



## Checklist for Digital Outputs Assessment - REF 2021

Author	Arianna Ciula (Deputy Director and Senior Research Software Analyst, King's Digital Lab, King's College London)
Reviewed by	King's College London: James Smithies, Ginestra Ferraro, Miguel Vieira, Sam Callaghan, Stacey Moon, Patrick Ffrench, Simon Tanner School of Advanced Study, University of London: Jane Winters
Date	July-August 2019
Contact(s)	Stacey Moon (REF Coordinator, Faculty of Arts & Humanities), Patrick Ffrench (Vice Dean Research, Faculty of Arts & Humanities)
Department or Client organisation	Faculty of Arts & Humanities
Funder & Scheme	Internal

*The purpose of this document is to provide a guiding checklist for colleagues across the Faculty of Arts and Humanities who are in the process of assessing whether to submit digital research outputs to REF 2021. Its pragmatic scope is to provide a checklist that would support holistic reviews of digital outputs beyond strictly field-specific academic criteria and enhance the submitted products.*

*Evidence from REF 2014 analysis suggests that digital research resources “were given equal recognition to other forms of research output and that research datasets may be identified as notably worthy of recognition” (Tanner 2015).<sup>1</sup>*

Digital artefacts which can be submitted to REF (*Guidance on submissions* 2019: 110-112) encompass a broad spectrum and are not limited to specific types of digital outputs; however, as any 4\* and 3\* research outputs, they are expected to be respectively world-leading or internationally excellent in terms of originality, significance and rigour. The output must also meet the REF definition of research (*Guidance on submissions*, Annex C,

---

<sup>1</sup> Note that digital outputs can feature in a REF submission as research outputs on their own, as research underpinning impact case studies or material supporting impact case study or the narrative of environmental statements. With respect to impact case studies in the REF 2014 submission: “Many arts and humanities departments (whatever the discipline) offered evidence of impact using their digital outputs or digital engagement as part of the mix of evidence offered. At King's, digital humanities collaborations underpinned a large number of impact case studies in a diverse range of subjects and they all seem to have done extremely well in the context of the REF mode of measurement” (Tanner 2015).

2019: 90). This checklist focuses mainly on website content and research databases but includes considerations applicable to other types of digital artefacts.

Digital artefacts which include personal or sensitive data are not considered in this checklist. The assumption is, however, that any digital output submitted to REF has been subjected to relevant ethical clearance for the institution submitting it.

It is important to remember that the content and technical structure of a digital output are interdependent; this checklist focuses on research software engineering best practices with the understanding that any review of digital outputs ought to be combined and integrated with the REF academic assessment criteria.

In addition to the preliminary check list presented here, every research output should be assessed against specific criteria depending on its context, discipline, type and format (e.g. criteria to assess significance, originality and rigour should be tailored and vary for a digital monograph as opposed to a prosopographical database or an annotated georeferenced map).

## CHECK LIST

<b>Credits</b>	<p><i>Are the people who produced the digital output credited appropriately on e.g. the website? Are institutional logos and/or funding credentials present?</i></p> <p>This is particularly important for REF as it relates to work attribution and to the roles different people had in conceiving and building the digital output. A statement of authorship should relate to content but also, for example, data model, information architecture, interface design etc.</p> <p>It is crucial to match the output to the submitting researchers and institutions: “A digital resource should not be viewed as the creation of, for example, a single lead applicant or project director, but as a collaborative exercise to which all members of a team have made unique contributions.” (AHRC and ICT Strategy project, 2006: 26)</p> <p>Example: <a href="https://pleiades.stoa.org/credits">https://pleiades.stoa.org/credits</a></p>
<b>Licences, Ownership and Copyright</b>	<p><i>Does the digital output or its submission clearly state who owns the data or the resources (e.g. images used on the site or digital archive, texts produced as part of a digital edition)? Does it contain copyright statements and information about licences (e.g. for data re-use)?</i></p> <p>With respect to REF submissions, not only is ownership of digital outputs essential but their alignment with principles of research integrity and reproducibility (e.g. availability of the data supporting the findings in research) as well as their potential for reuse<sup>2</sup> (e.g. availability of data for sharing with</p>

---

<sup>2</sup> See Beagrie (2019).

other users or justified restrictions) is also crucial. If research data are produced as part of RCUK funding, compliance with Research Council common principles on data policy should be followed (RCUK, 2015).<sup>3</sup>

