THE THEATRE OF MEMORY:

Welsh print online

THE INSPIRATION

The Theatre of Memory: Welsh print online will make the printed record of Wales and the Welsh freely available, to be searched and read by anyone on the internet. Its aim is to lay bare the memory of our country so that it is seen, as it were at a glance, by anyone in the giant auditorium that is the World Wide Web.
What will it achieve?

What does it aim to do?

The aim of the programme is to digitise as much as possible of the printed material published in Wales and about Wales and the Welsh people.

This information will be made freely available and easily findable to anyone with an internet connection – giving Wales and the Welsh an enhanced online identity on a worldwide stage.
Why is it so important that this Welsh material is digitised?

Mass digitisation of print materials is now commonplace. But the dominance of Anglo-American texts on the Web has been, and continues to be, a concern not only for smaller nations but also for larger European cultures. It is expected that some Welsh interest and Welsh language texts will be digitised as part of larger collections involved in projects such as Google Books, but it is almost inevitable that these will be incidental. The National Library of Wales believes that it is its responsibility, along with a variety of partners, to ensure that Wales and Welsh printed culture is comprehensively represented on the Web.
What will be included?

The programme aims to encompass as much as possible of the printed record of Wales, and will include material:

- On any aspect of Wales and the Welsh: on any subject, in any genre, about Wales as a nation and as a geographical area, and about Welsh people in Wales and beyond;

- In all printed formats: books and pamphlets, official publications, music, ephemera, periodicals, magazines, and newspapers;

- From every period, from the earliest days of printing to the present;

- In English, in Welsh, and in any other language through which Wales and the Welsh are discussed;

- At every level: from children’s literature to advanced research, from the highly local to the international;

- Both in and out of copyright.

Although the aim will be to be as comprehensive as possible, it may be necessary to exclude certain categories of material, for instance, items that in their original form are rare, fragile or otherwise unsuitable for digitisation, or if satisfactory agreements cannot be reached with copyright holders.
How much material will be digitised?

It is estimated that some three hundred thousand books, hundreds of periodicals, journals and newspaper titles and ballads from the National Library’s collections will be included. These could amount to around forty million pages of print.

However, to achieve the ultimate aim of digitising the whole printed record of Wales and the Welsh, it will be necessary to look beyond the Library’s collections and forge partnerships with repositories in other parts of Wales and beyond.
What are the benefits?

The Theatre of Memory: Welsh print online will provide access to a colossal repository of knowledge about Wales and its peoples in a way never achieved before. It will help bring Wales to the attention of the world, serve as a further introduction to the curious, wherever they may be, and provide the Welsh and those passionate about Wales with a depth of information and knowledge that can enrich their self awareness and sense of identity.

All the material included in the programme is likely to be used much more frequently and much more intensively in an online digital form than it would be if it remained trapped in its covers in a physical library.

Books long out of print and available only in few research libraries – which may be physically, culturally or economically remote – will be easily searchable and available in all libraries and in every home with access to the internet. With the co-operation of publishers more recent publications will also be available at the click of a mouse.

The Theatre of Memory: Welsh print online will provide the world with the largest ever corpus of digital heritage relating to Wales that will be available for use, reuse and exploitation, as a basis for personal interactions and the development of new technologies.

Specific information, but will also satisfy those wishing to browse more arbitrarily.

Almost the entire body of published writing in the Welsh language will be within the reach of everyone. This will end a period when Welsh has been all but invisible in the online world, and will help consolidate the language in a host of different areas of discourse. The Theatre of Memory: Welsh print online will provide the world with the largest ever corpus of digital heritage relating to Wales that will be available for use, reuse and exploitation, as a basis for personal interactions and the development of new technologies.
Who will benefit?

