

The Ryerson Index

INDEXING FROM NEWSPAPER WEBSITES

Table of Contents

Introduction	2
Saving the Notices	4
Ryerson Notice Repository (RNR).....	4
Updating the RNR	5
The Indexing Process.....	5
1. Gathering the Notices.....	5
2. Indexing the Notices	5
Naming Saved Notices	6
Saving Notices from the NewsCorp Website.....	7
The Windows Snipping Tool	11

Version 4 – 15 July 2020 – rewritten to include the new procedure for NewsCorp notices

Version 5 – 24 Aug 2020 – include News Corp’s “Digital Only” classification.

Introduction

For more than 21 years, Ryerson had resisted using newspaper websites as a source for indexing. Wherever possible, we used either a copy of the printed paper, or a .pdf of the paper.

There were two reasons for this.

The first reason was that we found, when checking the print version against the newspaper website, that there were discrepancies between the two sources. Sometimes notices which appeared on the newspaper website were not included in the print edition. Sometimes notices in the print edition did not appear on the website. There appeared to be no logical reason why some notices were missed, but they were – hence the website was considered an unreliable source when compared to the print edition (which, unlike the website, would still be available many years hence.)

The second reason developed slowly, as newspapers began to concentrate more and more on their websites. We found that notices were being listed on the website with the “published” date shown being a date on which the paper was not printed. We tracked this down to the newspapers (most particularly APN papers) considering that a notice was “published” when it was added to the website, not when it appeared in print. Because we could not rely on the publication date being correct, we could not index any of these notices.

That situation remained the case until April 2020.

The COVID-19 worldwide pandemic caused huge disruption to most countries. In Australia (and many other countries), businesses were forced to close, personal movement was severely restricted, and working from home became the norm for a large number of people - all with the aim of stopping the spread of COVID-19.

Large-scale business closures, and the reduction in spending by the general populace, had a severe impact on the newspaper industry. Almost overnight, the advertising income stream for most publications dried up. During the second week of April 2020, the two major newspaper publishers (News Corp and Australian Community Media (ACM)) announced the suspension of many of their smaller mastheads until at least the end of June. The stated intention of both publishers was that the "suspensions" would only be temporary.

News Corp was the first to break ranks, announcing on 28 May the closure of a number of regional and suburban publications, and the switch of many more to become digital only. Thus a very large number of print mastheads have ceased to exist, with in many cases their final print edition being in the second week of April 2020.

At this time, ACM has re-opened many of its suspended papers, with others scheduled to re-open between now and September 2020. **So far, no ACM paper has been converted from print to digital-only.** For that reason, this document concentrates on saving and indexing digital notices from NewsCorp only.

To compensate for the closure of so many papers, the Ryerson Committee took the decision that, for digital-only papers, we would

1. Save each notice published on the newspaper website in the RNR, and
2. Index from the saved notices.

Our reasoning was that this switch from print to digital was only the beginning, and eventually most papers would follow suit. If we stuck to our previous position of not indexing from a newspaper website, then Ryerson would wither and die, as the number of notices appearing in print gradually reduced to nil, or very close to nil. We considered our action of saving each notice to the RNR would provide us with a source which we could produce on request from a researcher if, sometime in the future, the particular notice had disappeared from the newspaper website. It could even be possible that sometime in the future, the RNR would be the sole repository of some notices.

With the conversion of many NewsCorp papers from print to digital-only, the structure of their online database of notices was also changed. Notices were now stored in two different ways:

1. For papers still in print, each notice showed the paper and the publication date, eg

Publication: Herald Sun Date listed: 15/7/2020

2. For papers which became digital-only, the publication is identified as "Digital", eg

Publication: Digital
Date listed: 24/8/2020

It is now possible to select all digital-only notices.

Saving the Notices

To control what could potentially become a very large collection of notices, some standardisation of the process is necessary, both in how the notices are saved, and how they are stored.

Ryerson Notice Repository (RNR)

The RNR is a folder stored in Dropbox and accessible to all indexers on a read-only basis. It contains all the notices we have saved over the years – approximately 2 million – and is also the place where notices saved from newspaper notice databases will be stored.

The RNR has a rigid structure of folders and sub-folders. The **standard** three-level folder structure is:

Newspaper Name
Year
Month

As an example,
Sydney Morning Herald
2019

2019-01
Notices for each day of Jan 2019
2019-02
Notices for each day of Feb 2019
.....
2019.12
Notices for each day of Dec 2019
2020
2020-01
Notices for each day of Jan 2020
etc

To cater for digital notices, where we have no newspaper name, we have created a “bucket” called “NewsCorp Digital Notices”, which will contain ALL digital-only notices from the NewsCorp database.

