

Wolfson Digital Research Cluster



Inaugural Meeting

Wolfson Digital Research Cluster



What could a Digital Research Cluster do for the College?

Donna Kurtz

Professor of Classical Art and Beazley Archivist

Why clusters?

- ✓ Promote interdisciplinary research
- ✓ Raise profile of the college
- ✓ Attract funds for scholarships
- ✓ Attract funds for DRC activities,
- ✓ *e.g., Workshops, Summer Schools, etc.*



Why Wolfson?

- ✓ Graduate
- ✓ Large
- ✓ Arts and Sciences

Other clusters in college:
transfer of benefits



CLAROS – an example of a collaboration

- Humanities
 - Social Sciences
 - MPLS
 - Museums and Collections
-
- Kurtz and Shotton, GBFs
 - Membership across college



OeRC

www.oerc.ox.ac.uk

OeRC is the institutional base
for CLAROS

OeRC can coordinate new
projects created by DRC



OII

www.oii.ox.ac.uk

OII could serve this role
for:

- Politics
- Economics
- e-governance



What is Digital Research?

Almost any type of research

Expressions of interest have come from:

- Medical Sciences
- Politics
- Social Sciences and Law
- Humanities



Wolfson Digital Research Cluster

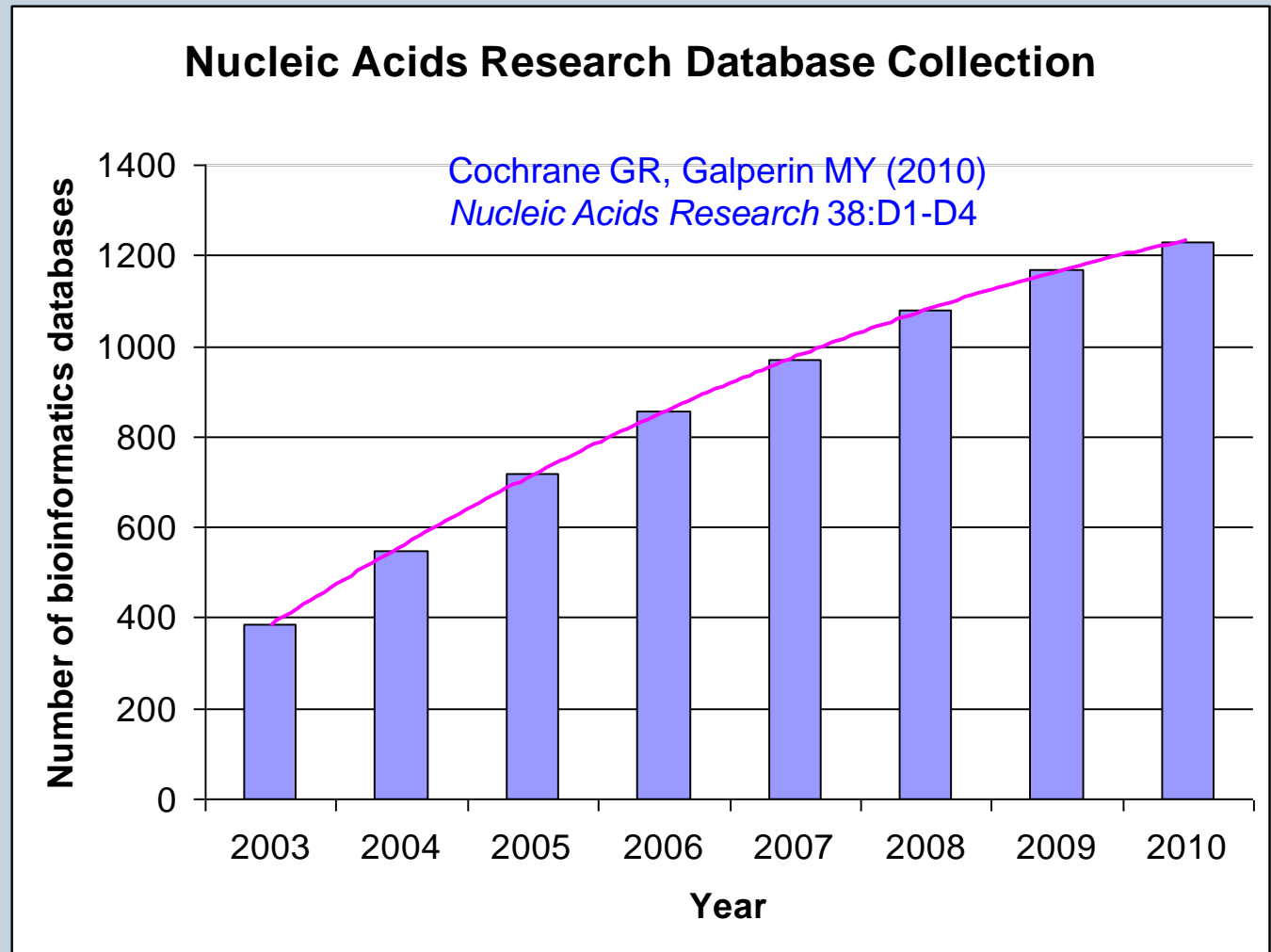


Biological data management and semantic enhancements of scholarly publishing

David Shotton

Reader in Image Bioinformatics

The challenges of the biological data deluge



OpenFlyData sources: *Drosophila* gene expression data









FlyTED: the *Drosophila* Testis Gene Expression Database

Home Browse by Gene Name Browse by Strain Browse by Expression Location Links

Welcome to FlyTED

Search: Simple Search

- Browse by Gene Name
- Browse by Strain
- Browse by Expression Location
- Related Links
- About the Database

BDGP in situ Homepage - Mozilla Firefox

File Edit View History Database Parameters Tools Help



Patterns of gene expression in *Drosophila* embryogenesis

Index 1 (March 2007)

Project overview:

We use high throughput [BDGP in situ](#) to determine patterns of gene expression during embryogenesis for *Drosophila*. Genes represented in our database are *Drosophila* ESTs (CGC1 and CGC2). At the end of each [BDGP in situ](#) experiment, gene expression patterns are determined by taking a large number of [digital images](#) of individual embryos. The quality and clarity of the captured image data are verified by independently derived [BDGP in situ](#) images of gene expression using Affymetrix OneChip technology (microarray data from [BDGP](#)). Gene expression patterns are [associated with gene names](#) for developmental analysis of *Drosophila* embryogenesis. Image, microarray and association data are stored in a modified version of [Gene Ontology](#) database and the entire dataset is available on the web in a browsable and searchable form (see below) or using [Gene Ontology](#) tools. So far we measured expression of 4336 genes and documented 5418 of them with 7441 digital photographs. [Gene Ontology](#) summarizes all genes that have in situ images.