Potentially, millions of people, young and old – anyone with any interest in finding out about Wales and the Welsh people – will gain. Because the scope of the project is so wide, in subject, language and period, there will be something for everyone, whatever their background and interest:

- A school student pursuing a project on the Rebecca Riots will be able to locate contemporary newspaper accounts, works by later historians, and novels;
- A family historian can trace references to her ancestors in the printed literature with confidence, rather than following hunches and false trails;
- A journalist will be able to trace a story about a sporting event hidden, and previously almost impossible to locate, in a local newspaper;
- A social historian can trace the development of a village or urban community through time by using place names for keyword searching;
- A government researcher may track reports and statistics on economic development in the Valleys over the last thirty years;
- A Web user will be able to work directly with the digital information to create new and exciting ways of using the data through ‘mash ups’ with other data sources;
- A linguistic scholar or lexicographer can trace the history of a Welsh word across the centuries as it appears and changes in the printed record;
- A writer or critic will easily be able to compile an anthology of Welsh poetry, or follow a theme such as motherhood or pacifism in Welsh fiction, or research the background for an historical novel;
- An artist or writer will draw inspiration from and reuse the programme’s content to create new works of art;
- Digital businesses and technical innovators could use the data set to develop new technologies.
Technical issues

What does digitisation entail?

There are five basic steps that will be taken to develop The Theatre of Memory: Welsh print online:

1. Printed volumes will be selected, category-by-category. Most will be sourced from within the National Library of Wales’s collections.

2. Each volume will be scanned, and a digital image of each page will be produced. The text on the page will then be read electronically using optical character recognition (OCR) technology to convert it into searchable text.

3. Limited descriptive data (‘metadata’) will be attached to the digitised texts, to aid searching and retrieval.

4. The whole corpus of digitised texts will be made available for users to explore online, using a search engine that is simple to use but sophisticated in the way it retrieves and presents search results.

5. The mass of data will be stored and preserved in a way that will ensure its accessibility for future generations.

The Library has already developed software that allows most of these steps to be automated, thus ensuring that this large-scale digitisation process is both efficient and reliable.
What about copyright?

Copyright has proven to be an important issue for mass digitisation projects worldwide. *The Theatre of Memory: Welsh print online* will take into account developments elsewhere and their effect on practices closer to home, as well as conducting research on the effect of copyright on this programme’s aims.

Under the current UK law, only books and other printed material published before about 1868 are out of copyright and can be copied without permission. Their full texts will be available for viewing, downloading and printing online.

In the UK, printed works within copyright – generally those published within 70 years of the author’s death – cannot be digitised without the permission of the copyright owner. Permission to digitise these items will be sought from publishers and other copyright owners.

To what extent it will be possible to view these copyright-protected texts will depend on the nature of an agreement or licence arrived at with the copyright owner. Where possible the project will aim for a licence that allows wide online dissemination of texts without harming the essential interests of the owner. Agreements that allow limited access or time embargoes will also be considered.
How will *The Theatre of Memory: Welsh print online* be realised?

It is inevitable that the *Theatre of Memory: Welsh print online* will be realised in sections, with each main category of material being digitised in turn: books and other single items, newspapers, periodicals etc. It is also inevitable that funding for this undertaking will also come in tranches; various types of material and content will appeal to funders at different times.

Hitherto we have undertaken four major projects involving different material types:

**Welsh Biography Online - [http://wbo.llgc.org.uk](http://wbo.llgc.org.uk)**

Welsh Biography Online includes approximately 5,000 biographies of eminent Welshmen and women who died before 1 January 1971, originally published as the Dictionary of Welsh Biography. It is an invaluable resource to all who are interested in Welsh history and culture. It is the result of a partnership with Biography’s publisher, the Honourable Society of Cymmrodorion, and an investment by the National Library of Wales in text conversion and pioneering XML encoding. The full text of the five printed volumes, together with new articles, has been combined here in one live bilingual electronic database.

**Development Cost: £150K**

**Welsh Journals Online - [http://welshjournals.llgc.org.uk](http://welshjournals.llgc.org.uk)**

Welsh Journals Online offers access to a selection of 50 modern periodicals and 300 historical titles in both Welsh and English including learned publications, literary journals, science and social science titles, the transactions of county historical societies and denominational publications, and popular publications discussing current affairs in Wales and beyond. This amounts to an estimated 1.2 million pages of rich cultural heritage. The historical titles will be available online from 2011.

**Development Cost: £750K** (£600K JISC funding, £150K NLW funding) for the 50 modern titles. The historical titles were funded as part of the Welsh Newspapers Online project.

