



Preservation In Practice - What Are The Options? A Case Study

Decoding the Digital
British Library Preservation Advisory Centre

27 July 2010



- o Joint archive service for the 5 West Yorkshire metropolitan authorities:
 - o Bradford City Council
 - o Calderdale Metropolitan Borough Council
 - o Kirklees Metropolitan Borough Council
 - o Leeds City Council
 - o Wakefield Metropolitan District Council

- o Largest local authority record office outside London:
 - o Over 4200 cubic metres of traditional archive material
 - o In 2008/09: more than 13000 personal visits, 25000 document productions, answered nearly 26000 enquiries

- o Public service points in each West Yorkshire district, plus:
 - o Commercial records management storage and administrative headquarters in Morley (near Leeds)
 - o Three archive outstores, including purpose-built electronic records room at Morley



“A decidedly analogue approach to a digital problem”



- Nobody has all the answers
- Doing something is better than doing nothing – start small
- Knowing what you’ve got (or are about to get) is a vital first step
- “If we try we may fail, if we don’t try we will certainly fail”

WYAS Digital Archives Working Group

Digital Archives Policy

available at <http://www.archives.wyjs.org.uk/wyjs-archives-policies.asp>

fostering awareness and support

creating a framework for practical work

Practical Digital Archives Work

Software testing, using sample objects

Survey of digital records already held by the Service (first documented digital accession received 1999)

Digital deposit form and guidelines

A digital media recognition guide for staff

Advocacy and training

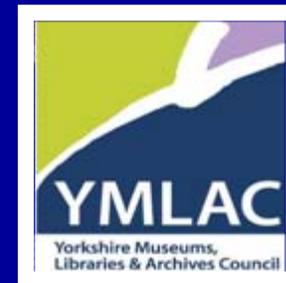
Small, non-specialist, staff working group

3 archivists, plus records manager

Flexible, but limited, support from West Yorkshire Joint Services' IT department

MLA Yorkshire: A Case Study

- Regional Museums, Libraries and Archives partnership organisation for Yorkshire
- Wound up December 2008
- Developed out of the Yorkshire Museums Council (YMC), with a continuous history back to 1963
- 1.4 m³ paper, 1963-2008
- 80GB+ digital, c2002-2008



Negotiating a Hybrid Deposit

- Very short timescale – no time to try out software tools
- Depositor had not considered digital records to be ‘archives’
- Digital content may not be held locally (website, e-bulletins)
- ‘Self curation’ of electronic records by staff - closer to personal archive than organisational
- Confidentiality
- Conflicting interests
- Positives:
 - No obsolete hardware
 - Predominantly current office file formats
 - Helpful IT support
 - ‘Dowry’ to pay for boxes, portable hard drive etc.



Collecting Digital Content

- New digital deposit documentation
- 'Grab & Run' copying - no time to complete appraisal prior to transfer
- Transfer using 1TB USB external hard drive & FTK Imager Lite (<http://www.accessdata.com/downloads.html>)
- Capture 'original order' & generate checksums
- Problems encountered:
 - Long file paths/file names
 - Need all users to be logged out of corporate network
 - Don't underestimate the time needed to transfer large amounts of data

Processing Digital Content: Procedure

- Transfer to standalone PC
- Authenticity verified using checksums
- Virus check – twice – using AVG Anti-Virus (Kaspersky on corporate network)
- Use spreadsheets to analyse file listings / check for duplicates
- Appraisal

Processing Digital Content: Tools



- Pre-accession appraisal tools
 - Karen's Directory Printer
<http://www.karenware.com/powertools/ptdirprn.asp>
 - Directory structures
<http://www.paradigm.ac.uk/workbook/record-creators/capturing-directory-structures.html>
- Forensics software
 - FTK Imager Lite
<http://www.accessdata.com/downloads.html>
- Checksums
 - MD5, SHA-1
- Virus checking
 - Free options: AVG Anti-Virus, Avast! etc.
 - Software and procedures
- File format identification
 - PRONOM
<http://www.nationalarchives.gov.uk/pronom>
 - DROID <http://droid.sourceforge.net>
- Paradigm Workbook
 - <http://www.paradigm.ac.uk/workbook>

