



The National Library of Ireland has selectively archived the web since 2011. Web Archiving ensures the preservation of Irish websites and social media in the National Collection.



The National Library of Ireland is the library of record for Ireland, our collections are permanent, and accessible by all. Established in 1877, our mission is to collect, protect and make available the recorded memory of Ireland and share the story of Ireland with the world through our unique collections. We do this for the benefit of the public.

# What is web archiving?

Web archiving is the process of collecting portions of the World Wide Web, preserving the collections in an archival format, ensuring that it remains accessible in the future. The archived website or social media account forms part of the national collection and is made available for research and private study.

# Why archive the web?

The web plays an integral role in all aspects of modern life. It has become the key place to find and share information, but it changes rapidly. It is estimated that a web page changes, moves or disappears in 90-100 days. Without actively carrying out web archiving, information about our society and culture will be lost at an unprecedented rate.

# How is the web archived at the NLI?

The NLI carries out web archiving with a technical partner. There are a number of steps involved in this process:

## **Step 1: Selection and Identification**

Websites and social media accounts are identified for inclusion in the web archive. These "seeds" are selected in line with our collection development policy.

## **Step 2: Notification**

Website and social media account owners are notified of the NLI's intent to include them in the web archive. They can decline at any point to being included in the web archive.

## **Step 3: Archiving**

The NLI schedules a web crawler to archive the website. This crawler downloads code, images, documents and other files that are essential to reproducing the web content as similarly to its original form as possible.

## **Step 4: Quality Assurance**

All archived websites and social media undergo QA to ensure the highest quality of capture possible.

## **Step 5: Access**

Metadata is attached to each archived website and social media account before it is made available from <https://archive-it.org/home/nli>

# Technical Limitations

Web archiving is an emerging technology which is constantly responding to new developments in web design. There are some known limitations and while this list is not exhaustive, it is indicative of the type of websites or content that cannot be archived.

- Search boxes
- Drop down boxes or menus
- Content behind paywalls or log ins
- Intranets
- Websites that rely heavily on Javascript or Flash
- Websites using *Post*
- Google Maps
- Social media feeds or hashtags
- Facebook

# Collaborating with the NLI

The NLI is committed to working with subject specialists, academics and community groups to enrich and develop the NLI web archive. We also welcome suggestions from the public. Below are some guidelines that might inform the collection of websites and social media accounts that you suggest.

What we *can* get:

- Websites
- Twitter Accounts
- Instagram Accounts
- Youtube Accounts

Please note:

The website/social media account must be:

- Irish or of Irish interest
- Publicly accessible
- Free to view

After the websites/social media accounts are selected, inclusion in the web archive is subject to the notification process. Websites may be excluded if they do not meet technical requirements.

Suggestions of websites can be emailed to [webarchives@nli.ie](mailto:webarchives@nli.ie)

# Frequently Asked Questions

**HELP! The archived website I'm trying to access won't display properly!**

It is important to bear in mind that although an archived website looks like a live website, there can be some differences in how they behave.

**Why are some archived websites incomplete or displayed incorrectly?**

Unfortunately there are limitations to what can be captured during the web archiving process.. Embedded videos, Flash content and other dynamic features like Google Maps may not capture.

**Why can't I access all links on the website?**

Sometimes links on a website will not be captured because they were 'out of scope' for the web archiving process. This means they were external to the website URLs specified for capture. For example if the main NLI website [www.nli.ie](http://www.nli.ie) is specified for a crawl then the page <http://www.nli.ie/en/list/current-events.aspx> will be captured but <https://blog.nli.ie> will not be crawled because its URL is out of scope.

# Frequently Asked Questions

## If an archived website is displaying in mobile view

Try adjusting it by moving the frame to the left.

## An archived website won't load

This happens when some websites rely heavily on pop-ups or advertising banners. Try installing an ad blocker as an extension for your web browser. This should help the issue.

## There are lots of embedded videos I can't access! Why is this?

Embedded YouTube videos cannot be archived at present. This will result in an empty black box or an error message. We hope that in time, the technology will improve that will allow us to capture these features.

## There is a Twitter feed on the archived website, but it wasn't captured?

Embedded Twitter feeds cannot be captured when archiving a site. Often a live Twitter feed will appear and this will not reflect the time and date that the main website was archived.