Search tools:

Gene:

Browse controlled vocabulary used to annotate in situ expression patterns



BDGP in situ Homepage - Mozilla Firefox

File Edit View History Database Parameters Tools Help

Change to look for: e.g. this, cell adhesion, receptor, apoptosis, adhesion, CGC147, etc.

Searching for CGC147 through 16769 annotations produced 1 hit:



[cdm147 \(key: 147\) \(FlyBase:FBgn0000025, FBgn0000025, 147001, 147001\)](#)

Annotation: [CDM147 \(key: 147\) \(FlyBase:FBgn0000025, FBgn0000025, 147001, 147001\)](#)

BDGP in situ Homepage: [BDGP in situ Homepage](#) (see [BDGP in situ Homepage](#) for details) [BDGP in situ Homepage](#) (see [BDGP in situ Homepage](#) for details)

CDM147 has 2 BDGP in situ images (see [BDGP in situ Homepage](#) for details)

Tissue	mRNA Signal	Present	Call Enrichment	Alt. call
Brain	1 ± 0	0 of 4	0.00	Down
Heart	0 ± 0	0 of 4	0.00	Down
Thoracic/abdominal ganglion	1 ± 0	0 of 4	0.00	Down
Salivary gland	3 ± 1	0 of 4	0.04	Down
Crop	0 ± 0	0 of 4	0.00	Down
Midgut	0 ± 0	0 of 4	0.00	Down
Tubule	1 ± 0	0 of 4	0.00	Down
Midgut	0 ± 0	0 of 4	0.00	Down
Ovary	0 ± 0	0 of 4	0.00	Down
Testis	1140 ± 36	4 of 4	13.90	Up
Male accessory glands	0 ± 0	0 of 4	0.00	Down
Virgin spermatheca	0 ± 0	0 of 4	0.01	Down
Mated spermatheca	1 ± 0	0 of 4	0.01	Down
Adult carcass	1 ± 0	0 of 4	0.00	Down
Larval salivary gland	0 ± 0	0 of 4	0.01	Down
Larval midgut	1 ± 0	0 of 4	0.01	Down
Larval tubule	0 ± 0	0 of 4	0.00	Down
Larval hindgut	0 ± 0	0 of 4	0.00	Down
Larval fat body	6 ± 43	3 of 4	0.60	None
Whole fly	61 ± 7	4 of 4	0.60	None

FlyBase Gene Report: Dmef1schuy - Mozilla Firefox

File Edit View History Database Parameters Tools Help

FlyBase

Home Tools Files Species Documents Resources News Help Archives

Gene Dmef1schuy

Profile Manager:

Gene Information	Species	D. melanogaster
Symbol	Dmef1schuy	Dmef1schuy
Name	Dmef1schuy	Dmef1schuy
Protein type	protein, coding, gene	Protein type
Gene Model Status	Current	Created/Updated
Characterization Status	Uncharacterized	2003-12-01/2003-12-01
Gene Location	3L	18,079,177-18,081,272(+)
Cytogenetic map	18C4-18C4	18C4-18C4
Map	18C4-18C4	18C4-18C4
Hybridization	18C4-18C4	18C4-18C4
Summary Information	<p>The gene Dmef1schuy is referred to in FlyBase by the symbol Dmef1schuy (CG11735, FlyBase:FBgn0000025). It is a protein-coding gene from <i>Drosophila melanogaster</i>. It has the biological role of being a gene. Its sequence location is 18,079,177-18,081,272. It is located on chromosome 3L. The biological processes in which it is involved are not known. Gene details in FlyBase: the phenotype data is available. It is a gene annotated in FlyBase and is available in FlyBase.</p>	



Same data integrated in a single OpenFlyData window

Search *D. melanogaster* Gene Expression Data by Gene

gene name:

E.g. schuy, CG17736 or FBgn0036925 (case doesn't matter)

gene expression levels by tissue

found 1 matching probe from flyatlas.org (retrieved on 2008-09-16) for gene **schuy** ...

probe	tissue	mRNA signal	present call	enrichment	affy call
1626032_at	whole	81.9 ± 7.9	4 of 4	-	-
	brain	1.2 ± 0.6	0 of 4	0	Down
	head	0.4 ± 0	0 of 4	0	Down
	crop	0.8 ± 0.2	0 of 4	0	Down
	midgut	0.6 ± 0	0 of 4	0	Down
	hindgut	0.6 ± 0.1	0 of 4	0	Down
	tubule	1.5 ± 0.7	0 of 4	0	Down
	ovary	0.5 ± 0.2	0 of 4	0	Down
	testis	1140.4 ± 36.8	4 of 4	13.9	Up
	acc	0.9 ± 0.1	0 of 4	0	Down
	l_tubule	0.6 ± 0.2	0 of 4	0	Down
	l_fatbody	68.5 ± 43.0	3 of 4	0.8	None
	ta_ganglion	1.4 ± 0.5	0 of 4	0	Down
	carcass	1.3 ± 0.3	0 of 4	0	Down
	sgland	3.5 ± 1.3	0 of 4	0.04	Down

references (flybase)

found 8 references from flybase.org (FB2009_02) for gene **schuy** (FlyBase report: [FBgn0036925](#)) ...

