The South Australian Red Cross Information Bureau

Andrew Piper
Manager, Online Projects
State Library of South Australia

Abstract:
The State Library of South Australia (SLSA) is honouring the memory of those who served at the front lines on World War 1 and on the home front, through a series of projects. The principal project is the creation of a new web resource of the records of the South Australian Red Cross Information Bureau (SARCIB). The web resource will utilise volunteers and provide opportunities for community created content. By leveraging existing tools such as MARC, the project will lead the way for future delivery of our collections and provide an opportunity for SLSA to apply Linked Open Data. It is anticipated this resource will be available early in 2015.
Introduction

From 1916 until 1919, the South Australian Red Cross Information Bureau (SARCIB) performed the service of undertaking research into 8,033 enquiries from family and friends of missing Australian Imperial Force (AIF) personnel who fought in World War 1 (WW1). An envelope was created for each enquiry containing:

- information about each soldier enquired upon
- letters from the requesting family
- eyewitness statements about the soldier’s last known whereabouts
- and, in most cases, the resulting information on the fate of the soldier, expressed to the enquirer.

The State Library of South Australia (SLSA) is undertaking the phased development of a web resource for this information that will enable other institutions, researchers and family historians worldwide to interrogate, contribute to and harvest the database, and highlight South Australia’s contributions to WW1.

Along with providing SLSA with an opportunity to be a part of the international centenary of WW1 programs, it will contribute to the fulfilment of SLSA’s Strategic Plan 2014-2017 (SLSA 2014) by:

- delivering a new online family history resource, that helps to reposition the SLSA family history service
- assisting current and future generations of South Australians to access and learn from the South Australia collections
- providing opportunities for South Australians to know and value their heritage
- creating another data set that contributes to the South Australian Government’s desire to make its data available to its citizens (Weatherill 2013).
- providing opportunities for volunteering.

It will also:

- lead the way for future delivery of the Library’s SA Memory program (SLSA’s online delivery of South Australiana collections)
- contribute to the development of a content release plan for SLSA
- enable SLSA to explore platforms for new content management systems
- provide SLSA with a starting point to engage with Linked Open Data (LOD)

A critical and inadvertently innovative aspect is the data entry and database.

SLSA will use its volunteer program to create a team of volunteers who will work under controlled conditions onsite to extract details from copies (PDFs) created from
the records and key data into a spreadsheet. The spreadsheet in turn will be converted to MARC records using the MarcEdit open source program and loaded to SLSA’s Library Management System (LMS), Millennium. The capture of details to a MARC record is an interim stage enabling data entry to commence, as it is our intention to export the details from Millennium to a purpose built web resource during 2014.

The project is managed and led by Online Projects, a small SLSA team fluctuating between 1-2 people (depending on funding), that provides project management services to SLSA. Planning commenced in February 2012 and the project plan and schedule were signed off in December 2012, after much discussion about funding and resourcing. Midway through this process, the team was advised that seeking external funding for staff and other resources would not be supported and that the project had to be achieved from internal budgets with existing resources.

The project will be achieved in five broad, overlapping phases:

Phase 1 - December 2012-December 2013
- Stakeholder identification and management
- Development of business requirements
- Digitisation of records

Phase 2 - July 2013-October 2014
- Develop database requirements and data structure for data entry
- Engagement and training of volunteers
- Data entry, conversion of spreadsheets through MarcEdit to MARC and then loaded to Millennium

Phase 3 - March – November 2014
- Identification and procurement of Content Management System (CMS)
- Building of web resource in CMS, including the introduction of 3rd party resources and community engagement elements
- Migration of metadata to new CMS web resource
- Verification of associated locations and addition of geo-codes

Phase 4 - January 2015
- Go Live
- Promotion

Phase 5 - June 2014-April 2015
- Exploration and potential implementation of LOD
Review requirement to keep metadata in Millennium

SARCIB forms part of a wider Centenary of ANZAC program developed by SLSA.

History

In September 1915, it was reported that a Bureau of “like character” to those in London and Cairo was to be opened in Adelaide (The Advertiser 1915, p.4).

By early 1916, it was announced that:

Sir Josiah Symon, K.C. has completed the organisation of an information bureau in Adelaide for the purpose of supplying information to relatives and friends in the State concerning the wounded, sick and missing soldier. (The Advertiser 1916, p.8).

The legal profession bore the cost of running the Bureau and the co-ordination of activities (The Advertiser 1916, p.8). A ‘staff of ladies’ giving their services ‘free and freely, being glad to do their bit for the Empire’ undertook the office work (Saunders 1917, p.8).

A series of notices placed within SA newspapers (Figure 1), invited those seeking news of the sick, wounded or missing, to contact the Bureau.

Figure 1 (Edmunds 1916, p.6)

The SA Bureau co-ordinated the enquiries between the military, families and other Bureaux nationally and internationally, analysing the results of those searching the hospital wards and convalescent homes for eyewitnesses who could tell them something about the missing. Often, the frank eyewitness statements show the realities of war, and sometimes they were included in the letters that the families ultimately received (Schneider 1997). Information on prisoners of war was gathered through ‘neutral intermediaries’ (Schneider 1997, p.229).

