READ ME File For 'Shaping the Quantum Internet'

Dataset DOI: 10.5258/SOTON/D2556

Associated Grant: EP/L016117/1: EPSRC Centre for Doctoral Training in Web Science Innovation

Date that the file was created: March, 2023

GENERAL INFORMATION

ReadMe Author: Juljan Krause, University of Southampton [ORCID ID: 0000-0002-1956-6720]

Date of data collection: Q3/Q4 2021

Information about geographic location of data collection: EPI global

Related projects:
n/a

SHARING/ACCESS INFORMATION

Licenses/restrictions placed on the data, or limitations of reuse: CC-BY

Recommended citation for the data:
Dataset in support of the Southampton doctoral thesis 'Shaping the Quantum Internet'

This dataset supports the publication: n/a

Links to other publicly accessible locations of the data: n/a

Links/relationships to ancillary or related data sets: n/a

DATA & FILE OVERVIEW

This dataset contains:

This .zip file contains
•   six R files for building ERGMs
•   two R files for descriptive and summary network statistics for US and Chinese data
•   two R files for rendering the US and Chinese HiveR plots
•   two Matlab files for the computation of the core of the coalitional game

The raw data cannot be made available as they are proprietary. They were retrieved from the European Patent Office's Global Patent Index.

Full references and acknowledgements of the packages, manuals and documentation used can be found in the Appendix of the PhD thesis.

Relationship between files, if important for context: n/a

Additional related data collected that was not included in the current data package: n/a

If data was derived from another source, list source: n/a

If there are there multiple versions of the dataset, list the file updated, when and why update was made: n/a

----
METHODOLOGICAL INFORMATION

Description of methods used for collection/generation of data: EPO GPI Data Retrieval Manual:
https://www.epo.org/searching-for-patents/technical/espacenet/gpi.html

Methods for processing the data: standard data cleaning and wrangling strategies, methods for creating network objects (directed graphs) from .csv files

Software- or Instrument-specific information needed to interpret the data, including software and hardware version numbers: MS Excel, OpenRefine, R, Matlab

Standards and calibration information, if appropriate: n/a

Environmental/experimental conditions: n/a

Describe any quality-assurance procedures performed on the data: manual check of each record prior to inclusion in final dataset

People involved with sample collection, processing, analysis and/or submission: author

----
DATA-SPECIFIC INFORMATION <Create sections for each datafile or set, as appropriate>

Number of variables: 10,000+

Number of cases/rows: 10,000+

Variable list, defining any abbreviations, units of measure, codes or symbols used: please see R files

Missing data codes: n/a

Specialized formats or other abbreviations used: n/a