Barreau et al., 2008, Development 135(11): 1897--1902
Post-meiotic transcription in Drosophila testes. [[FBrf0205264](#)]

Dickson et al., 2007.7.18, RNAi construct and insertion data submitted by the Vienna Drosophila RNAi Center
RNAi construct and insertion data submitted by the Vienna Drosophila RNAi Center [[FBrf0200327](#)]

Benson et al., 2006, A. Dros. Res. Conf. 47: 494A
The Drosophila testis gene expression database. [[FBrf0189116](#)]

Benson et al., 2005, Europ. Dros. Res. Conf. 19: G615
The Drosophila testis gene expression database. [[FBrf0194732](#)]

Benson, 2004.2.24, Helping FlyBase: ADRC-10142.
Helping FlyBase: ADRC-10142. [[FBrf0178789](#)]

Benson et al., 2004, A. Dros. Res. Conf. 45: 398B
Epigenetic Regulation by the aly-class Meiotic Arrest Genes. [[FBrf0173441](#)]

in situ hybridisation in embryos

found 8 matching images from embryofly.org (retrieved on 2008-10-30) for gene **schuy** (BDGP report: [CG17736](#)) ...

stage 11-12

no staining;



stage 13-16

gonad;



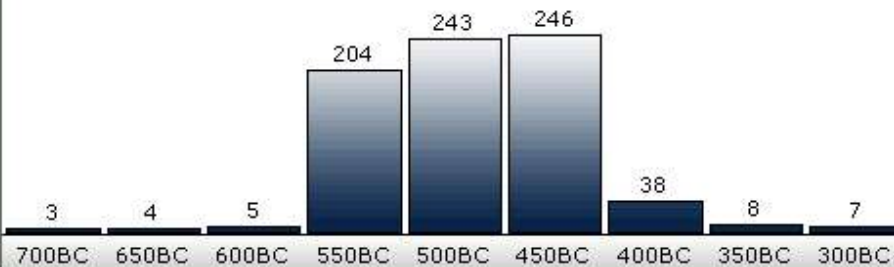
in situ hybridisation in testes

found 11 matching images from www.fly-ted.org (retrieved on 2008-12-03) for gene **schuy** (Fly-TED reports: [schuy](#)



The CLAROS Explorer classical art browse interface

The timeline shows the number of occurrences in each period. Click on the bar to show the distribution within the period or [click here to view the distribution for all periods](#).



Distribution of heron for all time periods

Click on the marker ▼ on the map and a balloon will pop up with the site name and number of occurrences of the name found at that site.



Page 13



40008233, Edinburgh Tassie, 2203, EROS HOLDING A HERON BY THE NECK AND RESTING HIS BOW ON THE GROUND, Unpublished Tassie, TRAY 35.2



40000953, Poniatowski, T915, DIOMEDE AND ULYSSES PROCEEDING AS SPIES AT NIGHT TO THE TROJAN CAMP [CONDUCTED BY THE HERON WHICH MINERVA HAD SENT TO GUIDE THEM], KROMOU, Kromos, Prendeville, J.: Explanatory catalogue of the proof-impressions of the antique gems possessed by the late Prince Poniatowski and now in the possession of John Tyrrell, Esq. (1841), 915, Cornelian

Web page search results 1 to 4 of about 4 for heron

[Dexamenos - Classical gems - Gems](#)

... 21mm. GGFR no. 467. Blue chalcedony scaraboid, from Kerch (Crimea). A flying **heron**. ...

A **heron** preening and catching a locust with his foot. Signed Dexamenos. St. ...

<http://www.clarosnet.org/gems/styles/classical/dexamenos.htm> - 8k - 2008-12-11

Other Claros references for this period



Pottery 102155 records



Rundplastik 37905 records



Relief 25488 records

Semantic enhancement of an exemplar research article



PLOS **NEGLECTED
TROPICAL DISEASES**

a peer-reviewed open-access journal published by the Public Library of Science

[Home](#) [Browse Articles](#) [About](#) [For Readers](#) [For Authors and Reviewers](#)

RESEARCH ARTICLE

OPEN  ACCESS

Impact of Environment and Social Gradient on *Leptospira* Infection in Urban Slums

Renato B. Reis^{1#}, Guilherme S. Ribeiro^{1#}, Ridalva D. M. Felzemburgh¹, Francisco S. Santana^{1,2}, Sharif Mohr¹, Astrid X. T. O. Melendez¹, Adriano Queiroz¹, Andréia C. Santos¹, Romy R. Ravines³, Wagner S. Tassinari^{3,4}, Marília S. Carvalho³, Mitermayer G. Reis¹, Albert I. Ko^{1,5*}

 [Download Article XML](#)

 [Download Article PDF](#)

 [Download Citation](#)

 [E-mail this Article](#)

 [Order Reprints](#)

 [Print this Article](#)

Citation: PLoS Negl Trop Dis 2(4): e228. 2008 doi:10.1371/journal.pntd.0000228

Page 14

Received: January 22, 2008; **Accepted:** March 27, 2008; **Published:** April 23, 2008



Current 'data care' R&D projects

- The ADMIRAL Project

- helping researchers in Zoology to locally manage research data, then publishing selected datasets in the Oxford DataBank

- The Dryad-UK Project

- publishing biological datasets related to peer-reviewed journal articles in the Dryad data repository

- The Open Citations Project

- publishing bibliographic citations as Linked Open Data

- Open Research Reports

- publishing structured summaries of infectious disease articles as open access mini-publications, to get round subscription barriers that block access to biomedical information in the developing world



