

Weaving the net: developing selective digital collecting principles at the National Library of New Zealand
Rhonda Grantham
17 December 2024

Introduction

All works published in Aotearoa New Zealand are in scope to be collected under the legal deposit provisions of the National Library of New Zealand (Te Puna Mātauranga o Aotearoa) Act 2003, including works published by New Zealanders on websites and hosted on overseas self-publishing platforms.

The National Library of New Zealand (Te Puna Mātauranga o Aotearoa) Act 2003¹ and the National Library Requirement (Electronic Documents) Notice 2006² authorise the National Librarian to make a copy of any electronic documents in scope for legal deposit.

Unlike the situation with physical format documents, where the obligation is on the publisher to deposit copies of all eligible documents, the copying of eligible electronic documents by the Library has always been intended to be selective. This has been the operational principle of the National Library's selective web archiving programme since its inception, but to date has not been deliberately applied to all other areas of digital legal deposit collecting.

Context

The Library does not have unlimited resourcing to collect legal deposit material, and available resources are further stretched by collecting in both physical and digital formats where material is made available in both. To build a digital collection which is representative of New Zealand's published documentary heritage, we need to be intentional and transparent in where we focus our resources.

Selective digital collecting acknowledges the reality that we cannot comprehensively collect online content. Having publicly available principles communicates that we will be making choices and provides guidelines to explain those choices.

We are seeing in our collecting that we are collecting very similar content in particular areas, and we would like to collect a representative sample of such publications. From this we have developed a principle that supports us to limit our collecting in areas of high coverage.

We have considered whether we are able to provide a general definition of sample and representation. At these early stages we have found that we can only provide a meaningful definition of what a sample is, once we have done an in-depth analysis of what has been published in a particular area.

The expansion of publicly available digital content has blurred the boundaries between the National Library's legal deposit collecting and the collecting scope of other institutions. Considering our place within an international, national and institutional knowledge system,

¹ National Library of New Zealand (Te Puna Mātauranga o Aotearoa) Act 2003 (<https://legislation.govt.nz/act/public/2003/0019/latest/DLM191962.html>).

² National Library Requirement (Electronic Documents) Notice 2006 (<https://legislation.govt.nz/regulation/public/2006/0118/latest/DLM381515.html>).

we have developed principles with the intention to reduce duplication of collecting to work towards our goal of sustainable collecting.

The Principles

We have developed nine selective collecting principles. These will provide a framework for more detailed operational guidelines for Library staff. To inform publishers and library users we plan to publish the principles on our website. This also indicates to groups we have not previously worked with that we have an intention to address gaps in our collection.

