

Digital Library Activities at Oxford



OXFORD
UNIVERSITY
LIBRARY
SERVICES

Michael Popham

<http://www.odl.ox.ac.uk/>

Libraries at Oxford

- Oxford University Library Services (OULS) established in 2000
 - Now c.40 major University libraries; excludes Colleges
- ODL arose as the result of a scoping study 1998-99 funded by the Andrew W Mellon Foundation
- Proposal for the establishment of the Oxford Digital Library
 - \$775K from Mellon (2001-04)
 - £3.1M from OULS Capital Campaign (2003-08)

Goals during the establishment phase

- “Conversion of the former Bodleian Photographic Studio into a self-funding, on-demand digitization service with an OULS-wide remit.”
- “A major new line of library support for the teaching and research activities of the University.”
- “A significant production stream of digitized material, based on Oxford holdings, for subsequent re-use in the ODL and in a variety of developmental contexts of local and general benefit.”
- See <http://www.bodley.ox.ac.uk/librarian/visionplus/vision2008.htm>

Who and Where?

- Osney Mead (c.10 full-time)
 - Administration, metadata and technical specialists
 - Full-text capture (e.g. EEBO-TCP)
- New Bodleian library (16-20 full-/part-time)
 - Imaging Service (digital and analogue)
 - Studio, picture research

Services

- Provide a framework for the creation and management of digital resources
- Digital resource creation
 - on a cost-recovery basis; in-house /out-sourced
- Advice across OULS
- Research and development, proof-of-concept
- Management of digital assets – throughout their life-cycle

Activities

- Oxford Digital Library Development Fund
 - Majority of the \$775K from Andrew W Mellon Foundation
- Legacy data
 - Early MSS at Oxford, ILEJ, Toyota Imaging ...
- External partnerships etc.
 - EEBO-TCP (University of Michigan, ProQuest)
 - PARADIGM, MAT (JISC funded)
- Broadening access
 - ODL/OULS/Oxford websites, RLG-CM ...

ODL Development Fund

- Creating content for the ODL
 - Librarian-academic partnership
- Exploring different data types for the ODL
 - Images, full-text, others...
 - Standards and technologies
- Scholarly resources for use in teaching and research – at Oxford and elsewhere

ODL DF

- Editorial Board
 - Chair: Professor Kathryn Sutherland (St Anne's)
 - Local and national representation
 - Academics and librarians
- Three Funding Rounds
 - May 2002, May 2003, March 2004
- Total allocation to projects = £264,726
 - <http://www.odl.ox.ac.uk/collections/>

Round One (May '02)

1. John Johnson Collection: C18th Entertainment Ephemera
2. Prints from the Curzon Collection
3. ODL for Forestry (demonstration project)
4. Flora Graeca 2000 (demonstration project)
5. Full text recognition of Bodleian Broadside Ballads (demonstration project)
6. Key C17th, C18th, and early C19th geological literature
7. Cobbett's Parliamentary History

Total commitment =£79,396



Round Two (May '03)

1. Seeing is believing...
2. ODL for Forestry: Stage 2
3. Flora Graeca in the 21st Century
4. Publications of Obadiah Westwood
5. Digitization of MS SJC 17
6. John Gould's ornithological works
7. Full-text capture/OCR Study

Total commitment = £95,344

Aug



Mator hab morante alexandria papi
tui.

Thule
Arccon. e. septimo.
Infirmitas
Francia.

Membris. i. maridat. i. infir.

DE AVOTA FERIA INQUIRITIO IN UNIVERSA DIE.

