the resources used in research activities, e.g., staff and funding. Most of the information they contain is considered confidential. However in terms of publicly available data, they contain top level data (such as titles of projects and names of researchers.)

Respondents explained that these systems were useful but they would like them to be connected.

Matrix: grant application management and workflow ***Research Services***

Matrix contains all applications (successful and failed) for external funding. It is a requirement for every researcher or department who is submitting a bid to send it to Research Services. Research Services provide support to researchers when writing their applications. It has formal responsibility for reviewing and approving applications for funding and for accepting successful awards, in respect to those applications. Research services send monthly reports, generated by matrix, to all academic departments with the status of their research activities.

The BRIL team understands from interviews that Matrix contains data such as names of projects and researchers, project start/end dates and financial information. However, it does not have the details of the actual data fields held in Matrix.

Matrix does not contain information about research activities which do not use external funding.

ORA: Oxford University Research Archive, a repository for research outputs ***Oxford University Library Services (OULS)***

Conceptually ORA¹² and BRIL are connected in two ways. BRIL will provide the contextual data for research outputs held in ORA in order to populate item records and will also use ORA to

¹² OULS is developing the Oxford DAMS (Digital Asset Management System) as the foundation for the storage and preservation of Oxford's digital library collections. ORA, DataBank (<http://databank.ouls.ox.ac.uk>) and the RIL are built on the DAMS. All these services use shared architecture and similar semantic web technologies. BRIL grew out of a need to obtain contextual data

obtain information about researchers and their outputs. ORA is a complementary system to Symplectic (see below).

A few interviewees mentioned ORA, all of them administrators. All of them agreed in that ORA is potentially a useful tool for helping them maintain researchers' profiles. However some of them stated that uploading information into ORA was too large a task for them and not their responsibility.

Symplectic: Research data Management & R.E.F. reporting Central University

Only a few administrative staff mentioned Symplectic, and referred to it as the REF reporting system.

Interviewees involved with the development of the Social Sciences system were more aware of the characteristics of Symplectic. They are designing ways to take advantage of the information that will be held in Symplectic and at the same time did not want to duplicate any of its functionality.

The BRII team is investigating how the two initiatives will work together as complementary systems.

Database of Experts

Information Press Office

The Information Press Office has a database containing the names, contact details and areas of expertise of academics who are willing to be contacted by the press. They use this database to identify people willing to be interviewed about issues or events which are in the news. As there is no other source like this in the University, they receive requests for information from other areas of the University and for purposes that are not news related.

Business Intelligence – Data warehouse

Business Processes Services

The planning of this data warehouse is still at an early stage. The aim is to collect mainly finance related information and then develop tools which explore that information for strategic purposes. For example discovering patterns across time. There is a slight overlap between this system and BRII as both aim to include information about research funding bodies. The Business Intelligence – Data warehouse system and the BRII project share the same beliefs about reuse of data which provides increased efficiency and consistency of data.

e-Humanities database

Oxford e-Research Centre (OeRC)

A project to create a database of people, projects and events in the Humanities and which has strong synergies with BRII. BRII will provide digital objects for entities such as people and projects that form part of this database and BRII project staff are working with e-Humanities staff on areas such as controlled vocabularies.

Other systems

There are other systems within the University environment which appear to be relevant to the BRII project. The degree of validity and influence of those systems will be studied after having first hand information from the people responsible.

3.3. An overview of research activities

This section is not an exhaustive examination of research activities in Oxford. From the effort made so far some data have been collected which shows a wide variety of interpretations. However these same data suggest that more information is needed to obtain a

for research outputs held in ORA in order to populate item records. The resulting research information infrastructure will use ORA to obtain information about researchers and their outputs, and provides an additional beneficial means for wider use of that data across the University.

comprehensive account in this area. As the BRII project aims at building a pilot the following overview explains the data samples used.

In this report we use the term research activity to represent the processes of inquiry or examination which seek to extend knowledge. Traditionally, research activities tended to be bounded by research disciplines (e.g., chemistry, management, and genetics.) Nowadays research activities have more blurred boundaries as they intersect with other disciplines. These intersections are commonly known as multidisciplinary research. They usually occur in the form of collaborations between experts in different disciplines¹³. Once intersections are established they become new research fields in their own right. Similarly, once areas become less attractive or less relevant they fade from the University research spectrum. However evidence is left of them taking place. This evidence develops into an important and valid (historic and reference) source of information for current endeavours.

Research activities define and are defined by the nature of research cultures. Traditional research fields which have existed for a long period of time tend to be slow moving and their outcomes are cited by subsequent researchers for a long time. An example of this is philosophy. New research fields, particularly in the area of science, are fast moving. They adapt quicker to changes in their environment as do their research activities.

