Jisc

Research Data Management:
The Numbers

90% 90% of the world's data has been created within the last two years

2.5 **Billion** gigabytes of new data is generated **EVERY DAY**
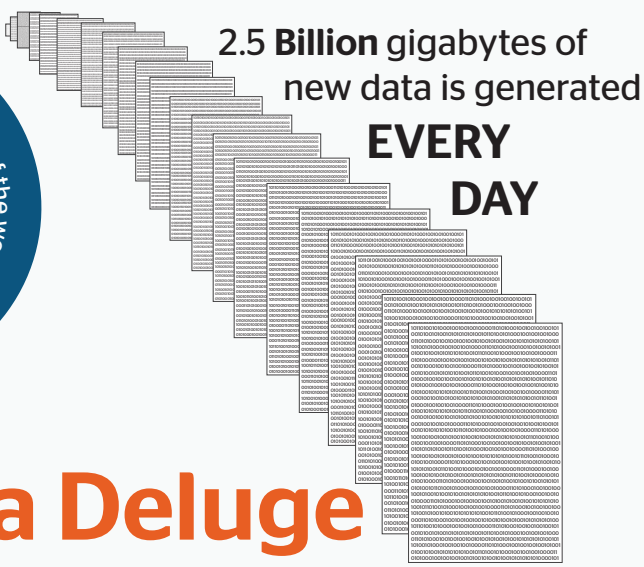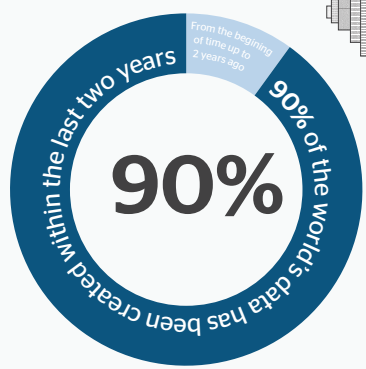
500,000 research datasets in the Figshare repository

450 PB of research data is held in UK Universities

University

The Data Deluge

$

3,000,000,000,000

(That's 3 trillion) Estimated value of open data to the world economy

500,000 employees working in UK companies built on open data

182,000 new data-related jobs created in the UK by 2020
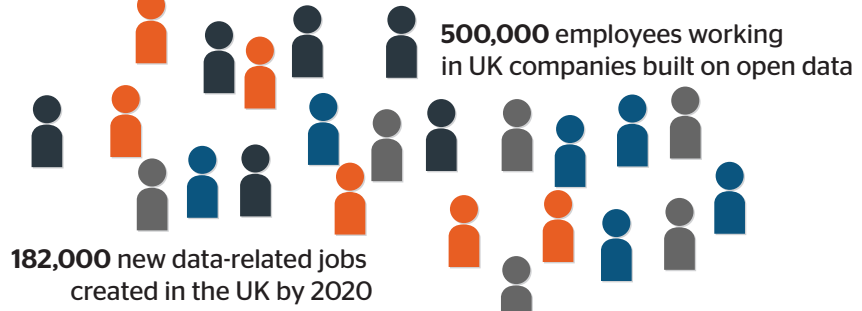
The Value of Open Data

Closed data
Closed data
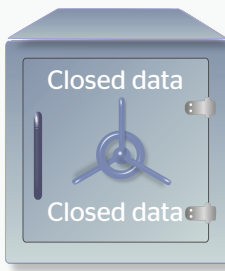
Less than 50% of clinical trial results are currently published and shared

Only 20% of scientific data from the 1990s is still available

Lost

Research Quality and Efficiency

?.top

Only 1/3 of findings in high-ranked Psychology journal papers can currently be reproduced

11 %

Only 11% of results in 'landmark' oncology papers can be replicated

UK Investment in Research

10:1

RESEARCH COUNCILS UK

8 Great Technologies built upon data

Regenerative Medicine

Agri-science

Advanced Materials

Energy Storage

Big Data

Synthetic Biology

Robotics & Autonomous Systems

Satellites

Every pound invested by Research Councils UK generates a 10 fold Return on Investment

£26 billion UK government investment in research and scientific infrastructure over the next 5 years

Changing culture

50% of researchers are already sharing data!

Over half of the top-ranked science and social science journals either mandate or strongly suggest researchers share the data that relates to the publication
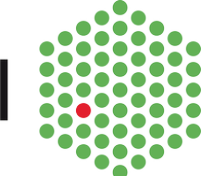
61% of recent papers funded by the Engineering and Physical Sciences Research Council were linked to the underlying research data

75% of researchers strongly favour formal citation of data as a reuse indicator

The Human Genome Project's open research data contributed $1 trillion to the US economy
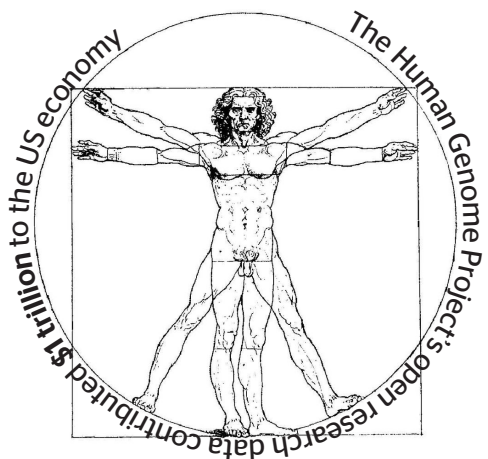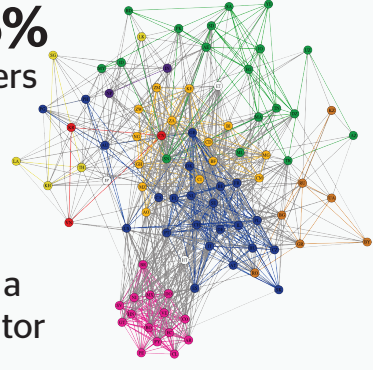
EMBL-EBI

£920 million per annum

Estimated return on investment on future research enabled by the European Bioinformatics Institute

UK Data Service

The UK Data Service generates **£100 million** per annum wider benefit to the UK.

The value of Research Data

How big is big?

1 plain text character can be thought of as (approximately) 1 byte of data.

a

Typically, in written English, 1000 characters would equate to about 175 words. This can be thought of as 1 **kilo**byte of data

So 1 normal page (a little more than 525 words)...

...is about 3 kilobytes of data. A typical page is about 0.1 mm thick—the width of a hair

333 pages (a stack about 33 mm high—a little more than half a ream) equates to 1 **mega**byte

That's the equivalent of a medium sized photo and would fill an old 5 ¼ inch HD floppy disk

333,000 pages equates to 1 **giga**byte. Our stack is now about 33 m—the length of a blue whale—high. That's about 1 ½ CDs of mp3 music tracks (about 150). These days even basic smartphones have a capacity in the 10s of gigabytes.
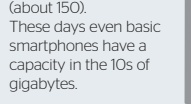
At 1 **tera**byte our paper stack is 33 km high.

We've penetrated the Stratosphere. Modern entry level computers typically have 1 terabyte hard disks

1 **peta**byte makes a stack 33,300 km high

We're now 10% of the distance to the moon. CERN produce 30 petabytes of data a year.

1 **exa**byte makes a stack 33.3 *million* km high. We're talking interplanetary scale distances now—we've got to Mars. A 1 **zeta**byte stack is 3.33 x 10⁹ km high (or 223 Astronomical Units (AU)—1 AU is the distance from the Earth to the Sun). **We've left the solar system!** In 2013 the world generated about **3.5 zettabytes of data**