The folder Structure will be:

NewsCorp Digital Notices
Year
Month
Day
All notices for a single day

This structure will enable indexing to be carried out in the normal way, ie by day, because all notices for one day will be grouped within the same folder.

Updating the RNR

Because indexers only have read-only access to the RNR (at 250+Gb, it will not fit into the usual free Dropbox space allocation), then we needed to create a structure to allow indexers to “update” the RNR.

We did this by creating an RNR update folder within dropbox, called “Ryerson_Gathering”. This folder is shared by all indexers who are saving either copies of the print notices, or digital-only notices. The folder is emptied at the end of every month, to avoid it getting too large and impacting on an individual indexer’s Dropbox space allocation. The contents are copied to the RNR daily.

The Indexing Process

The entire process is broken into two stages: gathering the notices and indexing the notices.

1. Gathering the Notices

A team of seven gatherers will control the saving of notices – one person per weekday. This means the gathering can be done anytime during the week, not necessarily on the day.

An important point to note is that gathering cannot take place until after midnight on any given day – because notices are being added to the database continuously. So gathering of the notices with a Friday publication date can be done anytime from Saturday onwards – but NOT on the Friday of the publication date.

The creation of a daily folder within the Dropbox structure outlined previously is the responsibility of the gatherer. Folders are named in a standard format of yyyy-mm-dd. A gatherer MUST create a daily file for each day being gathered, regardless of whether or not there are any notices. An empty file will indicate to the indexer that a NIL file can be created.

2. Indexing the Notices

A team of seven indexers will control the indexing of digital-only notices – one person per weekday. As now, indexing can be done at a time to suit the indexer.

It is anticipated that at least some days will be done by the one person carrying out both the gathering and indexing. Where we have separate indexers and gatherers, a degree of co-ordination between the two people will be necessary so that an indexer doesn’t waste time looking for a daily folder which has not yet been created. It is the responsibility of the gatherer to agree with the indexer on a suitable schedule.

The File code for indexing NewsCorp digital-only notices is DN.

Naming Saved Notices

There are four items of information which are vital when we are currently retrieving a notice following a lookup request:

- The newspaper
- The publication date
- The type of notice (death or funeral)
- The name of the deceased

Consequently each notice should be named in the following format:

Newspaper code_publication date_notice type_surname

(ie with each field separated by an underscore)

While it may seem that the newspaper code is redundant information, given that all digital notices have the same code, it is there purely to identify the file as being a digital notice

The publication date should be in the format yyyy-mm-dd, to ensure notices are displayed within the folder in chronological sequence.

The notice type will be either D or F.

The surname should be in the same case as you would use when indexing (eg SMITHERS, de BORTOLI)

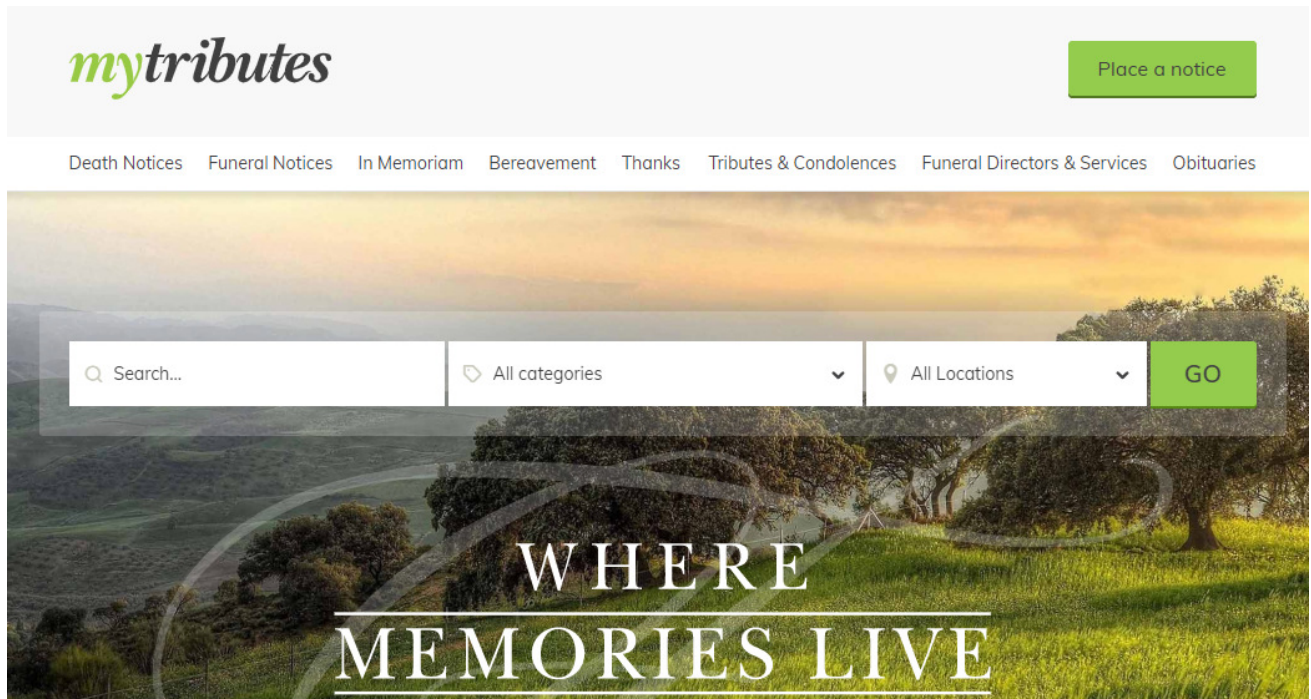
In the rare case of these four fields not providing a unique filename, add a suffix of _#n to the filename, where n starts at 2.

