Caring for Sharing: Improving Self-Deposit Repository Publishing Systems

Anca Vlad

UK Data Service

University of Essex

IASSIST & CART 2018

Once Upon A Data Point: Sustaining Our Data Storytellers

May 29 – June 1, 2018

UK Data Service

Introduction & Summary

- ReShare is the self-deposit data repository of the UK Data Service, now in its third year and holding some 1305 collections of data, spanning a wide disciplinary range of research assets.

- This presentation will:
  - look at the main features of ReShare
  - provide an overview of data collections ingested
  - summarise the ReShare ingest process
  - demonstrate tracking metrics collected via Google Analytics
  - discuss possible improvements

UK Data Service

ReShare overview

- UK Data Service's online data repository, where researchers can archive, publish and share research data: http://reshare.ukdataservice.ac.uk

- Extensively customised version of Eprints open-sources repository software;

- It allows depositors to create data collections themselves - upload data and documentation files;

- Once completed by depositors, collections go into a review area, where they are checked for disclosure risk, copyright breaches, validity of file formats and level of documentation;

- Is an easy-to-use, step-by-step system, with a user-friendly interface designed to made the deposit process as simple and fast as possible;

- Metadata profile is based on the DDI schema;

- Access controls are customised for each collection;

- Studies indexed using the Humanities and Social Sciences Electronic Thesaurus (HASSET);
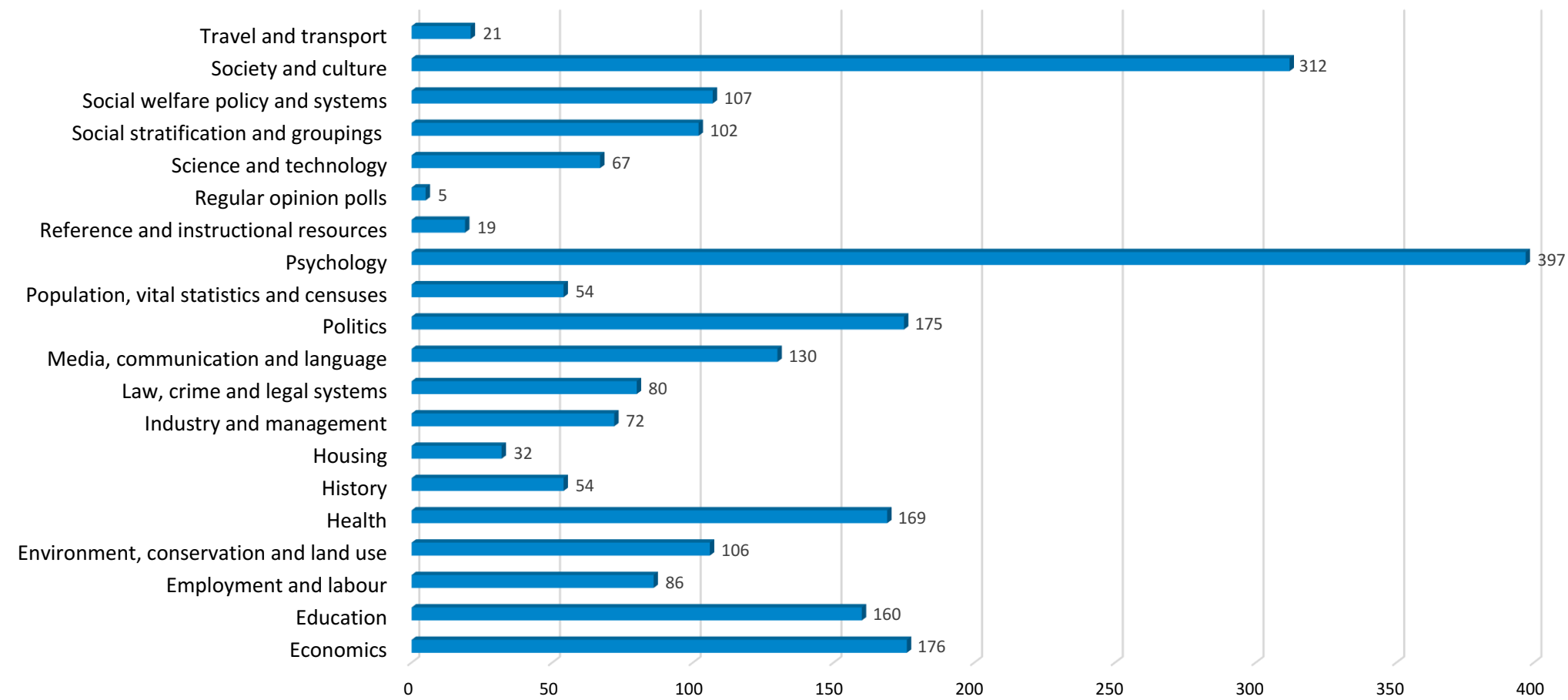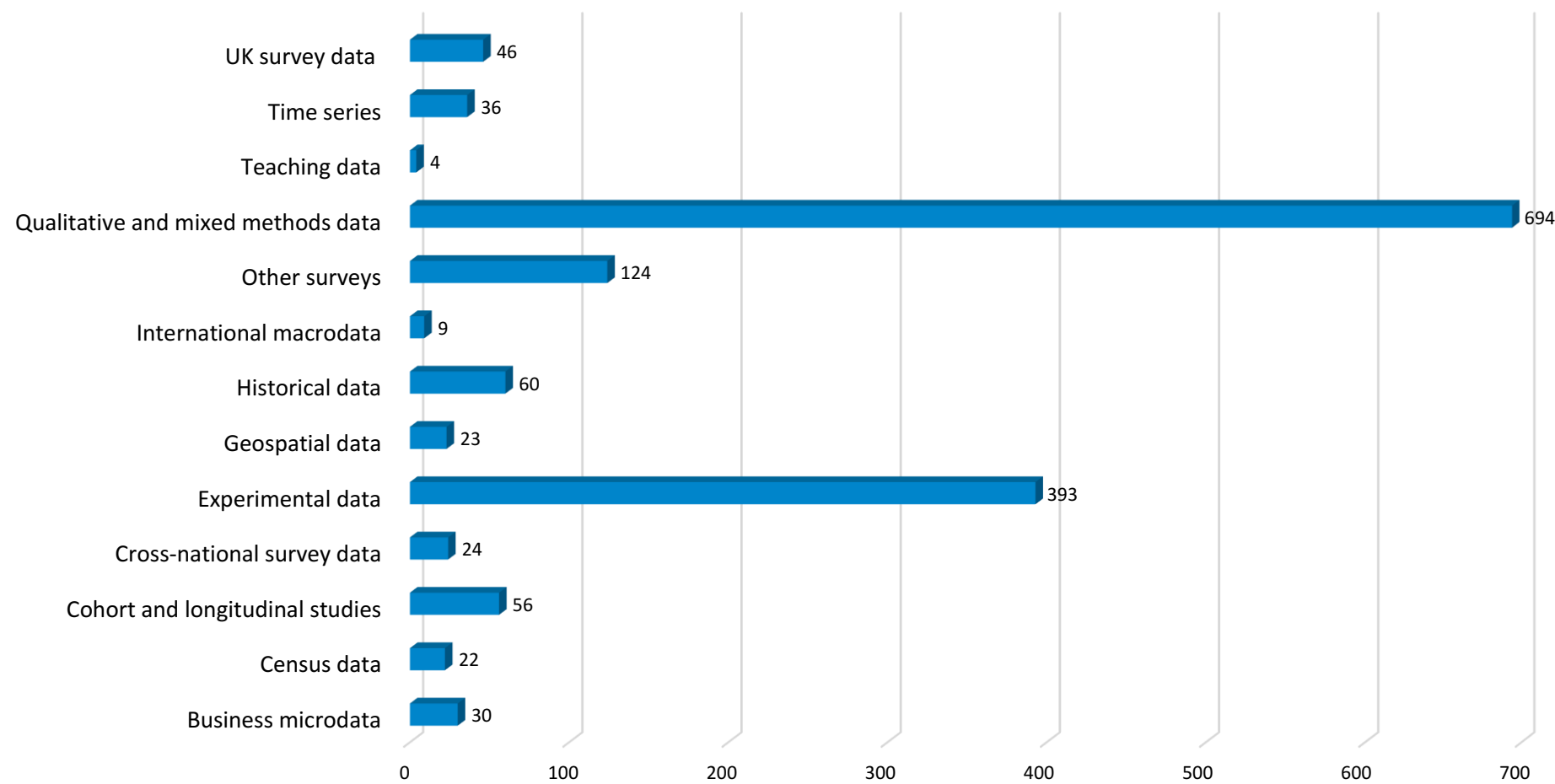
ReShare overview

