

The best of all possible worlds

The AHDS has been working in collaboration with the English Heritage National Monuments Record (EHNMR), to include the 'National Inventory' for England within ArchSearch, the Archaeology Data Service catalogue. William Kilbride reflects on the experience

I disagree with Adam Smith. In the digital age, demand always exceeds supply. This claim can be supported through the biography of ArchSearch where the prodigious expansion of resources available seems permanently exceeded by the stunning escalation of expectation.

Since its launch in September 1997, ArchSearch has processed in the region of 13,750,000 requests. The data has grown from zero to about 90 gigabytes: 40 gigabytes are available instantly and another 50 gigabytes can be accessed on request.

It contains information on 969,141 archaeological sites, monuments and research projects from about 40 countries ranging in date from the hominins of Africa six million years ago, to the abandoned cruise missile bunkers at Greenham Common in Berkshire. The growth continues: some recent additions give an idea of the direction things are moving.

In April this year in collaboration with the English Heritage National Monuments Record (EHNMR), we were able to include the 'National Inventory' for England within ArchSearch.

This is the primary record of England's archaeological and architectural sites and contains just under 400,000 records. It includes archaeological, maritime, architectural and historical sites from earliest times to the present, and the records included within ArchSearch link to English Heritage's own PastScape service and in this way opens up further opportunities for research. It is the third of three complimentary data sets that English Heritage delivers to researchers through ArchSearch, it

In this newsletter...

Digital Picture results	2
Open Source? Low-cost digitisation	7
The Learning Index	8
Developing Linguistic Corpora	9
Diasporas, Migration & Identities	11
New AHDS collections	12

AHRC ICT Programme: Find out more about the Methods Network & ICT Strategy Projects from pages 4 - 6



The English Heritage National Monuments Record includes records on renowned English buildings such as Salisbury Cathedral

joins the Excavation Index for England and the Index to Microfilmed Archaeological Archives.

Although distinct, these data sets can be used together, rewarding users with really thorough insights into England's rich historic environment and generations of research into it.

Because they are incorporated into ArchSearch, these data sets can be used in conjunction with other related datasets, such as Canmore, the online National Monuments Record of Scotland, The Northern Ireland Sites and Monuments Record or the National Museums and Galleries of Wales.

continued on page 3



In the last edition of the AHDS Newsletter, we introduced the Digital Picture, a national initiative to establish a national overview of the issues, and potential solutions, relating to the use and impact of digital images within visual arts, Higher Education institutes and associated organisations.

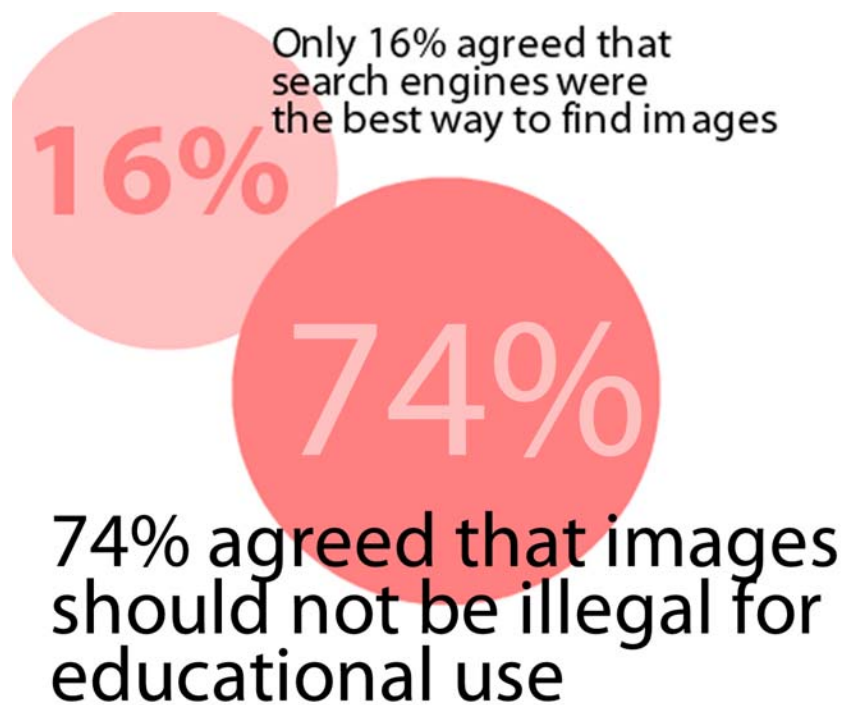
Mike Pringle's summary from the report is published here, along with some of the responses to the particular questions the Digital Picture asked.

It is clear, from the results of the Digital Picture, that there are serious concerns within the arts education community relating to the impact of digital technologies on arts education and, in some ways, on the arts in general.

Although the central focus of the Digital Picture was on digital images, it soon became apparent during the consultation with the community that discussion of the problems could not be constrained by such a label: the real fears relating to digital images are as much to do with educational processes and pedagogical values as with the quality of JPGs or the latest imaging software.

The vast majority of respondents to this survey feel that the increase of digital images in arts education has had an impact on them. For an extremely heartening three quarters of people, this impact materialises as an improvement in the learning and teaching environment, with two thirds believing that the rise in digital images helps them to be more computer literate.