Some examples:

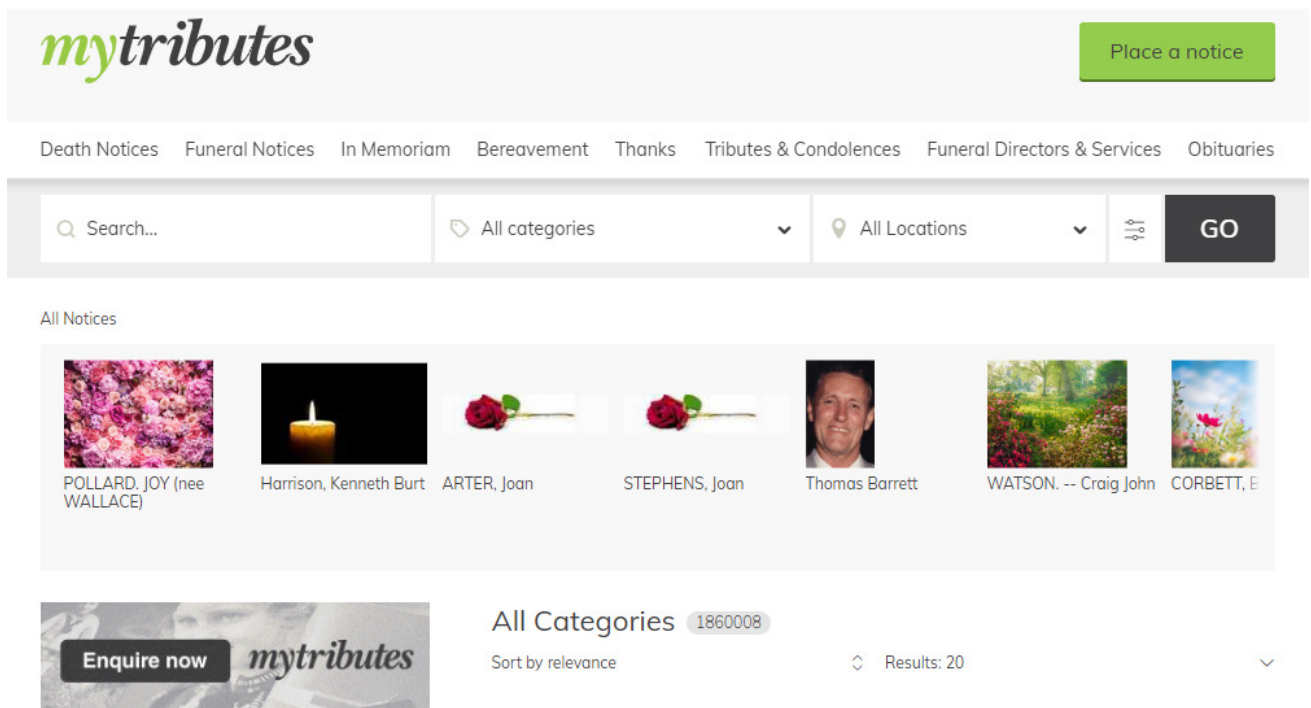
```
DN_2020-07-15_D_De BORTOLI
DN_2020-07-15_D_SMITHERS
DN_2020-07-15_F_SMITHERS
DN_2020-07-15_F_SMITHERS_#2
```