Portrait of a Gentleman Techie, after Thomas de Keyser

<http://www.flickr.com/photos/notionscapital/2496369741>

Appraisal in a Digital World

Microsoft Excel - G_and_H_complete_file_List.xls

File Edit View Insert Format Tools Data Window Help

Save As... 75% Arial

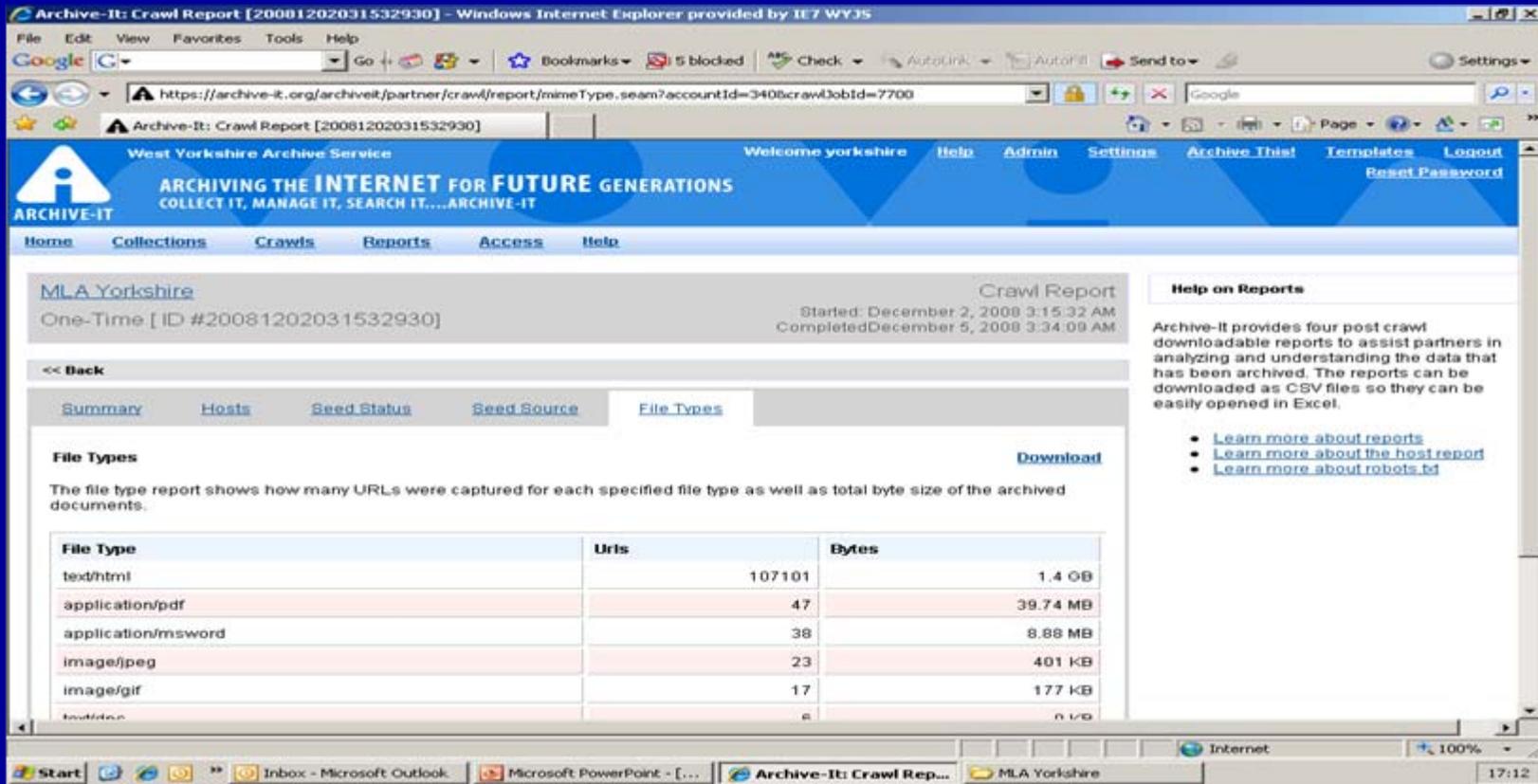
A6824 FILE

1	A	B	C	D	E	F	G	H	I	J
		Filename/Filepath	Extension	Size	Date Created	Date Last Modif	Date Last Access	File Vers	MD5 Hash	SHA-1 Hash
6803	FILE	Digital Strategic Leadership Paper.doc	doc	46,592	12/19/2007 16:30	12/19/2007 16:30	17/12/2008 09:40		222965C730198C7E50200E138D4F90CB	6A4DF29DB4
6804	FILE	Big Day De-brief.doc	doc	35,329	13/11/2007 16:20	20/11/2007 12:55	17/12/2008 10:06		222A318D05906EFCF8BA92178789AE330	F3F69F5DC6
6805	FILE	Draft M&G Educ Strat PAC 0204_JV.doc	doc	59,392	23/03/2008 10:19	23/03/2008 14:19	17/12/2008 10:00		222B1D209F3D2775F8870198E8727E70	92284419AEE
6806	FILE	Learning Managers Meeting 15 June 06.doc	doc	97,040	14/08/2007 15:05	26/06/2006 17:02	17/12/2008 09:37		2223D027A46B2E80462C96D63648E3E	6EFB9802B8
6807	FILE	Feeling the V'sth.doc	doc	138,752	19/08/2007 19:03	22/12/2004 10:55	17/12/2008 09:41		2232F4FC8A19ECC3C9060042455F2C6A	1048E1176E87
6808	FILE	Accreditations 2007.doc	doc	56,320	21/11/2007 12:10	23/11/2007 12:21	17/12/2008 10:23		2237C74348CC64044EBDCC8E2A508BA6C	56417769D3F1
6809	FILE	ymiaedc1atreport-part3(2mar).doc	doc	295,696	19/08/2007 18:58	13/03/2006 08:12	17/12/2008 09:36		2238D7B4055733827928558D4EBDF44F	11D0F790AF6
6810	FILE	Health policy Brief.doc	doc	186,880	19/08/2007 19:02	17/09/2006 15:36	17/12/2008 09:41		223E59DBF73DD0B06D476CB53D000494	D0B96480E9
6811	FILE	Transfer to MLA Control description 10.doc	doc	31,744	19/11/2008 08:27	19/11/2008 08:27	17/12/2008 09:35		2240905E83BA408698E2A76F05E58A03	30948402630
6812	FILE	SVC Agenda 25Nov04.doc	doc	38,400	17/11/2004 00:03	19/11/2004 09:29	17/12/2008 10:23		2240462C3044E996B633DEF1C659B91	40F4F4842E1
6813	FILE	SVC Agenda 25Nov04.doc	doc	38,400	17/11/2004 00:03	17/11/2004 00:03	17/12/2008 10:23		2240462C3044E996B633DEF1C659B91	40F4F4842E1
6814	FILE	SVC Agenda 25Nov04.doc	doc	38,400	24/09/2007 11:46	19/11/2004 09:29	17/12/2008 09:41		2240462C3044E996B633DEF1C659B91	40F4F4842E1