However, less than a third of the community remains unconcerned at the corresponding loss of traditional products, skills or knowledge.



Much of the fear lies in the speed with which change is happening, and in concerns that there is a lack of forethought and consideration for any negative impact that such change brings.

Furthermore, the new technology seems to threaten livelihoods and the essential qualities of 'human' skills, with an apparent shift away from traditional expertise and resources towards increased 'virtual learning' and reliance on online materials with all attendant prob-

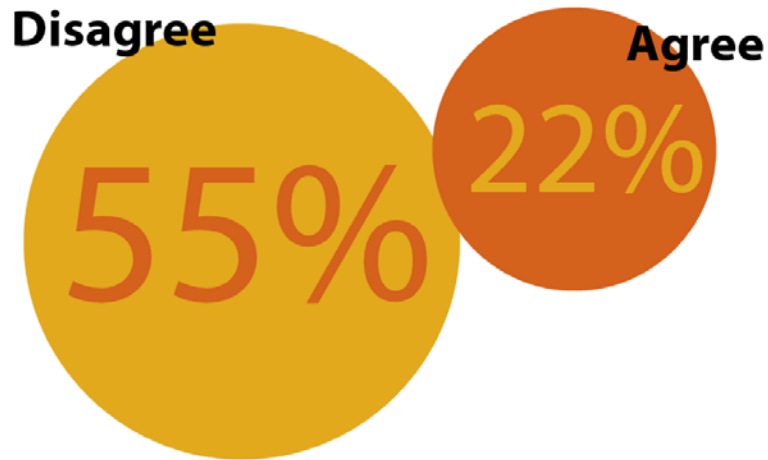
lems.

The community has problems with finding the right images for study; problems of reliable provenance and image quality; problems of ownership and licensing costs; and problems relating to educational agendas being led by technological advances. And, for tutors and students alike, there is a perceived lack of support and investment for the new technologies across the sector.

In particular, there are fears that, although a balance of old and new is to be desired, financial and strategic imperatives mean that this is unlikely to happen. Institutes will not, the community believes, maintain 'out-of-date' equipment and practices because of issues with cost, space, expertise and, increasingly, health & safety.

Despite such concerns, the consultation has clearly demonstrated that, overall, the community is keen to embrace the new technology, but in a measured and informed way, and in a manner that would compliment, rather than replace, traditional art teaching and practice. Sculptor Antony Gormley, when asked about digital imaging by AHDS Visual Arts Director Mike Pringle, at the Association of Art Historians' conference in April, perhaps sums up the art education community's feelings about digital media. He stated that his studio

Digital images are never good enough for studying closely or printing



practice now made use of digital technologies and that "the Angel of the North could not have been made without digital space", but also emphasised the importance to him of taking his work "out of the virtual and into the real".

See <http://thedigitalpicture.ac.uk/> for the full report.

Mike Pringle, AHDS Visual Arts, University College for the Creative Arts (Farnham campus)

All for the best ...

continued from the front page

In March we launched a new library of unpublished excavation reports. The research community often struggles to keep up with the fast turn around in archaeological research in the UK because much of the research is published informally or in very short print runs.

It has been estimated that there are perhaps 9,000 of these 'grey literature' reports produced every year in the UK, but access is notoriously difficult and quality is varied. The new library provides access to some 340 of these reports, many of them produced in the last twelve months.

That total is set to grow and grow: it has already doubled since March. This constant flow of information will greatly improve the life of researchers in the short term, and in

the medium term the increased visibility is likely to improve the quality of the reports too.

These are just two examples of how the data available through AHDS continues to increase in both breadth and depth. By any standards, the AHDS hosts and delivers a significant digital resource, but it's still only a fraction of what we could do and it's only a fraction of what we'd like to do.

Moreover, if a recent survey of users is to be taken seriously, it's only a fraction of what the user community wants.

"The best of all possible worlds will come about when the whole of the academic domain is accessible online and in digital form."

Two trends are obvious: a desire for ever more detailed and diverse resources, and a lack of information about the many resources already available. Expectation is such that it is hard for AHDS ever

to feel smug about our undoubted achievements.

For researchers who choose to rely on the internet, the best of all possible worlds will come about when the whole of the academic domain is accessible online and in digital form.

For librarians and those who support our information needs, the best of all possible worlds will come about once the user community is trained to articulate and resolve its information needs with self adapting and mature tools. Both are technically possible though both are practically remote.

Two realities prevail: the need to redouble our efforts in communication; and the need to extend the available resource. For AHDS, whose mission is about training and enhancing digital creativity, the challenge is already the best of all possible worlds.

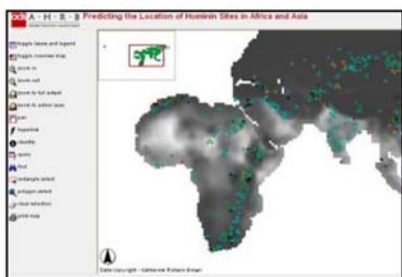
William Kilbride, ADS / AHDS Archaeology, University of York

AHRC's ICT Programme

The AHRC ICT Methods Network, a major initiative to promote and develop the use of ICT for advanced research methods in the Arts and Humanities, was officially established on 1st April, 2005.