Similarly the information representing research activities changes and adapts according to the nature of their research cultures. A good example of this is project websites. Data from interviews show that within Medical Sciences division, new websites are created and information is uploaded and updated on a regular basis. The amount of information needed to be managed is so big that dedicated staff are appointed for those tasks.

A research activity can be defined in many ways depending on the research culture. Some activities are clearly defined, for example where (external) funding is involved, staff tend to use the terms projects or grants¹⁴. They have a title, a start and an end date; they have a budget, resources and staff appointed for their duration. Information is generated about them, in internal systems and in websites. Additionally outputs such as publications also constitute references to those activities.

¹³ Researchers carrying out multidisciplinary activities have sometimes more than one affiliation depending on the departments where their activities are based.

¹⁴ Another variation of a research project is the clinical trials which are research activities which test or develop on previous projects and which have direct implications on practice.

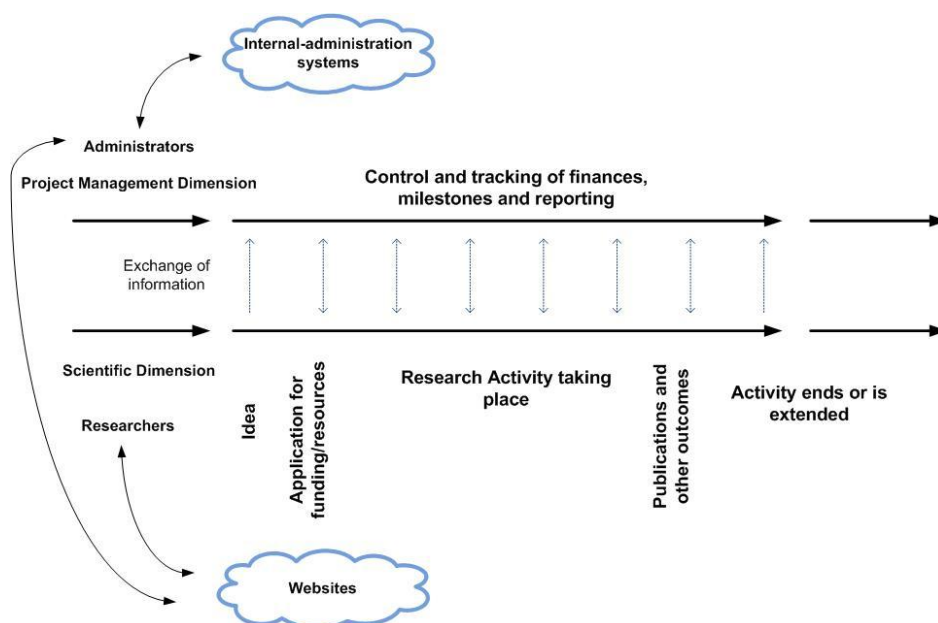


Figure 3 Simplified life cycle of a research activity

Academics and administrators work together during the life span of the research activities. This is more evident where these have a clearly defined existence (e.g. research projects.) Figure 3 shows a view of a typical research project with a start and an end. The research project is seen as having two dimensions: scientific and project management. The diagram illustrates that although they work together, researchers and administrative staff possess very different perspectives on the activities.

Other activities like writing books (in the humanities for example) are not clearly defined. Some of them are not sponsored by purposed funding. Researchers believe these activities are included in their normal workload and that they are already being paid for them. Little evidence apart from the actual outputs (e.g. books or papers) is created about the activity.

Between the clearly and not so clearly defined activities there is a huge spectrum of shapes and colours of research activities. As mentioned above, this document will not report on all of them but some examples will be highlighted.

There are some research activities which comprise minor activities with assigned funding, i.e., they exist as projects in internal systems. When asked, some principal investigators would like to refer to the umbrella activity rather to the smaller, funded activities. Their project and personal websites reflect that view. The opposite is also true. Some funded projects will be split in smaller tasks (not reflected in information about grants) to facilitate the execution of research. Some researchers emphasise the smaller tasks rather than the overall project. In some instances, these smaller activities are part of more than one initiative, i.e., their outcomes will be used in many other activities. Again project and personal websites reflect these views.

Information in divisional or departmental websites tends to be more standardised than personal or research group sites, perhaps because administrative staff control them and because they need consistency in the presentation of the varieties of research carried out.

As a consequence information about research activities is presented in different ways on project, personal and departmental websites, depending on the perspective of the person who manages them. These perspectives are all also influenced by the expected audience of the website (see Audience, under Factors of Influence in the previous category.)

