

The Safari Platform

Smart XML web publishing for books, documents and articles

Bernard Struyf

Bureau van Dijk Electronic Publishing
Avenue Louise 250 – 1050 Brussels
bernard.struyf@bvdep.com

Abstract

When two leading publishers of technical information wanted to transform concept into reality, they tapped into the expertise of Bureau van Dijk Electronic Publishing (<http://eps.bvdep.com>) for development, support and hosting. The Concept? Using the Web to provide a virtual library of technical books. The reality? The highly successful Safari Books Online, LLC, a joint venture of O'Reilly & Associates and Pearson Technology Group, has become the premier electronic reference library for Information Technology (IT) professionals.

1 The Products

Safari Books Online (<http://safari.oreilly.com>) is a service that offers IT professionals instant access to the full text of thousands of leading technology books right at their desktops. When you subscribe to the Safari service you can:

- Access the latest titles from the world's foremost IT publishers.
- Get and immediate answers to specific questions thanks to BlitzHit, the contextual XML search engine
- Copy, paste, bookmark, annotate and download secure chapters.
- Personalize your library.

Similarly, *SafariX Online* (<http://www.safarix.com>) is service for college students offering subscriptions to WebBooks using the Safari platform. Based on the same platform, the *Springer/Kluwer Academic Publishers Online Reference works* (encyclopedia, handbooks, ...) are available for academic users on <http://reference.springerlink.com>.

2 The Safari Platform

- **Publish and Subscribe.** Upload XML contents and let thousands of users subscribe online.
- **Browse and Search.** Offer virtually unlimited browse and search possibilities.
- **Customize and Develop.** Use the XML API to build new portals with XSL and exploit web services.

2.1 XML Repository

Smart native XML storage ensures optimal flexibility and performance. Index generation and dynamic rendering can be highly customized.

2.2 Advanced XML Search

BlitzHit, Safari's search engine enables powerful and fast retrieval featuring full-text and fielded metadata. An advanced query language offers multiple possibilities.

2.3 Online Browsing and Reading

Instant access to any document using navigation controls, multi-criteria lists and hierarchical structures. Safari's dynamic XSL transformations offer tailor-made rendering.

2.4 Secure Offline Reading

Predefined contents can be dynamically converted to secure, encrypted PDF with watermarking.

2.5 Personalization

Safari offers personal bookshelves, bookmarks and private / group / public annotations. It will keep track of the recent browse and search activity.

2.6 Back Office & Reporting

Extensive digital and access rights management for portals, group accounts and individual users. Detailed financial, marketing, usage and product reports in XML, HTML or Excel.

2.7 E-Payment

Safari offers secure online or batch processing of subscription payments, with the generation of monthly billing and royalty reports.

2.8 Product Lines

Documents can be organized in collections, which in turn can be combined and presented through portals and accounts.

2.9 XML API and XSL

Safari's XML API enables highly customized portal creation with XSL. Empower your applications using Safari's Web services.

2.10 Performance & Scalability

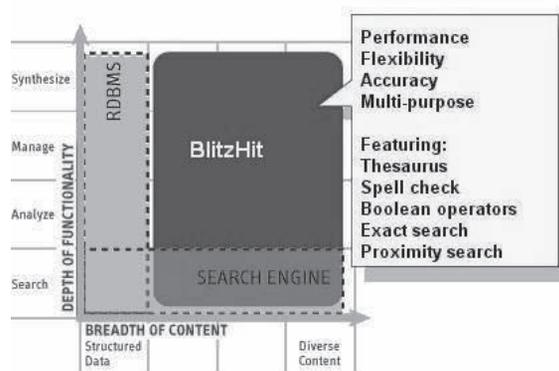
Safari will robustly serve high volumes of data to countless users. Load-balancing, recovery, clustering are part of Safari's SLA.

3 BlitzHit

BvDEP has developed a highly effective, scalable, relevant XML search engine, called BlitzHit. BlitzHit indexing and ranking agents use a large set of methods to mark relevancy for every word in every context, using word frequency, word location, contextual data derived from the XML tagging, and a set of other variables that have been determined through fine tuning to contribute to overall relevancy.

BlitzHit features customised contextual searches, including full text and metadata search, through sophisticated search syntax with Boolean, context and proximity operators. Since BlitzHit not a classical RDBMS, its performance and accuracy outrange most commercial search engines.

BlitzHit dynamically builds structured views on the result set, based on categorisation of the entries. This concept allows interactive drilling down into any data set without preliminary knowledge of the data structure. It also features spell checkers and the ability to integrate a thesaurus for synonym search or inflected forms.



4 The Company

Figure 1 BlitzHit features

Bureau van Dijk Electronic Publishing (<http://eps.bvdep.com>) is the Technology Partner behind the Safari Books Online platform and many other online products such as Yellow Pages, Directories, Magazines and Newspapers.

With its international presence (20 companies) and its strong team of experts committed to long-term partnerships, BvDEP helps publishers to remain focused on their core businesses by providing flexible state-of-the-art platforms to deliver content electronically.