



## WEBSEARCH – Suntek Chinese/English Search Engine

Search is the core of our and your business.

Main Features	
✓ Search English, BIG5, GB, and Hong Kong Supplementary Character Set (HKSCS/GCCS)	✓ Customized index phrases and synonyms
✓ Automatic code detection	✓ Search HTML, ASCII, RTF, WORD, EXCEL, POWERPOINT, PDF and XML files
✓ Powerful spider administration system	✓ Search all standard SQL databases
✓ Search remote web sites and local files	✓ Scalable to multiple CPUs and servers
Query Functions	
✓ Support both Chinese and English search keywords in the same query	✓ Order output by document attributes or similarity to query
✓ Support Boolean operators: AND, OR, NOT	✓ Search full text, metatags and document fields
✓ Support phrase and synonym search	
Applications	
✓ Document management on intranet	✓ Internet marketing and promotion
✓ Online product catalogs	✓ Online newspapers and magazines
✓ Home shopping and electronic commerce	✓ Data mining for your business
Add-On Features	
• Access Control • Same-Sound Search • Advertisement Server • GB/BIG5 Conversion	
Platforms	
Windows XP • Windows 2000 • Windows NT • UNIX (Solaris, AIX, Tru64, HPUX) • Linux	

Websearch is a bilingual full-text search engine for your web site. Whether you host your web site on your company network or use a Web hosting service, Websearch allows your customers to directly access the information they need with the convenience of keywords instead of following layers of links. You can use Websearch to index web pages on Internet to create resource databases that are critical to your business, or use it as a tool to manage your corporate information within your intranet and to disseminate information both inside and outside your company.

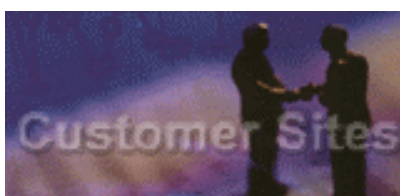
The spider administrator is a unique, indispensable component of Websearch. It allows you to:

- Schedule the spider to start at any time and to run periodically
- Index different sites or different parts of a site at different frequencies and number of levels
- Launch several spiders simultaneously
- View information on sites and URLs indexed

Suntek's three-tier architecture and unique index design give you speedy response time and high scalability. In addition, the search engine can index new information at near-constant time – an important factor to consider if you want to make your time-critical information immediately searchable to your users.

### Company Profile

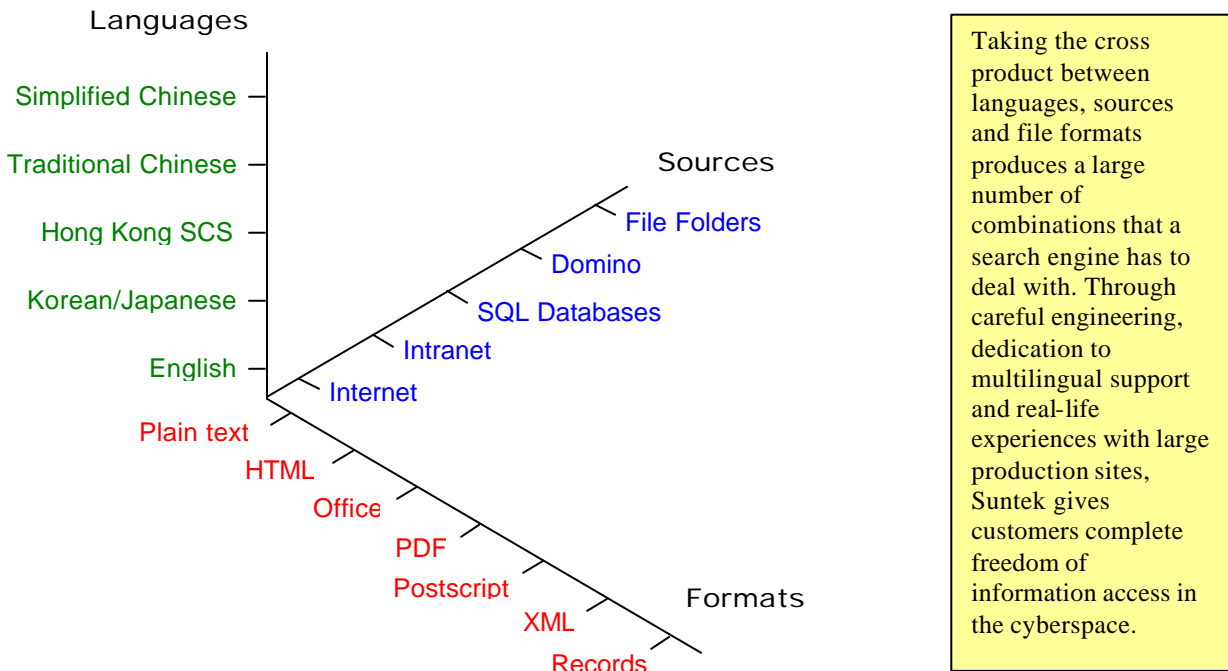
Suntek is a software company specializing in search and content-management solutions for Asian languages, especially Chinese. Owing all the source codes, Suntek can provide high-quality, fast turnaround support for and customisation of its products. Major customers include *The Hong Kong Police Force, The Hong Kong Government Information Center, Independent Commission Against Corruption (ICAC), Hong Kong Securities and Futures Commission (SFC), TVB.COM, The University of Hong Kong, Hong Kong University of Science and Technology, City University of Hong Kong, Hong Kong Institute of Education, and Higher Education Resources Organizer.*