PRIMA FERIA.	II. FE RIA.	III. FE RIA.	IV. FE RIA.	V. FE RIA.	VI. FE RIA.	Sabbu
I. B. II.	III. B. III.	IV. B. IV.	V. B. V.	VI. B. VI.	VII. B. VII.	VIII. B. VIII.
IX. B. IX.	X. B. X.	XI. B. XI.	XII. B. XII.	XIII. B. XIII.	XIV. B. XIV.	XV. B. XV.
XVI. B. XVI.	XVII. B. XVII.	XVIII. B. XVIII.	XIX. B. XIX.	XX. B. XX.	XXI. B. XXI.	XXII. B. XXII.
XXIII. B. XXIII.	XXIV. B. XXIV.	XXV. B. XXV.	XXVI. B. XXVI.	XXVII. B. XXVII.	XXVIII. B. XXVIII.	XXIX. B. XXIX.
XXX. B. XXX.	XXXI. B. XXXI.	XXXII. B. XXXII.	XXXIII. B. XXXIII.	XXXIV. B. XXXIV.	XXXV. B. XXXV.	XXXVI. B. XXXVI.
XXXVII. B. XXXVII.	XXXVIII. B. XXXVIII.	XXXIX. B. XXXIX.	XL. B. XL.	XLI. B. XLI.	XLII. B. XLII.	XLIII. B. XLIII.
XLIV. B. XLIV.	XLV. B. XLV.	XLVI. B. XLVI.	XLVII. B. XLVII.	XLVIII. B. XLVIII.	XLIX. B. XLIX.	L. B. L.
LXI. B. LXI.	LXII. B. LXII.	LXIII. B. LXIII.	LXIV. B. LXIV.	LXV. B. LXV.	LXVI. B. LXVI.	LXVII. B. LXVII.
LXVIII. B. LXVIII.	LXIX. B. LXIX.	LXX. B. LXX.	LXXI. B. LXXI.	LXXII. B. LXXII.	LXXIII. B. LXXIII.	LXXIV. B. LXXIV.
LXXV. B. LXXV.	LXXVI. B. LXXVI.	LXXVII. B. LXXVII.	LXXVIII. B. LXXVIII.	LXXIX. B. LXXIX.	LXXX. B. LXXX.	LXXXI. B. LXXXI.
LXXXII. B. LXXXII.	LXXXIII. B. LXXXIII.	LXXXIV. B. LXXXIV.	LXXXV. B. LXXXV.	LXXXVI. B. LXXXVI.	LXXXVII. B. LXXXVII.	LXXXVIII. B. LXXXVIII.
LXXXIX. B. LXXXIX.	XC. B. XC.	XCI. B. XCI.	XCII. B. XCII.	XCIII. B. XCIII.	XCIV. B. XCIV.	XCV. B. XCV.
XCVI. B. XCVI.	XCVII. B. XCVII.	XCVIII. B. XCVIII.	XCIX. B. XCIX.	C. B. C.	CI. B. CI.	CII. B. CII.
CIII. B. CIII.	CIV. B. CIV.	CV. B. CV.	CVI. B. CVI.	CVII. B. CVII.	CVIII. B. CVIII.	CIX. B. CIX.
CLX. B. CLX.	CLXI. B. CLXI.	CLXII. B. CLXII.	CLXIII. B. CLXIII.	CLXIV. B. CLXIV.	CLXV. B. CLXV.	CLXVI. B. CLXVI.
CLXVII. B. CLXVII.	CLXVIII. B. CLXVIII.	CLXIX. B. CLXIX.	CLXX. B. CLXX.	CLXXI. B. CLXXI.	CLXXII. B. CLXXII.	CLXXIII. B. CLXXIII.
CLXXIV. B. CLXXIV.	CLXXV. B. CLXXV.	CLXXVI. B. CLXXVI.	CLXXVII. B. CLXXVII.	CLXXVIII. B. CLXXVIII.	CLXXIX. B. CLXXIX.	CLXXX. B. CLXXX.
CLXXXI. B. CLXXXI.	CLXXXII. B. CLXXXII.	CLXXXIII. B. CLXXXIII.	CLXXXIV. B. CLXXXIV.	CLXXXV. B. CLXXXV.	CLXXXVI. B. CLXXXVI.	CLXXXVII. B. CLXXXVII.
CLXXXVIII. B. CLXXXVIII.	CLXXXIX. B. CLXXXIX.	CCL. B. CCL.	CCLI. B. CCLI.	CCLII. B. CCLII.	CCLIII. B. CCLIII.	CCLIV. B. CCLIV.
CCLV. B. CCLV.	CCLVI. B. CCLVI.	CCLVII. B. CCLVII.	CCLVIII. B. CCLVIII.	CCLIX. B. CCLIX.	CCLX. B. CCLX.	CCLXI. B. CCLXI.
CCLXII. B. CCLXII.	CCLXIII. B. CCLXIII.	CCLXIV. B. CCLXIV.	CCLXV. B. CCLXV.	CCLXVI. B. CCLXVI.	CCLXVII. B. CCLXVII.	CCLXVIII. B. CCLXVIII.
CCLXIX. B. CCLXIX.	CCLXX. B. CCLXX.	CCLXXI. B. CCLXXI.	CCLXXII. B. CCLXXII.	CCLXXIII. B. CCLXXIII.	CCLXXIV. B. CCLXXIV.	CCLXXV. B. CCLXXV.
CCLXXVI. B. CCLXXVI.	CCLXXVII. B. CCLXXVII.	CCLXXVIII. B. CCLXXVIII.	CCLXXIX. B. CCLXXIX.	CCLXXX. B. CCLXXX.	CCLXXXI. B. CCLXXXI.	CCLXXXII. B. CCLXXXII.
CCLXXXIII. B. CCLXXXIII.	CCLXXXIV. B. CCLXXXIV.	CCLXXXV. B. CCLXXXV.	CCLXXXVI. B. CCLXXXVI.	CCLXXXVII. B. CCLXXXVII.	CCLXXXVIII. B. CCLXXXVIII.	CCLXXXIX. B. CCLXXXIX.
CCCL. B. CCCL.	CCCLI. B. CCCLI.	CCCLII. B. CCCLII.	CCCLIII. B. CCCLIII.	CCCLIV. B. CCCLIV.	CCCLV. B. CCCLV.	CCCLVI. B. CCCLVI.
CCCLVII. B. CCCLVII.	CCCLVIII. B. CCCLVIII.	CCCLIX. B. CCCLIX.	CCCLX. B. CCCLX.	CCCLXI. B. CCCLXI.	CCCLXII. B. CCCLXII.	CCCLXIII. B. CCCLXIII.
CCCLXIV. B. CCCLXIV.	CCCLXV. B. CCCLXV.	CCCLXVI. B. CCCLXVI.	CCCLXVII. B. CCCLXVII.	CCCLXVIII. B. CCCLXVIII.	CCCLXIX. B. CCCLXIX.	CCCLXX. B. CCCLXX.
CCCLXXI. B. CCCLXXI.	CCCLXXII. B. CCCLXXII.	CCCLXXIII. B. CCCLXXIII.	CCCLXXIV. B. CCCLXXIV.	CCCLXXV. B. CCCLXXV.	CCCLXXVI. B. CCCLXXVI.	CCCLXXVII. B. CCCLXXVII.
CCCLXXVIII. B. CCCLXXVIII.	CCCLXXIX. B. CCCLXXIX.	CCCLXXX. B. CCCLXXX.	CCCLXXXI. B. CCCLXXXI.	CCCLXXXII. B. CCCLXXXII.	CCCLXXXIII. B. CCCLXXXIII.	CCCLXXXIV. B. CCCLXXXIV.
CCCLXXXV. B. CCCLXXXV.	CCCLXXXVI. B. CCCLXXXVI.	CCCLXXXVII. B. CCCLXXXVII.	CCCLXXXVIII. B. CCCLXXXVIII.	CCCLXXXIX. B. CCCLXXXIX.	CCCLXXX. B. CCCLXXX.	CCCLXXXI. B. CCCLXXXI.