Wolfson Digital Research Cluster

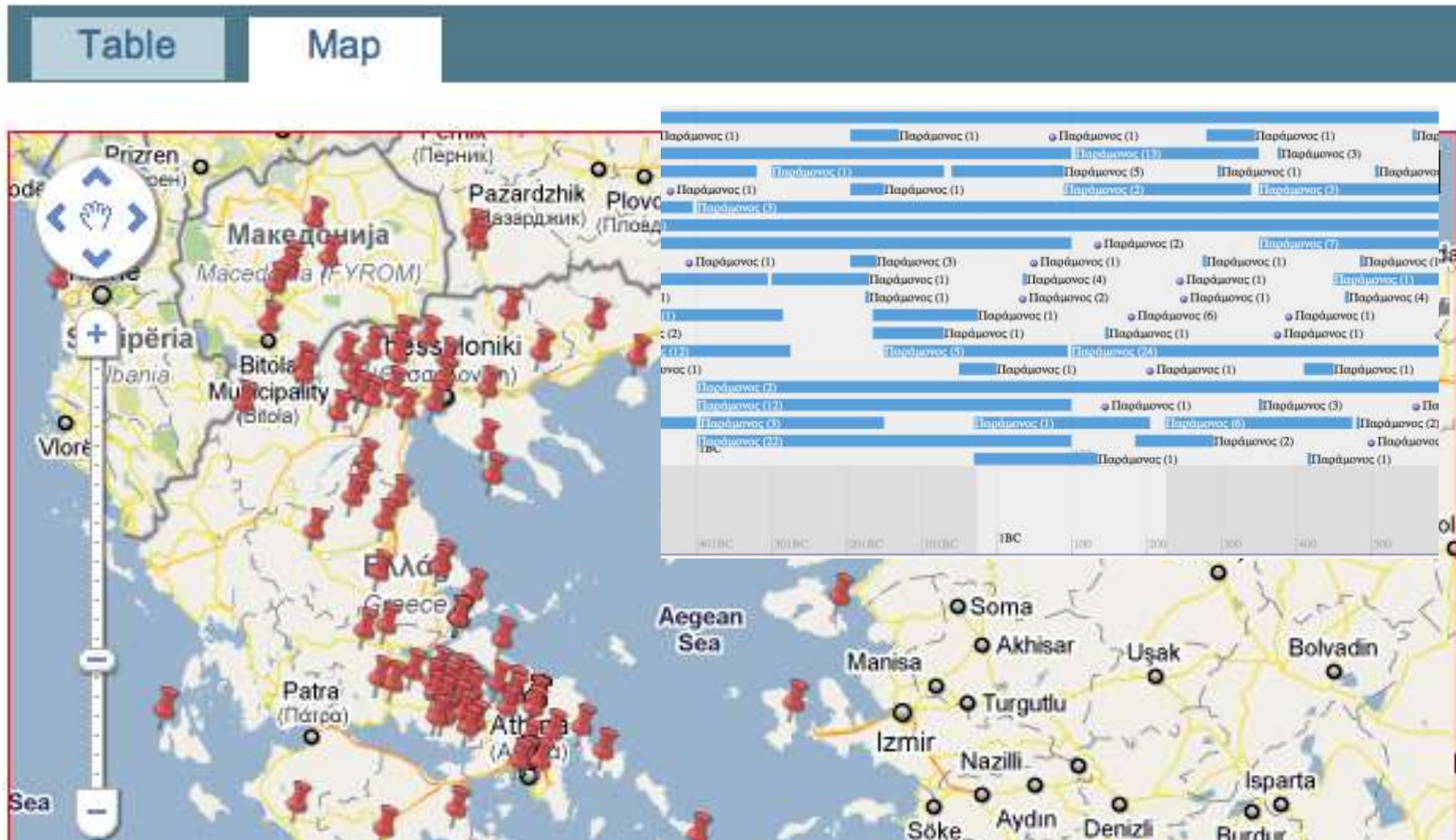


Places ancient and modern

Sebastian Rahtz

Head of Information and Support Group, OUCS

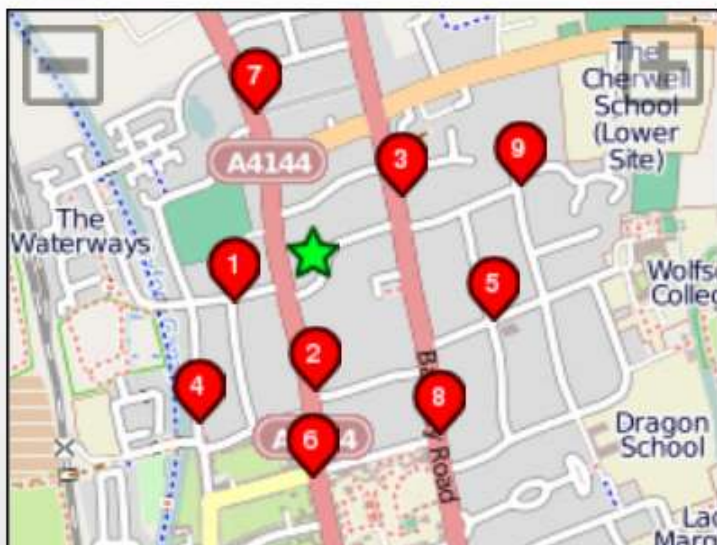
The name Παράμονος



Mobile Oxford



Found 9 post boxes near 25 Staverton Road



Rapid freezing, freeze fracture, and deep etching / edited by Nicholas J. Severs and David M. Shotton

Publisher: New York ; Chichester : Wiley-Liss, 1995

Description: xi, 372 p ; ill. : 27 cm

ISBN: 0471014338

Library holdings

Please note that all libraries have admission and borrowing policies; inclusion in this list does not imply access.



William Godwin's Diary

About ▾ Diary ▾ People ▾ Events ▾ Reading ▾ Writing ▾ Meals ▾ Meetings ▾ Search

Dep. e.203, fol. 25v

[Image List](#) -- [Full Image](#) -- [Diary Text](#)