It will run for three years as a national forum for the exchange and dissemination of expertise in the use of ICT for Arts and Humanities research across the whole range of subjects covered by the AHRC.

The aims and objectives of the Methods Network are to focus on new developments and advanced methods, and the critical engagement with digital resources by scholars in the arts and humanities.



The Methods Network will want to exploit the potential for technologies such as GIS (Geographical Information Systems) such as used in the Hominin Sites in Africa and Asia resource hosted by ADS / AHDS Archaeology

The Methods Network will also address the research processes and collaborations that can facilitate innovative work, by researching, reporting and addressing the strategic issues and questions that underpin the use of advanced ICT methods in the Arts and Humanities.

To this end, the Methods Network is implementing a three-year programme of activities to train scholars and practitioners, foster new modes of collaboration, build capacity, and create new research opportunities throughout the disci-

The Arts and Humanities Research Council launched its ICT Programme in October 2003, with funding of £3.8m, and it is designed to run until 2008.

Included here are two articles on the some of the main outputs of this project. Further below Stuart Dunn talks about the ICT Strategy Projects.

Meanwhile Lorna Hughes, talks about the Methods Network, which is supporting advanced use of computer technology in the arts and humanities

plines.

The first year's programme of activities and events was launched at a reception at the Digital Resources for the Humanities conference in September 2005 at the University of Lancaster.

Core activities and publications

The Methods Network has developed a broad and inclusive programme of core activities which invite the involvement of the community as a whole: subject specialists, humanities computing experts, postgraduates, and the staff of libraries and cultural heritage organizations.

While the AHDS concentrates on resource creation, delivery and preservation, the Methods Network "focuses on new developments and advanced methods"

The key activity for the first year will be a series of five "expert seminars" on the subject of advanced ICT methods for Linguistics, History, Music, Practice-based Arts, and Literature.

The presentations and discussions

resulting from these events will subsequently be published in a series of "Methods Guides".

In addition, the Network will develop a series of advanced workshops, adopting a "train the trainers" approach to build capacity around the community, and to disseminate training materials broadly.

Workgroups

The Methods Network is also developing a programme of research into the ways in which ICT for advanced research is used and supported in the Arts and Humanities. This research will primarily be conducted by means of convening a series of workgroups each year to address core strategic issues.

The workgroups will explore specific topics, report on key issues, and make recommendations, both for specific Methods Network activities in the workgroup's area of interest, and at a strategic level. Workgroups will, therefore, provide an important forum for members of the Arts and Humanities community to participate in an ongoing discussion that will consolidate a great deal of existing expertise.

Regular reports and updates on the progress of these groups, as well as final reports and recom-

AHRC ICT

METHODS NETWORK

mentations, will be disseminated broadly.

See the panel on the right for the first year's workgroups.

Tools Development

Over the past ten years, several initiatives have been launched to address the dearth of tools to support ICT methods in the Arts and Humanities. The question of tools development in humanities computing is presently being addressed by a number of international groups, which are addressing the financial, social, and political factors that hinder the development of such tools.

This workgroup will actively collaborate with these international initiatives, and seek input from the worldwide humanities and arts computing community on recommendations for supporting tools development.

Call for proposals

In addition to these core activities, the Methods Network is inviting the community as a whole to submit proposals for activities and publications that would serve to address specific training needs, discussion topics, or publication. Funding of up to £5,000 is available for proposals funded under this initiative.

A call for proposals for such activities is available on the Methods Network website; however, given that the Methods Network is conceived as a service that is flexible and responsive to the needs of the community, we welcome suggestions for proposals that may be outwith the parameters defined in the call for proposals – please contact us to discuss any ideas that you may have.

The Network Administrative Centre

Methods Network Workgroups

Postgraduate Training

This group will look at case studies to explore strategic issues, and assess the feasibility and desirability of developing curriculum models, specifying 'core' curricular components and developing common training materials.

Collaboration Modes and Recognition

The development of ICT methods for the Arts and Humanities is contingent on new models of working. Specifically, most initiatives must be collaborations between partner institutions and academics in order to be successful, and to convince colleagues of the value of the work that is produced. What are the implications of this new model of scholarship, and how is such work developed, supported and facilitated? How is it rewarded, assessed and reviewed?

Training and Awareness for Reviewers

As new models of working are developed, and ICT methods are increasingly used for Arts and Humanities research, it is imperative that the reviewers are able to competently review, understand and evaluate such work. This group will review the present situation and make recommendations for future training and awareness of reviewers.

All Methods Network activities are supported by the Network Administration Centre, which is based in the centre for Computing in the Humanities at King's College London.

*Lorna Hughes, Method's Network,
King's College London*

Stuart Dunn talks about the AHRC's new ICT Strategy Projects

The AHRC has launched a major new scheme to strategically survey, test and build the use and capacity of ICT across its subject domains.

Operating under the auspices of the ICT in Arts and Humanities Research Programme, the AHRC ICT Strategy Projects Scheme <http://www.ahrcict.rdg.ac.uk/activities/strategy_projects.htm> consists of 12 projects with individual funding of up to £140,000, which will each run for one or two years. The Programme has spent just under £1m on the Scheme.