An SLSA index for the collection records the Bureau’s explanation of its process:

When an inquiry is received a card is immediately opened with the full particulars of the soldier enquired for, on which card is also recorded from time to time a précis of the correspondence and results of enquiries made. Attached to and filed with this card is a carbon copy of
all correspondence with a duplicate of the original enquiry instruction block. This is all placed in a packet carefully numbered and indexed. On receipt of all official lists, Casualty, Hospital, Returned Soldiers etc., so far as they concern soldiers enlisted from this State, are entered upon a card system with a cross reference thereon to the enquiry packet, if any, concerning such soldier’ (SLSA, n.d, p.1.).

In 1919, after conducting 8033 enquiries, the South Australian Red Cross Information Bureau closed its doors and donated its records to SLSA (The Register 1919, p.4).

SLSA’s acquisition is most likely attributable to Sir Josiah Symon, who was instrumental in the creation of the SA Bureau and a significant benefactor for SLSA.

SLSA’s holdings are unique in Australia. While the Australian War Memorial (AWM) holds copies of most of the Australian enquiry records from the London bureau and may include some of the South Australian information, its collection does not include the letters from families or friends that initiated the search. SLSA’s collection includes all the original documents from the SA Bureau, including the enquiry letter, making this a more complete representation.

Although available to the public since receipt, access to the enquiries has been through a microfiche version of the original card index (Figures 2 & 3) which is arranged by surname of the soldier enquired upon. Each card records a series of numerical codes that, depending on their location on the card, indicate regimental rank or number, name, unit, enquiry file number, and links or references to other series within the records. Interpretation of the handwritten numbers is onerous at the best of times and made more so by use of the microfiche.

Figure 2
The proposed web resource aims to enable researchers to search across a range of indexed metadata and retrieve newly digitised scans of the original packets and, for the first time, enable searching by a range of characteristics and by eyewitness names.

The original packets

Archival records such as these are arranged at SLSA in five categories: business (BRG), private (PRG) and society record groups (SRG). Single items such as diaries, letters and maps are held as separate sequences. The records of SARCIB are categorised as SRG 76, and made up of a number of series. Series 1 of this record group comprises the packets (Figure 4) relating to enquiries into missing soldiers and are the subject of the proposed web resource. These packets may contain some or all of the following:

- a summary of the enquiry - Form C. (Figure 5)
- the Red Cross Bureau's Searchers' report
- letters sent to the Bureau by relatives or friends (Figure 6)
- eyewitness accounts (Figure 7) by men who knew the soldier concerned or statements by men who served in the same unit or the same action as that soldier. These statements may include circumstances of the death or wounding, or the place of burial and physical description of the soldier
- the Bureau's replies to enquiries
- correspondence between the Bureau and the wounded or missing person
- correspondence between the Bureau and the AIF (Australian Imperial Force).

Each packet is numbered consecutively from 1 to 8033.

Because the Bureau commenced operations in January 1916, the gathering of eyewitness statements for the Gallipoli campaign did not occur until well after the ANZAC evacuation at Gallipoli in December 1915 (Bean 1938, pp. 853-910) and, in general, these records contain less information than those of the Western Front era, when statements and reports were more contemporary. Form C (Figure 5) summarises the events regarding the soldier in question and is the basis for the data structure described later.
FORM C.

No. 702.
Surname. 
Christian Names. Brunel John
Date. 15/4/16

Rank. Private
Regiment, &c. 10th Battalion

File No. G/L 43-14.

OFFICIAL REPORTS:

Officially reported


Information from Major Hands - Pen Records.

Melbourne killed in action 3rd May 1916. Result by Court of Inquiry.

Other Information:

Left Australia for Egypt in 20th Oct 1914 by steamer.
Not heard from since 25th April 1915 (died Feb 1915 at Sallpoli).

Held for Trusty Agricultural as prisoner.
Not been reported as dead.

Weekly to be a prisoner.

Inquirer's Name and Address.
Mr F. H. Hook
15, Oxford Street
London

Instructions:

Write to Military authority in Red Cross Society in Egypt.

Call Again.
YES. 
NO.
Cost of Cable, &c.

Figure 5
Anne St.
Smynhon
June 26th 1915

The Superintendent
Red Cross Bureau
Adelaide

Dear Sir,

I am anxious to know if you can give me any information as to the whereabouts of my Son as I have not received any letters from him for six months. In the last, he was in Egypt; after leaving Gallipoli, he stated that probably I would not receive any news of him for six weeks. It is now that number of months and he used to write so often.

I saw in the Register the Bureau’s kind offer to help those seeking information. Trusting you will be successful and accept the thanks of

Yours faithfully,
A Webbing

P.S. My Son’s last address enclosed.

Figure 6
Phase 1 December 2012-December 2013

Stakeholders

A project of this nature has both internal and external stakeholders. The SA Branch of the Red Cross was consulted at the beginning of the project, and is one of our conduits for community engagement programs in the coming years, particularly with regard to the 2014 celebrations commemorating 100 years of the Red Cross in Australia.