6815	FILE	SVC Agenda 25Nov04.doc	doc	38,400	24/09/2007 11:46	17/11/2004 00:03	17/12/2008 09:41		2240462C3044E996B633DEF1C659B91	40F4F4842E1
6816	FILE	Framework 4 Future YMLA Response F.doc	doc	31,744	19/08/2007 18:56	24/03/2004 10:47	17/12/2008 09:35		224462A98F2EAF8AA82B9B95C3AF2	BDC204E1DF
6817	FILE	PDNet Gen booking form.doc	doc	25,600	19/08/2007 18:50	05/02/2006 10:47	17/12/2008 09:34		2246086DFDA8650A5A71228C3E95019	F2C50F78B1
6818	FILE	PDNet Gen booking form.doc	doc	25,600	19/08/2007 18:50	05/02/2006 10:47	17/12/2008 09:34		2246086DFDA8650A5A71228C3E95019	F2C50F78B1
6819	FILE	Finance assistant.doc	doc	59,392	02/11/2004 13:45	20/05/2008 13:20	17/12/2008 10:02		224AB824DC73EF033087190140B1C0D5	6827477A181F
6820	FILE	Teenagers and stories project Yorkshire.doc	doc	26,112	14/08/2007 15:04	19/09/2006 07:48	17/12/2008 09:37		2252E8D75D1300C58229748FBEADD039A	D6D0AA803F
6821	FILE	Strategic Commissioning 2004-6 MAP1.doc	doc	31,232	03/05/2005 08:27	29/07/2004 08:50	17/12/2008 10:02		22585D7F6402AADF7C5D12F367D695F7	8C358775D0F
6822	FILE	Strategic Commissioning 2004-6 MAP1.doc	doc	31,232	14/08/2007 15:00	29/07/2004 08:50	17/12/2008 09:39		22585D7F6402AADF7C5D12F367D695F7	8C358775D0F
6823	FILE	CMSC009Mens.doc	doc	34,816	14/08/2007 15:01	05/01/2005 17:12	17/12/2008 09:39		225CA8D35A8BC8E88D759322F2E17482	3E7F1658BAE
6824	FILE	Digital Film Archive for Yorkshire CD A1.doc	doc	45,568	03/12/2008 11:12	19/12/2005 11:34	17/12/2008 09:59		227003FDC45E35E44A3CA201261C897	560EAE3885
6825	FILE	Digital Film Archive for Yorkshire CD A1.doc	doc	45,568	03/12/2008 11:14	19/12/2005 11:34	17/12/2008 09:59		227003FDC45E35E44A3CA201261C897	560EAE3885
6826	FILE	application ChangeUp.doc	doc	122,880	19/08/2007 19:02	07/09/2005 14:50	17/12/2008 09:39		227E327567950D1A00C36917EEA8C88	F5A8E8CCB
6827	FILE	Membership Info & Criteria May03.doc	doc	214,528	04/06/2003 10:03	14/05/2003 15:07	17/12/2008 10:24		22825926CB3045FBE6C13387EA0E202	DF590682CF
6828	FILE	07.12.05 notes.doc	doc	24,576	30/05/2008 10:42	07/12/2007 08:29	17/12/2008 09:42		2287BC1F8104FE7DB179C4674DB78284	9462CFD28C
6829	FILE	07.12.05 notes.doc	doc	24,576	07/12/2007 09:24	07/12/2007 09:29	17/12/2008 09:42		2287BC1F8104FE7DB179C4674DB78284	9462CFD28C
6830	FILE	funding and sustainability PSG paper Sept.doc	doc	142,848	19/08/2007 18:57	01/09/2006 08:57	17/12/2008 09:36		2289AC4579E8AA027A026A211F5ECADD	320458CE283