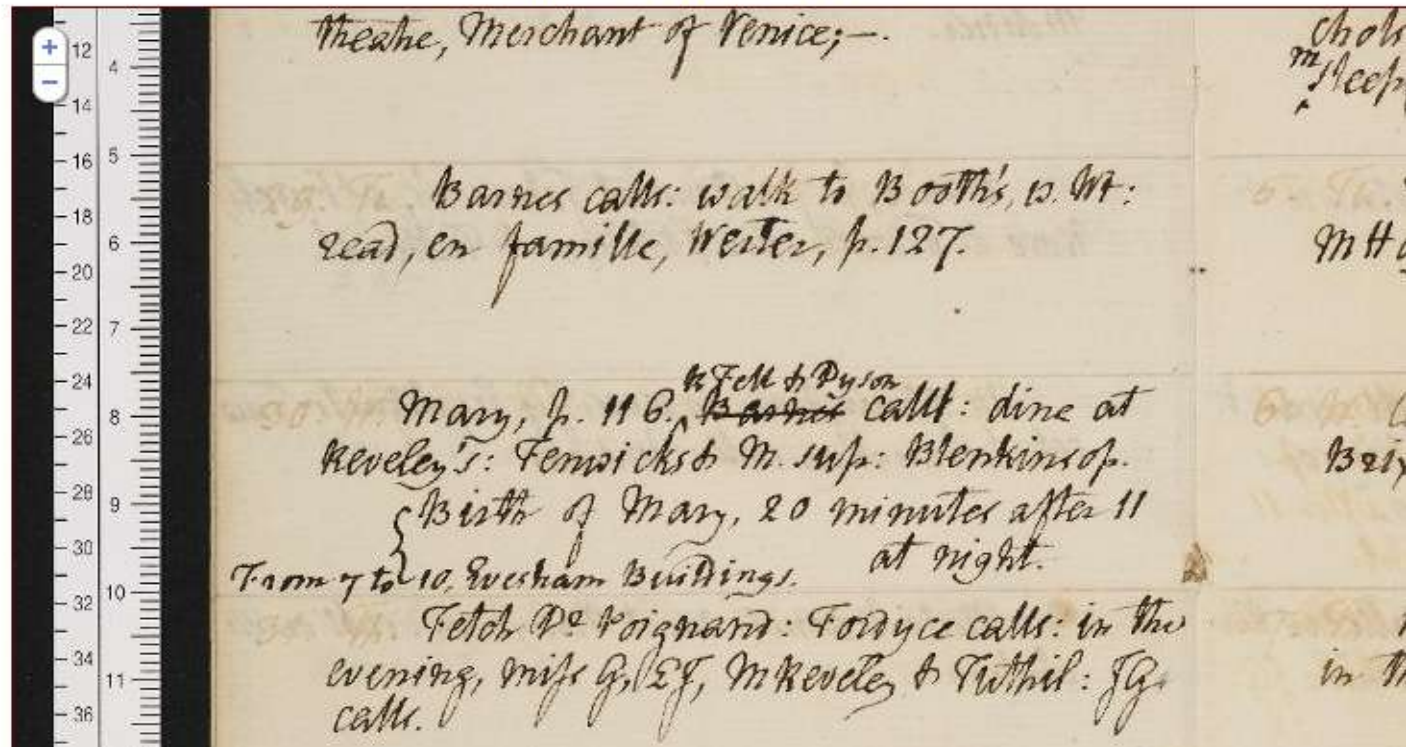
← [Prev. Fol.](#)



← [Prev. Vol.](#)

Volume List:

Jump to: e.196,
e.197, e.198, e.199,
e.200, e.201, e.202,
e.203, e.204, e.205,
e.206, e.207, e.208,
e.209, e.210, e.211,
e.212, e.213, e.214,
e.215, e.216, e.217,
e.218, e.219, e.220,
e.221, e.222, e.223,
e.224, e.225, e.226,
e.227, and e.273.



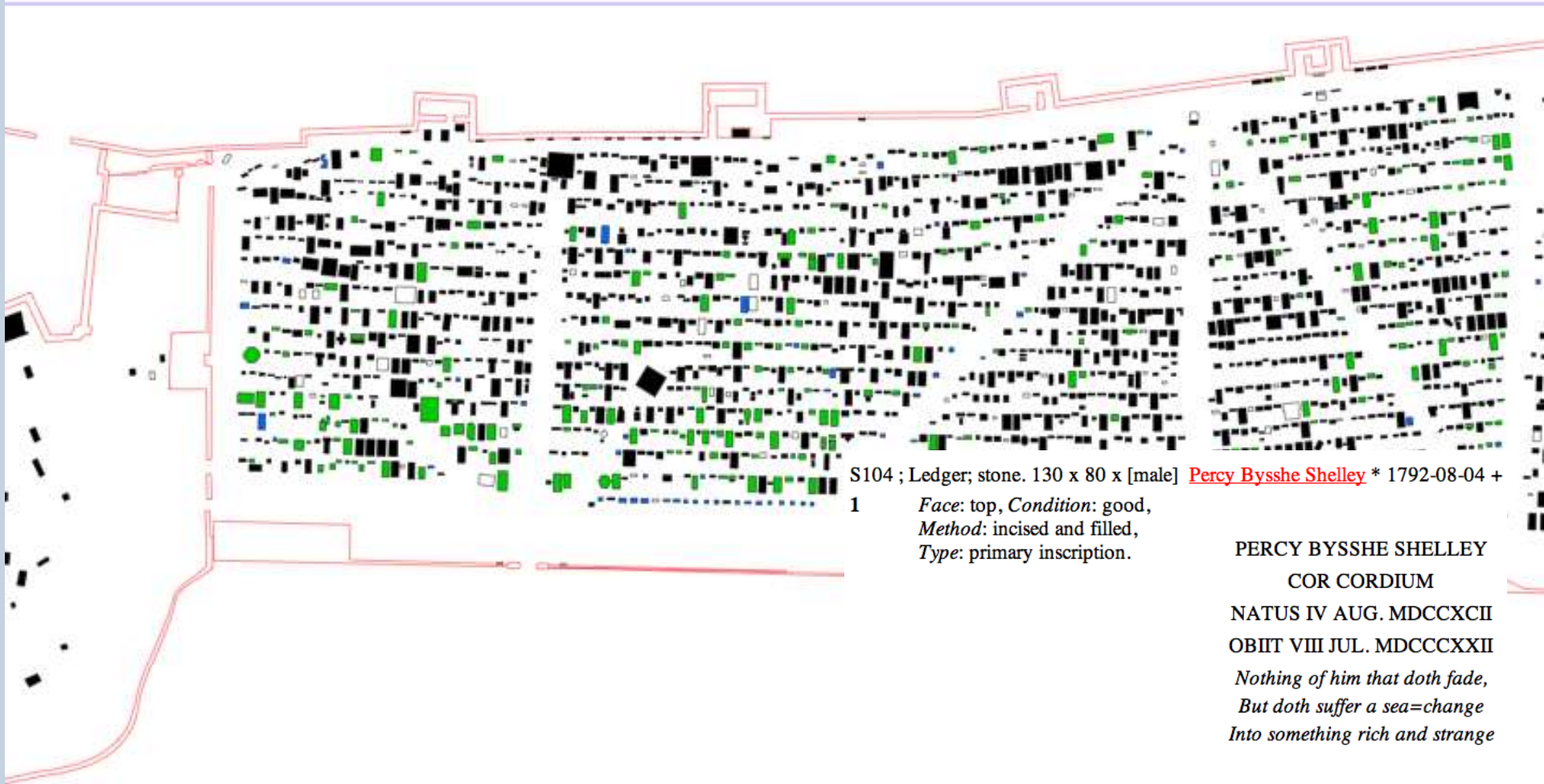
theatre, Merchant of Venice;--

Barnes calls: walk to Booth's, 12. Mt:
read, en famille, Webster, p. 127.

to Felt & Dyson
Mary, p. 116. Barnes calls: dine at
Revelley's: Fenwick & M. sup: Blenkinsop.
Birth of Mary, 20 minutes after 11
from 7 to 10, Rusham Buildings. at night.
Fetch Dr Poignard: Fordyce calls: in the
evening, Miss G. & F., M. Revelley & Tutthil: J. G.
calls.