**Welsh Ballads Online - [http://www.llgc.org.uk/baledi](http://www.llgc.org.uk/baledi)**

Welsh Ballads Online comprises approximately 4,000 digitised ballads, mainly dating from the eighteenth and nineteenth centuries, from the collections of the National Library of Wales and Cardiff University Library. The ballads were digitised as a joint project between these two institutions, with, with the support of the Universities of Bangor, Swansea and Lampeter. The project was funded by the JISC as part of its Enriching Digital Resources programme.

**Development Cost: £75K** (JISC funding)

**Welsh Newspapers Online**

The National Library of Wales’s entire holdings of pre-1910 Welsh newspapers: a magnificent resource of everyday knowledge estimated to contain over 1 million pages and 200 newspaper titles covering every corner of Wales. This project is partly funded by the Strategic Capital Investment Fund and will be available online from 2012.

**Development Cost: £3M** (£2M SCIF funding, £1M other funding)
These projects highlight a number of important principles and lessons that have been learnt:

**Prioritisation**
Content has been prioritised according to what users need and what funders are willing to finance. The copyright status of the material can also influence its prioritisation.

**Partnership working**
Welsh Biography Online was funded and completed exclusively by the Library, but it was a relatively small project. Welsh Journals Online utilised the Library’s holdings, but was undertaken in partnership with a number of higher education libraries that were represented on the project board. Welsh Ballads Online was a collaborative project that provides access to material held by the Library and that of Cardiff University, and was funded by the JISC. Similarly, material from other libraries in Wales is being used to enhance the content of Welsh Newspapers Online. The success of *The Theatre of Memory: Welsh print online* will be dependent upon similar working partnerships.

**Copyright**
The Library recognises that copyright issues are integral to the successful delivery of this ambitious programme. Some of the existing projects have deliberately focused on out-of-copyright material, whilst others have necessitated securing consent to copy from authors and publishers. The Library recognises that its users require access to recent material, and *The Theatre of Memory: Welsh print online* will investigate ways in which this expectation can be realised.

**Technical approaches**
The National Library aims to complete as much digitisation work as possible in-house in order to gain, retain and enhance experience and expertise. The technical aspects of these existing projects have varied greatly; the implementation of Welsh Biography Online was dependent on human intervention, whilst this was kept to a minimum during the Welsh Newspapers Online project. All OCR work was originally outsourced, but some is now undertaken in-house.

**Future developments**
The pioneering University of Wales Chair in Digital Collections has been established in the Library and is expected to undertake, and academic research on the latest developments in digital resources, as well as applying the findings to large digital collections, including *The Theatre of Memory: Welsh print online*. 

The JISC’s Enriching Digital Resources programme, and a grant awarded from the Welsh Assembly Government’s Strategic Capital Investment Fund has been used to part-finance the Welsh Newspapers Online project. It will not be possible to fully realise *The Theatre of Memory: Welsh print online* programme without external funding.
Who will develop *The Theatre of Memory: Welsh print online*?

The programme will be owned and developed by the National Library of Wales.

However, the Library will never realise such an ambitious vision by itself. It is inevitable that other institutions will have an interest in this programme:

- The Welsh Assembly Government - as a proxy for a wide range of beneficiaries, as a stakeholder, and as a possible funder
- The Welsh Books Council - as the central public body responsible for the publishing industry and with a possible interest in digital preservation and copyright issues
- Higher education bodies - as potential users, as owners of much of the contributing content and as partners in funding bids
- Representatives of the Welsh publishing industry and authors - representing copyright holders but also with a possible interest in digitisation and digital presentation of texts
- Commercial partners - as service providers
- Large-scale donors and sponsors

What will *The Theatre of Memory: Welsh print online* cost?

This programme is very ambitious and will require significant investment. The project costs outlined above illustrate the cost of the existing projects. It is inevitable that the programme costs will staff costs (project management, rights clearance, digital capture, metadata, and technical issues), equipment costs (digitisation, data storage, access) and the ongoing cost of digital preservation. However, it is impossible to estimate the cost of the whole programme, not only because its technological context is in constant flux, but also because each project element will be unique in its scope, content and method of delivery.

How will the *Theatre of Memory: Welsh print online* be funded?

The finance to achieve this vision *Theatre of Memory: Welsh print online* is likely to come from a variety of sources, some public and others private. The National Library of Wales always welcomes new partners who share its vision to provide open, free access to the printed record of Wales.