6831	FILE	Charities Commission Direct initial lette.doc	doc	26,112	07/06/2008 12:09	07/06/2008 13:39	17/12/2008 10:02		228DBA8E3705E2AF76A6E01D7F6349	588F1926F47
6832	FILE	Hubled notes.doc	doc	33,280	19/08/2007 19:02	05/07/2003 10:32	17/12/2008 09:42		228B4C76B48AD7438537D635FFD03F63	319F3A1A1F61
6833	FILE	Agenda 21 May 02.doc	doc	20,992	02/12/2007 17:54	14/05/2002 16:04	17/12/2008 10:00		229D5A4DC24723593231289C5EDC761E	3CB808A709
6834	FILE	Minutes 08.05.doc	doc	59,904	16/06/2008 08:42	17/06/2008 14:40	17/12/2008 09:43		22A068AD685950C0537005CECE09847	6D71706C58C
6835	FILE	Logistics.doc	doc	24,064	06/03/2006 15:19	06/03/2006 15:19	17/12/2008 10:03		22A19E759C8CF2E54D9228E61AFC452	C3E3517F2F
6836	FILE	Accred Introduction letter.doc	doc	796,160	19/08/2007 19:11	17/09/2006 09:21	17/12/2008 09:42		22A509F4A44F4333297AE6596823493	0C361E163A
6837	FILE	Treasurers report re subscriptions.doc	doc	22,528	01/12/2008 16:36	15/11/2001 16:53	17/12/2008 09:59		22A67F8E77CC01362F845E0A39A96629	10D910E596F
6838	FILE	Treasurers report re subscriptions.doc	doc	22,528	01/12/2008 16:33	15/11/2001 16:53	17/12/2008 09:59		22A67F8E77CC01362F845E0A39A96629	10D910E596F
6839	FILE	26.05.06 LAFS Consultation Letter.doc	doc	28,624	19/08/2007 19:13	31/05/2006 09:35	17/12/2008 09:46		22A9EF6D9D0229CE2918EF6240D19AA	5F50E378917
6840	FILE	Item 15 Dates of future meetings.doc	doc	24,576	23/03/2008 17:31	04/03/2004 11:27	17/12/2008 09:59		22B932F8E1F5F02544D408F564887F	37318EAF48
6841	FILE	Raw minutes 10.3.06.doc	doc	29,696	14/09/2007 14:13	14/03/2006 16:24	17/12/2008 09:42		22BE683E2952C3DC58527699D97F7C50	722EACE88A
6842	FILE	A taxonomy of MDO activity and interve.doc	doc	35,840	19/08/2007 19:11	14/05/2007 11:52	17/12/2008 09:44		22C4CB498E8E7EA9F79F330F0BF98A2	82AA2F58A5
6843	FILE	jmevents.doc	doc	20,480	04/11/2008 13:03	02/08/1999 10:43	17/12/2008 09:59		22C8C915007F797CEC63143DF1ABE00F	1E488B7E6E
6844	FILE	jmevents.doc	doc	20,480	30/08/2008 13:39	02/08/1999 10:43	17/12/2008 10:22		22C8C915007F797CEC63143DF1ABE00F	1E488B7E6E
6845	FILE	Performance Review Documentation1.doc	doc	63,488	19/08/2007 19:13	21/12/2005 10:41	17/12/2008 09:46		22DA1C0B4E77827E420294DEAD27069A	D903F710FC