## Integrated Access to Your Diversified Information Sources

Information vital to your business is spread across many diversified information sources. The information may come from a large variety of sources on your intranet: web servers, SQL databases, Domino servers, and file folders on shared drives. Of course, it may also come from your business partners' web servers or XML records received on the public Internet.

Further complicating the problem is the large number of file formats that useful information is embedded in: plain text, HTML, Word, PowerPoint, Excel, PDF, Postscript, etc.; each of them may contain text in different languages. Of particular concern to customers in the Asia Pacific region are English, Traditional Chinese (including the Hong Kong Supplementary Character Set), and Simplified Chinese.



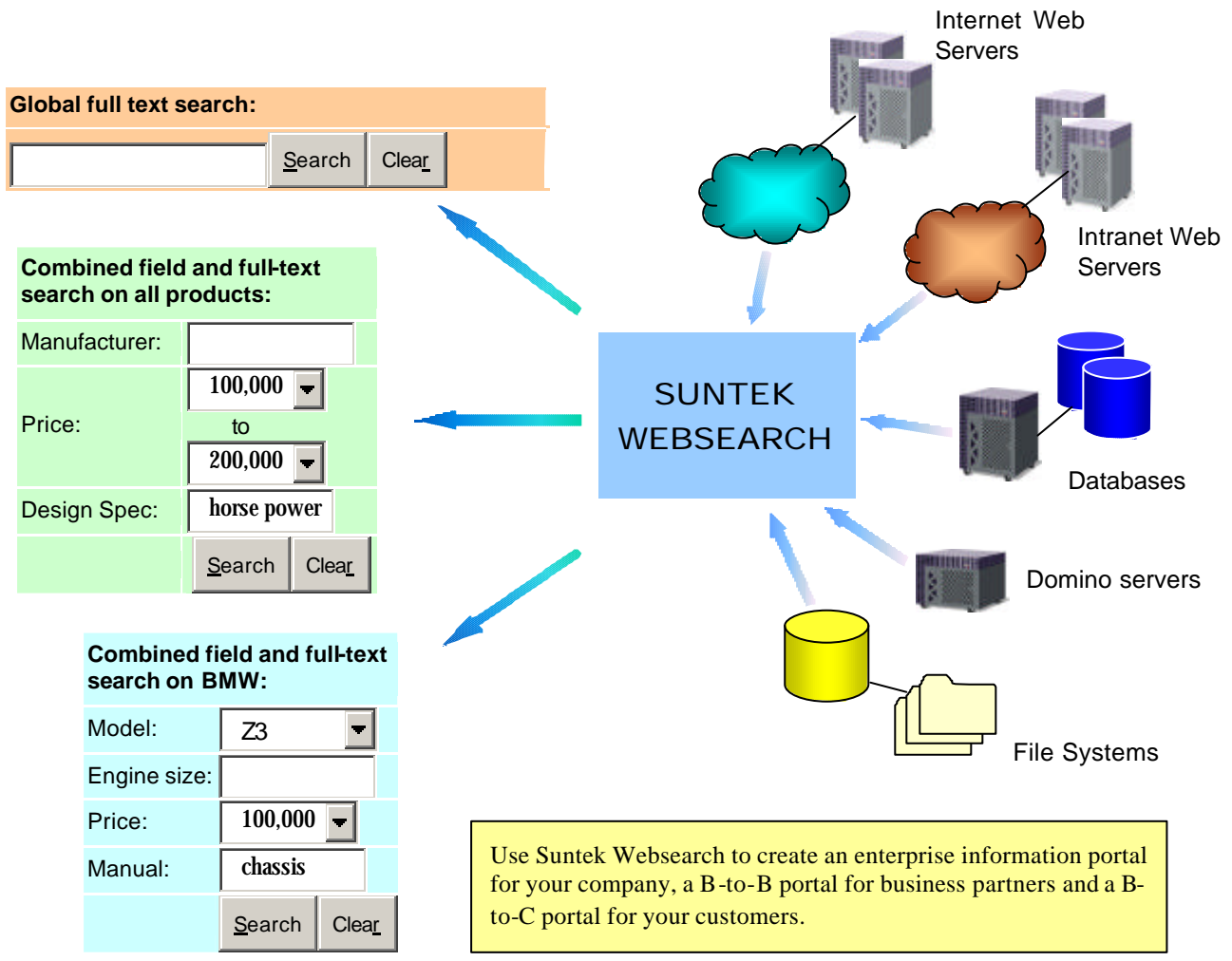
The matrix above illustrates the three dimensions that you should expect a search engine to handle *at the same time*. Many search engines depend on the operating system's locale and can handle one of the preset languages at one time. On the other hand, systems implemented on Unicode, while capable of handling multiple languages, fail at local character sets, such as the Hong Kong Supplementary Character Set, that are common in Asian languages. Through expert knowledge in Asian languages and real-life experiences on a large number of production web sites, Suntek offers a search solution that squarely meets all of the requirements in this three-dimensional space.