Protestants buried in Rome



S104 ; Ledger; stone. 130 x 80 x [male] [Percy Bysshe Shelley](#) * 1792-08-04 +

1 Face: top, Condition: good,
Method: incised and filled,
Type: primary inscription.

PERCY BYSSHE SHELLEY
COR CORDIUM
NATUS IV AUG. MDCCXCII
OBIIT VIII JUL. MDCCCXXII
*Nothing of him that doth fade,
But doth suffer a sea=change
Into something rich and strange*

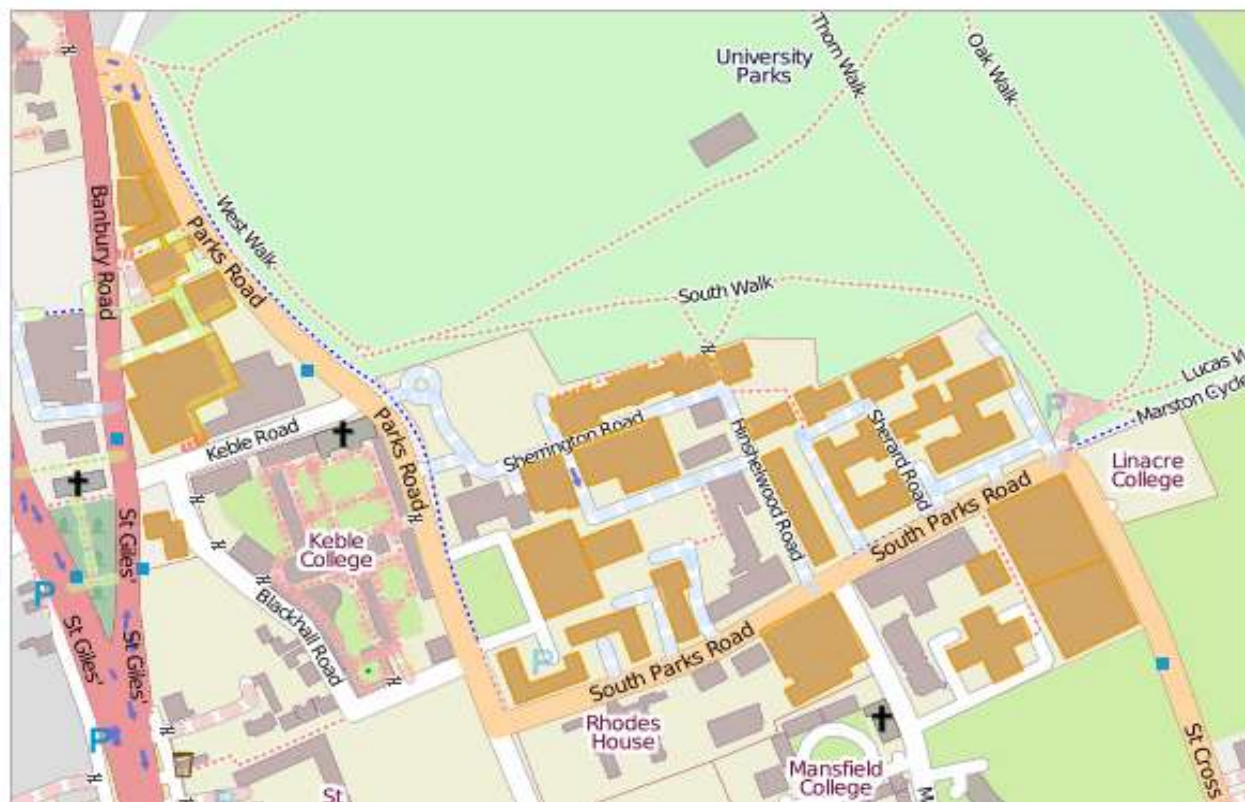


Oxford Science Area

[home](#) [datasets](#) [explore](#) [sparql](#) [contact](#)

University Science Area Map

A re-imagining of [the official science area map](#) using OpenStreetMap and OxPoints data.



Department of Biochemistry
Rodney Porter Building
New Biochemistry Building
Rex Richards Building
Department of Chemistry
Laboratory Chemical Biology
Department of Engineering Science
Thom Building
Holder Building
Jenkin Building
Engineering And Technology Building
Information Engineering
Inorganic Chemistry Laboratory
Department of Materials
Mathematical Institute
The Peter Medawar Building for Pathogen Research
Oxford University Bioinformatics Centre
Pathology
Rex Richards Building
Department of Pharmacology
Plant Sciences Library
Physical and Theoretical Chemistry Laboratory
Department of Earth Sciences
Pitt Rivers Museum
University Museum of Natural History
Medical Sciences Division
Astrophysics
Department of Physics
Particle Physics
Radcliffe Science Library
Centre for Cognitive Neuroscience
Department of Human Anatomy and Genetics