Ready

Start | Inbox - Microsoft Outlook | Microsoft PowerPoint - [...] | Final Directory Readings | Microsoft Excel - G_a... | 16:05

MLA Yorkshire Website



The screenshot shows a web browser window displaying a crawl report from Archive-It. The browser's address bar shows the URL: <https://archive-it.org/archiveit/partner/crawl/report/mimeType.seam?accountId=340&crawlJobId=7700>. The page header includes the Archive-It logo and the text "ARCHIVING THE INTERNET FOR FUTURE GENERATIONS COLLECT IT, MANAGE IT, SEARCH IT...ARCHIVE-IT". The main content area is titled "MLA Yorkshire" and "One-Time [ID #20081202031532930]". It shows a "Crawl Report" with the following details:

- Started: December 2, 2008 3:15:32 AM
- Completed: December 5, 2008 3:34:09 AM

Navigation tabs include Summary, Hosts, Seed Status, Seed Source, and File Types. The "File Types" tab is active, showing a table of file types and their corresponding URIs and byte sizes.

File Type	UrIs	Bytes
text/html	107101	1.4 GB
application/pdf	47	39.74 MB
application/msword	38	8.88 MB
image/jpeg	23	401 KB
image/gif	17	177 KB
text/css	6	0.00 MB

On the right side, there is a "Help on Reports" section with the following text: "Archive-It provides four post crawl downloadable reports to assist partners in analyzing and understanding the data that has been archived. The reports can be downloaded as CSV files so they can be easily opened in Excel." Below this text are three links: "Learn more about reports", "Learn more about the host report", and "Learn more about robots.txt".

Archive-It - <http://www.archive-it.org/>

UK Web Archiving Consortium <http://www.webarchive.org.uk>

MLA Yorkshire Website

MLA Yorkshire - Windows Internet Explorer provided by IE7 WYJS

File Edit View Favorites Tools Help

Google G Go Bookmarks 5 blocked Check Look for Map AutoFill Send to Settings

http://wayback.archive-it.org/1237/20081202031535/http://www.mlayorkshire.org.uk/webdav/harmonise?Section/@id=4511

MLA Yorkshire

You are viewing an archived Web site, archived on 3:15:35 Dec 02, 2008, that is part of a collection of archived websites created using [Archive-It](#). The information on this web page may be out of date. External links, forms, and search boxes may not function within this collection. [[hide](#)]

Search

MLA MUSEUMS, LIBRARIES AND ARCHIVES YORKSHIRE

Home
News & Events
The Region
Archives
Libraries
Museums
Resources & Publications
Training
Careers
Networks & Groups
About MLA Yorkshire
Contact Us

MLA Yorkshire is the lead strategic agency for museums, libraries and archives in the region. We work to improve people's lives through access to these collections and resources – building knowledge, supporting learning, inspiring creativity, and celebrating identity.