Saving Notices from the NewsCorp Website

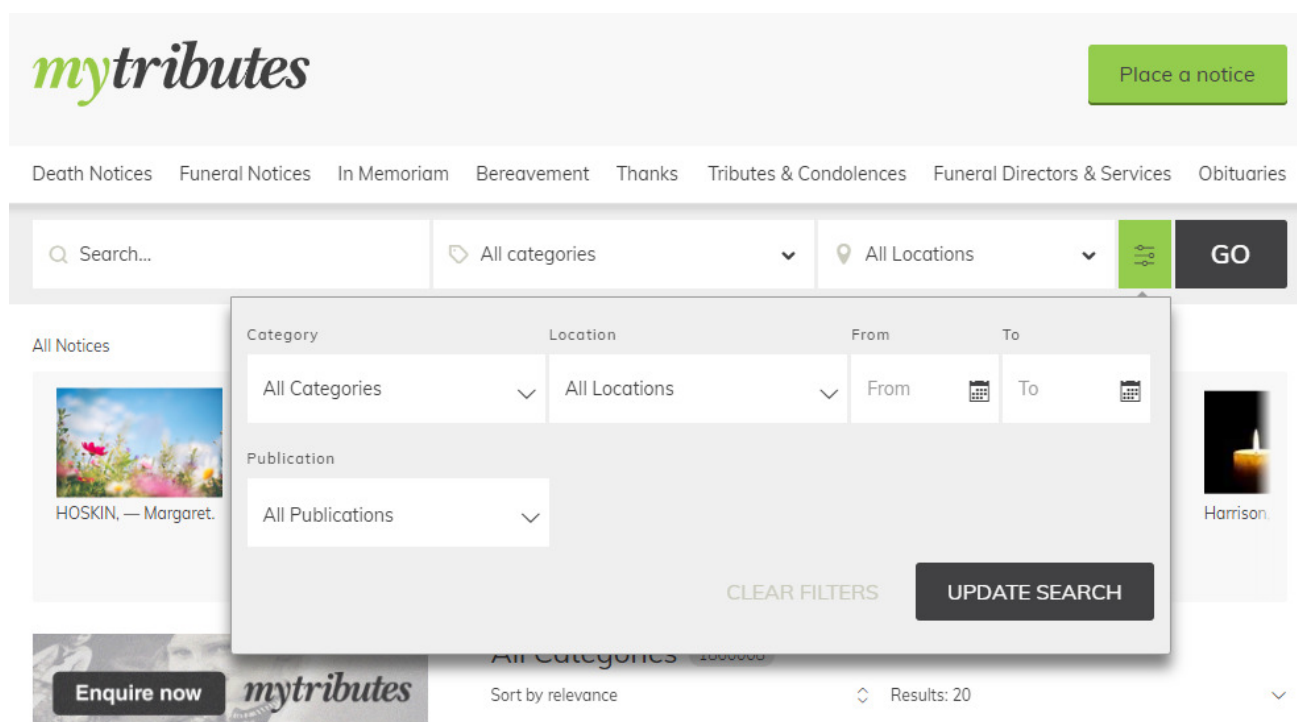
The NewsCorp notice database is found at <https://www.mytributes.com.au/> and looks like this:



The key item on this screen is the green GO button. Forget the other fields on this screen. Just click the GO button. You will see this screen:



The key to opening up the search options is the odd little symbol to the left of the GO button. Click that symbol, and the following screen should appear:



To search for notices:

1. Use the drop-down menu in the “All Categories” box to select “Death Notices”.
2. Set the FROM date and the TO date to be the same date – the date of the day you are saving.
3. In the “Publications” box, select “Digital Only”.
4. Click the UPDATE SEARCH button.

The result will be a list of digital-only death notices published on the selected date. It is now a simple task to page through the list, saving each notice.

The default number of notices to the page is 20, but you can increase this to 50 if you wish by finding the “Results: 20” icon and changing it to 50 results per page, eg

All Categories 1860012

Sort by relevance


Results: 20

When all the Death Notices have been saved, repeat the process for Funeral Notices. Click the icon to the left of the GO button, change “Death Notices” to “Funeral Notices”, check the two dates are correct, check “Digital Only” publications are selected and click “Update Search”.