The Returned and Services League South Australia Branch (RSL) and SLSA regularly liaise to share developments, with regard to their ambitious plan to create a web page for every service man and woman (with an SA connection) who went on active service for every conflict. There are also plans to link those names with data sources from Australian War Museum (AWM), National Archives of Australia (NAA), Commonwealth War Graves Commission (CWGC) and Veterans SA, as well as State Records South Australia (SRSA) and SLSA.

As our project activities gain profile, organisations and individuals make contact with SLSA to share their plans, projects and desires. As of November 2013, no state-wide organisation of centenary programs exists, and neither the project team nor SLSA is currently resourced to take that role, although we do recognise our public position and try to connect people where possible.
SLSA has made contact with the *International Encyclopedia of the First World War* to promote the project and to determine potential relationships and as of November 2013, this will most likely result in SLSA supplying images for the Encyclopedia’s researchers to use.

SLSA has been successful in making connections with *World War Onelink* and the Imperial War Museum’s (IWM) Centenary partnership program.

SLSA is a member of National and State Libraries Australasia (NSLA), a coalition of libraries with aims to work collaboratively and share information. It is expected that 2014 will see more discussion between these agencies with regard to WW1.

In April 2013, SLSA project staff also visited AWM, NAA and NLA to share project information and establish relationships.

### Development of business requirements

The consultations confirmed SLSA’s approach to produce a web resource that will comprise three searchable elements:

1. The original “packets” of information related to the 8,033 requests made through the Bureau, in a digitised format so that the community can view and download copies of the original documents, providing the ability to search across a range of indexed metadata.

2. A place for members of the community to contribute additional information. This may range from family photos of the subjects to personal reminiscences of loved ones; transcription of letters within the packets; tagging.

3. Content from third-party sources, such as the NAA, CWGC, the IWM and the National Library of Australia’s (NLA) Trove, that directly relate to the soldiers who are the subjects of the Bureau’s enquiries.

The web resource will also:

1. Be constructed to allow data to be shared
2. Provide a history of SARCIB and relationship to the AWM collection
3. Provide the ability to engage in LOD
4. Provide curated stories for some of the soldiers contained within
5. Provide links to other relevant collections.

### The Packets - Digitisation

The packet content is paper based and is a mix of handwritten and typescript documents. Most of the typescript is the carbon copy generated at the time, and the handwritten material is generally ‘Form C’ (Figure 5) and the original enquiry letter. Digitisation of the 8033 packets was completed by September 2013.
Prior to scanning, the packets were assessed by volunteers, identifying missing boxes or packets, double sided documents, conservation requirements, and ensuring the chronology of the packets was in order. This allowed reformatting staff to concentrate on the quality of the scans rather than assessing physical attributes.

Fragile documents were placed in a Mylar sleeve with a backing sheet prior to feeding through the document scanner, thereby eliminating reflection from the Mylar, as well as supporting the document. It also increased legibility for the more transparent and fragile pages by providing a solid, opaque background.

While this addition to the workflow slowed the scanning progress, prevention of damage through handling was paramount. The Mylar sleeve also provided support for some damaged documents, thereby avoiding the need to wait for conservation.

The project was fortunate to have one staff member assigned to digitising 4 days per week and to be the first opportunity for SLSA to use the new high Kodak i1440 document scanners, which simultaneously captured and output both uncompressed TIFF, for preservation, and searchable multipage PDF for display.

The Kodak CapturePro software also applied Optical Character Recognition (OCR) during the scanning process. This will enable the keyword searching of the typescript documents delivered through the new web resource.

PDF sizes were tested for speed of download before arriving at a file size that did not cause problems for the end user. This was achieved using Adobe Acrobat to resize the PDFs to a file size of approximately 2MB, well within SLSA’s preferred size for publicly-downloadable content, while also maintaining the quality and legibility of the documents.

Sorting and reformatting commenced in early December 2012, and was completed ahead of schedule in September 2013.

**Phase 2 July 2013-October 2014**

**Database requirements**

One requirement incorporated into the planning was the possibility that the details captured by those doing data entry would also be added to Millennium, regardless of where the metadata and data was going to be stored.

SLSA holds published and unpublished (archival) collections, both of which are managed through Millennium. Archival records are described through a series of MARC bibliographic records, and hierarchy is achieved through a series of linkages between these records, based on the unique archival identifying number (for example SRG 76/1/116).

We initially explored keying information directly to Millennium; however, because the Millennium interface did not sufficiently meet the desired discovery requirements and would not allow for the crowd-sourcing or harvesting elements, it was not seen as a suitable solution at that time.
While there is a possibility of funding to contract a web designer later to create a search interface and delivery screens for public use, there was no funding available during phases 1 and 2 of this project to build the database. The requirement to commence data entry grew and the time constraints are clear (delivery in early 2015).

Throughout 2013, Online Projects was also leading other projects, such as the creation of The Bradman Collection app (SLSA 2013), released in November 2013, for the collection SLSA holds of Sir Donald Bradman’s personal memorabilia, and the continued development of delivery of digital assets through the catalogue (in particular, photographic albums).

