**# Repository description**

This OSF repository contains data and code for the manuscript, ‘Big team science as a response to urgent societal developments’.

**## Note on computational workflow**

To begin, we recommend reviewing btmpact24e\_FileOrganization.png, which illustrates the project’s computational workflow. This includes the initial pulling of articles (e.g., btmpact24e\_api\_cgpt.Rmd) to making the paper figure (btmpact24e\_figure.Rmd) to compiling key results (btmpact24e\_paperrefs.Rmd).

Session info for the main computing device used to develop and run the code is available in btmpact24e\_session.info.

**## Note on proprietary nature of Altmetric data**

The Altmetric data used in this paper is proprietary. We only have permission to share for private peer review.

https://help.altmetric.com/support/solutions/articles/6000254871-faqs-for-scientometric-researchers

Data were made available during the initial submission, and we are firmly committed to sharing with others who wish to engage in post-publication review. In fact, we believe that \*not\* doing so would violate data retention rules at our host institution (see DataRetentionGuidelines\_012625.pdf).

To facilitate data sharing for legitimate private peer review cases, we have uploaded encrypted copies of Altmetric data to this OSF repository. Please contact the corresponding author for the password.

To further support efforts to assess the computational reproducibility of our workflow, we have uploaded *scrambled* copies of Altmetric data to this OSF repository. These scrambled data do not require a password to use.

**# Description of major files**

**## Main directory**

\* btmpact24e\_SI.Rmd:

\* btmpact24e\_altmetric\_scramblescript.Rmd: This code scrambles proprietary Altmetric data for sharing.

\* btmpact24e\_api\_cgpt.Rmd: This code uses the OpenAlex API to compile papers with keyword ‘ChatGPT’

\* btmpact24e\_api\_covid.Rmd: This code uses the OpenAlex API to compile papers with keyword ‘COVID-19’

\* btmpact24e\_api\_terrorism.Rmd: This code fetches papers from the OpenAlex API using keyword ‘terrorism’.

\* btmpact24e\_combine.Rmd: This code combines OpenAlex and Altmetric data.

\* btmpact24e\_figure.Rmd: This code creates the main figure for the paper.

\* btmpact24e\_paperrefs.Rmd: This code compiles the numbers quoted in the paper.

\* btmpact24e\_readme.docx: This code compiles information for the Supplemental Information.

**### admin folder**

\* btmpact24e\_FileOrganization.png: This figure illustrates the project’s computational workflow, highlighting the main scripts and data sets that were used in our analysis.

\* DataRetentionGuidelines\_012625.pdf: Copy of Data Retention Guidelines at the corresponding authors’ institution.

\* btmpact24e\_session.info.txt: Session info for the main computing device used to develop and run the code is available.

**### data folder**

\* altmetric.proprietary.encrypted.zip: Encrypted folder containing datasets (both raw and processed) with proprietary Altmetric data. Files can be shared for private peer review. Contact corresponding author for password.

\* btmpact24e\_codebook.xlsx: This spreadsheet is a codebook with variable descriptions for each of the spreadsheets within the data folder.

**#### cgpt subfolder (for papers linked to keyword ‘ChatGPT’**

\* btmpact24e\_data\_raw\_cgpt.Rds: Raw data for papers pulled from OpenAlex.

\* btmpact24e\_data\_binded\_cgpt.Rds: Processed data for papers pulled from OpenAlex.

\* btmpact24e\_data\_doi\_cgpt.csv: This spreadsheet contains the OpenAlex ID and DOIs for papers pulled from OpenAlex. DOI’s were used for manual Altmetric data extraction.

\* btmpact24e\_altmetric\_cgpt\_scrambled.csv: This spreadsheet contains DOI, policy document mention, and news output mention data manually extracted from Altmetric. Data are proprietary, so the data are scrambled.

**#### covid subfolder (for papers linked to keyword ‘COVID-19’)**

\* btmpact24e\_data\_raw1\_covid.Rds: Raw data for papers pulled from OpenAlex.

\* btmpact24e\_data\_raw2\_covid.Rds: Raw data for papers pulled from OpenAlex.

\* btmpact24e\_data\_binded1\_covid.Rds: Processed data for papers pulled from OpenAlex.

\* btmpact24e\_data\_binded2\_covid.Rds: Processed data for papers pulled from OpenAlex.

\* btmpact24e\_data\_binded\_covid.Rds: Processed data for papers pulled from OpenAlex.

\* btmpact24e\_data\_doi\_covid.csv: This spreadsheet contains the OpenAlex ID and DOIs for papers pulled from OpenAlex. DOI’s were used for manual Altmetric data extraction.

\* btmpact24e\_altmetric\_covid\_scrambled.csv: This spreadsheet contains DOI, policy document mention, and news output mention data manually extracted from Altmetric. Data are proprietary, so the data are scrambled.

**#### terrorism subfolder (for papers linked to keyword ‘terrorism’)**

\* btmpact24e\_data\_raw\_terrorism.Rds: Raw data for papers pulled from OpenAlex.

\* btmpact24e\_data\_doi\_terrorism.csv: This spreadsheet contains the OpenAlex ID and DOIs for papers pulled from OpenAlex. DOI’s were used for manual Altmetric data extraction.

\* btmpact24e\_altmetric\_terror\_scrambled.csv: This spreadsheet contains DOI, policy document mention, and news output mention data manually extracted from Altmetric. Data are proprietary, so the data are scrambled.

**#### combined subfolder**

\* DF.oa.combined.Rds: Combined OpenAlex data for papers linked to keyword “ChatGPT”, “COVID-19”, or “terrorism”.

\* DF.alt.combined.scrambled.Rds: Combined Altmetric data for papers linked to keyword “ChatGPT”, “COVID-19”, or “terrorism”. Data are proprietary, so the data are scrambled.

\* DF.full.combined.scrambled.Rds: Combined OpenAlex and Altmetric data for papers papers linked to keyword “ChatGPT”, “COVID-19”, or “terrorism”. Some variables are pulled from the proprietary Altmetric dataset. Those data are scrambled.