Choices of licences (e.g. Creative Commons) with respect to content, software and data in a digital output should be explicit and justified. Enabling data discoverability and alignment with FAIR (Findable, Accessible, Interoperable, and Reusable) principles (Hagstrom, 2014) are increasingly becoming requirements for digital outputs in research across disciplines and domains. Inevitably, however, contexts and disciplines define what level of reproducibility and re-use is appropriate and to what aspects of the research (e.g. for a specific output, such as linguistic analysis, it might be crucial to have access to the research data or the corpus the argument is based on, while for other types of output, such as an exhibition in practice as research domains, documentation on the research process might be more important than providing access to ephemeral data).<sup>4</sup>

<b>Editorial Context and Documentation</b>	<p><i>Does the digital output contain information about its editorial context?</i></p> <p>This is particularly important for REF as it provides a description and essential information on the digital output's scope, limitations, date of public release and intended audiences, in order to assess its alignment to scholarly and research software engineering practices. Editorial introductions or equivalent sections of the output or submission should illustrate how the design of the digital output was consciously positioned within its wider intellectual context of relevance.</p> <p>Panel members should be able to assess the reasons why the digital output qualifies as a 3* or 4* output from its editorial context. This should be used to clarify and document the content of the output and the decisions made in all key steps of its creation, including, for example, what thesauri or controlled vocabularies are used in searching, the scholarly standards applied in the selection of material to digitise, the digitisation workflow (e.g. quality assurance of reproductions and subsequent handling or manipulation of the material in the case of a digital edition or digital archive project), transcription policy and editorial rules, when relevant.</p> <p>Editorial context and documentation should not only be limited to the academic editorial practices but be complemented by and ideally integrated with information on technical development (e.g. in the form of a technical overview describing technical infrastructure including software created or used, development approach, community technical standards adopted for the creation of the digital output, access to datasets including texts and archival records). The choice of technologies as well as modifications of community</p>
--	--

---

<sup>3</sup> With respect to data management and data deposits, see the Library guidelines on research data management: <https://www.kcl.ac.uk/library/researchsupport/research-data-management/index>

<sup>4</sup> While in the Arts & Humanities researchers do not necessarily recognise a division between the output and supporting data, "REF is moving towards a more flexible view of what the output is, so that for a practice researcher the whole lifecycle of research can be captured and the process seen as equally important as the outcome" (Beagrie 2019, 31).

standards implemented within the output should be commented on and documented if undocumented elsewhere.<sup>5</sup>

Contextual information should also include anything else that would be appropriate for any other research outputs to support its 3\* or 4\* assessment, including reviews (still rare for digital outputs, unfortunately), conference presentations and derivative publications.

At a minimum, the editorial context of the digital output should include information on “its scope, its creators, its methodology, its development and future maintenance” (AHRC and ICT Strategy project, 2006: 20).

This information should be up to date prior to REF submission (e.g. systematic check of broken links; correct tense to avoid out of date information e.g. ‘this project will...’) and packaged for the submission if the digital output is *in fieri* (evolving) or updated regularly (e.g. as a first edition or numbered version or release).

If the editorial context on the website itself does not provide sufficient information to assess why the digital output constitutes a 3\* or 4\* output for the submitted researcher, the additional statement accompanying the submission should explain this (see also ‘added value’).<sup>6</sup>

Examples of technical overviews:

<https://ncse.ac.uk/cms/about/technical-overview/>;

<https://janeausten.ac.uk/edition/technical.html>

#### **Edition Version**

*If the site is dynamic and forthcoming updates are foreseen, is there a way to retrieve the edition submitted to REF during the assessment period?*

Best practices implemented by King’s College London, Department of Digital Humanities and other institutions in 2014, included the creation of site ‘clones’ especially packaged for REF submission and including contextual information targeted to that scope. For each submitted site, freeze copies acting as a snapshot for that site were created, made accessible via a new URL and kept live for a year for the REF panel to examine. Whether hosted by KDL or another organisation, it takes time and effort to create such copies so this work should be planned in advance for the REF submission and in consultation with faculty.

The collection formats required for the different output types are detailed in Annex K of the *REF Guidance on submissions* (2019: 102-113).

---

<sup>5</sup> See Rockwell (2012): “annotation and interpretation takes place in the sphere of digital scholarship in ways that are different from the print world where interpretation often takes the form of an article or further book. Evaluators should ask about the depth of annotation and the logic of such apparatus”.