As a result of these projects, there had been a review of practices within SLSA regarding the description of pictorial collections, for which there was agreement for a consistent, structured use of some of the MARC 21 500 fields, in particular the 598 field. This repeatable field is now used to record the content of multiple photographs attached to one bibliographic record. For example, one bibliographic record is used to describe one cricket bat in the Bradman Collection, the 598 field being used to describe each photograph of the bat, for example front, side, back (Figure 8); or one bibliographic record describes an album of photographs, and repeatable 598 fields contain the captions for each photograph in that album.
The content of the 598 field is delimited using Library specific subfields separated by delimiters and subfields, which allows scripts to be written to extract the data following each delimiter/subfield in any combination necessary.

This will enable us to use bibliographic records as the basis for harvesting metadata from Millennium and re-use it in other systems to deliver new discovery services, including apps.

This change led to the realisation that a MARC record in the LMS might be the best way to proceed for data entry for this project. Millennium records could now allow relationships between records, and between metadata in records, to be expressed and interconnected, and we could always export the metadata to another database if
necessary. Archival items like an SRG 76/1 packet need to be considered as part of larger structures in terms of their contextual relationships.

The 598 MARC field has been applied to individual enquiry packets within the Red Cross records, where one new bibliographic record is created for each soldier enquired upon; and the 598 MARC field is used in each bibliographic record to identify eyewitness, and enquirer (Figure 9).

Parallel to this is the development of a repository at SLSA, which further assists in the representation of data and content outside the catalogue.

Capturing data in Millennium is considered an interim step towards developing the web resource as it provides the project with the ability to commence data entry in advance.

Because a separate web resource will be built and the metadata exported to it, SLSA will review the necessity of keeping this information in the catalogue in 2015.

Leveraging the use of MARC and library metadata standards for data entry has meant a cost saving which allows any project funds to be allocated to the build of the
web resource, and development of the crowd-sourcing and harvesting aspects. Work on this will commence in 2014.

Data Structure

Considerable time and debate was spent on the choice of data structure and database. It must be simple enough for volunteers, yet meet library standards; enable sharing of the data and metadata. SLSA is also desires the capacity to provide SLSA with an opportunity to explore LOD as a way of further sharing its collections.

To determine the field structure, SARCIB packets were surveyed to identify the range of information contained. This in itself has been a challenging process, as the type of enquiry, or answer to that enquiry, was so varied.

Packets range from a simple request for an address of the soldier, to a request to undertake research to find a missing soldier, find a wounded soldier, whereabouts of a burial and a host of other questions.

The results of the enquiries also vary from discharged; killed in action; died of wounds; wounded, no trace; no information available; hospitalised and/or returned to unit and/or discharged; returned to Australia; POW, POW died; repatriated POW. We will not know the full range until every packet is read.

The variations made it difficult to select a structure that would be easy for volunteers to complete, and makes it difficult to write clear rules for volunteers to follow.

In developing the data structure, and subsequent instructions for volunteers, the team undertook a month of test indexing to verify our approach, help document instructions and to identify issues with data entry.

The content ultimately selected for indexing and data entry was defined after consideration of:

- what information is essential
- what information is nice to have
- what is feasible for volunteers
- what is achievable in the timeframe
- eliminating interpretation of information contained in the packets.

Following test indexing with volunteers, the content to index was determined to consist of:

1. Soldier as subject of an enquiry – Person type A
   - Name of soldier enquired upon
   - Enquiry number
• Service number
• Rank
• Unit
• Date range of the enquiry
• NAA URL reference to enlistment papers
• Locations associated with this person, from specific (Mouquet Farm), to broader (France), to even broader term (Western Front). Geo codes to be added later.
• Burial: place of burial as recorded in the packet (if documented)

2. Eyewitness – Person type B (potentially more than one)
• Same information recorded as Person A

3. Enquirer – Person type C (potentially more than one)
• Name
• Relationship: Wife, Mother, Father, Sister, Brother, Friend, Fiancé
• Address, Geo code will be added later.

A fourth element, which described the incidents within the packet, was considered. This again was to be a repeatable field and was instrumental in leading SLSA to reconsider the MARC record and usage of the 598 field.

This element contained:
• Date of incident
• Incident (controlled vocabulary based upon survey of content): Killed in action; Missing; Wounded; POW; Repatriated POW; Hospitalized; Died; Returned; Ill; Deserted; Unknown.
• Location (as above).

Ultimately, this fourth element was removed, because identifying the information required too much interpretation by the volunteer. Further, indexing this information added significantly to the indexing time per record as we saw time jump from an average of 20 minutes to one hour. This is unacceptable, given the timeframe for delivery.

We did consider recording only the final action within the packet, rather than index multiple incidents. Again, the time taken to interpret the information was too long. For example, many last actions in the packets were dated 1919 and stated “No Trace – Germany”, whereas the preceding document recorded that the soldier was Killed in Action or Died of Wounds. Did this mean that they were now listed as missing, presumed dead?
There were many instances requiring interpretation of this nature, and the decision was to remove this element and focus on identifiers for individuals rather than incidents or events.

On completion, indigenous soldiers will be identified using a list obtained from Bill Denny AM, Director of Veterans SA and Chair of the ANZAC Day Committee and the ANZAC Day Commemoration Council.

Volunteers

Just as the Bureau used a volunteer workforce, SLSA is continuing this tradition by encouraging volunteers to contribute to the discovery of these missing South Australian soldiers again.