Sodali committe sic consilium clauem vel. clauem consilium tota sic committe sodali
cu volis ipsis mysteria vestra notatis
Non quinq; vider leges et si doctior extet
Hinc meritis utens liber. discrimine gratos
Ad celare viam qui monstrari tibi tunc

Round Three (March '04)

1. Scenes from an extraordinary life (the Disraeli exhibition)
2. Digitization of Champollion and Rosellini
3. Flora Graeca in the C21st. – Stage 3
4. Bringing Laxton to life: a unique insight into feudal England
5. The Gough Map – gateway to medieval Britain
6. Maps on the Web – placing the Bodleian's existing collection of cartographic slides and filmstrips on the Library website
7. Prints from the Curzon
8. Arthur Evans archive (Knossos excavations)
9. Writing Blanks, Board Games, and other Educational Games...John Johnson and Harding Collections

Total commitment = £89,986

General lessons learned

- Clear ToR for the Editorial Board
- Need robust agreements with project partners
- Schedule reviews of procedures and policies, and make revisions periodically rather than progressively
- Be prepared to commission more projects
- Better management of internal publicity and communications
- “Accentuate the positive” vs “Learning from our mistakes”

Specific lessons learned

- Lack of good metrics for costing and monitoring digitization projects
- Selection can be difficult, time-consuming, and expensive
- Metadata – most of our librarian partners still don't really get it
- Partners don't seem to have appreciated the likely on-going effort required to maintain and update online resources

Life after the Development Fund

- Concentration on bringing legacy collections into the new ODL framework
- Changing financial situation
- Need to reassert and re-sell the strategic role of the ODL
- Oxford-Google digitization partnership