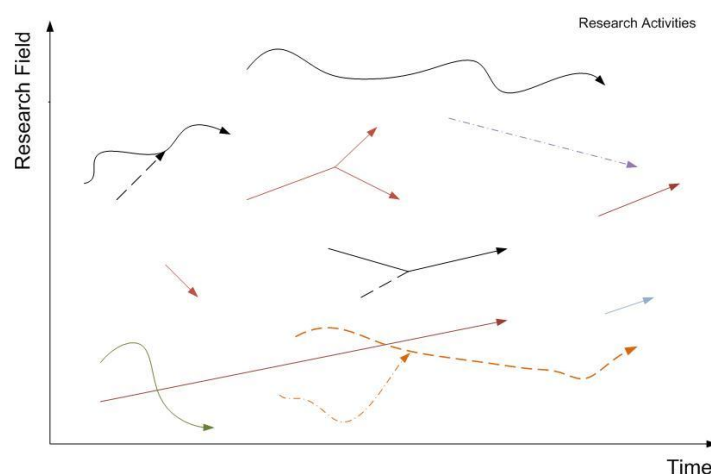


Figure 4

Figure 4 depicts the nature of research activities as understood from the data gathered for this study. Seen from the eye of an outsider research activities could be seen as a chaotic universe. Each line represents a different research area during its life, and illustrates that they can merge or split. They can take a long time or be very short. Information about these activities can be found in traditional media such as publications or presented in academic conferences. However not all of them have a web presence (continuous and dotted lines.) Overlaps might exist but researchers are not always aware of them, particularly when the activities take place in different research fields. Finding information which reflects an overview like the one shown in the picture is currently an impossible task in this University.

In addition to information about the nature of University research activities, data were collected which reflect the way stakeholders deal with those activities. The following section presents an overview of some of the strategies that are employed to organise and carry out activities.

Ways for organising and carrying out research activities

Research is a highly competitive business. To be successful a researcher has to demonstrate rigour and originality. Researchers have to be aware of and produce cutting edge research in their field. Collaborations are currently a popular and sometimes necessary route to developing revolutionary research. Also researchers have to keep up to date with what others are doing and seize the momentum as soon as they see an opportunity. The amount of information generated by a fast moving research field may require the help of support staff to manage it. This section discusses the strategies most frequently mentioned by interviewees that researchers use to carry out their research.

- a. **Collaborations** can take many shapes. They might be discipline bounded involving the participation of researchers in the same field, from other, possibly international, institutions; or they can be multidisciplinary involving the participation of (possibly

international) researcher experts in different fields. The purpose of research collaborations, as emphasised by the interviewees, is to access knowledge and expertise which they lack. Some departments like the National Perinatal Epidemiology Unit (NPEU) do not collaborate within the University, but collaborate externally with every maternity and neo-natal unit in the UK and internationally. Other departments like Phonetics and Linguistics have ongoing collaborations with other areas within the University, in their case with Babylab in the department of Experimental Psychology.

Collaborations can be formal, as in funded research projects involving multiple areas or institutions or informal as in networking activities or book writing. They can have a set duration (start and end date) or they may exist in the background and produce intermittent outputs.

In terms of building collaborations, interviewed researchers stated that they develop networks of contacts with peers whose work is relevant and interesting. Collaborations usually start from these networks as they are known and trusted sources of expertise. However when new expertise is required, researchers tend to expand their networks by scrutinising peer reviewed publications. They do this by focusing on particular keywords which represent the researchers' interests and expertise. Through peer reviewed publications researchers know who and where an expert is based, as well as being able to gain an in-depth understanding of their previous work. These publications reflect research activities which have finished and whose outcomes have been proven and validated.

On the other hand, finding information about collaborations in progress or which have no published outputs yet may be a harder task. Although some researchers publish information about work in progress in traditional media (e.g., journal articles), most do not do this until they feel satisfied with their achievements. However, information about research in progress can be found online, i.e., project websites, or sponsors' websites, if the activity is externally funded. Other sources of information about work in progress are departmental databases (e.g., spreadsheets containing administrative information) and the Research Services system. Both provide only a limited amount of information restricted to the department or area which owns the departmental database or the department which request that information (from Research Services.)

As researchers are usually interested in one specific research area and familiar with its terminology (interests and keywords), finding potential collaborators within that research area is a natural task for them. However, for non-experts and in particular, those who need to be aware of a number of activities, looking for collaborations and collaborators can be a tedious task (consider the chaotic research universe shown in figure 4.) In order to write about research collaborations in certain areas, one strategist said they would have to assemble teams whose main task was to search for research publications in such areas and look up the names and affiliations of authors by opening every file one by one.

- b. Research themes** are one way of organising research activities which cross disciplines, but which have one strong commonality between them. Use of research themes is a strategy which is popular in the Medical Sciences Division. One interviewee from that Division said:

“Themes have no money. They are different from institutes and centres. Themes are purely a way of helping to sell their research in a way, showing where their strengths are in this University.”

