



NLA uses Flax to build index of 5 years of UK newspaper articles

The Newspaper Licensing Agency (NLA) is the latest organisation to install the Flax enterprise search system as part of a major online news development known as ClipShare.

ClipShare will allow journalists online access to a searchable database of national and regional newspaper content in text and PDF format from a central resource, the first time direct access to articles from multiple titles in image form has been available. Five years of content from all the UK's major newspapers will be available.

Andrew Hughes, Managing Director, NLA Digital says 'NLA is pleased to be improving the ability of journalists in national and regional titles to research and access articles in text and image form with ClipShare. The new service builds on and expands the national

newspaper data exchange service, adding article images, a wider content set, earlier delivery, and web based access technology for the first time. NLA is owned by publishers and our goal is to help the titles we represent by improving access to content and cutting the cost of development. ClipShare is a significant step towards this'.

The Flax technology, developed by Lemur, builds on established open-source software. 'Flax is much more than just a search engine' explained Charlie Hull, Managing Director of Lemur. 'We've taken a world-beating core – used by organisations from the German newspaper Die Zeit to Cambridge City Council – and added software allowing it to build indexes of hundreds of different kinds of files, emails and databases.

We've also added several different types of front end so that Flax will fit seamlessly into a company's existing infrastructure.'



newspaper licensing agency
Respect for Copyright

'The days of closed-source, low-performing search engines are gone,' says Hull. 'If your system doesn't cope reliably with millions of documents, for a sensible price, then it simply isn't suitable for today's business environment. If the system is a closed-source 'black box' you have no way of verifying that it does what the vendor says, and maintenance and improvement is more difficult. Flax's performance and reliability has been well demonstrated and is improving all the time.'

Xapian 1.0 release

Lemur is pleased to be a long-standing contributor to the open source Xapian project which released version 1.0 of its search engine library last month. This marks major improvements in the features and ease-of-use of Xapian, and Lemur will continue to offer full support and development services to users of this leading open source information retrieval system.

The 1.0 release includes the capability to handle Unicode text, a better indexing strategy, a faster and more compact database format, better documentation and a cleaner API, along with many other improvements, representing the culmination of eight years of collaborative development effort.

Usage of Xapian continues to increase, with the traffic on the mailing lists growing steadily and a wider community of users beginning to form. This ensures the project's ongoing development with

many new features planned for the future and work has already begun on the next release!

www.xapian.org

About Lemur

Lemur is highly active in the information retrieval market with international clients from sectors including academia, public relations, e-commerce, government and private businesses.

Lemur's Flax™ project delivers a cutting-edge enterprise search solution, using the power of open source software to drive down costs and provide world beating search performance with no software licence fees.

Lemur's Bamboo™ software is a collaborative, web-enabled, multimedia database system used by academics and researchers.

Lemur also provides general expertise and development in technologies such as Java, C++, Linux, Windows and XML.