SLSA has a volunteering community of people who work a maximum of 15 hours each per week. Many of these volunteers have expressed an interest in this project, and we are fortunate that several volunteers are recently-retired SLSA Archivists, who require little training and are willing to act as leaders amongst the volunteers onsite.

Students from the University of South Australia’s (UniSA) Business Information Management and Library Information Management course have also volunteered.

Workstations onsite at SLSA have been quarantined for the project's exclusive use and new wide screen monitors have been purchased to assist with reading the documentation and accessing the data entry spreadsheets (Figure 10).
Presentations to public libraries have provided an opportunity for the project manager to reduce the burden on the volunteer workstations onsite, by signing up selected public libraries as regional “hubs” for data entry. The public libraries identify interested volunteers from their Friends of the Libraries groups, and provide computers for them to enter data.

The first on board for this is the Barossa Council Library in Nuriootpa, where volunteers will commence indexing from March 2014.

Throughout 2014, Online Projects staff will visit the Barossa and other public libraries every four to five weeks, to work with volunteers while they gain confidence in indexing. Staff will also be available online and by telephone to provide advice as necessary.

More regional hubs will be established as we progress.

Throughout 2014, we will also host regular weekend indexing “blitzes” for SA public librarians interested in contributing.

**Data entry**

As it is not possible to have volunteers key data directly into Millennium, a spreadsheet (Figures 11-13) was developed for volunteers to enter the metadata described in the Data Structure section of this paper. This spreadsheet is then converted to a MARC record and added to Millennium.

Volunteers were employed to test the indexing process and the instructions for indexing developed by project staff. This was a critical element in determining the final data structure as described earlier, testing the workflow and trying to determine an average time to index one record.

Data entry commenced in December 2013, and the data entry spreadsheet is made available on Google Drive and will be split across the volunteer hubs, e.g. Onsite 1-999, Barossa 1000-1999, etc. Volunteers access the PDFs of the enquiry packets remotely via a web server.

As some fields for each record will contain the same information across the collection, the spreadsheet will be pre-populated and those fields hidden; volunteers will only key unique metadata for that record.

To aid consistency, the preferred form of names as subject heading or added name is available from a drop down menu within the spreadsheet. For example, the South Australian-raised 10th Battalion appears regularly in the packets, having been amongst the first to land at Gallipoli and then serving on the Western Front in Europe. For this to appear as a subject heading in MARC in its correct form, it needs to be entered in the spreadsheet to produce the appropriate form of authorised subject heading.

This process is also applied to information that may not be turned into a subject heading or added name yet requires consistency and aids data entry. Information such as the rank of the soldiers or relationship of the enquirer uses this approach.
The NAA records will be used to verify both the names of the soldiers enquired upon and eyewitnesses, ensuring that names match the NAA record. Eyewitness names are often represented by initial, surname, service number and unit. The full name of the eyewitness can be found using NAA.

Locations associated with each soldier will be recorded as well. To assist the process of data entry, volunteers will be asked to record all the geographic names (for example, Mouquet Farm; Somme); and venue names (for example, Edgehill Clearing Station; London General Hospital, Denmark Hill) associated with a soldier as they appear in the documentation. For example, record number 6037 has the following locations (with variant spelling) associated with it:

- Edgehill Clearing Station
- Morlencourt
- Beurre
- Revermont
- Buire
- Somme
- Ribermont
- Morlancourt

Validation of these names and addition of a geographic code will be added in Phase 4, where consideration will be given to what gazetteer and geocodes to use.

The spreadsheet will be downloaded from Google Drive weekly by project staff and the project team’s metadata co-ordinator will convert the spreadsheet to a MARC Record to load into Millennium.

The converted spreadsheet is loaded to the LMS through the Data Exchange module in Millennium using a load profile table developed specifically for this project. Loading the record from the spreadsheet is done automatically and controlled using the unique archival identifier number (for example SRG 76/1/116).

Once loaded to Millennium, the records will remain suppressed from public view for validation. This will be done by staff working from a list, generated by Millennium, of new subject headings created.

Validating authority controls is also undertaken by staff at this stage. During the course of the war, new battalions appeared to be created ‘on the fly’. It is impossible for SLSA to know the full extent of the battalion names contained within the packets and therefore to have pre-prepared the authority controlled names for these. In these cases, volunteers will be instructed to key in the names as they appear, and our cataloguing staff will identify the new names and subject headings and then undertake the appropriate authority work.