- Simplifies data deposit by enabling the upload of multiple files, either individual or zip bundles;
- Addresses problems researchers may have when publishing their data*:
  - Limited or poor metadata (abstract/methodology);
  - Lack or insufficient supporting documentation;
  - Poor file names, not representative of content;
  - Data in formats not recommended for long term preservation.
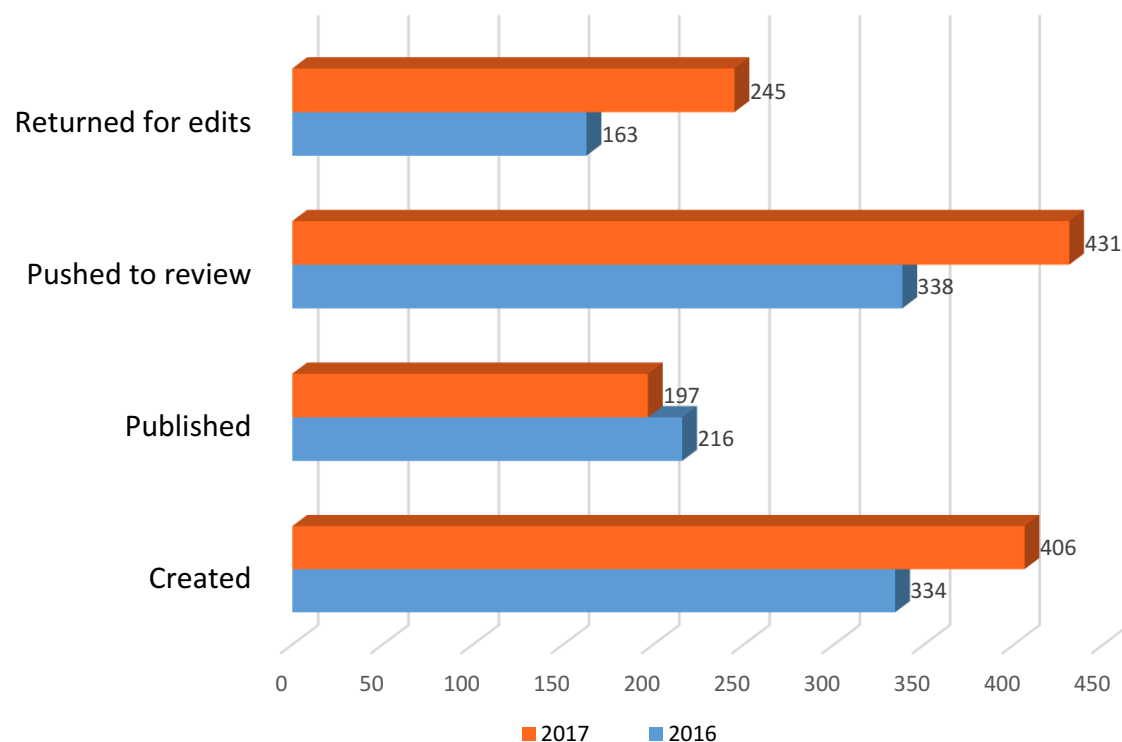
UK Data Service