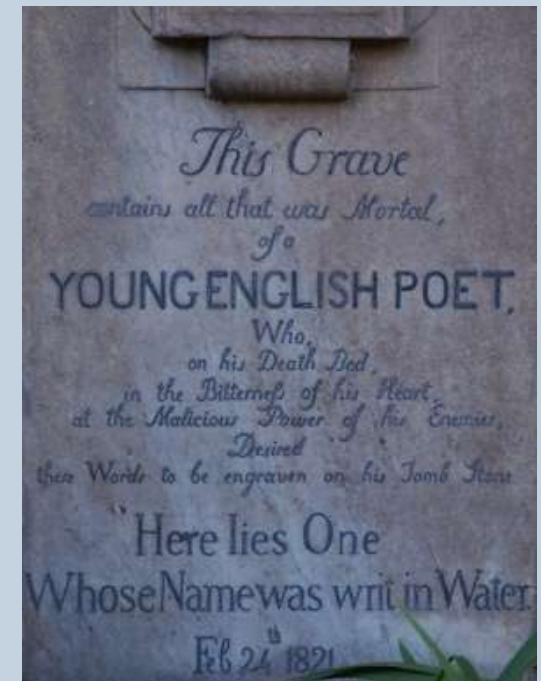
Underlying issues

- The importance of place
- The importance of time
- The importance of people
- Multiple sources and types of data

```
<person>
  <sex>1</sex>
  <persName>
    <forename>John</forename>
    <surname>Keats</surname>
  </persName>
  <death when="1821"/>
  <nationality key="UK"/>
  <occupation>poet</occupation>
</person>
```

Some answers

- Common vocabularies
- Common data formats
- Resources, not answers



The Metamorphoses Project aims

- Combine many sources of geolocation data (CLAROS, Pleiades, Yahoo, Geonames, Getty)
- Find points of commonality by name, location and date
- Model change of name and place over time
- Provide geolocation lookup service for Oxford
- Produce visualisations and mashups
- Use the same technologies and vocabularies as modern open data



5-star data

	http://www.w3.org/DesignIssues/LinkedData.html
★	Available on the web (whatever format), but with an open licence
★★	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
★★★	as (2) plus non-proprietary format (e.g. CSV instead of Excel)
★★★★	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
★★★★★	All the above, plus: Link your data to other people's data to provide context



Wolfson Digital Research Cluster



A database for the Roman Economy

Miko Flohr

Assistant Director, Oxford Roman Economy Project

The Oxford Roman Economy Project

- AHRC funded first phase (2005-2010), continuing until end 2012 at least.
- Studying economic history of the Roman world through large, quantifiable datasets.
- Aims:
 - Understanding economic geography
 - Understanding change over time (growth and decline)
 - Studying local phenomena to greater depth
- Using known (and published) archaeological and textual data



The OxREP databases, an inventory:

- Roman Shipwrecks
- Roman mines
- Roman villas with multiple sets of wine and olive presses
- (all) Roman cities
- A complete inventory of the Karanis Tax Rolls
- Land property lists
- Alimenta lists
- An inventory of machines mentioned in literature
- An inventory of canals in the Roman world
- Roman coins found in India

... and (many) more to be added in the future



What we needed

- A powerful, flexible database
- A good back-end website for updating the database
- A user-friendly front-end website for browsing, searching and querying the database, including:
 - Mapping facility
 - Chart facility
 - Facility to store queries
 - References to scholarly literature



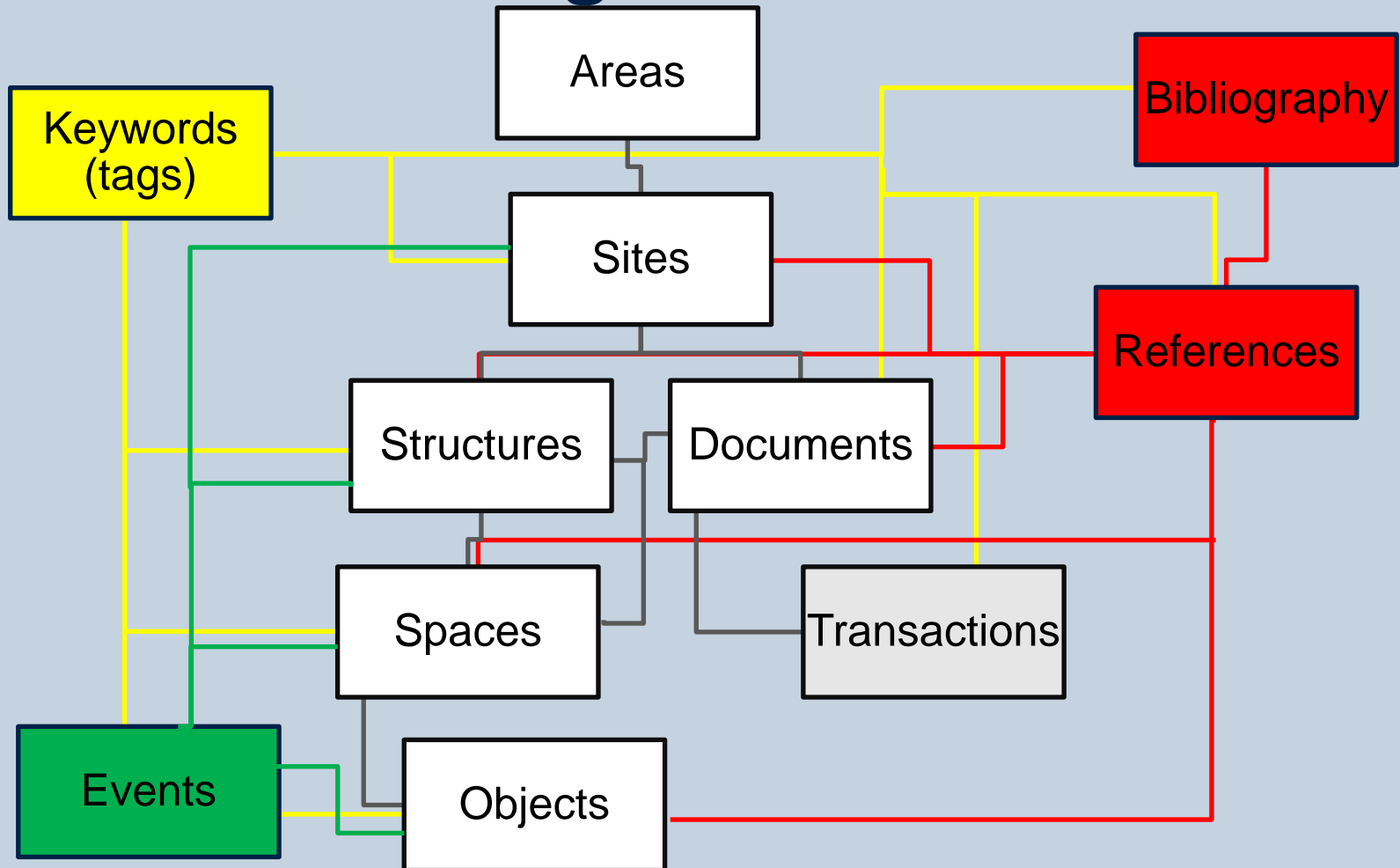
Database design: what we had...