Random testing of data entered and weekly reports will be run by the project lead with corrections made as required.
<table>
<thead>
<tr>
<th><strong>Subject - Unit</strong></th>
<th><strong>Soldier - Date of Death</strong></th>
<th><strong>Soldier - Locations</strong></th>
<th><strong>Soldier - POW</strong></th>
<th><strong>Enquirer 1</strong></th>
<th><strong>Enquirer 1</strong></th>
<th><strong>Enquirer 1</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Military unit of the soldier being enquired upon. Leave the unit blank if it is not listed.</td>
<td>If relevant</td>
<td>All locations associated with the soldier indexed as written in the packet, including burials and spelling variations. e.g. 2nd Casualty Clearing Station; Merrij, Borre Hospital; Hazebrouck; Borre near Hazebrouck; Longuenesse Cemetery; ST Omer</td>
<td>Was the soldier a POW? e.g. Y or N</td>
<td>Name of enquirer Use format: Surname, First name(s) e.g. Hanley, Alice Rose or Hanley, A., Mrs or Hanley, Mrs or Hanley, Alice, Mrs</td>
<td>Relationship to soldier Select from drop down list</td>
<td>Address Indexed as written in the packet e.g. Bailey Street, off Gouger Street, Adelaide</td>
</tr>
</tbody>
</table>

**Figure 11**

<table>
<thead>
<tr>
<th><strong>Subject - Unit</strong></th>
<th><strong>Soldier - Name</strong></th>
<th><strong>Soldier - Service No.</strong></th>
<th><strong>Soldier - Rank</strong></th>
<th><strong>Soldier - NAA Barcode</strong></th>
<th><strong>Soldier - Unit</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Enter the soldier's name Use the format Lastname, Firstname(s) e.g. Hanley, Daniel Henry (Check NAA for correct spelling)</td>
<td>Enter the soldier's service no. e.g. 337</td>
<td>Select from drop down list or add in the rank if it is not listed.</td>
<td></td>
<td>From the NAA war service record e.g. 4255190</td>
<td>Enter the soldier's unit Select from drop down list or add the unit if it is not listed</td>
</tr>
</tbody>
</table>

**Figure 12**
<table>
<thead>
<tr>
<th>EYEWITNESS 1</th>
<th>EYEWITNESS 1</th>
<th>EYEWITNESS 1</th>
<th>EYEWITNESS 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of EYEWITNESS</td>
<td>Service no.</td>
<td>Rank of EYEWITNESS</td>
<td>Unit of EYEWITNESS</td>
</tr>
<tr>
<td>Use format:</td>
<td></td>
<td>Select from drop down list or add in the rank if it is not listed.</td>
<td>Select from drop down list or add in the unit if it is not listed.</td>
</tr>
<tr>
<td>Surname, First name(s) or initials e.g. Whiting, Albion William</td>
<td></td>
<td></td>
<td>Indexed as written in the packet</td>
</tr>
</tbody>
</table>

(Check NAA for correct spelling)

Figure 13
Phase 3 March – November 2014

SLSA is undergoing a review of its web presence and intranet. In 2014, SLSA will undertake a review of CMS requirements for the corporate website and the SA Memory program, seeking a solution for these as well as this project. While still to be defined, SLSA is keen to exploit LOD as a way of further sharing our collections.

Planning for requirements is due to commence early 2014.

The structure developed for data entry using the MARC record for this project will be reviewed and consideration given to separating content captured in the MARC 598 fields where required.

Initial screen mock-ups (Figures 14-15) have been created for the web resource to assist with the development of the SARCIB web resource and contain the following business requirements for the web resource:

- Search and advanced search capacity
- Curated stories
- Links to relevant institutions
- An honour roll that displays the name of those who died on a particular day.
- History of SARCIB – written and researched through in kind support by Professor Melanie Oppenheimer, whose centenary history of Australian Red Cross will be published in 2014.
- Display of third party resources
- Ability to add community created content. Consideration will also be given to displaying records where no community content exists.
- Ability to engage with LOD
Community participation

This element of the project is scheduled for planning from March 2014, and intends to provide an opportunity for the public to contribute and create content.

Community participation is an activity yet to reach maturity at SLSA; however, small forays in this area have occurred with the development of the crowd-sourcing activity South Australians of World War 1: Share their story Flickr group (SLSA 2013).

The ability to engage with the public and offer the community opportunities to contribute more information is attractive to SLSA and a requirement of the Strategic Plan.

The intention of this element of the resource is to provide an opportunity for the public to:

- contribute personal stories, reminiscences and photographs
- offer tagging of each packet
- crowd source transcription of the letters and other documents contained within the packets.

SLSA had considered the Trove model for data entry, where the citizen archivist may register to undertake the indexing of content according to the data structure outlined in Phase 2. This was reconsidered because the requirements for data entry are such that a level of control is necessary to ensure the quality of metadata entered. Providing tagging options allows for less structured indexing.

An associated activity intended for 2014 is the use of third party applications to exploit crowd-sourcing opportunities for transcription of SLSA content, in particular, a collaboration with History SA to serialise war diaries. As planning and development occurs, consideration will be given as to how to apply this to SARCIB, especially for the handwritten enquiry letters. The American Memory’s Abraham Lincoln Papers at the Library of Congress is an example of another initiative that provides searchable, transcribed documents (Library of Congress, 2002).

Third Party Content

An element of Phase 3 of the project is to incorporate third party content into the developed web resource.

Information related to AIF personnel is contained within many different institutions across Australia, and internationally. A researcher needs to visit many different institutions to bring together the experiences or stories about a soldier. Fortunately, most institutions such as NSLA libraries, AWM and NAA present much of their material online, or at the very least, an index to their material, so that a soldier can be identified through some common points of reference, such as Name, Rank, Service Number, Unit or Conflict (for example WW1).
The SARCIB web resource will provide access to the personal dossiers for servicemen held in the NAA and, using the Trove API, bring together obituaries from *The Chronicle* newspaper available in Trove.