plicam sententia in qua dicitur ut possint comprehendere cunctis: scilicet que sit latitudo. Longitudo.
 sicut inquit. p[ro]funditas. q[ui]a et potentia n[on] inu[en]it[ur] e[ss]e cognoscere. siue dicendo q[uo]d de ea alibi legitur.
 ut addendo que docent[ur] xpi gra[m]a inde extorgant pot[est]at[em]. paucis placet enotare. Remouenda: q[ui]a nob[is]
 in p[ri]ncipalib[us] debet ex[em]plum. ut hoc ap[osto]l[us] n[on] quia uo[le]t de[m]onstrare. q[ui]a fideles ad monerem quere
 que sit altitudo. q[ui]a que longitudo. l[at]itudo. p[ro]funditas. p[er] hanc sententia[m] s[ed] de s[er]p[ente]
 t[er]ra[m] t[er]rebat dicit. Altitudine q[ui]a latitudine. p[ro]funditas. q[ui]a s[er]p[ens] dicitur. V[er]u[m] uolens saltim. n[on] d[icit]ur. f[er]re. q[ui]a
 hanc u[er]ba sententia n[on] inuenit[ur] cruce dicitur. Accipit. sine cui[us] mysterio nulli filios. Ad p[er]uenit addit. cui p[er]sona
 sententia sic potest ap[er]iri. ut latitudine dicitur illud cunctis ligni crucis. p[er] q[uo]d brachia d[omi]ni se p[ri]ma s[er]uato
 exentia. Longitudine dicitur accipiam illa p[ar]te al[te]r[us] ligni que t[er]ra subterfuit ligno usq[ue] ad terram. ubi max[ime]
 p[ar]t[em] d[omi]ni corporis p[er]p[et]uit. altitudine d[icitur] illa p[ar]te que sup[er] emineat t[er]ra usq[ue] ad terram. ubi max[ime]
 sub terra fuit sup[er] intell[ect]um. Hec ad l[at]itudinem. V[er]u[m] h[ec] p[er] latitudinem carnis. p[ro]funditate illud ad
 u[er]u[m] n[on] accipit. p[ro]funditate. tolerancia lat[er]is que longitudo uniuersa. p[ar]te sup[er]ioris intell[ect]us. p[er] s[er]uato
 limine. altitudo. p[ri]ncipal[is] celestiu[m] deservat. p[ro]funditas in p[er]p[et]uitate iudicis d[icitur] p[ro]funditas. ut s[er]uato
 timor declaratur. q[ui]a s[er]uato. Ad q[uo]d p[er] l[at]itudinem dicitur altitudo d[icitur] p[ro]funditas. ut s[er]uato
 p[er]p[et]uitate d[icitur] p[ro]funditas. Ad p[er]p[et]uitate d[icitur] p[ro]funditas. ut s[er]uato p[er]p[et]uitate d[icitur] p[ro]funditas. ut s[er]uato
 p[er]p[et]uitate d[icitur] p[ro]funditas. ut s[er]uato p[er]p[et]uitate d[icitur] p[ro]funditas. ut s[er]uato p[er]p[et]uitate d[icitur] p[ro]funditas. ut s[er]uato

latitudinem implentem. p[er] i[st]a p[ar]te carnis
 inuenit[ur] diligit p[er] d[omi]n[u]m. Hec mag[is] de
 carnis confutatur. id est: tanta spe celestium
 carnis. nisi quoniam illud ad q[uo]d d[icitur] ad
 p[er]p[et]uitate n[on] ipse d[icitur] qui p[er]p[et]uitate
 s[er]uato d[icitur] p[er]p[et]uitate p[er]p[et]uitate. Hec. v[er]u[m] p[er]p[et]uitate
 s[er]uato d[icitur] p[er]p[et]uitate s[er]uato. s[er]uato. s[er]uato. s[er]uato.

