Citation Capture: Enhancing understanding of the use of unique and distinct collections within academic research and the research outputs produced as a result.

Phase 1: Standardisation of citation

Invitation to Tender
Title: **Citation Capture: Standardisation of citation**
Deadline for receipt of tender proposals: 5pm Monday 15th January 2018
Return to: Melanie Cheung, RLUK, who will manage and administer this contract:
Melanie.Cheung@rluk.ac.uk

1. INTRODUCTION

This invitation to tender specifies the requirements for the completion of a research and guidance piece examining the citation of repositories holding Unique and Distinctive Collections (UDCs) within academic publications.

The following further details are outlined in this document:

- The service to be provided
- The information to be provided by tenderers
- The terms of business relating to the award of any contract
- Anticipated timetable for the programme

2. BACKGROUND

2.1 Organisation

Research Libraries UK (RLUK): represents 37 of the leading and most significant research libraries in the UK and Ireland. We aim to optimise the contribution that research libraries and collections make to the economic, technological and cultural success of the UK and Ireland.

We are achieving this by working with our members, external institutions and our partners in the community, both nationally and internationally. We are investing in and developing strategic projects, reports and innovations that are fomenting much greater engagement with and services for the research community we serve. These are exciting times for the research sector. We believe there is so much more that we can do together to shape collections and services in support of academic excellence.

For the avoidance of doubt, RLUK is the commissioning body for this piece of work, with whom the successful candidate will enter into contract.

For more information: www.rluk.ac.uk

The National Archives: is the official archive and publisher for the UK government, and for England and Wales. We are the guardians of some of our most iconic national documents, dating back over 1,000 years. Our 21st-century role is to collect and secure the future of the government record, both digital and physical, to preserve it for generations to come, and to make it as accessible and available as possible. We are an accredited archive service, archive sector leader in England, and an Independent Research Organisation (IRO).

The National Archives' business plan, *Archives Inspire*, and *Archives Unlocked* identifies both the archive sector and academia as two of four key audiences. Our approach to the leadership role is a collaborative one, supporting partnership working to encourage a sustainable and innovative archive sector.

For more information: www.nationalarchives.gov.uk

Jisc: Jisc's vision is for the UK to be the most digitally advanced higher and further education and research nation in the world. Its mission is to enable people in higher education, further education and skills in the UK to perform at the forefront of international practice by exploiting fully the possibilities of modern digital empowerment, content and connectivity. To achieve this vision, each day Jisc will continue both to power teaching, learning and research and inspire its digital transformation.

Jisc will achieve the vision in 5 key ways:

1. By continuing to develop our Janet network, the foremost national research and education network in the world;
2. By focusing our products on the areas that make the most difference and offering solutions tailored to the unique challenges of the education and research sectors;
3. By making membership of Jisc meaningful through two-way relationships with individual members and the HE, FE and skills sectors;
4. By continuing to build a strong, financially sustainable, position;
5. By operating to high standards of efficiency and effectiveness.

In relation to libraries, special collections and archives, Jisc will ensure that UK researchers, teachers and learners have excellent access to rich and wide ranging content via sustainable and convenient discovery mechanisms and high quality data. The services and other resources that Jisc provide will deliver benefits to universities and colleges in three ways. Firstly, they will enhance the capability of libraries, archives and special collections to manage their collections more efficiently and effectively. Secondly, they will allow researchers, teachers and learners to more readily find and access a greater variety of materials that will enhance their work and shape new forms of scholarship. And thirdly, they will make a tangible contribution to the overall quality of data that makes its way into the discovery data ecosystem, through the promotion of more standardised data formats and by fostering good data creation practices.

For further information: www.jisc.ac.uk

2.2 Citation capture background (see also appendix 1)

Research Libraries UK, The National Archives, and Jisc are delighted to be working together to commission a significant piece of research into the use and presentation of unique and distinctive collections in academic publications as recorded through citations.

Unique and distinctive collections (or UDCs) is a term used to encompass the very varied collections held by archives, special collections, libraries and museums throughout the UK and further afield. The term can encompass collections including "medieval manuscripts, modern archives, printed books, pamphlets, periodicals, and maps. They (UDCs) increasingly also include non-print formats such as audio, film and digital files as well as print collections which are remarkable as a result of their breadth and comprehensiveness."[1] This is an inclusive term which does not denote one section of the archive or library sectors. UDCs are held throughout the archive, library and wider heritage sectors, whether in county record offices and publically funded archives, university archives and special collections, businesses, and national libraries and archives. They are also held within museums and a broader range of heritage and cultural organisations.