The projects will function as a co-ordinated and carefully targeted programme with two fundamental aims.

Firstly, they will gather knowledge: conducting surveys, scoping needs and assessing current use of a range of ICT methods and resources.

Secondly they will develop capacity, tools and resources, adding to the corpus of applications available to arts and humanities researchers at points where there are currently critical and / or strategic deficiencies. The scheme is divided into two strands which reflect these aims.

The seven projects funded under the knowledge-gathering strand are all one-year undertakings, scheduled to report in time for the Fundamental Strategic Review which the ICT Programme will be conducting in 2006.

Survey work on needs will be carried out by *Dr Lesly Huxley* (Bristol) in **Gathering Evidence: Current ICT Use and Future Needs for Arts and Humanities Researchers**. Other aspects covered are more technical. **LAIRAH: Log Analysis of Internet Resources in the Arts and Humanities** (*Dr. Claire Warwick*, UCL), will assess

use by the community of digital resources funded by the AHRC (and its partner organisations).

User Requirements Analysis for Portals in the Arts and Humanities (Professor Mark Greengrass, University of Sheffield) will consider what the Arts and Humanities need from portal technology. The outcomes of these two projects will be intimately linked with the future

measures can be taken to assess data resources to the same rigorous standards as those long in place for non-digital publications. Interdisciplinarity is, of course, a key issue.

Dr Mike Pringle (University College for the Creative Arts) looks a methods of cross-disciplinary mapping in his project **The Hunt for Submarines in classical art: mappings**

searchers.

Focusing on the strategically important areas of 3-D visualisation and data mining, **Making Space: A Methodology for Tracking and Documenting the Cognitive Process in 3-dimensional Visualisation-based Research** (Professor Richard Beacham, University of Warwick) and **ARMADILLO: Information Mining in Distributed Research Datasets in the Arts and Humanities** (Professor Mark Greengrass, University of Sheffield) will produce new trans-disciplinary resources for others to use.

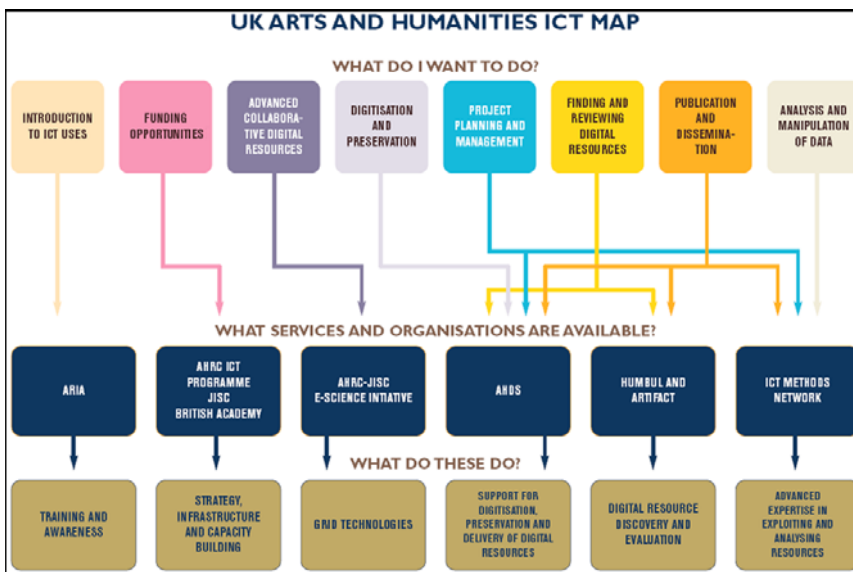
The topical and important issue of linking online archives with e-publications is being addressed by Professor Julian Richards (**Making the LEAP: Linking E-Archives and E-Publications**).

Lexical Searches for the Arts and Humanities (Professor Jeremy Smith, University of Glasgow) is investigating new ways of searching texts by drawing on The Historical Thesaurus of English, while Professor Robert Zimmer (Goldsmith's) is focusing on a select portion of the Tate's existing Online Event Archive to create an intelligent system for archiving and retrieving video and audio recordings of events about contemporary and modern art (**Creation of a flexibly searchable streaming media archive of contemporary and modern art theory and practice**).

Collectively, these projects will lead to a suite of tools which will guide the AHRC in new directions of capacity development for ICT.

For further information please contact the Programme: s.e.dunn@rdg.ac.uk

Stuart Dunn, AHRC, University of Reading



Conscious of the growing number of bodies designed to support humanities computing in the UK, the AHRC has designed its ICT map, giving a simplified representation of each organisation's functions and the way they relate to each other. For more information see <http://www.ahrict.rdg.ac.uk/ictmap/>

direction of the ICT Programme itself.

Similarly **Scoping e-science and e-social science developments and their value to the arts and humanities** (Sheila Anderson, King's College London) will underpin the ICT Programme's planning for the E-Science Initiative which it is overseeing in partnership with JISC (see <http://www.ahrict.rdg.ac.uk/e-science/>).