SLSA also intends to utilise records from CWGC.

During the data entry stages of this project, the names of missing soldiers and soldier eyewitnesses are verified using NAA records. As each soldier is searched and identified in that resource by the volunteers, the barcode number to the digital copy of the personnel dossiers (i.e. the attestation papers, service and casualty forms, military correspondence) is copied to the SARCIB data entry spreadsheet (see Data Entry) with the intention of providing access to the dossiers in the SARCIB web resource.

SLSA has indexed more than 30,000 newspaper references to personal notices from 1915-1919 published in *The Chronicle* newspaper (Figure 16). Nearly 3,000 notices contain a portrait. The intention is to bring the notices displayed in Trove into SARCIB’s search results for soldiers who have a personal notice in *The Chronicle*.

The references for *The Chronicle* have also been released as a data set on data.sa - South Australian Government Data Directory.

![Figure 16: Example of a personal notice (Chronicle 1918, p.12)](image)

Tim Sherratt, Trove Manager at NLA, built a proof-of-concept for an aggregated search interface to records about Australian WW1 personnel in his *Discontents* blog (Sherratt, 2013). This searches:

- service records held by the NAA
- the Roll of Honour at the AWM
- the First World War Embarkation Roll at the AWM
- the Red Cross Missing and Wounded files at the AWM
- the Honours and Awards database at the AWM
- the CWGC database

Ideally, SLSA would like to build on this proof of concept and bring some of the same resources into the SARCIB web resource as part of Phase 3 during 2014.
Phase 4

Go Live

SLSA intends to launch the new web resource in early 2015, in time for the centenary commemorations for the April landing at Gallipoli. Regular promotion will commence from August 2014 to coincide with the 100th anniversary of the declaration of war and the Adelaide opening at SLSA of the Remember me: the lost diggers of Vignacourt exhibition at SLSA.

Promotion

Promoting a project that has captured the public’s imagination has been one of the simpler, though more time-consuming, aspects of the project.

Potential stakeholders were advised through a forum in April 2013. Representatives from the Red Cross, RSL SA, NAA, SRSA, History SA, Veterans SA, Legacy and the Genealogy and Heraldry Society attended and were encouraged to present their plans.

The event resulted in SRSA and SLSA collaborating to release the South Australians of World War 1: share their story Flickr group, a crowd-sourcing project encouraging the public to add or correct the information both institutions hold on individual portraits of soldiers or add their own images and information.

Presentations for public and school librarians’ forums have resulted in invitations to speak at individual public libraries throughout the state. As this occurs, we are creating a network of volunteers to assist with data entry.

His Excellency Mr Reha Keskintepe, Ambassador of the Republic of Turkey, and South Australian Premier, Jay Weatherill, have both attended presentations in 2013.

To provide a consistent message and reduce the pressure on the project team, the South Australian Red Cross Information Bureau film was made to summarise the project and was released in June 2013 on SLSA’s YouTube channel.

The Word Adelaide festival also partnered with SLSA and the State Theatre Company to produce The Red Cross Letters. Two performances based around 20 enquiries were hosted by SLSA in August 2013, and more performances are intended closer to the release of the web resource.

We have also joined World War Onelink, which aims to encourage collaboration and reduce duplication of effort between projects across Australia. Internationally, we have joined the IWM’s centenary partnership program and are writing a feature article for publication by the National Trust UK in April 2014.
Phase 5

Linked Open Data

Where possible, the metadata will be expressed in the resulting web resource as LOD, enabling linking of this information with other WW1 projects. For SLSA, these ways of linking knowledge are unprecedented. This functionality is scheduled for later in the project, as it is an area of developing knowledge for SLSA and one where we have much to learn from others.

We recognise that the stories contained within the SARCIB project are just a slice of a much wider experience. The soldiers contained in this collection had lives before and, if lucky, after the war. They were people who left their normal lives to participate in an extraordinary global event. In some instances, individuals may already have a web presence in some way such as official service records, photographs, letters and diaries in cultural institutions' catalogues or mentions in newspapers.

The soldiers in the Red Cross records provide SLSA with an opportunity to explore LOD further. As the AIF is a structured organisation, with each soldier having a range of identifiers such as Name, Rank, Service Number and Battalion, we expect a common vocabulary to emerge from international institutions working in this area, to match these elements with those in other data structures from the AWM, NAA and IWM. This stage of the project will require SLSA to invest more thoroughly in research into LOD and open discussions with these agencies. The project team's initial meeting with NLA, AWM and NAA in April indicated that rigorous discussion in 2013 was premature.

One of the many challenges we face is our own developing knowledge of LOD and how to apply it. We do not yet have this skill, and are seeking advice from colleagues nationally and internationally. Quite often, when discussing this project with prospective vendors we received blank looks, which did not fill us with confidence. Even when interviewing for web positions, our hope that LOD had permeated into the psyche and local industry revealed that our hopes were unfulfilled.

During Phase 3, SLSA’s Online Strategy team will be seeking a Content Management Solution (CMS) to export the records created in Millennium for delivery to the public. One of the criteria for selection will be the ability of that CMS to accommodate LOD.