> Latest News

- MLA Appoints New Area Director North Accreditation - Final Update
- groupthing.org: Engaging Young People with Reading
- Their Past Your Future – Final Round Year 3 Grants
- Strategic Commissioning - The Future

Subscribe to our **E-Info Service**

> Resources and Publications

- Free Family Learning in Museums Resource Pack
- Yorkshire Museums Support Consultation - Final Report
- Building a Diverse Workforce: Libraries and Archives in Yorkshire
- Yorkshire GSO Pilot Report and Case Studies

> Conferences & Seminars

- Do Games Tell A Story? 2 December 2008
- EUCLID: Cultur€uro Seminar, 2 December 2008
- Tactile Graphics 2008: 3-5 December 2008
- MA: Designing for Disability, 5 December 2008
- Just Playing? The Role of Play in Children's Lives: 14 January 2009

> Training Events

- SINTO Courses - Promoting Reading to Young People 04-12-2008
Sheffield Hallam University
- Object Handling Essentials 11-12-2008
Yorkshire Museum, York
- Disability Awareness Workshops 2009 14-01-2009

Programmes

- 2012 Olympics
- Inspiring Learning for All
- Museums & Archives Alive!
- Museums Development
- Skills for Life
- Widening Participation

Choose a project...
Go

Inspiring Yorkshire

Download the latest edition of our [newsletter](#)

Done

Start Inboxes - Microsoft Outlook Microsoft PowerPoint - [...] Archive-It: Access - Way... MLA Yorkshire

Internet 100%

MLA Yorkshire - Wind... 17:13

MLA Yorkshire E-Bulletins - A Practical Example



MLA Yorkshire Cultural Olympiad Update
Mar 08



Arts Council England, Yorkshire Culture, MLA Yorkshire and Screen Yorkshire invite you to :

"YORKSHIRE'S CULTURAL OLYMPIAD - AN EMERGING FRAMEWORK"

A briefing for the region's cultural organisations on the national framework for the UK Cultural Olympiad for London 2012, progress and opportunities in the region.

The briefing will be led by Tessa Gordziejko, Creative Programmer (Yorkshire) for London 2012 and include speakers from the key partner agencies. The agenda will cover :

2008_03_20_Copy of Cultural Olympiad update.pdf - Adobe Reader

File Edit View Document Tools Window Help

1 / 2 100%

You are viewing the document in PDF/A mode.

Cultural Olympiad News from MLA Yorkshire Page 1 of 2



MLA Yorkshire Cultural Olympiad Update
Mar 08



Arts Council England, Yorkshire Culture, MLA Yorkshire and Screen Yorkshire invite you to :

"YORKSHIRE'S CULTURAL OLYMPIAD - AN EMERGING FRAMEWORK"

A briefing for the region's cultural organisations on the national

Start | Internet Explorer - Microsoft Outlook | Microsoft PowerPoint | ... | ... | PDF Archive Files | 2008_03_20_Copy of ... | 10:46

“I never have these types of problems with digital”

- **Don't panic at the format and overcomplicate – look for simple solutions that are ‘good enough’**
 - Make multiple copies
 - Research software tools and preservation services (eg PRONOM <http://www.nationalarchives.gov.uk/pronom>)
 - Archival theory is still relevant. Develop new procedures alongside traditional workflows.
- **Don't start with long obsolete material – it'll be much harder**
- **Deal with reality, not with ‘what ifs’.**
- **Do not underestimate your expertise with regards to cultural factors/interpersonal skills vital to digital preservation** (eg privacy concerns, depositor liaison)

The Preservation Chain

“Digital information lasts forever, or five years – whichever comes first” *Jeff Rothenberg*

You don't need to aim for 'forever'.

Just the next generation of recordkeepers.



Alexandra Eveleigh
AHRC Doctoral Award Holder
Email alexandra.eveleigh.09@ucl.ac.uk

or

West Yorkshire Archive Service

PO Box 5

Nepshaw Lane South

Morley

Leeds

LS27 0QP

Email archives@wyjs.org.uk