

Digital Preservation at Oxford and Cambridge

A collaborative research project to evaluate and provide sustainable recommendations for our digital preservation programmes

Institutional risk and born-digital content: the shutdown of DCist #IDPD17

Posted on [30 November, 2017](#) by [Sarah](#)

Another post for today's International Digital Preservation Day 2017. Outreach and Training Fellow, Sarah, discusses just how real institutional risk is and how it can lead to a loss of born digital archives — a risk that digital-only sites like DCist have recently proven. Read more about the Gothamist's website shutdowns this November.

In today's world, so much of what we create and share exists only in digital form. These digital-only creations are referred to as born-digital — they were created digitally and they often continue in that way. And so much of our born-digital content is shared online. We often take for granted content on the Internet, assuming it will always be there. *But is it?* Likely it will at least be captured by the Internet Archive's [Wayback Machine](#) or a library web archiving equivalent. But is that actually enough? Does it capture a complete, usable record? What happens when a digital-only creation, like a magazine or newspaper, is shut down?

Institutional risk is real. In the commercial world of born-digital content that persists only in digital form, the risk of loss is high.

Unfortunately, there's recently been a very good example of this kind of risk when the [Gothamist shut down](#) its digital-only content sites such as the DCist. This happened in early November this year.



The Washington Post
@washingtonpost

Gothamist local news network, including DCist and DNAinfo, abruptly shut down wapo.st/2A5r6wh

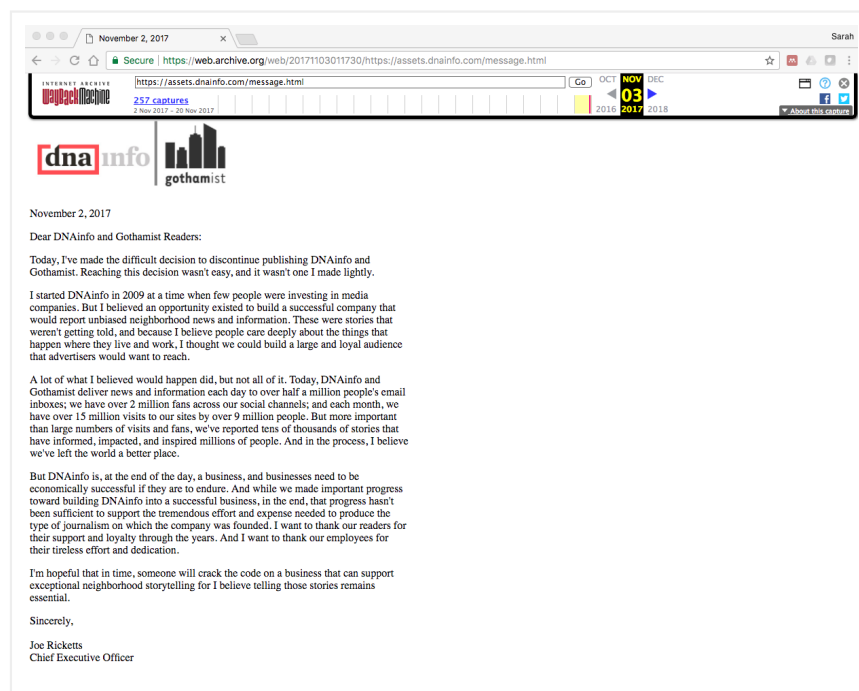
38 06:05 - 4 Nov 2017

Gothamist local news network...

The owner says that while they had strong readership numbers, the websites could not turn a washingtonpost.com

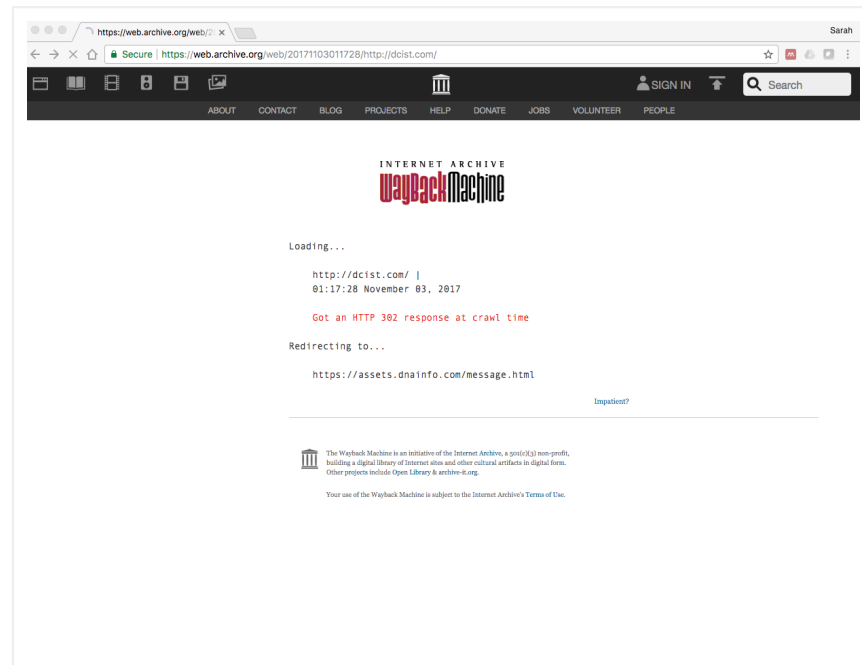
[28 people are talking about this](#)

The sites and all the associated content was completely removed from the Internet by the morning of 3 November. Gone. Taken down and replaced with a letter from billionaire CEO, Joe Ricketts, justifying the shutdown because despite its enormous popularity and readership, it just wasn't "economically successful."