The academic, research, and scholarly communities are major users of UDCs across the UK, and wider world. Over a period of considerable change, academic users have remained at the core of the use and better understanding of UDC material, in multiple disciplines. From historians to heritage scientists, information professionals to computer scientists, academic users feature heavily in search rooms across the country and are at the forefront of the digital use and discovery of UDC material. In recent surveys archivists state that academic users are a potential growth audience and are a significant voice in support of archives.

As a result of this use, every year, tens of thousands of citations are made to UDC holding repositories, whether archives, libraries, or museums, within the footnotes and endnotes of academic publications. These citations frequently appear as an acronym formed of the letters of that institution, and thus these denote the holder of UDC material and where it can be found by other researchers. Increasingly, citations are also appearing to materials held and presented digitally, as DOIs (digital object identifiers) and url links. Producing accurate and consistent citations to UDC repositories and the material they hold, whether contained within analogue or digital formats, is central to good academic practice. Citations underpin trust in academic research and allow other researchers to follow the same research path. They help expose the association and interlinkages between collections, held across multiple repositories, and how these can be associated with one another. Accurate citations encapsulate the transparency of the research process and the professionalism and accountability of the researcher.

Understanding how academics use and cite materials held by UDC repositories will be of great value to the institutions themselves. Academic publications are often one of the most visible of discernible outputs produced using UDC collections, changing the scholarly landscape, and adding to an incredibly diverse and exciting corpus of published research in individual fields. Yet whilst academics remain a major user of UDC collections, UDC repositories frequently struggle to gain oversight of these outputs in which they are cited.

[1] Definition as given in RLUK's UDC report: http://www.rluk.ac.uk/wp-content/uploads/2014/12/RLUK-UDC-Report.pdf, accessed 1st November 2017.

Citation Capture aims to furnish the archivist, special collections librarian, and information professional with a better understanding of how academics cite UDC repositories, the nature and volume of the published outputs produced based on UDC research, and how (and what) UDC material is referenced within these. The continued development of the Research Excellence Framework (REF) places a premium on original and high-visibility research by academics, whereas there is a need for more robust measurements of user engagement and output from within the archives, library, and information sectors.

This first phase of *Citation Capture* will establish current academic practice in the citation of UDC repositories and the possibilities of introducing greater standardisation in citation practice. As leaders in their respective fields, **Research Libraries UK**, **The National Archives**, and **Jisc** are well positioned to undertake this work.

2.3. Components of *Citation Capture*

Citation Capture means the counting and assessment of references to a particular UDC repository in publications. This capture might be of two types:

a) All references in all published media, digital or hard copy. This includes blogs and other social media, as well as non-academic books;

b) Counts of use in specifically academic publications, either papers or monographs.

In practice, the distinction is difficult to sustain using pure metrics. Increasingly, social media are becoming part of the academic landscape, while assessment of non-academic publication (however this might be judged) is both an important tool of dissemination and a marker of impact.

2.4. Benefits of Citation Capture

Creating a better understanding of how academics use UDC collections will bring a number of potential advantages.

Collection development: To help inform archivists, librarians, and information managers regarding the use of particular collections by academics and which are the most cited collections. Such information could help inform strategic decision making regarding collecting habits or cataloguing prioritisation.

Audience intelligence and development: To inform archivists, librarians, and information managers regarding the use of collections by academics, both currently and retrospectively, which can help to plan future events or projects.

Advocacy and impact: Provide valuable metrics regarding the use and visibility of UDC collections in national and international publications.

2.5 Citation Capture: Standardisation of citation project

Citation Capture: Standardisation of citation will be the first in a series of projects which will aim to help address the challenges outlined above.

One of the key challenges faced when looking to survey academic citations to UDC material within publications is the inconsistency with which individual UDC repositories are referenced. An agreed and coherent system of references to UDC repositories does not exist. A single repository may be referenced in multiple ways, which sometimes might be identical to other repositories. Performing a quantitative assessment of citations to a single repository can therefore be challenging due to the absence of an accepted and uniform system of referencing repositories. This inconsistency is not surprising. A 2015 survey, conducted by The National Archives, revealed that only a small percentage of archives proactively gave guidance to their users regarding how the latter should reference their archival repository. This project would seek to undertake more detailed research regarding what guidance is currently being given by UDC repositories to their users regarding citation styles, the means of establishing a nationally consistent system of referencing archival repositories, and to suggest a system of doing so.

The project will:

- Undertake detailed research into what guidance is currently given to users regarding citation style
    - Research whether this is online guidance, written, verbal, or a combination and the correlation between the method of guidance and standardisation of citations across the institution
    - Research whether any institutions offer online widgets to automatically generate citations, If so:
        a. What technology are they currently using?
        b. How easy is it for a user to find online citation generators?
        c. Does the widget create only one citation or multiple types (such as APA & Chicago)?
        d. If multiple citation options:
            i. Do they recommend use of a particular one?
            ii. Do they measure how compliant their users are at following the recommendation even with a citation generator?

