How the Wayback Machine is trying to solve the web's growing linkrot problem

As more sites collapse and shutter, the internet has increasingly become full of roads to nowhere.

By Nilay Patel, editor-in-chief of The Verge, host of the **Decoder podcast**, and co-host of **The Vergecast**.

We've been talking a lot about the future of the web on *Decoder* and across *The Verge* lately, and one big problem keeps coming up: huge chunks of the web keep going offline. In a lot of meaningful ways, large portions of the web are dying. Servers go offline, software upgrades break links and pages, and companies go out of business — the web isn't static, and that means sometimes parts of it simply vanish.

It's not just the "really old" internet from the '90s or early 2000s that's at risk. A recent study from the Pew Research Center found that 38 percent of all links from 2013 are no longer accessible. That's more than a third of the collected media, knowledge, and online culture from just a decade ago — gone. Pew calls it "digital decay," but for decades, many of us have simply called it linkrot.

Lately, that means a bunch of really meaningful work is gone as well, as various news outlets have failed to make it through the platform era. The list is virtually endless: sites like *MTV News*, Gawker (twice in less than a decade), *Protocol*, *The Messenger*, and, most recently, *Game Informer* are all gone. Some of those were short-lived, but some outlets that were live for decades had their entire archives vanish in a snap.

But it's not all grim. For nearly as long as we've had a consumer internet, we've had the Internet Archive, a massive mission to identify and back up our online world into a vast digital library. It was founded in 1996, and in 2001, it launched the Wayback Machine, an interface that lets anyone call up snapshots of sites and look at how they used to be and what they used to say at a given moment in time. It's a huge and incredibly complicated project, and it's our best defense against linkrot.

Mark Graham, director of the Wayback Machine, joins me on the show this week to explain both why and how the organization tries to keep the web from disappearing. (A quick note: the Internet Archive just lost an appeal in a lawsuit over a short-lived book-lending initiative it launched at the start of the covid-19 pandemic. We don't get into the details of that in this episode, since it happened after we recorded, but we wanted to mention the news.)

The answers are fascinating. There's the literal hardware side, where you'll hear Mark explain how the Internet Archive goes through pallets of hard drives. And then there are the choices that go into preservation: not everything necessarily *merits* preserving, and not everything is technically accessible, especially now as more of the online world moves to private platforms and communication.

Making those choices — not just preserving the internet, but *curating* it — is a complicated proposition that hits on every *Decoder* theme there is. The idea of running a library that stores the internet's history is a puzzle worth solving.

Decoder with Nilay Patel /

A podcast from The Verge about big ideas and other problems.