Of particular interest is the Out of the trenches: Linked Open Data of the First World War, project by the Pan-Canadian Documentary Heritage Network (PCDHN). The project was a proof-of-concept, showcasing a subset of multiple partners' resources using LOD and the semantic web. Fortunately for us all, the final report explains the ontologies/element sets used in their metadata mapping (PCDHN 2013, pp. 6-7). This provides SLSA and other Centenary of WW1 contributors with a starting point. The ontologies (Figure 17) contained in the report will be explored by SLSA during this phase.
Discussion with Eero Hyvonen (Professor, Alto University Semantic Computing Research Group), (Hyvonen, 2013, pers. comm. 8 August) suggested joining the Linked Open Data approaches to the Great War forum. As a result, it would appear at this stage of SLSA’s project, that there is still a way to go before agreement on a uniform set of ontologies is reached. Further research is required during Phase 5.

While further research will be undertaken during Phase 5, a brief scan of the literature has occurred and there other papers of note, in particular, *World War 1 as linked open data*, which describes the construction of the WW1LOD dataset. This dataset gathers events, places and agents (for example battalion) which are made available for indexing through API’s and a range of interfaces (Tornroos et al 2013, pp.1-6).

Although the SARCIB project records places and agents, the main focus is on personal names.

The decision to collect the descriptive data within a MARC record enables SLSA to take advantage of schemas and ontologies in use in the library profession, in particular the *Library of Congress Authorities* website (Library of Congress, 2012) and those outside such as provided by Schema.org (Schema.org, n.d.).

**Related activities**

Along with *South Australians of World War 1: Share their story* Flickr group, SLSA has a number of other online projects related to the WW1 centenary.

History SA and SLSA are collaborating on a web presence serialising selected war diaries throughout 2014-2018. Similar to their *Bound for South Australia* site (History SA 2011), this will be primarily directed at the education sector.

We expect to provide more opportunities for volunteers by creating a range of opt-in digital volunteering activities, which include:
• Text correction of the *Heroes of the empire* personal notices in *The Chronicle*.

• Utilising Wikisource to provide crowd-sourced transcriptions of diaries to be used in the History SA serialisation project.

In time, SLSA will use these strategies for activities outside of SARCIB.

Other activities include:

• Continuing to provide mash-up opportunities through the 20 data sets released for *Unleashed*, the Adelaide node of Govhack. Three of these data sets are WW1 specific and are available on data.sa.gov.au.

• Developing an App using source code for the Open Culture App developed by Europeana.

**Conclusion**

The Centenary of ANZAC provides SLSA with an opportunity to contribute to the State Library’s Strategic Plan by delivering more South Australiana content online for the public to learn from and reuse. It is also an opportunity to deliver on public expectations to be involved, add their knowledge to our collections and create new content. This in turn, enriches the value of our heritage resources.

We have also been successful in having centenary of WW1 requirements included in the new SLSA Strategic Plan 2014-2017.

Further, the project represents an interesting opportunity to leverage the developments of LOD. The application of Linked Open Data will be a bonus and will pave the way for other projects to follow.

The financial and resourcing constraints have required the project to look internally at what tools we have at our disposal and work creatively with them. While the capture of the metadata into a MARC record may be an interim measure to commence data entry, it represents an ability to think laterally.

As the rest of SLSA, and perhaps the SA Government, develop other programs related to the Centenary of ANZAC, there is the risk of scope creep or reducing the resources available at SLSA for this project. The project manager is continually exploring funding and staffing options for components of this project.

The recruitment of volunteers for data entry is a natural alignment for this project, neatly creating a full circle that commences with volunteers in 1915 and ends with volunteers in the 21st century.

The SARCIB project is a long-term project, and some of what has been written here is still to be developed or implemented. Changes in intentions and plans which may occur during the project, will require flexible approach management. This will also not be the only project managed by Online Projects during 2014-2018.
Commencing planning in 2012, Online Projects has been able to take the necessary time since to refine requirements and to change approach to the database development, without jeopardising the entire project.

There are several measures of success for this project, which are represented by the phases of this project. We have always said that, at the very least, an important part of our collection has been digitised and made available through the SLSA catalogue.

Achieving this goal demonstrates the fulfilment of a planned digitisation program, the engagement of volunteers in useful work, the innovative use of resources available to us and the pride associated with improving staff skill levels to achieve those goals.

The ultimate success, of course, is the delivery of a dignified web resource that delivers on the business requirements to deliver indexed records from SARCIB, takes advantage of user generated content and utilises third party resources.
Abbreviations

AWM Australian War Memorial
CWGC Commonwealth War Graves Commission
IWM Imperial War Museum
LOD Linked Open Data
NAA National Archives Australia
NLA National Library of Australia
NSLA National and State Libraries Australasia
RSL Returned and Services League, South Australian Branch
SLSA State Library of South Australia
SRSA State Records South Australia
References


Bean, C 1938, The official history of Australia in the War of 1914-1918: Volume II the story of ANZAC from 4 May, 1915 to the evacuation, Angus & Robertson, Sydney.


State Library of South Australia, n.d., South Australian Red Cross Information Bureau [finding aid], p.1, viewed 10 September 2013,