... a quantity of databases:

- made by different people, some of whom had already left Oxford
- In different file format (access, excel)
- With different levels of detail, complexity and accuracy
- Sometimes with internal inconsistencies
- High reliance on text fields (and thus, inevitably, many typos)
- Different ways of referencing



Database design



The Oxford Roman Economy Project Database

 You are logged in as oxrep. [Logout](#)

General

[OXREP Home](#)
[Introduction](#)
[Map](#)
[People](#)

Databases

[Roman Mines](#)
[Shipwrecks](#)
[Wine and Olive Presses](#)
[Karanis Tax Rolls](#)

Data Browser

 Select Site ▼

 Browse by Keyword ▼

 Arts & Humanities
Research Council

Shipwrecks Database

Browse Data Query Builder (BETA)

[Browse by country](#) > [Turkey](#) > Select from 47 wreck sites... ▼
[Show List](#)

Shipwrecks List

This search returned 78 wrecks

	Site	Wreck	Description
1.	Akkale	none	
2.	Anataş Adacik	Anataş Adacik (Palapetra)	
3.	Antalya	Antalya	Unconfirmed reports from sponge divers indicate there may be as many as 4 copper ingot wrecks in the area. One of these seems to be earlier than the Cape Gelidonya A wreck.
4.	Arap Adasi	Arap Adasi	
5.	Bodrum	Bodrum	
6.	Bozburun	Bozburun Byzantine	Dating changed since Parker 1992 as the site was re-examined by the INA from 1995-98. The hull remains provide some of the earliest evidence for the arithmetical design and construction methods that later dominated ship building in the Mediterranean.
7.	Bozburun	Bozburun Julio-Claudian	A large amphora mound measuring 15 x 5m although it is hard to discern a stacking pattern as the upper amphorae are tumbled and there is a large amount of marine growth covering the site.
8.	Bozburun	Bozburun Late Antiquity anchor	An oval shaped amphora deposit measuring 9 x 3m with 2 collections of anchors about 4m to the west and northwest of the amphorae.
9.	Bozburun	Bozburun Late Republican	A badly damaged mound of amphorae measuring 10 x 8m, rising to 0.5-0.75m, although many are still in their upright position in the sand, suggesting that the hull may be preserved underneath.
10.	Bozburun	Bozburun tile wreck	The site is a small mound measuring c 21 x 10m comprising both cover and pan tiles still in their stacked positions. A trail of amphorae starting about 10m from the site and continuing 600m away contains at least 3 concentrated deposits of numerous large
11.	Bozukkale	Bozukkale	
12.	Çakil Burnu	Çakil Burnu	Amphorae paralleled in 1st c AD Coan forms.
13.	Çamalti Burnu	Çamalti Burnu 1	Over an area 42m x 18m, at least 800 stamped amphorae have been found on a 25m long merchant ship that had 2 masts and a half deck. The amphorae on this wreck have been dated to the thirteenth century, placing them among the last amphorae in



Wolfson Digital Research Cluster



Web-scale music analysis, linking, and life-writing

David De Roure

Professor of e-Research, OeRC

National Director of Digital Social Research, ESRC



Digital Music Collections

Structural
Analysis of Large
Amounts of Music
Information

Crowdsourced
ground truth



Community
Software

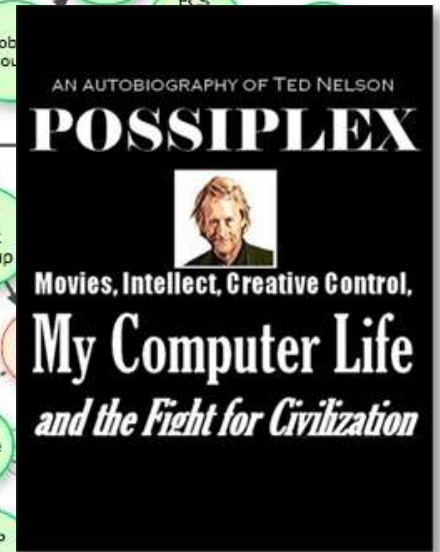


Supercomputer

Linked Data
Repositories



What's the
score at the
Bodleian?





Digital Social Research
Harnessing Technology for Social Science Research



Oxford e-Research Centre

innovative technology
accelerating research

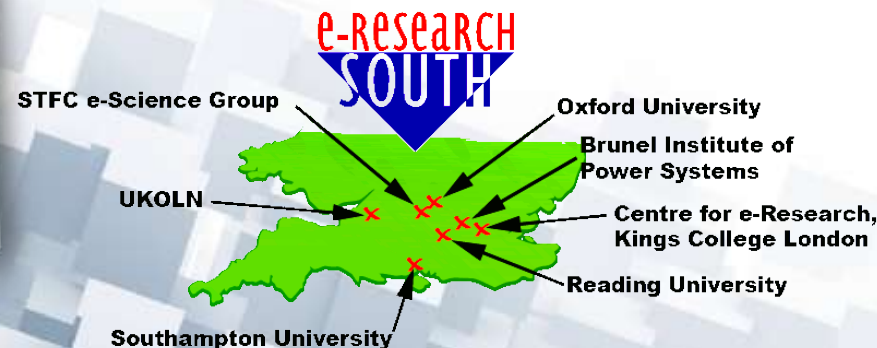
www.oerc.ox.ac.uk



Developing a coordination
platform:
digital.humanities.ox.ac.uk



Collaborating and supporting
Humanities projects



Wolfson Digital Research Cluster



Discussion

David Robey

Arts and Humanities Consultant, OeRC