*As reported in Van den Eynden and Corti (2017)

Data ingested - Subjects

Subject	Value
Travel and transport	21
Society and culture	312
Social welfare policy and systems	107
Social stratification and groupings	102
Science and technology	67
Regular opinion polls	5
Reference and instructional resources	19
Psychology	397
Population, vital statistics and censuses	54
Politics	175
Media, communication and language	130
Law, crime and legal systems	80
Industry and management	72
Housing	32
History	54
Health	169
Environment, conservation and land use	106
Employment and labour	86
Education	160
Economics	176

UK Data Service

Data ingested - Data type

UK survey data — 46
Time series — 36
Teaching data — 4
Qualitative and mixed methods data — 694
Other surveys — 124
International macrodata — 9
Historical data — 60
Geospatial data — 23
Experimental data — 393
Cross-national survey data — 24
Cohort and longitudinal studies — 56
Census data — 22
Business microdata — 30

0   100   200   300   400   500   600   700

UK Data Service

Metrics

Ingest process:

Once a collection is completed by the depositor, it arrives in the review area, where we performs our standards quality checks.

If any issues are found, we contact the depositor via e-mail and send the collection back to their work area to resume editing.

Once the collection is submitted for review once again, it is published in our catalogue Discover and a DOI is minted.

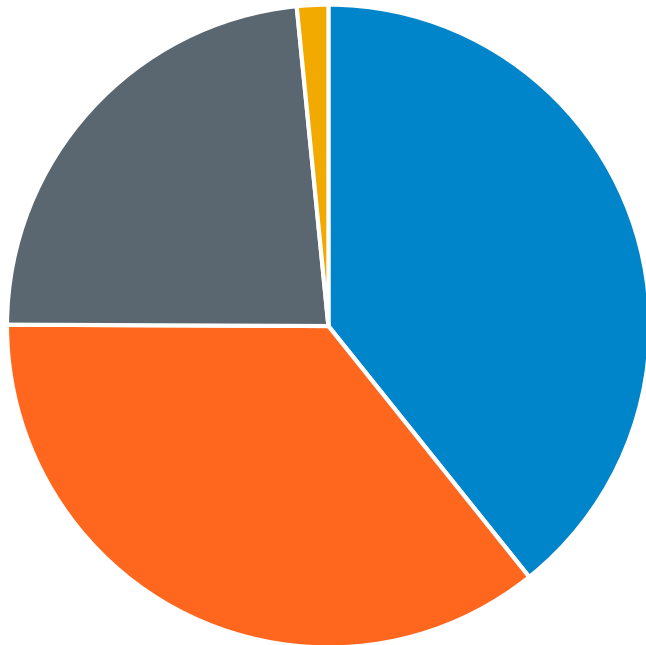UK Data Service

Users and sources

Users (per month)
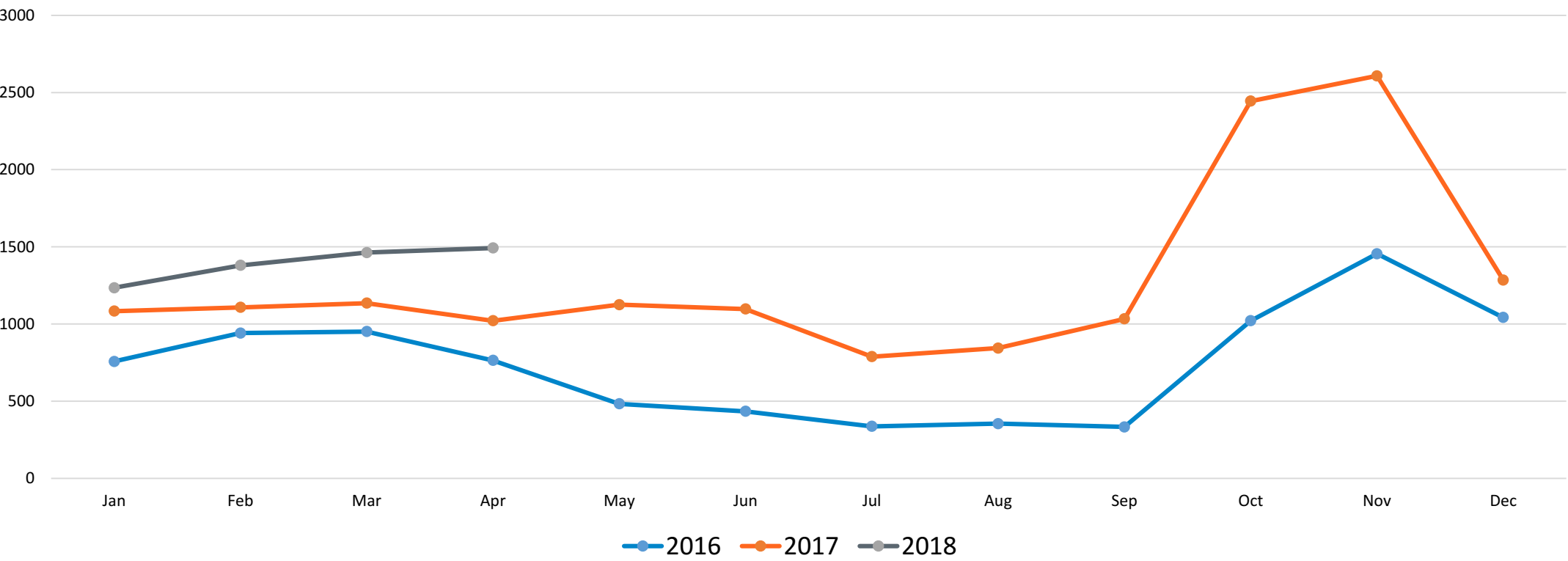
Users and sources

Traffic

- ■ Organic - Search Engine Listing (39%)

- ■ Reffered by links on other websites, e.g. DOI (36%)

- ■ Typing the URL directly into browser(23%)

- ■ Social networks (2%)

UK Data Service

Locations and devices

Location*

Country		Users ↓	New Users	Sessions
		14,646 % of Total: 100.00% (14,646)	10,131 % of Total: 100.00% (10,131)	23,079 % of Total: 100.00% (23,079)
1.	🇬🇧 United Kingdom	8,212 (55.65%)	4,652 (45.92%)	14,455 (62.63%)
2.	🇺🇸 United States	958 (6.49%)	786 (7.76%)	1,194 (5.17%)
3.	🇨🇳 China	541 (3.67%)	502 (4.96%)	660 (2.86%)
4.	🇻🇳 Vietnam	385 (2.61%)	380 (3.75%)	399 (1.73%)
5.	🇩🇪 Germany	380 (2.58%)	320 (3.16%)	464 (2.01%)
6.	🇮🇳 India	303 (2.05%)	256 (2.53%)	344 (1.49%)
7.	🇪🇸 Spain	173 (1.17%)	141 (1.39%)	233 (1.01%)
8.	🇸🇬 Singapore	172 (1.17%)	79 (0.78%)	409 (1.77%)
9.	🇧🇷 Brazil	170 (1.15%)	166 (1.64%)	183 (0.79%)
10.	🇳🇱 Netherlands	165 (1.12%)	123 (1.21%)	262 (1.14%)

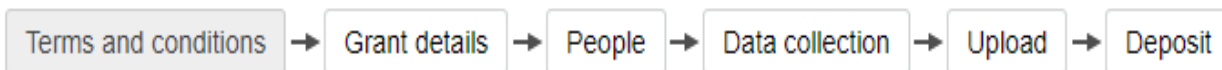## Devices used (new users* count)