Research themes comprise groups of researchers, whose work overlap in one overarching topic. These researchers belong to several research disciplines from within the University. Some departments define the research carried out by their researchers by the research themes their staff are involved in. In one particular instance, when asked about their research centres, one departmental administrator stated that researchers in their department belong to four major themes within medical sciences. He explained that as they are not self-sufficient (i.e., they do not possess all the expertise and resources needed) they have to work with people involved with research under the same broad theme both within and outside their division. He added that interactions between those working within the different themes are also necessary because of the diversity of the research topics in their department.

When asked about their affiliations, some researchers mentioned the research theme(s) under which they work instead of their departments. Information in some of their online profiles reflects this as well.

New research themes emerge in response to funding bodies’ initiatives, which seek to encourage research in areas that are not related to the existing organisational structure of the University.

To be effective, research themes require strong information support in order to disseminate and advertise their activities. This is mostly achieved via theme websites. As most of these sites focus on audiences such as funding bodies, they need to display the expertise of the researchers and progress of the research in a consistent way. Themed websites such as Cardiovascular Science¹⁵ contain researchers’ profiles and publications collated from a variety of sources. For the purposes of the website, these profiles can only include information relevant to the theme, for example, listing the three top publications of the researcher in that particular field. If the researcher has published in other areas, those publications will not be listed in the theme website, but possibly on the departmental website.

- c. Research Strategies:** every academic department has a research strategy which is manifested in various formal or informal ways. Research strategies delineate future activities aimed at improving their reputation and research impact. Some departments have a research committee that oversees research from different points of view. One example is the Department of Physiology, Anatomy and Genetics

¹⁵ www.cardioscience.ox.ac.uk

(DPAG) research committee which looks at recruitment and research skills with a view to how they can change the way science is done in the department. The DPAG research committee also looks at the development or acquisition of new equipment. Other departments such as Nuffield Department of Clinical Science (NDM) have a grant advisor who collects and looks for patterns in their research projects and for sponsors who might fund them. These patterns reveal research strengths, but also weaknesses which can then be tackled by providing training and support to research staff on, for example, writing applications for funding.

There are common needs in both of the above examples: the need to access large amounts of information and the means to manipulate that information in different ways, e.g., by generating statistics. This information is not always easily accessible. For example, financial information hosted in University systems which are not connected to internal systems, information from funding bodies, which is public but which is dispersed in every sponsor's website, information about successful and failed applications which is treated as sensitive by most academics.

- d. **Funding bodies** or sponsors are popular sources of money in some areas of the University such as the medical sciences. External funding is sought to sponsor activities of all kinds: individual, group, interdisciplinary, national or international collaborations, from six months to more than five years.

Funding bodies can impose conditions on researchers and organisations whose research they fund. Conditions vary from funder to funder. From the application process onwards, researchers and supporting staff have to comply with the formats and processes of each sponsor. Funders also influence the way research activities are designed. In general, funders sponsor projects with a defined start and end, and with specific set of activities, resources and outcomes. Researchers and administrative staff have to report on progress, finances and final results. In some instances the information generated is made publicly available on sponsors' websites.

Some sponsors focus on specific areas and are therefore well known in the corresponding University departments. Accordingly, such departments have their preferred or most targeted sponsors. For example one third of NPEU funding comes from the Department of Health.

Data from the interviews suggest that in some cases, funders set out the directions in which they wish research in a particular area is to proceed. Departments wishing to obtain funds from these sponsors need to be reactive (as opposed to proactive) and adapt by inventing new research avenues which fulfil the preferred topics of the funders. One way of doing this is by creating research themes.

Funders' websites are good sources of information on the research activities they fund. Administrative support staff and strategists use them to collect information on current activities or to look for new calls for research funding.

3.4. *Current issues on data*

This last category is complementary to the others. It focuses on the practicalities of generating and using research activity data, mostly from the point of view of administrative staff, but not limited to them. There are six issues mentioned by respondents:

- a. **Responsibility** for producing and maintaining research activity data varies from department to department. For internal University information systems the responsibilities are clearly defined, for example for the financial officer, departmental administrators or those in similar positions that are responsible for updating information in Oracle. Responsibilities for updating publicly available or online information are divided. Information on research activities such as researchers' profiles or project overviews are generally produced by the researchers themselves in various formats. However due to their time constraints, they are not always in a position to upload and maintain it on the web. Research administrators or research facilitators help researchers by setting up systems which facilitate the maintenance of websites. Different examples of systems were collected during the interviews. The three most representative are shown here:
 - A request is sent to all researchers to send updated research profiles and project overviews to the departmental administrator. Researchers produce that information during their limited free time. For example they write on a piece of paper in between surgeries. All information is then collected, reformatted and uploaded onto a website.
 - A member of the administrative staff is assigned the task of updating research profiles on a themed website. He searches for information on project and researchers' personal websites. He then gathers and organises that information and asks researchers whether that information is accurate. Researchers either agree or send new updated information.
 - All research active staff in a department have access to a computer system through which they can update information directly on their online profile. There is a member of administrative staff in charge of verifying that information periodically, however responsibility for updating information falls on the researchers.
- b. **Coverage of data** refers to the proportion of data from the totality of all University research activity data that will be accessible through the RII. Data from interviews shows that to be trusted and used, a research information infrastructure should contain an important amount of information, which saves users the task to search in other sources of data and which assures them they are finding everything that is relevant. While talking about access to research publications, one principal investigator stated that "one has to evaluate where the biggest proportion of important publications is." He added that if fewer than 95% of all important publications are covered by a particular source there would be a need to try and identify where the missing 5% is. However if the coverage is better, there will not be a need for something new and academics can keep doing what they are already doing (e.g., accessing online bibliographic indexing systems).

- c. **Duplication of data** is an issue that concerns some administrative staff particularly those who need to search for or manage large amounts of information. Duplication of data can be seen from two perspectives. The first is duplicating data by re-keying information from one system to another, something that administrative staff would rather avoid but need to do in order to do their job. The second is information available online which may be duplicated. Sometimes it is not clear from online data whether two records in two sources correspond to the same person or not. Some interviewees stated that they needed a way to disambiguate or normalise researchers' profiles, as some would have their names written in several ways (e.g., Prof J. Smith, Dr John Smith, John R Smith)
- d. **Formats:** most administrative staff interviewed said that they keep departmental information in spreadsheets because that gives them control over the data. A lot of re-keying and duplication is needed to maintain these spreadsheets.
- e. **Research activity objects** would be the equivalent to research activity fields in a database. Respondents were not asked direct questions as to which fields they would require to do their job. The words fields or objects were not mentioned to create this list. This list was made up from comments or answers to other questions. They refer to the actual data that respondents use (from internal systems or the web) or to "things" they search for, do or know about, and which are important for their work.
- Researchers: titles, names, departments, esteem indicators, activities, conferences
 - Publications and citations
 - Projects: title, start date, end date, overview, scientific information, collaborators
 - Contracts, spin-off companies, links with industry
 - Resources: by projects, by department, by researchers
 - Project outputs, software, methodologies, tools
 - Images
 - Collaborations: institutions (academic, industry), departments and people involved
 - Keywords, categories, subcategories: researchers, projects,
 - Methodologies
 - Funding bodies
 - Grant applications
 - Number of staff submitted to the RAE

The following is a subset of the above and are considered sensitive by respondents

- Project dependencies
- Progress against milestones
- Research data
- Failed and fine details of grant applications
- Budget
- Financial information in general

Some strategists and disseminators mentioned the following kinds of information they need and currently generate manually:

- Analyses and statistics
 - Participation in schemes or programmes offered by funding agencies: rate of success
 - Brochures to publicise research
 - Country profiles
- f. **Quality of data** refers to how accurate and up to date data are. This was not referred to as an issue by most interviewees. Some stated that they were happy as long as they could find enough information about research activities that could lead them to the original sources (e.g., a journal article). Obtaining the original sources and contacting the lead investigators is sometimes necessary to double check the validity of data. Types of information referred to in this category are researchers' profiles, lists of publications or project overviews. One example of checking the quality of data was given by the Press Office. They regularly collect information about research activities from varied sources, including online sources, in their search for interesting stories. When they find something appropriate they always contact the lead researcher to obtain first hand information. The Press Office stated that this can be a time consuming task as it has to be done on a one to one basis. This category contrasts with the "coverage of data" category (letter b above) which refers to academic publications. Academics prefer sources where they know (nearly) all of the most relevant publications are available (indexed or full text). This helps to reassure them that they are not missing anything important to their research.

4. Conclusions

The findings of this study (presented in the previous section) are extensive and highly informative. Careful selection ensured that every academic division, at least one college and one Central University unit were represented. These findings show how varied and disparate research activities and research related staff are at the University. They provide the BRIL project with a clearer picture of the context, uses and users of the new Research Information Infrastructure despite the fact that only a small proportion of all academic and administrative units participated in the interviews.

This last section highlights the most important issues which emerged from the findings. These will be presented in three parts. The first part presents aspects of the findings which are key to understanding the uses of the RII. The second presents some implications for development. Some ideas for web services are presented which are based on the issues previously mentioned. The final part summarises the user needs emerged from this study. This list of needs confirms the necessity for an infrastructure such as the RII to aid academic and administrative processes related to research activities as well as to increase the research visibility of the University and individual academic units.