<sup>6</sup> See the *REF Guidance on submissions* (2019, 113) on supplementing material: “For non-text outputs, practice-based outputs or any other output where the research dimensions are not evident within the output/representation of the output itself: a written description of the research process and/or content should be provided. Wherever possible this should be submitted in REF2 in the ‘additional information’ field (maximum 300 words). Only where necessary to enable the panel to assess the research dimensions of the output, a fuller written description of the research process and/or content should be provided instead of the written description in REF2. The fuller written description should be included as part of an uploaded PDF, or on paper together with a physical output. b.F.” Panel D recommend the submission of a 300 word additional statement for most digital artefact output types, please see Annex C of the *Panel criteria and working methods* (REF 2019).

**Added Value**

*What is the added value of the digital output with respect to existing and comparable tools or resources? How has the potential of the medium been exploited? How innovative and/or significant is the resource?*

*Does it stand in scholarly and/or digital isolation? What has been taken from earlier works (e.g. printed editions); what is new? Are there features in this digital output that comparable tools or resources lack? Are there examples of searches to showcase research possibilities enabled by the digital output? What is the ‘added value’ of digital over print delivery in the context of the resource under discussion?*

If not spelled out in the editorial context, this information should be provided with the statement accompanying the submission (which is up to 300 words).

MLA (2012) offers some useful pointers to outline the innovative (or disruptive) aspect of digital outputs for those cases where this is appropriate:

- describe how the digital output “may blend, redefine, or render obsolete the traditional boundaries between teaching, research, and service”
- “describe the process underlying creation of work in digital media (e.g., the creation of infrastructure as well as content) and their particular contributions”
- “describe how work in digital media requires new collaborative relationships with clients, publics, other departments, colleagues, and students”

**Accessibility**

*How accessible is the site or mobile application if applicable? Does it include an accessibility statement?*

To meet legal requirements a website hosted by a public sector organisation should:

- meet the international *WCAG 2.1 Accessibility guidelines*;
- publish an accessibility statement that explains how accessible the website or app is.

New regulations, building on existing obligations towards people who have a disability, came into force for public sector bodies on September 2018 according to which any website or mobile app has to be made more accessible by making it ‘perceivable, operable, understandable and robust’. Content and design should be clear and simple enough so that most people can use it without needing to adapt it, while supporting those with impaired vision, motor difficulties, cognitive impairments or learning disabilities, deafness or impaired hearing (e.g. can the website be navigated using a keyboard? Does it contain inaccessible PDF forms that can’t be read out on screen readers? Is the colour contrast poor so that it makes the text difficult to read especially for visually

impaired people?).<sup>7</sup> If the site is hosted by KDL, the lab is in the process of updating accessibility statements to comply with legal regulations but it should be noted that accessibility of the content and related editorial practices should also comply and/or justifications provided in the accessibility statement. The responsibility of making content accessible falls on editors of and contributors to the content of the site. It is advisable to set up a content review process with this purpose.

The REF guidance states that the submitting HEI should ensure that any digital material is accessible from a range of devices (REF, *Panel criteria and working methods*, 2019: 92)

**User experience**

*How usable is the digital output? How fluid is the user experience of the site, for example? Is the site experience appropriate for its intended audience/s? Were user testing or design experiments conducted? Has been the design of the user interface thought through systematically?*

User interface and user experience are at the core of effective web design which encompasses functionality in tandem with aesthetics. Design assessment relates mainly to how well the resource is structured to enable user navigation (e.g. browsing) and interactivity with the components of a site or external elements as appropriate (e.g. easy to use search; stated and limited dependencies on external software; clear relationships with other related online resources; performance).

For example, at a very minimal level, if entries in an index or glossary do not lead to corresponding records in a database or textual occurrences, the user experience is poor.

Interaction happens at different levels and it also involves users consuming content by reading text (or listening to text with assistive technologies, eg. screen readers) and browsing media content. The choice of language should be appropriate for the targeted audience and take into account the diversity of contexts in which content is digested. For example, providing closed captioning for videos, transcriptions for audio files, alternative formats for data visualisations and maps, are mechanisms to allow users to choose a way of interacting with the content suited to their current situation.

**Citability**

*Is there information on how to cite the digital output and its sections (e.g. specific records)? Are persistent identifiers provided at the level of relevant site components? Are Digital Objects Identifiers (DOIs) provided for datasets or other relevant records?*

---

<sup>7</sup> A new public sector website created on or after 23 September 2018 needs to meet accessibility standards and include an accessibility statement by 23 September 2019. For websites published before 23 September 2018, the deadline is 23 September 2020 (but if substantial changes to the code were made, e.g. to create new features, it's likely that these will need to be fully accessible from 23 September 2019 i.e. by the same deadline as for new websites). Mobile apps need to be accessible by 23 June 2021. See <https://www.gov.uk/guidance/make-your-website-or-app-accessible-and-publish-an-accessibility-statement>