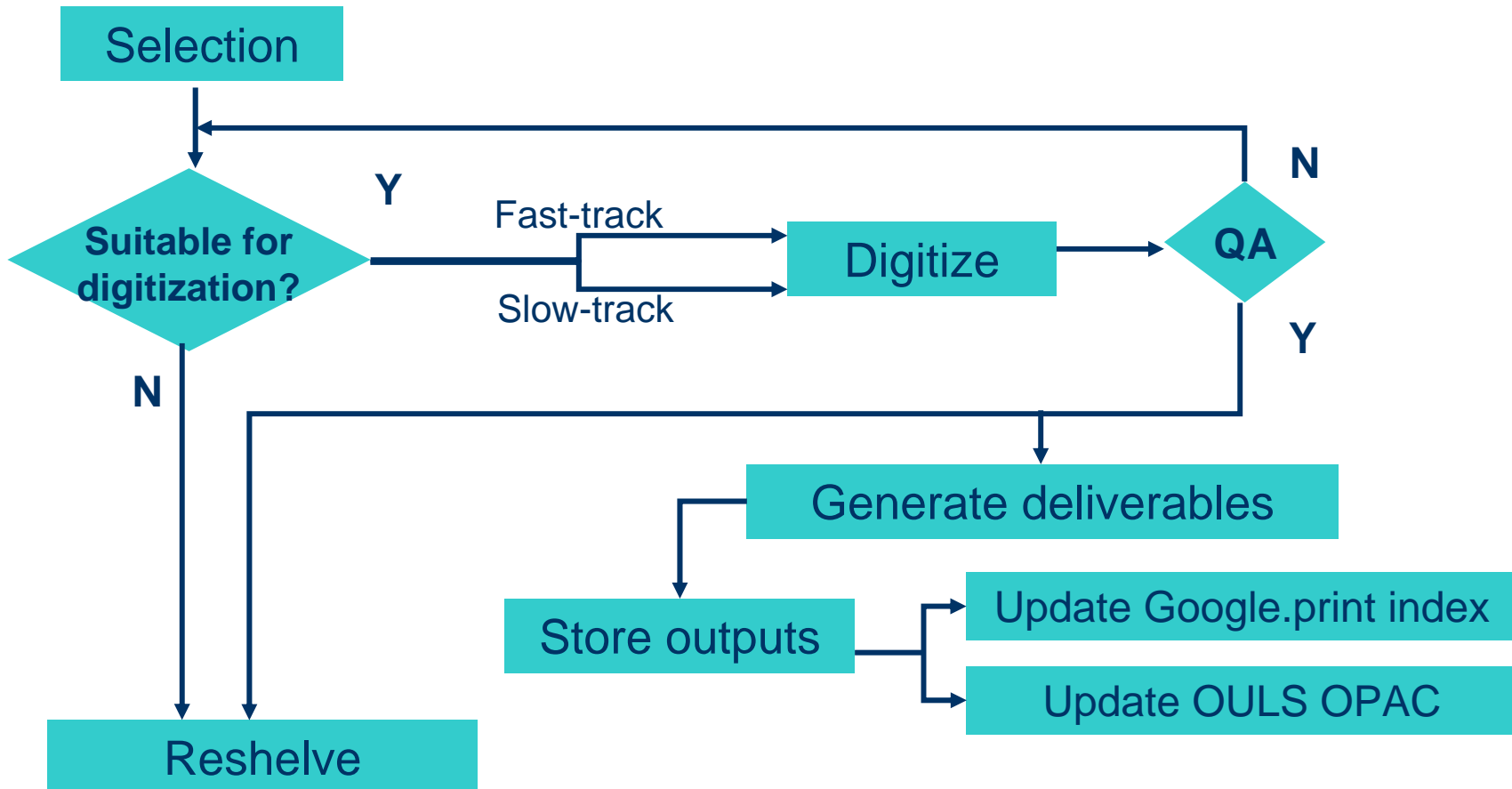
in p[er]p[et]uitate habet. ut p[er]p[et]uitate in o[mn]i
 latione fundat. fiduciam s[er]uato ad s[er]uato
 p[er]p[et]uitate roboret. ut n[on] ei ei pla
 p[er]p[et]uitate possit aduocare. h[ec] q[ui]a
 hanc crucis sup[er] s[er]uato. qua ad se
 p[er]p[et]uitate p[er]p[et]uitate p[er]p[et]uitate. h[ec]



Oxford-Google: what to digitize?

- Direct discussions with Google since 2003
- Win/win situation for both parties
- Extensive collection of out-of-copyright (and mostly out-of-print) material identified
 - Oxford differs from other partners in this aspect of our agreement
 - Decision made to begin with the 19th century material
 - Looking at approximately 1+ million items

Overview of workflow



Approach

- OULS staff will work closely with Google staff e.g. training on how to handle the material
- Each component of the workflow must be comfortable for both parties
- A large complex logistic operation that will not compromise the service to our users

Outputs and outcomes

- Large raw colour images from digitization process
- Per volume, OULS receives:
 - JPEG2000 (probably), and TIFFs
 - Uncorrected OCR
- Audit of production process
- There are quality control processes at Google & Oxford
- Deliverable images (to be hosted by Google in the first instance) linked to OPAC records
- Ongoing software/hardware developments to improve the process

Challenges that lie ahead...

- Building the local infrastructure to manage and deliver the Oxford Digital Copy of the data
- Investigating ways to exploit the data, e.g.:
 - Correcting OCR files, adding additional markup
 - (Re-)structuring the data – moving beyond a simple search and page-turning presentation
 - Completing/extending volumes and collections
 - Automatic collation, authorship attribution, plagiarism detection etc.....and many, many more(?!)
- Raising the barrier of what is possible, and end-users' expectations about what we can deliver

Useful links

- <http://www.bodley.ox.ac.uk/google/>
- <http://print.google.com/googleprint/library.html>