■ 2017   ■ 2016
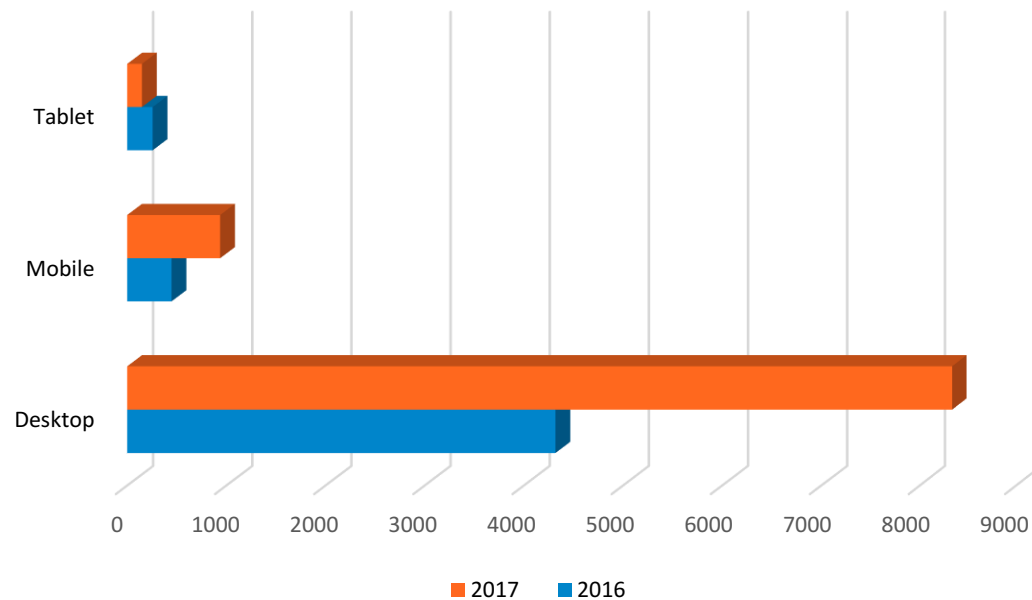
*last 12 months
**number of first-time users during the selected date range.

UK Data Service

Training

Edit collection: Data Collection #853213

Terms and conditions → Grant details → People → Data collection → Upload → Deposit

To deposit a data collection, you must accept the ReShare Terms and Conditions.

- I confirm that I am the owner of the copyright and associated intellectual property rights in the whole Data Collection or am otherwise lawfully entitled to grant this licence on behalf of each and every owner;
- I grant a non-exclusive, royalty-free licence to the UK Data Archive (a department of the University of Essex and not a separate legal entity) of Wivenhoe Park, Colchester, CO4 3SQ (the "University") to hold, make copies of, and disseminate copies of the Data collection, in accordance with the access conditions I will specify when uploading data files: open data accessible to users without registration; or safeguarded data accessible to users registered with the data services provided by the UK Data Archive.

* I agree to the ReShare data deposit terms and conditions

☑

Save for later   Cancel   Next >

Training is done via workshops, webinars and drop-in sessions.

The method which proved most effective is a walk through the system, with examples of how to create a data collection from start to finish, touching on certain details we sometimes receive questions about such as access conditions, licence, mandatory fields, metadata, formats etc.

To illustrate this we use a training area, which imitates the real system to the smallest details:

http://reshare-training.ukdataservice.ac.uk

UK Data Service

Looking ahead

Short term

- Documentation templates (Readme, data list, methods file, data deposit exemption form)

- ReShare website improvements

- Pop-up surveys (instant feedback, feed into FAQs)

- Instant prompts (access, licence)

Long term

- Ongoing DM training

- Negotiating access to meet 'as open as possible, as closed as necessary'

- QAMyData - will offer a free easy-to-use tool/service that automatically detects some of the most common problems in survey and other numeric data and creates a 'data health check'

UK Data Service

Contact

- ReShare: reshare@ukdataservice.ac.uk

- UKDS enquiries/ Help Desk:
  http://ukdataservice.ac.uk/help/get-in-touch.aspx

  help@ukdataservice.ac.uk

- Follow us on:
  https://twitter.com/UKDataService

  https://www.facebook.com/UKDataService

  https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=UKDATASERVICE

UK Data Service

Questions?

UK Data Service