*Are search results savable or is it possible to download search data or other data?*

*If it is a hybrid publication (print and digital), is it clear how they relate with respect to citation conventions?*

Example of citation recommendations:

<https://finerollshenry3.org.uk/content/book/citations.html>

Example of citation format provided at record level:

<https://ncse.ac.uk/periodicals/mruc/issues/vm2-ncseproduct1720/page/1/articles/ar00100/>

<b>Look &amp; feel</b>	<p><i>Is the digital output look &amp; feel appropriate for its intended audience/s?</i></p> <p>Presentation, layout and other elements of the <i>look &amp; feel</i> of a digital output are not only aesthetic but also functional aspects of its design. Depending on the scope and intended audience of the output, it might be irrelevant for its look and feel to be attractive; however, other elements of its user interface might be relevant for its assessment as 3* or 4* output e.g. whether it is overly complex or appropriately designed for its intended audiences or whether, if applicable, images are of sufficient quality to investigate research claims on the material.</p>
<b>Dynamic components</b>	<p><i>If the site allows for external users to import or manipulate data (e.g. lemmatise textual corpora) and analyse them dynamically (e.g. applying scripts or algorithms and create visualisations), is there any documentation on the data workflow to assess the quality of the software and of the process of generating and interpreting results?</i></p> <p>Some digital outputs are enhanced by the provision of technical interfaces like OAI-PMH, REST, APIs etc., which allow the reuse of the data in other contexts. Documentation on the data structures and schemas, for example, is essential to assess these dynamic elements of a research output and support claim of use with other tools. For example, if integration with content in other systems is an objective of the digital output, this information allows to assess how feasible it is to aggregate content from this output with other sources (see Sahle 2014). If the dynamic components of the output rely on proprietary software or plug ins as opposed to open source software, this choice should be justified whether in the editorial documentation, technical overview or the REF submission itself so as to give panel members enough elements to assess the dynamic components.</p>
<b>Sustainability</b>	<p><i>If the output is an access or reference resource claiming to support updates or post-publication activities and/or to be accessible and maintained in the long term, is there information on the site or its submission accompanying statement to assess its sustainability (e.g. link to Service Level Agreement or institution responsible for hosting and maintenance)?</i></p> <p><i>Is the institution hosting and maintaining the output a long-term reliable host with Research Software Engineering expertise?</i></p> <p><i>Is there a clear statement of the standards that have been used, and an explanation of their benefits and/or limitations? Does the output contextual information provide evidence that consideration to its sustainability and</i></p>

*preservation (e.g. through the adoption of open standards and file formats best suited for sustainability and accessibility)<sup>8</sup> has been given?*

As a REF submission, the creators of a digital output are expected to have considered issues around its longevity (e.g. what are the implications of the choice of an institutional as opposed to a commercial host?), security, reliability (e.g. is it backed up?) and long-term value.

Example of footer with link to SLA provider:

<http://www.historicalpageants.ac.uk/>

## **Usage and analytics**

*Are reports or data on usage provided with the REF submission?*

Information on data usage provided with contextual narrative for a submission of a digital output are an indicator of the attention paid to monitor use; they are useful to demonstrate how the output is relevant for its intended (or additional pool of) users and to assess its reach when relevant (especially for impact case studies). However, “levels of usage should not be viewed as a key indicator of the scholarly value, or even impact, of a resource” (AHRC and ICT Strategy project, 2006: 26).

Usage reports (where relevant or possible), on the other hand, can be useful to indicate high engagement with the output, wide-spread recognition in the scholarly or other relevant communities and international audience (Smithies 2012).

## **CONTACTS**

*For projects that are not hosted and maintained by King’s Digital Lab, the following institutions or laboratories might be able to help make a digital output eligible as a 3\* or 4\* REF submission, in line with the guidance provided in this document (especially important for overseas engineering teams). Please note that they might need a considerable notice period to take on the work.*

- DHI, University of Sheffield: <https://www.dhi.ac.uk/>
- Agile Humanities: <http://agilehumanities.ca/>
- Digital Humanities Lab, University of Exeter: <https://humanities.exeter.ac.uk/digital-lab/>

## **REFERENCES**

---

<sup>8</sup> See for example DANS recommendations: [https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats?set\\_language=en](https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats?set_language=en)

Arts and Humanities Research Council and ICT Strategy Project (2006) *Final report and recommendations of the 'Peer review of digital resources for the arts and humanities' project* [https://www.history.ac.uk/sites/history.ac.uk/files/Peer\\_review\\_report2006.pdf](https://www.history.ac.uk/sites/history.ac.uk/files/Peer_review_report2006.pdf)