This process might appear complicated at first, but after a few uses it will become second nature.

Saving an Individual Notice

Once a notice has been identified as a digital notice, it has to be saved. A sample notice could look like this:

 Gladstone, 4680

LAYER (nee Bailey) Norma Jean

Late of Katherine Court, Gladstone. Passed away peacefully at the Gladstone Hospital, Friday 10th July, 2020, aged 89 years.

Dearly loved Wife of the late Robert (Bob) Layer. Adored Mother of Ken and Kathy. Cherished Nana to Michael, Matthew and Tehlia. Loved Sister and Sister-in-law to *Wal (dec'd)*

[More...](#)

Publication: Digital Date listed: 15/7/2020

The first thing to do is to expand the notice by clicking the “More” link. This gives an expanded notice like this:

[All Notices](#) > [Funeral Notices](#)

LAYER (nee Bailey) Norma Jean

Date listed: 15/7/2020

Late of Katherine Court, Gladstone. Passed away peacefully at the Gladstone Hospital, Friday 10th July, 2020, aged 89 years.

Dearly loved Wife of the late Robert (Bob) Layer. Adored Mother of Ken and Kathy. Cherished Nana to Michael, Matthew and Tehlia. Loved Sister and Sister-in-law to Wal (dec'd) and Betty, Eddie and Irene (dec'd), Betty and Trevor (dec'd), and Aunty to their families.

Relatives and Friends are respectfully invited to attend Norma's Funeral Service. Due to covid 19 please contact Kathy 0448 010... or Betty 0437 727... to confirm funeral details.


In the care of
GLADSTONE VALLEY FUNERALS
Ph: 07 4972...

“Gladstone's Local Funeral Home”

Condolences: GladstoneValleyFunerals.com.au

Publications:

Gladstone 4680



This notice should now be saved, from the top line which identifies it as a Funeral notice, to the placename at the top of the map. That is all the relevant information of use to a researcher.

[All Notices](#) > [Funeral Notices](#)

LAVER (nee Bailey) Norma Jean

Date listed: 15/7/2020

Late of Katherine Court, Gladstone. Passed away peacefully at the Gladstone Hospital, Friday 10th July, 2020, aged 89 years.

Dearly loved Wife of the late Robert (Bob) Laver. Adored Mother of Ken and Kathy. Cherished Nana to Michael, Matthew and Tehlia. Loved Sister and Sister-in-law to Wal (dec'd) and Betty, Eddie and Irene (dec'd), Betty and Trevor (dec'd), and Aunty to their families.

Relatives and Friends are respectfully invited to attend Norma's Funeral Service. Due to covid 19 please contact Kathy 0448 010... or Betty 0437 727... to confirm funeral details.

In the care of

GLADSTONE VALLEY FUNERALS

Ph: 07 4972...

"Gladstone's Local Funeral Home"

Condolences: GladstoneValleyFunerals.com.au

Publications:

Gladstone 4680

Sometimes a notice will be too long to fit on your screen. In this case there are a couple of alternatives.

1. Reduce the screen size slightly (CTRL-Scroll) from 100% to say 90%.
2. If the notice is still too long, save to a Word document – a two-column table with narrow margins will usually do the trick. Save the Word document as a .pdf.

The Windows Snipping Tool

By mid-2020, the majority of indexers are using Windows 10, which has a very simple snipping tool as part of the standard OS. Where possible, indexers are encouraged to use this to save notices – but if you have another preferred snipping tool, keep using it – we are concerned about the final destination, not the journey!.

The simplest way to invoke the Snipping Tool under Win10 is to press Windows+Shift+S keys simultaneously.

An alternative method is to save it to your Windows taskbar – that way, it is accessible when required via a single mouse-click. To do this,

1. Click the Windows Start icon (bottom left)
2. Scroll through the list of apps until you find “Snipping Tool” or “Snip & Sketch”
3. Right-click on the app name
4. Select “pin to taskbar” (if not present, select “More”, then select “pin to taskbar”)

You will now see the little circular snipping icon in your taskbar. It looks like this:



To save a notice using the snipping tool,

1. Locate the notice
2. Click the snipping tool icon in the taskbar, or use the Windows+Shift+S keys
3. Click “New” – the window will go grey, and the cursor will change to a +
4. Outline the notice to be saved
5. Click the “Save As” icon
6. Select the destination folder, and a name for your file
7. Click “Save”
8. Close the snipping tool