

**Adroddiad i'r Grwp Rheolaeth Prosiect
Report to Project Management Group**

Martin Locock

1. Meetings attended

9/6/04 Presentation on ANW to Society of Archivists Wales, Swansea
 15/6/04 Interviews for PO5 post, Cardiff
 21/6/04 Linking Arms bid meeting, London
 22/6/04 Visit to Gwynedd Archives, Caernarfon
 23/6/04 Visit to Anglesey County Record Office, Llangefni, Conwy Archives Service, Llandudno, and Powys Archives, Llandridnod Wells

2. Staff

Two applicants were interviewed for the Project Officer (Cardiff) post (17 month appointment), and David Morris has been appointed, to start from 5th July.

The Aberystwyth team will shortly be back to full strength with the return of one member from long-term sick leave in July.

3. FLD creation

Currently there are 595 records live on the website (plus five sub-fonds descriptions¹. The database currently holds 2,587 records. The access points include 3,547 corporate names, 1,287 places (and a further 848 variant forms), 2,092 personal names and 3,730 subjects.

Partner	Target no. FLDs	Live	To be checked by partner	To be checked by Project Manager
NLW	813	123	20	204
Anglesey	150	172	6	0
Carmarthenshire	272	(24)	275	0
Conwy	27	37	0	0
Ceredigion	77	(4)	114	0
Flintshire	137	(33)	0	183
Denbighshire	142	(27)	0	161
Glamorgan	649	(37)	95	268
Gwent	680	(25)	0	153
Gwynedd	436	(36)	0	0

¹ In creating descriptions, the team has followed the collection or catalogue organisation of the repository to create fonds-level records. In some cases, after discussion with the partner, closely-linked records have subsequently been rationalised into fonds and sub-fonds records arranged hierarchically.

The creation of records for Anglesey, Conwy, Carmarthenshire, Ceredigion and Flintshire is complete. The target for work on Denbighshire is complete, but it is hoped that some further Minor Lists will be covered if time allows. Work on Gwent's minor lists held in the NLW has been completed. Work on Gwent's catalogues as they are supplied will proceed in parallel with work on Gwynedd (Caernarfon and Merionydd).

Batches of records for NLW and Glamorgan have been submitted to partners for checking prior to going live. Almost all of Anglesey's records are now live (there are 6 requiring final approval). ANW checking of Ceredigion and Carmarthenshire records is complete, and these have been sent to the partner for checking. Subject indexing of existing records is complete (new records are being indexed when created), apart from non-live records for NLW; these are currently being dealt with.

All the placenames created to date have been checked for consistency and upgraded to ensure that they can be output in MARC format; variant forms recorded in the standard sources have been recorded as NPTs (non-preferred terms).

4. Software

EAD mapping is nearing completion and preparations are being made for the upload of HE partner data from the Archives Hub; a meeting of the relevant partners will be held on 20th July.

5. Website

Period	Key statistics			English page hits	Welsh page hits	Unique users	Sessions/user
	Total hits/month	User sessions	Average hits/session				
September	35,589	1,555	22.9	31,389	4,200	1,081	1.44
October	35,120	1,853	19.0	33,086	2,034	1,437	1.29
November	35,234	2,357	14.9	32,112	3,122	1,434	1.64
December	30,032	2,359	12.7	27,149	2,883	1,583	1.49
January	45,004	3,156	14.3	40,682	4,322	2,326	1.36
February	55,548	2,952	18.8	50,338	5,210	2,386	1.24
March	87,020	4,561	19.1	75,184	11,836	2,359	1.93
April	98,471	6,407	15.4	80,847	17,624	2,810	2.28
May	114,132	6,438	17.7	93,306	20,826	2,926	2.20
June	188,484	7,093	26.6	144,981	43,503	3,007	2.36

The number of page hits continues to grow (65% increase month-on-month). Much of this traffic is derived from web-crawlers operated by search engines engaged in indexing the web content by following the links. Unlike most 'deep web' resources,² all of ANW's dynamic data can be reached via the html indexes (these are rebuilt overnight). Crawlers typically work in a four-week cycle; on their first visit to a site they will follow one level of links, and at the next follow the links from those links.