Elsewhere, familiar areas which have taken on new dimensions as a result of the proliferation of digital technology are addressed. Professor David Bates' **Peer review and evaluation of digital resources for the Arts and Humanities** (University of London) examines what

between scientific invention and artistic inspiration, while Dr. Alan Marsden (Lancaster) will survey technologies under development for searching and analysing audio and visual media (**ICT Tools for Searching, Annotation and Analysis of Audio-Visual Media**).

In sum, these projects will provide a framework of knowledge for and about arts and humanities computing which will be fundamental to the recommendations that the ICT Programme makes via the 2006 Review.

While this knowledge base is being developed, the Scheme's second strand, resource development, will increase the range of tools and capacity currently available to re-

Lower-cost digitisation?

OSS Watch

Randy Metcalfe **defines the meaning of open-source software, while in the panel, Elena Blanco suggests some cheap tools for digitisation projects**

Thinking open source is vital to one JISC-funded service. OSS Watch <http://www.oss-watch.ac.uk/> provides unbiased advice and guidance about free and open source software to UK higher and further education. Unbiased is important there.

This isn't about unthinking advocacy – open-source solutions may not be right in all situations. The important thing is to consider open source options at the beginning of a procurement process, or a development project, or, indeed a digitisation effort.

Two questions should be springing to mind. What is open-source software? Why should I be thinking about open source?

Open-source software is always released under an Open Source Initiative certified licence. There are more than 50 open source licences, all of which meet the 10 criteria of the Open Source Definition. Any software with an open-source licence will provide access to the source code – the code from which applications are compiled – and permit the modification and redistribution of that code.

What open-source software is may not be as significant as how it is. Open-source software development is usually associated with communities of (often voluntary) programmers. Open code invites scrutiny, experimentation and contribution.

Since open-source projects want their efforts to be re-used by others, they have a tendency to use open standards that support interoperability and open data formats that support longevity.

Why you should be thinking about open source will vary depending on who you talk to. Typically people think open source because they want to ensure best value for money.

standards and open data formats. Together – best value and open standards and data formats – these form a significant reason for thinking open source at the beginning of any software procurement, devel-

OpenOffice

<http://www.openoffice.org/>

OpenOffice is an open-source office suite that provides a raft of desktop applications including a word processor, a spreadsheet, a drawing program, a database, and a presentation manager. OpenOffice uses its own native file format that conforms to open standards but it includes excellent support for other file formats including Microsoft Office file formats. Its internationalisation support is excellent and a great deal of effort has been made to ensure the highest levels of compatibility with Microsoft Office.



GIMP

<http://www.gimp.org/>

The **GNU Image Manipulation Program (GIMP)** is a freely distributed program for image manipulation tasks such as photo retouching, image composition and image authoring. GIMP has many features including a full suite of painting tools, transformation tools and selection tools. It also has support for loading and saving in many different file formats and supports advanced scripting.



Other useful open-source programs include **PDF Creator** http://sector7g.wurzel6.de/pdfcreator/index_en.htm for developing files in the PDF format, **Nvu** for website design and creation <http://www.nvu.com/> and **FileZilla** <http://sourceforge.net/projects/filezilla/> a FTP (File Transfer Protocol) client for putting and extracting files from web servers.

However, when OSS Watch undertook its first national survey of universities, the number one reason why they said they were considering open source solutions was interoperability. That is consistent with the tendency to use open

development, or digitisation process.

For further information on free and open-source software visit <http://www.oss-watch.ac.uk/> or write to info@oss-watch.ac.uk

Randy Metcalfe,
University of Oxford

The Learning Index - putting pictures into context

Brenda Brinkley introduces new learning and teaching materials hosted by AHDS Visual Arts.

Everyone is familiar with the saying “a picture is worth a thousand words” and until recently AHDS Visual Arts has let its 50,000+ image catalogue speak volumes about itself.

Things are about to change, however, with the launch of the Learning Index, a new resource from AHDS Visual Arts that shows just how effectively words can be in bringing a picture to life.

Through a series of specially commissioned, richly illustrated learning and teaching materials, the Learning Index adds dimension to the AHDS Visual Arts image catalogue by interpreting the collections from an expert’s point of view.

Often, one of the primary barriers to the use of digital collections in higher education is the lack of time to engage with new resources on



A screenshot from the *Calligraphy: an education in letter form* module

offer. The Learning Index offers ready made packages of information that demonstrate the relevance of AHDS Visual Arts image collections to teaching needs therefore saving valuable time and effort.

Seven Learning Index resources, for instance, use over 800 high-quality images and text to enhance course curriculum and learning experiences in Art, Design and Communication.

Five of the Learning Index Modules

Thomas Becket’s stained glass windows at Canterbury Cathedral

This resource presents a sample of images from the photographic archive of the Corpus Vitrearum Medii Aevi (CVMA), an international research project dedicated to the publication of medieval stained glass.

Designing Britain 1945-1975: the visual experience of post-war society

A series of seven modules exploring the history of British post-war design using the historically important and extensive design archives at the Faculty of Arts and Architecture, University of Brighton

Britain in the age of the French Revolution

A tutorial in political imagery during the period 1789-1815. The primary content is a collection of political cartoons or satires relating to the French Revolutionary and Napoleonic wars. The cartoons are arranged thematically and chronologically.