- Research into the implementation of a nationally consistent system of referencing UDC repositories including examples from other sectors internationally and both quantitative and qualitative research methodologies
- A robust description of the challenge and potential benefit of establishing a national archival referencing system.
- Development of a plan for the establishment of a nationally consistent system of referencing archival repositories including suggested partners.

3 OBJECTIVES AND DELIVERABLES

3.1 Objectives

The project will seek to test the viability of establishing a national UDC referencing system.

The project will achieve this through quantitative and qualitative research methodologies:

- A detailed picture of current practices around citation style in the wider archive sector;
- Investigation into the possibilities of creating a nationally consistent system of referencing archival repositories and a plan for doing the same.

The contractors will work with Research Libraries UK, The National Archives, and Jisc to select participating repositories.

3.5 Deliverables

- A research piece on current citation style guidance amongst UDC repositories;
- Model citation style guidance which could be rolled out nationally;
- A robust description of the challenge and potential benefit of establishing a national referencing system;
- Research into the implementation of a nationally consistent system of referencing repositories to include examples and a review of any relevant current practices;
- Development of a plan for the establishment of a nationally consistent system of referencing UDC repositories including suggested partners.

4 CONTRACT MANAGEMENT

For the avoidance of doubt, the successful candidate will enter into contract with RLUK, who will manage this project on behalf of the partnership.

5 WORKPLAN/WORK PACKAGES AND MILESTONES

1	Project initiation	**February 2018**
2	Initial research into current citation style guidance	**March 2018**
3	Model citation style guidance	**May 2018**
4	Report on the challenge and potential benefit of establishing a national archival referencing system	**June 2018**

5 Report on the implementation of a nationally consistent system **August 2018**

6 Plan for implementation **September 2018**

6 BUDGET

The maximum available for this contract is £15,000, to include all travel and expenses. This amount is exclusive of VAT.

7 PROCESS FOR SUBMITTING PROPOSALS

6.1 Procurement timetable

Invitation to tender sent out	Monday 11th December 2017
Deadline for proposals	Monday 15th January 2018
Possible interviews	w/c Monday 29th January 2018
Appointment and initiation	w/c Monday 5th February 2018

6.2 Proposal content

The proposals should outline:

- Understanding of the project and deliverables
- Experienced suited to the project subject and proposed methodology
- Names and experience of individuals assigned to the project, to clarify their involvement with each phase or unit of the work
- Costs – including breakdown for costs of each phase or unit of work, day rate of each team member and other costs or expenses

6.3 Proposal submission

Please submit proposals by email to melanie.cheung@rluk.ac.uk

Proposals must be submitted by 5pm on Monday 15th January 2018

6.4 Evaluation of proposals

We will evaluate proposals using these criteria:

CRITERIA	WEIGHTING
Extent to which proposal demonstrates an understanding of the brief	20
Knowledge and experience relevant to the project, including understanding of unique and distinctive collections and the institutions in which they are held	25
Quality of methodology and experience in relation to academic and research practice	25
Proposed team composition and management	20
Cost	10

8. STANDARD TERMS AND CONDITIONS

RLUK, The National Archives, and Jisc reserve the right not to appoint and to achieve the outcomes of the project through other methods.

9. FURTHER INFORMATION

For an informal discussion about the work, please contact:

Dr Matt Greenhall, Head of Academic Engagement, The National Archives,
matt.greenhall@nationalarchives.gsi.gov.uk

Melanie Cheung, Executive Assistant, Research Libraries UK
Melanie.cheung@rluk.ac.uk

Appendix 1: Background paper

Archive Interest

A report by the Higher Education Archives Steering Group, based on 29 survey responses was published in 2015. It showed a fragmented response to the question of how to deal with citation, with most respondents not giving guidance on format to their users; nor did the majority have an established or recognised abbreviation. When it came to collecting published references dfr.jstor was the most popular source, with other sources including Google Scholar. The main reasons for collecting this data were:

a) Internal reporting
b) Monitoring of collection use
c) A performance indicator
d) Communications

Despite the lack of systematisation, the majority were attempting or planning collection; and there was an appetite for knowing more about the subject, with the virtually all 29 wanting to attend a workshop on the subject. This suggests that whatever the motivation of the archives staff – to demonstrate impact for funding applications, reassessing priorities, or defending budgets the usefulness of collecting data is known. The fact that for nearly all respondents university management was a key audience demonstrates this.

What measures can be taken to increase both citation by academics and citation capture itself?