4.1. Insights and outcomes of the analysis

These can be classified in:

BRIL stakeholders and their uses

BRIL stakeholders are everyone whose work is directly or indirectly related to research activities or who are interested in them; and who have a need to access or display research related information. The findings of this study show that stakeholders can be very different from each other. Aspects such as research field, type of job (academic, non-academic) and scope are crucial at defining their beliefs, behaviour and needs. Because of this, this study does not describe them individually (i.e., in a discrete way) but uses *spectrums of aspects*. Characteristics of stakeholders are explained at the extremes of the spectrums but it should be noted that most stakeholders are located between these extremes, presenting characteristics from either side of the spectrum at different levels.

The uses of the RII will be as diverse as its Stakeholders. Therefore one implication for the RII is the need to account for *variety*. Stakeholders' characteristics and behaviour are not clear cut, but vary along more than one spectrum. Similarly the RII should offer a variety of spectrums of uses to account for different perceptions of the same things and multiple ways for doing similar tasks. For example, research activities in some departments are described under themes, in others under institutes or the departments themselves. Another example is outcomes, which for some are connected to the people who produced them and for others are connected to the project where they originated. Searching for projects or outputs could be done by browsing through academic papers or sponsors' websites. This will vary according to what stakeholders think is more relevant and trusted. Having data in the RII which originate from these and other sources will account for these varied perspectives and facilitate stakeholders work.

Another factor to consider is the depth of content that stakeholders need from data. Depending on their roles, some users need overviews written in *easy to read* language, whilst others need detailed information written in scientific language. The implication for the RII is on the provision, where possible, of multiple views of the same research activity objects.

Research activity data

Aspects of research activity data that should be highlighted are heterogeneity and transformation. Sources of data will be heterogeneous because they will represent an heterogeneous universe of research activities. The University comprises an extensive array of research areas as presented in section 3 of this report. Each area has its own culture, its own people and its own way of seeing things. Therefore data harvested by the RII will be in no way consistent or compatible.

Transformation represents changes of data in time. Research activity data representing research activities, people and other resources will reflect the transformation of those in time. Information hosted in the RII will be likely to change not only in quantity but in form and content. So therefore it is likely that new attributes describing research activities will emerge and others will disappear. Implications for the RII are that the ways for displaying information

should not be bounded by current structures but should be open and allow variation. The RII and its services should evolve together with the University.

Similarly, another implication is that the means for searching specific information, whatever they represent, should evolve according to changes in the information held in the RII. For instance, when new categories of objects are added into the RII, means for searching by those new categories should be made available in its web services.

4.2. *Implications for development*

Although this is not a development specification document its outcomes will inform more detailed and specialised guidelines for development. From the discussion above some initial thoughts can be shared.

The RII should be seen as a complex network of research activity data objects complex because of the size of the University and the amount of research taking place in it. It is a network because the RII will enable links to be created between different sources where commonalities are found. These links are essential for stakeholders at all levels e.g., finding (potential) collaborations across institutions, departments or individuals.

One proposed output of the BRII project will be the Blue Pages, a directory of expertise. The Blue Pages will allow users to discover connections between objects which were previously unknown. Objects are those bits of information that represent parts or aspects of researchers or research activities and in which BRII stakeholders are interested in finding and/or sharing. As mentioned before, stakeholders belonging to different areas and positions will have different perceptions of those objects therefore different views of the same information will exist within the RII without requiring duplication of information.

Considering the above, the Blue Pages, should provide the means to access information from different angles (find activities through publications or through project websites? scientific explanation or an easy to read overview? current or past activities, national or international collaborations?) It should also provide information at different depths. For example, top level: a list of names of researchers and expertise in one research field, and bottom level: a detailed researcher's biography containing information about publications, methodologies used, projects and collaborators. Users should also be able to zoom in and zoom out information; i.e., go from top to bottom level and vice versa. For example for the pathology keyword, the Blue Pages could provide a list of research sub fields, or research projects under pathology. If the user is not interested in details, that list would give him an overview of the sort of research done there. If he is interested in detail, he can either search for specific names, or click on the links resulting from your Pathology search.

The findings of this study also reveal that websites containing information about research are created across the different academic units in the University. The creation and maintenance of these websites involve an important amount of time and resources to create, collect and/or reuse existing data. This is particularly evident in academic units with (or which want to have) strong web presence. Taking this into consideration, another implication for the RII is to facilitate this work by providing means to access and reuse research activity data in online sources. The BRII project will create a web service expected to be a themed website,

containing researchers' profiles, publications and research activities which are conducted within the field(s) of that particular theme.