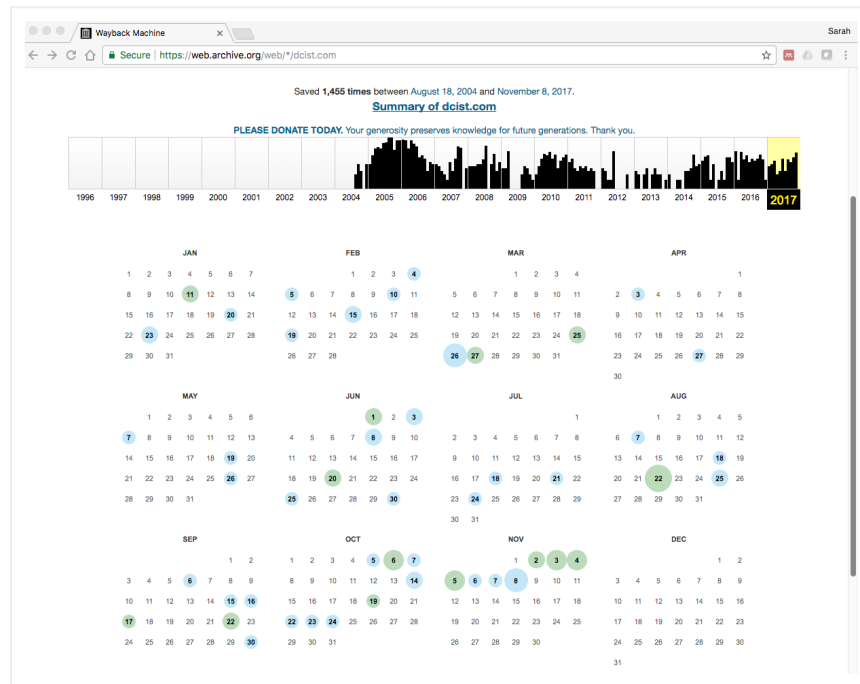
— Wayback Machine's capture of the redirect page and Ricketts' letter

The DCist site and all of its content was gone completely; readers instead were redirected to another page entirely to read Joe Ricketts' letter. Someone had literally pulled the plug on the whole thing.



- Internet Archive's 3 November 2017 capture, showing a redirect from the DCist.com page. DCist was gone from the Internet.

The access to content was completely lost, save for what the Internet Archive captured and what content was saved by creators elsewhere. But access to the archives of 13 years of DCist content was taken from the Internet and its millions of readers. At that point all we had were some web captures, incomplete records of the content left to us.



— The Internet Archive’s web captures for DCist.com over the past 13 years.

What would happen to the [DCist’s archive](#) now? All over Twitter people were being sent to Internet Archive or to check Google’s cache to download the lost content. But as Benjamin Freed pointed out in his recent [Washingtonian](#) article:

“Those were noble recommendations, but would have been incomplete. The Wayback Machine requires knowledge about URLs, and versions stored in Google’s memory banks do not last long enough. And, sure, many of the subjects DCist wrote about were covered by others, but not all of them, and certainly not with the attitude with which the site approached the world.”

As Freed reminds us “A newspaper going out of business is tragic, but when it happens, we don’t torch the old issues or yank the microfilms from the local library.” In the world of born-digital content, simply unplugging the servers and leaving the digital archive to rot means that at best, we may only have an incomplete record of the 1,000s of articles and content of a community.

If large organisations are not immune to this kind of institutional risk, what about the small ones? The underfunded ones?



Washingtonian

@washingtonian

The Death of DCist Wasn't Just a Labor Decision.
It Was a Digital Book

Burning [washingtonian.com/2017/11/03/the...](https://www.washingtonian.com/2017/11/03/the...)

8 02:12 - 7 Nov 2017

The Death of DCist Wasn't Ju...

Its owner restored the archives.

They'll be up as long as he's
feeling benevolent.

[washingtonian.com](https://www.washingtonian.com)

[See Washingtonian's other Tweets](#)

To be clear, I think web archiving is important and I have used it a number of times when a site is no longer available — it's a valuable resource. But it only goes so far and sometimes the record of website is incomplete. So what else can we do? How can we keep the digital archive alive? The good news is that while Ricketts has put the DCist site back up as an “archive” — it's more like a “digital graveyard” that he could pull the plug on again any time he wants. How do you preserve something so fragile, so at risk? The custodians of the digital content care little for it, so how will it survive for the future?

The good news is that the DCist archive may have another home, not just one that survives on the mercy of a CEO.



Pablo Maurer

@MLSist

Looking like [@dcist](#)'s archive may well live on permanently in the Washingtoniana collection at the MLK library downtown. ❤️

50 20:10 - 6 Nov 2017

[See Pablo Maurer's other Tweets](#)

The born-digital archives of the DCist require more than just a functioning server over time to ensure access. Fortunately, there are places where digital preservation is happening to all kinds of born-digital collections and there are passionate people who are

custodians of this content. These custodians care about keeping it accessible and understandable for future generations. Something that Joe Ricketts clearly does not.

What are your thoughts on this type of institutional risk and its impacts on digital preservation? How can we preserve this type of content in the future? Is web archiving enough or do we need a multi-prong approach? Share your thoughts below and on Twitter using the #IDPD17 hashtag.

SHARE THIS:



Share

This entry was posted in [born-digital](#), [digital preservation](#), [digital preservation coalition](#) by [Sarah](#). Bookmark the [permalink](#) [<http://www.dpoc.ac.uk/2017/11/30/1140/>] .

About Sarah

Digital Preservation Specialist - Outreach and Training:
Bodleian Libraries, Oxford University

[View all posts by Sarah](#) →

This site uses Akismet to reduce spam. [Learn how your comment data is processed.](#)