No standard form of citation

At present, staff are not giving a clear steer to academics and students on how to cite archival sources. Despite the presence of the ARCHON code and the example of a number of major archives (including TNA) the spread of good practice has been poor. This is a problem for two reasons:

a) Consistency in citation method across archives would have encouraged the use of this method by students and academics.
b) If a common method had become established then searching for citations in databases such as JSTOR would have been easier.

A standardised system?

A standardised system across the archive sector would assist familiarisation of academics and make them more willing to participate. It would also mean that searching for citations would be easier. Each archive would use a unique identifier. At present there is an ARCHON code (available through Find an Archive), which is a four-digit code; but there is much to be said for a three-letter code. This is more memorable and more easily recognisable by the average user, while covering a potential 17.576 institutions.

In the largely American literature on this subject e.g. Purdue, following an adapted version of one of the common manuals of style is usually recommended. For instance with the *Chicago Manual of Style* this would be:

Item, Date, applicable categories such as box and folder, collection name, repository name, place.

Combining this with a code might be something that takes account of local systems while providing a steer to standardisation.

Letter to Joe Bloggs 8/8/1947 WEA BT 255/692 Joe Bloggs Collection.

Digitised material would have a URL or DOI at the end.

Archives staff need to be rigorous in their advice and there is evidence they are not.

The Steering Group survey has revealed how unsystematic citation advice can be. This may be for many reasons; but one of the most important things staff can do is ensure the information is delivered to as many users as possible. This may be through advice on the Web e.g. TNA's page or it may be through the traditional bookmark or paper slip. But a sustained campaign would be straightforward win.

Measurement and Quantification

It is difficult to collect citations in a straightforward manner

In Google Scholar the use of "National Archives UK" yielded 52,500 citations, of which 18,300 date to 2012 onwards and 4,680 were published this year. While impressive, two problems become apparent:

a) "National Archives UK" is not the only way TNA could be cited; for instance, "United Kingdom National Archives" gives another 539 results and there are other possible combinations.
b) The collected citations need "cleaning". For instance, not all of this year's 4.680 turn out to be relevant to TNA – examples of this include the National Archives of Malta and of course the U.S. Nation Archives. The data needs rigorous checking before it can become usable.

A single source is not enough to capture the possible range.

Two sources that are often mentioned in the literature are

a) **dfr.jstor.org** Data for Research of JSTOR can interrogate its journals using a variety of order sequences, though subscription needed to view identified papers; however content, author, chronology and individual paper citation are still accessible.
b) **Google Scholar** is useful because it will often mark a particular index item as "citation" and will also categorise items chronologically.

Neither is perfect and neither is completely comprehensive. Dfr will only use journals and books stored on JSTOR and thus a number of biases (including regional) are created. Both of them also use only sources that have in some form or other been digitised.
For higher education archives Web of Science should also be added as a third source. Depending on the size and thoroughness of any capture project there are alternatives including the resources of Thomsons Reuter's ISI or Elsevier's Scopus. These are especially important when integrating social media.
Turnitin can be another key source, especially for accessing student usage. One criticism of it may be that the variety of citation practices amongst students is more variegated than within other academic groups; however, it is a significant resource for a specific area of the market.
A Jisc-ProQuest study on The Impact of digital collections that was carried out by the Oxford Internet Institute that sought to determine usage of *Early English Books Online* and *House of Commons Parliamentary Papers* downloaded data from Scopus, JSTOR, Google Scholar and ProQuest Dissertations and Theses. Each bibliometric source required its own method for downloading and exporting into an Access file, often using intermediate software to do so, and manual labour had to be expended on eliminating redundancies, cleaning the data and combining into a single file.
The Toolkit for the Impact of Scholarly Digital Resources at http://microsites.oii.ox.ac.uk/tidsr/welcome

There is no agreed significance for secondary citation
This is the citation of the citing paper by another publication. Unless the number of publications using an archive is low it may be tricky to track such secondary citations, not least because they may not include specific reference to the document/s in question.

What can TNA's role be?
Spreading good practice across the sector
Feedback has already indicated the willingness to accept best practice through workshops, which could also be complemented or substituted with online guides and webinars. This would be an important method for TNA to define what that best practice is.
Develop automated tools
The haphazard nature of much citation capture may be improved in the medium term, but for those with issues of scale and staff capacity getting a good picture of the level of use may be difficult.
An automated program, accessing key resources and using the full set of alternative names for archives, could be written Although this is a speculative suggestion such a program would use a set of alternative names (e.g. The National Archives, TNA, National Archives UK) download the references from four or more databases,, remove redundancies and load into a format of choice e.g. Access. For the foreseeable future, data cleaning will probably be largely manual.