4.3. User needs - summary

The findings of this study provide tangible evidence of:

- The need for an organised and easy to access store of information about research in the University of Oxford which at the same time represents its diversity. This will be particularly useful for disseminators and strategists as well as people outside the University such as potential research students.
- The need, particularly from researchers, for trusted and comprehensive sources of information which are relevant for every research field in the University.
- The need to search for information about people, activities and resources from across the University (e.g., find an expert in a particular field in Oxford) at administrative and academic levels across departments and disciplines.
- The need to discover and gather information from different sources (e.g., to create country profiles.)
- The need to reuse information for different purposes. (E.g., tracking activities and reporting.)
- The need to improve research visibility by creating websites or other sources which provide overviews of research activity (e.g., themed websites.)
- The need for tools to access information about research from a variety of perspectives and levels of aggregation.
- The need to reduce inefficiencies (e.g., time spent searching for and managing information, duplication of data and re-keying.)

From this list it is evident that an infrastructure with the characteristics of the RII¹⁶ will fit perfectly within the academic and research processes taking place in the University of Oxford.

¹⁶ A brief explanation of the BRII project is given in section 1.2 of this report. Please refer to <http://brii.ouls.ox.ac.uk> for functional and technical details of the RII.

5. Appendices

5.1. Appendix 1: List of interviewees

Area	Role of interviewee	
Medical Sciences Division		
Nuffield Department of Anaesthetics	1. Departmental Administrator	Liz Barnes-Moss E Liz.Barnes-Moss@nda.ox.ac.uk T 01865 231511
Department of Cardiovascular Medicine, Cardiovascular Science theme, Centre for Research Excellence (CRE)	2. Research Facilitator	Dr Ruth McCaffrey E Ruth.McCaffrey@cardiov.ox.ac.uk T 01865 234649
DPAG Department of Physiology, Anatomy and Genetics	3. Departmental Administrator 4. Associate Head of Department/Chair of the research strategy committee	Paul Warren E Paul.Warren@dpag.ox.ac.uk Dr Clive Wilson E Clive.Wilson@dpag.ox.ac.uk
Experimental Psychology – Babylab	5. Principal Investigator 6. Technical Officer	Prof Kim Plunkett E Kim.Plunkett@psy.ox.ac.uk Adam Roberts E adam.roberts@psy.ox.ac.uk
NDM Nuffield Department of Clinical Medicine	7. Departmental Administrator 8. IT Officer 9. Research Administrator/Grant Advisor 10.Principal Investigator (OCDEM) 11.Research Administrator (OCDEM) 12.IT Developer Officer (Pinfox system)	Darren Nash E Darren@well.ox.ac.uk Karen Valentine E Karen.Valentine@ndm.ox.ac.uk Dr Liesl Osman E Liesl.Osman@ndm.ox.ac.uk Prof Mark McCarthy E Mark.Mccarthy@drl.ox.ac.uk Edward Gibbs E Edward.Gibbs@ndm.ox.ac.uk Dr Claire-Lise Escher Kessler E Ckessler@well.ox.ac.uk
NPEU National Perinatal Epidemiology Unit	13.Director	Prof Peter Brocklehurst E Peter.Brocklehurst@npeu.ox.ac.uk
Dunn School of Pathology	14.Principal Investigator	Prof Simon Hunt E Simon.Hunt@keble.ox.ac.uk

Mathematical, Physical & Life Sciences Division

Computing Laboratory — Computer Science department	15. Project Manager/Research Facilitator	Sharon Lloyd E Sharon.Lloyd@comlab.ox.ac.uk
	16. Webmaster	James Welch E James.Welch@comlab.ox.ac.uk
	17. Principal Investigator and fellow of St Anne's college	Prof Peter Jeavons E Peter.Jeavons@comlab.ox.ac.uk
	18. Project Administrator	Janet Sadler E Janet.Sadler@comlab.ox.ac.uk

Social Sciences Division

Social Sciences Division – IT strategy and DIS new information system	19. Projects Officer	Dr Bridget Taylor E Bridget.Taylor@socsci.ox.ac.uk T 01865 614874
	20. Research Facilitator	Dr Samantha Sneddon E Sam.Sneddon@socsci.ox.ac.uk
Social Sciences Library	21. Librarian	Dr Bella Kogan E Bella.Kogan@ssl.ox.ac.uk

Humanities Division

Faculty of Linguistics	22. Head of Faculty	Prof Aditi Lahiri E Aditi.Lahiri@ling-phil.ox.ac.uk
	23. Post-doctoral researcher	Dr Allison Wetterlin E Allison.Wetterlin@ling-phil.ox.ac.uk
Faculty of Philosophy	24. Administrator and Secretary to the Faculty Board	Thomas Moore E Tom.Moore@philosophy.ox.ac.uk
	25. Assistant Administrator	Catriona Hopton E Catriona.Hopton@philosophy.ox.ac.uk