Crafts Study Centre

Six learning and teaching guides reviewing the key subject areas of the Crafts Study Centre’s collection. Eminent craft historians and curators, as well as practitioners, have written these unique guides; for instance Emmanuel Cooper on Bernard Leach; Ewan Clayton on lettering; Barley Roscoe on textiles and Margot Coatts on ceramics and a general introduction.

Tim Mara Artist’s Alphabet

The Artist’s Alphabet consists of 26 individual sections, one for each letter of the alphabet, each one highlighting and representing a different area or topic relating to his practice.

Resource Features

- * Artist biographies
- * Full catalogue records
- * Multiple image views
- * Glossaries
- * Image galleries
- * Links to record information
- * Assignments, essays
- * Case studies, reading lists
- * Bibliographies

These materials use the extensive Design Archives at the Faculty of Arts and Architecture University of Brighton to explore the history of British post-war design, including product design, sculpture, fashion and the crafts.

Another suite of materials further an understanding and appreciation of the link between the historic collection of the Crafts Study Centre and new developments in the crafts.

Simon Olding, Director of the Crafts Study Centre, comments on the importance of these Learning Index resource: "Each resource examines in detail the works held in the Crafts Study Centre's collection in relation to the techniques, use of materials, history and biographies associated with their makers, thus providing substantive context to a valuable body of material now globally accessible to students, researchers and en-

thusiasts of 20th-century and contemporary crafts."

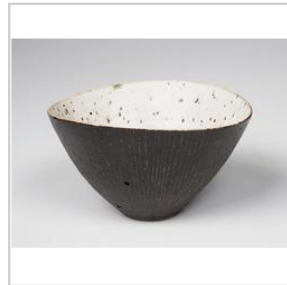
All Learning Index resources are quality assured having been written by subject specialist lecturers – often based on their own teaching experiences.

They employ a traditional walk through approach to learning us-

AHDS Visual Arts subject matter and work well as a supplement to coursework, as a means of exploring e-learning or simply as an introduction to AHDS Visual Arts collections.

All Learning Index resources are freely available and rights cleared for use in UK Education.

Introduction Liquid Colour Brush Decoration Textured Surfaces Bibliography Image Gallery



No. 13 Lucie Rie
Oval bowl, c.1955
P.74.23



[core record](#)
[larger image](#)

Ceramic Decoration: a selective guide to pots in the collection of the Crafts Study Centre

[Liquid Colour](#) > [Lucie Rie](#)

[back](#) | [start](#) | [next](#)

No 13: Oval bowl, c.1955 (P.74.23)

Bowls of this type were used either for serving food or for decorative purposes. The basic form is thrown in stoneware clay and then squeezed into a gentle oval ellipse at the rim, the long sides of the oval are raised slightly higher than the ends. The bowl's surfaces are treated in two strongly contrasting colours: the inner one coated in a smooth white tin glaze, heavily speckled with spots of manganese, the outer glazed in manganese pigment through which bold sgraffito lines radiate upwards from foot to rim.

A screenshot from the *Ceramic Decoration: a selective guide to pots in the collection of the Crafts Study Centre* module

ing a uniform design and an easy to use web interface. This uniformity of design allows the resources to be re-usable and more adaptable to a variety of learning environments – an essential element in today's resource intensive market.

Learning Index resources engage students and teachers with

The Learning Index is available at <http://visualarts.ahds.ac.uk/learning/>

Brenda Brinkley, University College for the Creative Arts, Farnham, AHDS Visual Arts

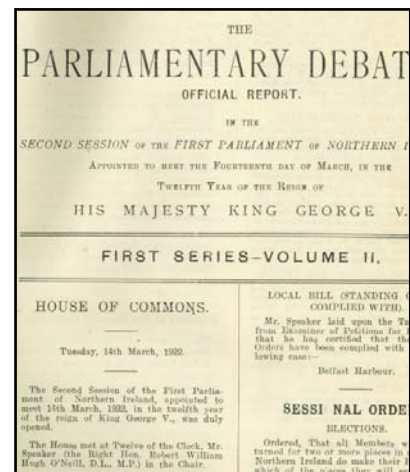
Stormont Papers Online

Along with the Centre for Data Digitisation and Analysis at Queen's University Belfast, the AHDS is creating an online version of the Northern Irish Parliamentary Papers.

The papers, consisting of 92,000 pages of transcripts of parliamentary discussion, cover the entire length of the first period of Parliament at Stormont, from Partition in 1921 to the advent of Direct Rule in 1972.

Digitised in Belfast, the digital files are being hosted by the AHDS, who are also building an interface to search, browse and display the papers. The papers are due to come online in 2006.

See <http://ahds.ac.uk/stormont-papers/> for more information.



Title page from the 1922 Parliamentary Proceedings

Developing Linguistic Corpora

The AHDS has published the latest in the series of its Guides to Good Practice, *Developing Linguistic Corpora*. Editor Martin Wynne, head of AHDS Literature, Languages and Linguistics, talks more about the book.

A linguistic corpus is a collection of texts which have been selected and compiled so that they can be studied on the computer. Corpus linguistics offers some of the most powerful new procedures for the analysis of language, and the impact of this dynamic and expanding sub-discipline is making itself felt in many areas of language study.

