

A #DataCreativities How-To Guide for Collaborative Attribution

This output was produced by members of the #DataCreativities team Kristal Spreadborough, Gene Melzack, and Emily Fitzgerald.

This output is a single PDF document containing information about the collective attribution for outputs created by the #DataCreativities collaboration.

Please cite this as:

#DataCreativities(2020): A #DataCreativities How-To Guide for Collaborative Attribution. <https://doi.org/10.26188/13325207>

Table of Contents

1 Context.....	2
2. Data and Outputs Citation - Citing us.....	2
2.1 Citing #DataCreativities data and outputs	2
2.2 Citing #DataCreativities outputs that are based on others data/outputs	2
3 Data and Outputs Citation - Us citing others.....	3
3.1 Citing an entire dataset.....	3
3.2 Citing part of a dataset.....	3
4 Individual and collective attribution	3

1 Context

This document has three goals:

1. To clearly define how the #DataCreativities collaboration outputs should be cited by others,
2. To clearly state how the #DataCreativities collaboration will cite others,
3. To make explicit how the #DataCreativities collaboration approaches individual and collective attribution.

We believe that such a document is necessary as data and output citation (especially when those outputs are Non Traditional Research Outputs which are common in the #DataCreativities collaboration) can often be ambiguous. Additionally, we wish to clearly articulate our approach to attribution since collective attribution is less common than individual attribution in the academic context. This document serves as a record of the #DataCreativities approach, and may be used as a blue print for other collaborations engaged in similar work.

2. Data and Outputs Citation- Citing us

2.1 Citing #DataCreativities data and outputs

If you are citing all or part of a dataset/output produced by #DataCreativities, please use the suggested citation we provide. If you cannot find the suggested citation, please follow the dataset citation guidelines for your preferred citation style. For example, see the [University of Melbourne referencing guide](#). Additionally, for citing subsets of a dataset/output produced by #DataCreativities, please also include the protocol (i.e. list of steps taken) used to obtain the subset and include this within your publication or metadata.

2.2 Citing #DataCreativities outputs that are based on others data/outputs

Sometimes, #DataCreativities produce outputs (e.g. visualisations) that are based on others data. For examples, see the "Data Citation - Us citing others" section below. To cite a #DataCreativities output that is based on others data, we recommend the following:

1. When talking/writing about the output, include a statement such as "This output was produced by #DataCreativities (provide relevant #DataCreativities citation) based on the work of (provide relevant secondary citation)". All outputs produced by #DataCreativities that draw on the work of others will provide the relevant secondary citation within the output. You may use this secondary citation in your talk/writing.
2. When curating a dataset, if you are using a dataset curated by #DataCreativities which consists of others data, include in your dataset metadata a statement such as "The #DataCreativities dataset is a curated output based on the following existing datasets (list the datasets #DataCreativities has drawn on, this information should be available in the #DataCreativities metadata) (Link to the #DataCreativities citation)".

3. When producing a visualisation or other non traditional research output, include in your output metadata a statement such as "The #DataCreativities output is based on the following existing outputs/datasets (list the outputs/datasets #DataCreativities has drawn on, this information should be available in the #DataCreativities metadata)(Link to the #DataCreativities citation)".

3 Data and Outputs Citation- Us citing others

3.1 Citing an entire dataset

Sometimes when producing visualisations or curating a new and novel dataset, the #DataCreativities team will use **an entire dataset** generated by someone else. In these cases, we will cite that dataset according to convention such as these examples from the [University of Melbourne Data Citation Guide](#). In line with attribution licensing practices, we will produce this citation on the published record where possible with the statement "This output was produced from the following dataset:". For example, citation within the image, video, or website. Where not possible, for example if data is being interacted with in VR environments, the citation will be produced in the abstract/description metadata with the statement "This output was produced from the following dataset:".

3.2 Citing part of a dataset

Sometimes when producing visualisations or curating a new and novel dataset, the #DataCreativities team will use **part of a dataset** generated by someone else. In this case, if the author of the original dataset provides a suggested citation for citing subsets of their data, then #DataCreativities will use this suggested citation and display as in "Citing an entire dataset" above. If there is no clear way to cite the subset, then #DataCreativities will:

1. Display the full dataset citation as in "Citing an entire dataset" above, with the statement "This output was produced from a subset of the following dataset:"
2. Include a protocol (i.e. a set of steps outlining) how the subset was extracted from the dataset in the output metadata. For example, if the output is on Figshare, then in the Figshare metadata/uploaded as an additional file with the output. If the output is on GitHub, included in the GitHub repository.

4 Individual and collective attribution

A strength of #DataCreativities is the collective nature of the group - the team has a very flat structure and each member has the opportunity to make their own unique contribution. In keeping with the emphasis on the collective, we produce a number of outputs under the banner of #DataCreativities and list individual author contributions under this banner. This is similar to existing practices of citing groups/corporate authors on publications. For examples of citations see the [University of Melbourne referencing guide](#). For an example of a Traditional Research Output published with a corporate author, but with individual attributions also included see [this book example](#) and this [journal article example](#). For Non Traditional Research outputs, a similar process is followed with #DataCreativities being

listed as author and individual attributions being listed as, for example, co-authors or in an authorship foot note.