Beagrie, Neil (February 2019) *What to Keep: A Jisc research data study*. JISC  
[http://repository.jisc.ac.uk/7262/1/JR0100\\_WHAT\\_RESEARCH\\_DATA\\_TO\\_KEEP\\_FEB2019\\_v5\\_WEB.pdf](http://repository.jisc.ac.uk/7262/1/JR0100_WHAT_RESEARCH_DATA_TO_KEEP_FEB2019_v5_WEB.pdf)

CASRAI *CRedit* (Contributor Roles Taxonomy) <https://www.casrai.org/credit.html>

DANS *File formats*

[https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats?set\\_language=en](https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats?set_language=en)

Hagstrom, Stephanie (3 September 2014) “The FAIR Data Principles.” FORCE11  
<https://www.force11.org/group/fairgroup/fairprinciples>

King’s College London Library Services *Research Data Management*

<https://www.kcl.ac.uk/library/researchsupport/research-data-management/index>

MLA (February 2012) *Guidelines for Evaluating Work in Digital Humanities and Digital Media*

<https://www.mla.org/About-Us/Governance/Committees/Committee-Listings/Professional-Issues/Committee-on-Information-Technology/Guidelines-for-Evaluating-Work-in-Digital-Humanities-and-Digital-Media>

MLA (February 2013) *Guidelines for Authors of Digital Resources*

<https://www.mla.org/About-Us/Governance/Committees/Committee-Listings/Professional-Issues/Committee-on-Information-Technology/Guidelines-for-Authors-of-Digital-Resources>

Open Knowledge Foundation, Open Definition project *Conformant Licences*

<http://opendefinition.org/licenses/>

Purdue University The Online Writing Lab (OWL). *MLA Works Cited: Electronic Sources*

[https://owl.purdue.edu/owl/research\\_and\\_citation/mla\\_style/mla\\_formatting\\_and\\_style\\_guide/mla\\_works\\_cited\\_electronic\\_sources.html](https://owl.purdue.edu/owl/research_and_citation/mla_style/mla_formatting_and_style_guide/mla_works_cited_electronic_sources.html)

Research Council UK (RCUK) (July 2015, updated 2018)

REF (January 2019) *Guidance on submissions*

[https://www.ref.ac.uk/media/1092/ref-2019\\_01-guidance-on-submissions.pdf](https://www.ref.ac.uk/media/1092/ref-2019_01-guidance-on-submissions.pdf)

REF (January 2019) *Panel criteria and working methods*

[https://www.ref.ac.uk/media/1084/ref-2019\\_02-panel-criteria-and-working-methods.pdf](https://www.ref.ac.uk/media/1084/ref-2019_02-panel-criteria-and-working-methods.pdf)

Rockwell, Geoffrey (2012) Short Guide to Evaluation of Digital Work *Journal of Digital Humanities* 1.4.  
<http://journalofdigitalhumanities.org/1-4/short-guide-to-evaluation-of-digital-work-by-geofrey-rockwell/>

Sahle, Patrick (June 2014) *Criteria for Reviewing Scholarly Digital Editions, version 1.1*. In collaboration with Georg Vogeler and the members of the IDE. English version 1.1:  
<https://www.i-d-e.de/publikationen/weitereschriften/criteria-version-1-1/>

Smithies, James (2002) Evaluating Scholarly Digital Outputs: The Six Layers Approach. *Journal of Digital Humanities* 1.4  
<http://journalofdigitalhumanities.org/1-4/evaluating-scholarly-digital-outputs-by-james-smithies/>

Tanner, Simon (2 February 2015) “3 reasons why REF2014 was good for digital humanities scholars” When the Data hits the Fan! The blog of Simon Tanner (@SimonTanner)

UK Government Digital Services (20 May 2019) *Sample accessibility statement (for a fictional public sector website)*  
<https://www.gov.uk/government/publications/sample-accessibility-statement/sample-accessibility-statement-for-a-fictional-public-sector-website>

UK Home Office *Accessibility posters*  
[https://github.com/UKHomeOffice/posters/blob/master/accessibility/dos-donts/posters\\_en-UK/accessibility-posters-set.pdf](https://github.com/UKHomeOffice/posters/blob/master/accessibility/dos-donts/posters_en-UK/accessibility-posters-set.pdf)

W3C (5 June 2018) *Web Content Accessibility Guidelines (WCAG) 2.1*  
<https://www.w3.org/TR/WCAG21/>