In this volume, a selection of leading experts offer advice to help the reader to ensure that their corpus is well-designed and fit for the intended purpose.

Increasing numbers of researchers are seeing the potential benefits of the use of an electronic corpus as a source of empirical language data for their research. However, it has not been easy to find out how to build a corpus.

“A corpus is a remarkable thing, not so much because it is a collection of language text, but because of the properties that it acquires if it is well-designed and carefully constructed.”

John Sinclair

There is a great deal of useful information available which covers principles of corpus design and development, but it is dispersed in handbooks, reports, monographs, journal articles and sometimes only in the heads of experienced practitioners.

This Guide draws together the experience of corpus builders into a single source, as a starting point for obtaining advice and guidance on good practice in this field. It aims to bring together some key elements of the experience learned, over

many decades, by leading practitioners in the field and to make it available to those developing corpora today.

This Guide is aimed at those who are at some stage of building a linguistic corpus. Little or no knowledge of corpus linguistics or computational procedures is assumed, although it is hoped that more advanced users will also find the guidelines here useful.

provided for a corpus, and gives examples of how these can be implemented using the Text Encoding Initiative guidelines.

Tony McEnery and Richard Xiao take on the tricky issue of encoding characters in languages other than English, giving an historical overview of the various solutions, leading to a discussion of how to use Unicode today in encoding corpus texts.

Developing Linguistic Corpora - Contents

Corpus and text: basic principles, John Sinclair

Adding linguistic annotation, Geoffrey Leech

Metadata for corpus work, Lou Burnard

Character encoding in corpus construction, Tony McEnery, Richard Xiao

Spoken language corpora, Paul Thompson

Archiving, distribution and preservation, Martin Wynne

It also has relevance for those who are not building a corpus, but who need to know something about the issues involved in the design of corpora in order to choose between available resources and to help draw conclusions from their analysis.

The authors set down principles and definitions in a clearer and more accessible form than they have done in previous publications. It is therefore likely that this Guide will become a key reference point for discussions of methodology in corpus linguistics, as well as a practical guide for corpus builders.

John Sinclair sets out ten principles for corpus design, plus a new definition of a corpus. Geoffrey Leech offers a taxonomy of types of annotations as well as clear guidelines and some provisional standards for annotation at various linguistic levels.

Lou Burnard explains the different types of metadata which can be

Paul Thompson draws on his experience in developing the British Academic Spoken English (BASE) corpus to set out the stages involved in the development and exploitation of a corpus of speech, covering data collection, transcription, markup and annotation, and access. In chapter six, I explain how good planning and design can help to ensure the ongoing availability and usefulness of a corpus.

This and other guides in the series are available from <http://ahds.ac.uk/creating/guides/> This also includes details on buying hard copy versions.

AHDS Literature, Languages and Linguistics is also host to many freely available corpora in several languages, including English, French, German, Italian and a variety of South Asian languages.

Martin Wynne, AHDS Literature,
Languages and Linguistics,
University of Oxford

In this edition: explore resources on

In this, edition of AHDS Update, we explore a theme that is the subject of an AHRC funding programme. Over the page you will find the latest available AHDS collections and some dates for your diary. Go to <http://ahds.ac.uk/newsletter> for direct links to all of the featured URLs.

Diasporas, Migration and Identities

Scottish Emigration Database (study numbers 5161 and 5193)

Category: History

The database seeks to identify patterns of movement from Scottish ports between 1890 and 1960 using information transferred digitally from the passenger manifests of the Board of Trade's Statistical Department (BT27) held at The National Archives in Kew, London. It contains the records of 19,113 passengers and adds a unique quantitative dimension to the scholarship of the Research Institute of Irish and Scottish Studies.

<http://www.data-archive.ac.uk/finding-Data/snDescription.asp?sn=5161>

Predicting the Location of Hominin Sites in Africa and Asia

Category: Archaeology

A recently released resource from the ADS / AHDS Archaeology features a database of the predicted locations of palaeolithic archaeological sites.

The data is the result of a project that formulated a method of predicting where well-preserved, early human archaeological and fossil sites might be located. Such sites are known in Africa from about 6 million years ago, and in Asia from about 1.8 million years ago.

A large number of variables were involved in the predictions including geomorphy, climate and vegetation. In assessing how well various sites would be preserved collagen loss was a variable thought to have particular significance.

The resource itself features, not only a searchable database, but also an interactive map and downloads.

http://ads.ahds.ac.uk/catalogue/resources.html?hominids_ahr_2003

SALIDAA: South Asian Diaspora Literature and Arts Archive

Category: Visual Arts

The SALIDAA digital archive aims to showcase the richness and diversity of contemporary South Asian literature and arts in England by digitising a variety of text-based and visual material accompanied by descriptive and contextual information.

The digital archive covers five main areas: literature, visual arts, theatre, dance and music and the material dates from approximately 1947 to the present. A wide range of materials are being digitised, such as excerpts from fiction, poetry and plays, images of art works, photographs, brochures and programmes of events, CD and record covers, and music scores.