University Administration and Services

Research Services	26. Associate Director	Dr Stephen Conway E Stephen.Conway@admin.ox.ac.uk
Business Services and Projects (BSP)	27. Business Intelligence Programme Manager	Nick Lupis E Nick.Lupis@admin.ox.ac.uk
Publics Affairs Directorate	28. Head of Internal Communications	Dr Susannah Wintersgill E Susannah.Wintersgill@admin.ox.ac.uk
Information Press Office	29. Head of Office	Ruth Collier E Ruth.Collier@admin.ox.ac.uk
	30. Deputy Head of Office	Clare Woodcock E Clare.Woodcock@admin.ox.ac.uk

Colleges

Keble	(also PI in Dunn School of Pathology)	Prof Simon Hunt
-------	--	-----------------

St Anne's	(also PI in the Computer Laboratory)	Prof Peter Jeavons
-----------	--------------------------------------	--------------------

Other areas

Student Recruitment	31. Senior Consultant	Dr Catherine Paxton E Catherine.Paxton@admin.ox.ac.uk
---------------------	-----------------------	---

International Strategy (International Support Services)	32. Director	Dr Heather Bell E Heather.Bell@admin.ox.ac.uk
--	--------------	---

The following were contacted by Sally Rumsey BRIL Project Manager

Oxford e-Research Centre (OeRC)	33. Arts and Humanities Consultant	Professor David Robey E David.Robey@oerc.ox.ac.uk
------------------------------------	---------------------------------------	---

Balliol College	34. Senior tutor	Dr. Nicola Trott E Nicola.Trott@balliol.ox.ac.uk
-----------------	------------------	--

NPEU	35. Webmaster & Design Co-ordinator	Andrew Kirk E Andy.Kirk@npeu.ox.ac.uk
	36. Senior Trials Programmer	Andy King E Andy.King@npeu.ox.ac.uk

5.2. Appendix 2: List of data contributors

The following is a list of departments whose research activity data may be harvested by the RII. The BRIL team will make every possible effort to include as many of these sources before completion of the project. Some staff members of these departments were interviewed for this study. However, other contributors were approached through other networking activities. The BRIL Project is grateful to these departments for their offer to participate in the project in this way.

- Computing Laboratory - Computer Science department: access to data in their object oriented database system has been granted by Sharon Lloyd, Departmental Administrator and James Welch, Web Master.
- Nuffield Department of Anaesthetics: access to data in their website has been granted by Liz Barnes-Moss, Departmental Administrator
- Nuffield Department of Clinical Medicine: access to publicly available data in the NDM Pinfox system was granted by Darren Nash, Departmental Administrator.
- Department of Cardiovascular Medicine, Centre of Research Excellence and the Cardiovascular Science theme: access to data in their websites has been granted by Ruth McCaffrey, research facilitator, after consultation with Hugh Watkins Head of Department and Centre.
- Babylab - Experimental Psychology: access to data in their website has been granted by Adam Roberts, Technical Officer, after consultation with departmental staff. (Note: the website is currently under development.)

- Department of Physiology, Anatomy and Genetics: access to data in their website has been granted by the Head of Department and academics overseeing the DPAG website.
- National Perinatal Epidemiology Unit: access to data in their website has been granted by Peter Brocklehurst, Director of NPEU.
- Social Sciences Division: although the system is in the early stages of development an agreement was made to share data with BRIL at the earliest opportunity. (Bridget Taylor, Projects Officer.)
- Faculty of Philosophy: information gathered for RAE 2008 was offered by Tom Moore, Departmental Administrator.
- Balliol College: the College Governing body granted permission to use the details of the College Fellows as displayed on the College website (Dr. Nicola Trott, Senior Tutor.)
- E-Humanities project: Prof David Robey is developing an e-Humanities website/database for which BRIL will provide the backend and store details of people and projects.
- A number of research councils and funding bodies, via their freely available online sources. (e.g., EPSRC, Cancer Research UK, Wellcome Trust.)

6. References

- Boland, R. J. J., (1985), 'Phenomenology: a preferred approach to research on information systems', in Mumford, E., Hirschheim, R. J., Fitzgerald, G. and Wood-Harper, A. T. (eds.), *Research methods in information systems : proceedings of the IFIP WG 8.2 Colloquium ... 1984*, Amsterdam, North-Holland.
- Hammersley, M., (1992), *What is wrong with ethnography?*, London, Routledge.
- May, T., (1997), *Social research : issues, methods and process*, Buckingham, Open University Press.
- Priest, H., Roberts, P. and Woods, L., (2002), 'An overview of three different approaches to the interpretation of qualitative data. Part 1: theoretical issues', *Nurse Researcher*, 10, no.1: 30-41
- Walsham, G., (1995), 'Interpretive case studies in IS research: nature and method', *European Journal of Information Systems*, 4, no.2
- Williams, M. and May, T., (1996), *Introduction to the philosophy of social research*, London, UCL Press.