² The 'deep web' or 'hidden web' comprises data sources available over the internet whose architecture (requiring logging on or typed querying of dynamic data) makes them invisible to search engines, which rely on indexing 'surface web' (hard-coded html) content.

Crawler hits will therefore be expected to grow significantly month-by-month (until all links have been followed), and also to increase directly in proportion to the amount of live content (probably reflected in the increase in hits/session).

It is difficult to establish precisely how much of the usage is the product of crawler activity. One indication is the number of occasions the 'robots.txt' file has been accessed, since this is visible only to crawlers (935 on the English site, 931 on the Welsh site, representing 18% and 45% of sessions respectively).³ Another measure is the identity of the ISPs listed as the top 30 users of the sites. Both lists are dominated by crawlers, with only 4 out of 60 representing human users (NLW (both lists), Gwynedd County Council and The National Archives); the listed crawlers are responsible for 35% of the English hits, and 80% of the Welsh hits. However, there clearly is a long tail of human users looking at fewer than 100 pages, making up most of the visits and a substantial proportion of the total hits.

The crawler activity is an important consideration in interpreting the statistics, particularly in terms of 'reach'. The slowest-rising statistic of all has been the number of users per month, which is perhaps the key statistic for the project. There are grounds to hope that this too may begin to rise more emphatically, owing to the greater visibility of the resource as a direct result of the crawling.

The search terms "archives" + "wales" now show Archives Network Wales in the first page of results in all major search engines (listed third by Google, Yahoo!, AltaVista, Lycos, All the Web and Search.com, fifth by Ask Jeeves, and eighth by Excite and Metacrawler; exceptions include Netscape (15th) and Looksmart (not shown)) (the precise ranking depends on the weighting applied by the search engine; in many cases, .gov domains are scored higher than others, so The National Archives comes top). In addition, searching for individual index entries will usually lead to the relevant ANW entry.⁴ Therefore web users looking for general or specific information should become aware of ANW at an early stage in their search. A total of 50 websites now include hyperlinks to the site, so searchers in related resources should also be led to ANW.

6. Outreach and publicity

The training days for partner staff have been finalised as University of Wales, Swansea (3rd August), Conwy Archives Services (5th August) and NLW (9th August). So far, about 6 participants have booked at each venue. An additional day, in South Wales, will be organised in September.

³ These figures are not precise because some crawlers ignore "robot.txt" instructions, while others may hit the file more than once in a session.

⁴ When the full indexes were live even if the index term was linked to no live data, most searches would show the index page for the relevant letter; since the indexes have been restricted to show only those linked to live records, the search shows the 'results page' listing the collections indexed by that term.

7. Budget and finance**Quarterly budget, April-June 2004**

	Projected	Actual to date
EXPENDITURE		
Project Manager	8800	8800
Project Officers 1-4	23889	23889
Project Officer	5815	7900 ⁵
Travel & Subsistence	186	566.06 ⁶
Equipment	0	0
Software	0	0
System Development	0	0
ICT Support	313	42.55
Administrative Support	325	0
Training	300	0
Publicity & Marketing	0	384.00
INCOME		
NLW	35,000	
Other partners	10,000	3999 (+ 2333 previous advance payments)
HLF Grant (74%)	29728	

Three partners have paid their full contribution for the project previously. Six partners so far have been invoiced for their contributions for the current financial year.

This quarter's claim will include invoices for ICT support and Administrative Support from NLW covering the project to date.

8. National meetings

A further Linking Arms meeting was attended in the run-up to submission of the bid on 30 June by TNA.

9. Project Management Group and Project Board

The next PMG will be a video conference on a date to be confirmed in late September, at NLW, Aberystwyth, and UWS, Swansea.

The next Project Board will be on 23rd July in Llandridnod Wells.

⁵ Includes 5730.33 covering January-March 2004, invoiced in May 2004

⁶ Includes 51.06 covering travel costs for PO5 recharged to ANW by RCT and claims covering Jan-Mar 2004.