

# Establishing a digital library of Dutch literature

Erasmus and Anne Frank in bits and bytes? In an ambitious attempt to preserve almost a millennium of literature and making it available to the greater public, the Digital Library for Dutch Literature and its Philippines-based digitization partner SPi Technologies are collaborating to build one of the largest online repositories of Dutch literature.



Photo: xlucas

By Sameer Raina

The Digital Library for Dutch Literature (Digitale Bibliotheek voor de Nederlandse Letteren, DBNL) was founded by the Society of Dutch Literature in 1999 with the aim of digitizing the main works from Dutch literary history, together with all kinds of secondary information like biographies, articles, portraits, audio files and illustrations.

## Executing on the vision

Since 2000, armed with funding from the Netherlands Organization for Scientific Research and later on from the Dutch Language Union, DBNL embarked on an ambitious journey aimed at preserving and making available to the greater public a vast body of Dutch language and literature.

The goal, in terms of scope, encompassed hosting on the internet a large spectrum of books and periodicals covering the field of Dutch humanities over a period of nearly 1,000 years – from the 11th century to the 21st century. Mr. René van Stipriaan, editor-in-chief of DBNL comments: “We started digitizing and posting the books online in late 2000 and expect to host approximately one-third of the whole universe of Dutch literature (30,000 books or 10 million pages) on our website [www.dbnl.org](http://www.dbnl.org) by 2012. This online database, when completed, would

be the largest collection of primary and secondary information of Dutch literary works and will provide a rich and profound insight in all the developing stages of Dutch history, society and culture. The ultimate goal, subject to funding, is to build, together with several partners including libraries and research institutions, a national network of databases that hosts the entire collection of Dutch literature ever to be published.”

Already, with over 4,000 books and periodicals available online, including works by well known Dutch authors such as Erasmus, Vincent van Gogh and Anne Frank, the DBNL website has become one of the most visited and most popular sites of its kind in the world. On an average day, the site attracts between 15,000-20,000 visitors who download between 100,000-140,000 pages. The loyal following includes people from all walks of life - researchers, school and college students, linguistic experts, and the general public from Dutch speaking regions at large.

### Scanning antique literature with modern search facilities

One of the reasons that has made the website popular is that DBNL offers all kinds of search facilities, like advanced search and full-text PDF searching, besides a navigational structure that gives an insight into the wealth of Dutch literary culture. Commenting on the underlying reasons for the popularity of the website, René van Stipriaan said: “The website is based on XML and the 1.5 million pages which it hosts are well structured and are almost flawless in their transcription. It is also this excellent quality of our text content that makes our website tremendously popular amongst our audience.” Currently DBNL prepares for a big book-scanning project of about 25,000 volumes of periodicals and books. Ultimately, the collection of PDF-files will enable DBNL to accelerate the production of reliable and structured XML, which would provide even more flexibility in repurposing content for multi-channel distribution.

### Literature's comeback

The DBNL user statistics prove the usefulness of a broad scope digitization initiative. Not only do canonical texts or well known authors find

their expected audience, even the forgotten or less important texts of the past centuries start a new career: for example - some really marginal text like collections of proverbs have found a huge new audience through DBNL.

Further, with an easy to use and mostly automated service to the general public, DBNL always tries to find new ways of making public interesting content in attractive presentation formats and compilations. In 2008, DBNL prepared a compilation called ‘The Basic Library, 1000 Key Texts from Dutch Cultural History’, and in 2009 DBNL will release a compilation called ‘1000 novels, the Golden Era of Dutch prose 1780-1980’.

The main advantage of this goal-oriented digitization is that it combines a sense of urgency with attractiveness for a big audience. The new book scanning prepares for another stage in the functioning of DBNL. This digital showcase will ever more be a good alternative for the classical library, not only for easy orientation but also for more elaborate reading.

### Partnership with SPi Technologies

With a budget reliant on donor support and an aggressive agenda, DBNL realized that in order to execute on its vision, it would have to look to Asia-based digitization service providers to meet its conversion needs. After talking to and testing multiple vendors, DBNL decided to select SPi as one of its main partners in the digitization process. “On evaluating various vendors, we felt very comfortable with SPi’s experience and track record in large scale PDF and XML conversion and digitization projects, especially as it related to converting and processing sensitive historical material”, said René van Stipriaan.

Mr. Bart Loevens, European Sales Manager at SPi comments: “Our goal is to be a service provider who understands the strategic objectives of DBNL and can provide the highest quality and accuracy at an attractive price. At the same time, our Philippines-based production team was very excited and enthusiastic about participating in this landmark initiative.”

To ‘test the waters’, DBNL and SPi started collaborating on low volume conversion and digitization projects to improve their respective workflows and capabilities. The first project was undertaken by SPi in 2003 and involved the conversion of 100,000 pages of 19th and

20th century Dutch language books into XML format. From 2003 to 2007 DBNL and SPi partnered on projects with increasing scope and complexity. In 2008, comfortable with the experience of the past few years, DBNL and SPi agreed to take their partnership to the next level by agreeing to digitize, scan and convert into searchable PDF over 25,000 hardcopy books with approximately ten million pages. While the relationship between SPi and DBNL has grown over the years, it has not been without its share of challenges and learning experiences. Some of the key learnings involved putting in place measures to improve data quality, smoothen volume fluctuations, and using best-in-class technology and tools to reduce turn-around time.

Reflecting on the engagement, René van Stipriaan says: “In SPi we found a partner that gives us all the essential feedback that is necessary to speed up and to prepare for the challenges of the next decade. For us, as a data provider, it is very important to be reliable, transparent, and also very user friendly. The only way to achieve that goal is to develop and maintain high quality standards and use the workflow systems that SPi provides. Communication and exchange of innovative ideas is crucial to the whole process and SPi has been a true partner in this regard.”

### contact



**Sameer Raina** is the Director of Global Marketing for SPi. He manages the strategic marketing and sales support function for SPi’s Publishing & Media

vertical and focuses on growth initiatives across the United States and Europe.

s.raina@spi-bpo.com  
www.spi-bpo.com