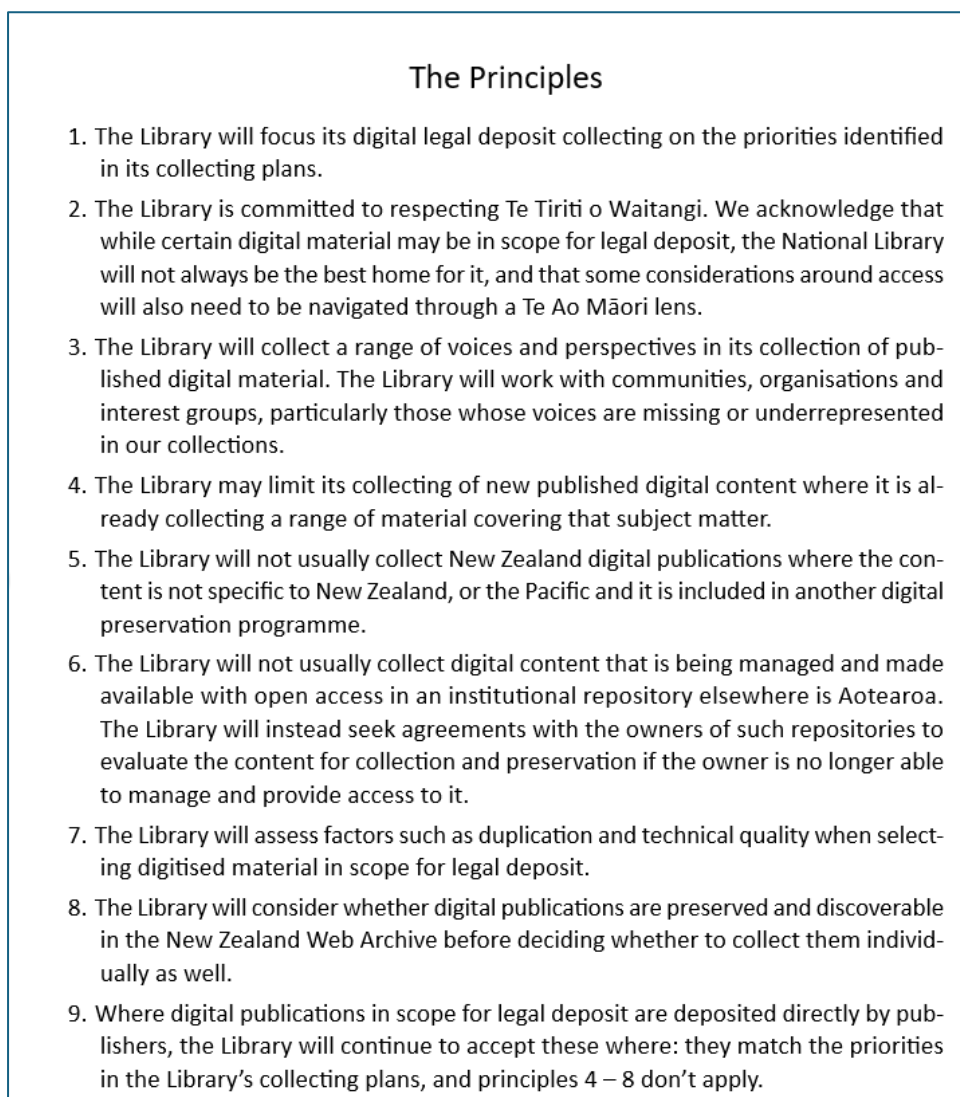


Fig. 1. Nine selective digital collecting principles.

Our first principle is that our collecting will be based on the priorities identified in the Library's collecting plans. At present the Library has a mix of subject and format-based collecting plans which signal ongoing and emerging priorities.

The second acknowledges that we will seek to work with Māori/indigenous content creators in a way that considers and is respectful of different cultural practices around the production, ownership, and use of information.

The third principle flags our intention to build a collection that is representative of New Zealand's population, "particularly those whose voices are missing or underrepresented in our collections". We are already skilled in developing and maintaining relationships with publishers, and proactive outreach into communities will be a further extension of this.

An example of the Library using Principle 4, to "limit its collecting of new published digital content where it is already collecting a range of material covering that area", has been an assessment of economic insight publications issued by New Zealand banks. In our assessment we not only looked at the bank publications, we also considered if similar content was issued in other sources, for example, the website harvest of the New Zealand Reserve Bank, or the business pages in newspapers. Our assessment identified some key resources we had been missing and guided us about where we could reduce our collecting of very similar publications.

One way we are working towards our goal of sustainable collecting is by reducing duplication of collecting with other institutions both in New Zealand and overseas. To that end we have developed three principles:

- The Library will not usually collect New Zealand digital publications where the content is not specific to New Zealand, or the Pacific and it is included in another digital preservation programme.
- The Library will not usually collect digital content that is being managed and made available with open access in an institutional repository³ elsewhere in Aotearoa.
- The Library will consider whether digital publications are preserved and discoverable in the New Zealand Web Archive⁴ before deciding whether to collect them individually as well.

The principle that considers whether a publication is in the New Zealand Web Archive looks toward further enhancements of website harvest discovery technology, and developing our practices for describing parts of websites, and resources within websites. This is a rich area for future collaboration with our cataloguing/metadata and technical colleagues.

Our collecting plan⁵ confirms that resources digitised by New Zealand Libraries and other institutions and made available to the public on the Internet are in scope for legal deposit and may be collected by the library. To provide library staff and content creators with more guidance we have developed a principle that says we will assess factors such as duplication and technical quality when deciding whether to collect digitised collections in scope for legal deposit.

We will continue to collect what has been deposited, but we have the option to decline material when other principles apply, or, if the content does not match with the priorities in our collecting plans. This principle was added to avoid risks to our relationships with publishers who have made the effort to proactively deposit. There is a risk that continuing to collect comprehensively what is deposited could skew areas of the collection. For example, high-volume publishers of a particular genre, or high-volume publishing models such as the use of generative AI to create publications, potentially flooding us with generic content. To

³ "An institutional repository is an archive for collecting, preserving and disseminating digital copies of the intellectual output of an institution, particularly research institution" (https://en.wikipedia.org/wiki/institutional_repository).

⁴ New Zealand Web Archive Te Pūranga Paetukutuku o Aotearoa (<https://natlib.govt.nz/collections/a-z/new-zealand-web-archive>).

⁵ New Zealand and Pacific Published Collections collecting plan (<https://natlib.govt.nz/about-us/strategy-and-policy/collections-policy/newzealand-pacific>).

mitigate this risk, we have refined principle nine to provide us with some critical flexibility to limit our acceptance of direct deposits.

Measurement

Seeking to build a digital collection which is representative of New Zealand's published documentary heritage, demands the question, how do we measure the effectiveness or impact of our efforts? Currently our impact is measured quantitatively (i.e. the number of items added to the collection), not by qualitative measures. We still need to test whether it is possible to make this shift.

Conclusion

There is a risk in perception that by selective collecting we are making value judgements and preferring one resource over another, which is at odds with the democratic underpinning of building a collection by legal deposit. However, we would argue that there is a risk in our current model. That in accepting what is offered or available, by doing what is expedient, we maintain biases in our collections. Deliberate selective collecting, supported by a range of engagement strategies with creators, publishers and communities, could mitigate the risk of biases and blind spots in our collecting. It should also enable us to create much-needed capacity in our collecting work, allowing us to explore emerging formats and publishing models and develop collections that reflect the diversity, complexity and range of voices, experiences, and events that shape New Zealand.