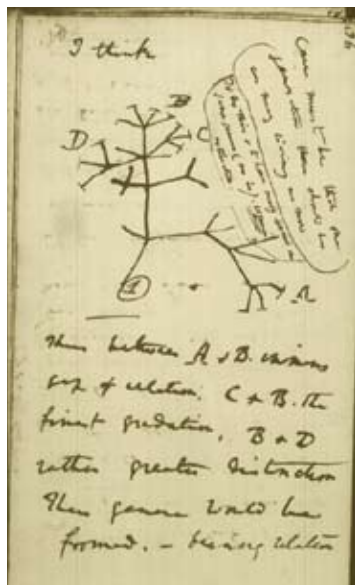


A reception study in the making?

John van Wyhe recounts the unprecedented reception of Darwin's private papers online.



A page from Darwin's theoretical notebooks.

On 17 April 2008 Darwin Online launched the largest collection of Charles Darwin's private papers ever published online. In terms of the amount of newly published material, the size and variety of the collection and the overwhelming public interest, the launch was a historic event in its own right. Never before had so much material by and about Charles Darwin been released to the public. The online collection of c. 20,000 items across c. 90,000 electronic images are scans from microfilms of thousands of the papers in the Darwin Archive at Cambridge University Library. It is only with the extremely generous cooperation of Cambridge University Library that such a release was possible. The words 'private papers' were chosen with care. For a start many lay readers would not understand 'manuscripts' and secondly the collection includes thousands of printed items, photos and maps that belonged or pertain to Darwin.

The public and media reception exceeded all expectations – marking an involvement with the history of science on a global scale.

The news was announced via press releases and interviews in most of the national newspapers and on Radio 4's Today Programme. Following the Today Programme there were four local and regional radio interviews, two on Radio 5, two in Australia and one in Ireland and Brazil. Subsequently there was an interview on National Public Radio in America and Voice of America radio. Hundreds of newspapers, magazines and news websites carried the story on the 17th and following days.

The traffic to the website was unprecedented. On the 17th the site recorded 7 million hits, bringing the server several times briefly to a standstill. On the 18th there were 8 million hits. On the 17th the site uploaded 1.86 terabytes of data (a terabyte is 1000 gigabytes) which was six times the data uploaded when Darwin's publications were launched online on 19 October 2006. In total it is estimated that over 20 terabytes of data were uploaded in April.

The Google news service blog search feature showed 17,000 hits for blogs discussing 'Darwin's papers' in the first 48 hours. Users made hundreds of thousands of searches within the site's manuscript catalogue, the first electronic union catalogue of Darwin papers around the world (still a work in progress), and thus found their way into the papers. Others browsed through whole volumes – clicking through sequences of hundreds of images – to read and explore.

Users were also accessing the rest of the site. On the 17th alone 14,000 copies of the pdf version of the 1st edition of *Origin of Species* were downloaded – and in the month of April 68,778 copies. This and the other downloads show that more copies of Darwin's works were distributed in 48 hours than in the whole of the 19th century.

Another interesting feature was the large percentage of traffic from non-English speaking countries. Hundreds of thousands of referrals were recorded by the site coming from translation services like Google and Babblefish.

The collection includes many items readers will have heard about or quoted without ever having seen the original manuscripts. Journalists like to have a small number of highlighted items to report. In this case it was almost impossible to choose any individual items that could be representative of such a varied collection. A combination of dramatic, representative and personal items were chosen.

One curious feature of the reporting and commenting of the site, which was observed when images of Emma Darwin's diaries were launched in March 2007, there were no references to the fact that the images are not in colour. This is perhaps the major drawback to using scans from microfilms, although many online projects use them. They offer, at present, unbeatable value. It would probably cost several million pounds to scan these materials afresh in colour. The greatest expense for Darwin Online was actually buying copies



of the microfilms. These were then sent to India and commercially scanned and cropped for c. £10,000. The resulting images still need a lot of work. They arrived back in the UK named 001.jpeg, 002.jpeg etc. The files needed to be named according to the archive catalogue. To date just over 10,000 items have been numbered by hand. (We are still looking for volunteers to help with this.) To make the remainder accessible and, importantly, to allow users to move through them without needing to go through the catalogue for specific items, browseable sequences of entire volumes were created. Hence for Darwin Archive 226 a single sequence of 781 images was created all bearing the name DAR226 and numbered automatically.

Many letters from readers were received. These express gratitude for being able to see Darwin's papers for the first time for free and excitement at the richness to be explored.

This is encouraging news for the history of science. While few figures generate as much fuss as Darwin, the general public is interested in accessing and using primary materials from the history of science. The possibilities before us for online research and public engagement with the history of science are limitless.

John van Wyhe
University of Cambridge
jmv21[a]cam.ac.uk
www.darwin-online.org.uk

Images © Cambridge University Library