<http://www.salidaa.org.uk/salidaa/site/Home>



Ghosh, Amal 2001, *Oracle*, Copyright Amal Ghosh
South Asian Diaspora Literature and Arts Archive
<<http://visualarts.ahds.ac.uk>>

North African Film and Filmmakers Dictionary

Category: Performing Arts

This database looks at films produced since independence in North Africa (the Maghreb) - in Algeria, Mauritania, Morocco and Tunisia, as well as those by film makers working in Europe who were born in the Maghreb or are of North African descent. Ten years ago these would have been separate categories, but now the boundaries are much less precisely defined, especially because of the diaspora of Algerian film makers caused by the political problems there.

There are five major entry points for an overview of this production activity: Algeria, Mauritania, Morocco, Tunisia and what is termed 'Immigrant Cinema'. Each of these entries is intended to list all film makers who have created a feature film and all the resultant films, with the ability to cross-reference to director names and film titles.

<http://ahds.ac.uk/performingarts/collections/north-african-film.htm>

Continued overleaf...

The Acquisition of Welsh Syntax

Category: Linguistics

One of the main aims of this project was to create a substantial corpus of speech from children in the early stages of the acquisition of Welsh as a first language.

The dataset contains transcriptions of audio recordings of spontaneous exchanges between adults and very young children. Regular recordings of 30 to 45 minutes were made of six children over a period of approximately nine months. The audio recordings were transcribed using the transcription system CHAT, of the CHILDES project (Child Language Data Exchange System).

Each transcription consists of headers which supply general details, the utterances of the participants, and comments. To gain access to this resource you will need to complete an online form.

<http://ahds.ac.uk/catalogue/collection.htm?uri=III-2427-1>

Rivenhall: investigations of a villa, church and village, 1950-1977

Category: Archaeology

This report traces the development of settlement in the rural parish of Rivenhall, Essex, from the prehistoric era to the present day, and attempts to relate this to the general settlement history of the surrounding area. The methodology involved in the recording and interpretation of the evidence is explained. The approach has involved the seeking of reasons why the various settlement components took the form they did, how buildings were planned and erected, and how they functioned in their contemporary contexts.

The report shows that once the locus of the principal settlement had been established at Rivenhall domestic occupation was maintained down to the present day, with no evidence of discontinuity.

<http://ads.ahds.ac.uk/catalogue/library/cba/rr55.cfm>

October 2005

15th - Closing date for applications to Autumn 2005 British Academy Larger Research Grants Scheme.

November 2005

11th-12th - Theory and Practice: CHArt Twenty-first Annual Conference

12th - Association for History and Computing: *Methods and Madness Conference*

24th - Closing date for applications to Autumn 2005 AHRC Resource Enhancement Scheme and Research Grants Scheme.

December 2005

1st - Closing date for Leverhulme Trust Research Project Grant applications

New AHDS History Releases

Growth of British Industrial Estates, 1900-1939 (study number 5191)

This dataset provides estimates of the number of industrial estates established in Britain in the early 20th Century. It examines their locations, the companies or other institutions that developed them, and the number of employees working in them at several 'benchmark' dates.

The main aims and objectives of the research project from which this dataset arose were:

- (1) To assess the contribution of industrial estates to the growth and location of new manufacturing enterprises in interwar Britain;
- (2) To examine the ways in which location of interwar industrial estates boosted firm growth;
- (3) To explore the contribution of industrial estates to fostering locational externalities for the firms which located on them;
- (4) To examine the regional impact of industrial estate development.

<http://www.data-archive.ac.uk/findingData/snDescription.asp?sn=5191>

Religion, Dynasty and Patronage in Rome, c.440-840 (study number 4630)

The project aimed to offer a new approach to the social history of Christianity in Europe in the formative period from the disintegration of the Western Roman Empire to the rise of the Carolingians. By compiling a relational database of patronage (both ecclesiastical and secular) in Rome from the mid-fifth to the mid-ninth centuries, it sought to make visible the social basis of the much-discussed process of 'Christianization.'

The database covers patronage transactions in the city of Rome (including the suburban basilicas) from the mid-fifth to the mid-ninth centuries taken from a range of different sources such as narrative sources, inscriptions, charters and letters.

<http://www.data-archive.ac.uk/findingData/snDescription.asp?sn=4630>

JISC Repositories Review

The Joint Information Systems Committee (JISC) Digital Repositories Programme initiated a programme of work to assist deployment of digital repositories within the learning and research communities. This review was intended to provide useful background information for participants in this call. The review was not intended to be comprehensive; the intention was to identify useful areas of activity for the programme rather than to prescribe activity in detail.

The final report is the result of a selective review of current digital repository activity. Information was gathered by interviewing stakeholders, holding a focus group, sending out an email questionnaire and undertaking a gap analysis.

The report was co-written by Sheila Anderson, Director of the AHDS, and Rachel Heery, of UKOLN.

The review is available to access on the AHDS website <http://ahds.ac.uk/preservation/preservation-reports.htm> and further information about the JISC Funded Digital Repositories Programme is available from the JISC website http://www.jisc.ac.uk/index.cfm?name=programme_digital_repositories.

The AHDS is funded by the
AHRC and the **JISC**

Please visit <http://ahds.ac.uk/newsletter> for all URLs