## One Search fits All

Now imagine having the power of accessing all of the diversified information through a single search engine, taking only a few keystrokes. And what it can do for your business, and how it would enhance decision-making and improve communication.

Yet different applications require different ways of searching. Suntek provides a wide variety of search capabilities that fit your applications. To enjoy the convenience of global search, you can use full-text search to look across all documents, including all document properties and text body.

Field search allows you to perform precise search by specifying conditions on each field, e.g., the document title shall contain "World Bank", the attachment, whether it is a WORD or PDF file, shall contain the words "WTO" and "real estate", and the publication date shall fall between 2000 and 2002. The search engine will return the precise results, sorted by any field you like. You may recognize these queries as typical database (SQL) queries, except that it goes beyond searching document fields to include attached documents.



Consider the grand picture above with Suntek Websearch pulling information from data sources across your intranet and from the Internet. You can search all of the information by typing a few keywords in a text input box. In addition, dedicated search interfaces can be created to search information in a finer manner. In the figure above, specific field-search interfaces can be defined on popular products (BMW or Toyota), and a separate interface to search across all products. Note in the example above the combined use of field-search and full-text search in the same search interface. For example, you can search across all BMW products that come with a manual, which could be a PDF or WORD file, containing the word “chassis”.

## All you Need for a Successful Search Portal

Besides powerful search functions and unsurpassed scalability and stability – which are well proven in our customer sites, Suntek provides you a suite of functions to ensure the success of your search portal, be it an internet site or an intranet enterprise portal.

*Suntek’s Message Server* allows you to deliver strategic messages to search engine users. The advertisement management function keeps track of advertiser information, keywords purchased and number of click-throughs. It allows you to charge based on the number of impressions or cost-per-click (CPC). For intranets, you may display daily news and action items to all users unconditionally or based on the queries they type in.

*Suntek’s Security Server* will tightly integrate into your application’s authorization system to ensure that unauthorized users will not find – or even know the existence of – privileged information.

*Suntek’s Same-Sound Server* supports match of Chinese characters that sound alike. It is available for both Pinyin and Cantonese pronunciation. This function is essential for bridging the gap between the terminologies used on your web pages and those that users are accustomed to.

*Suntek’s BIG5/GB Conversion Server* converts pages to and from BIG5 and GB, allowing users to search both character sets with the same query, and view the result in the encoding they use.

## Suntek Search Engine in Action

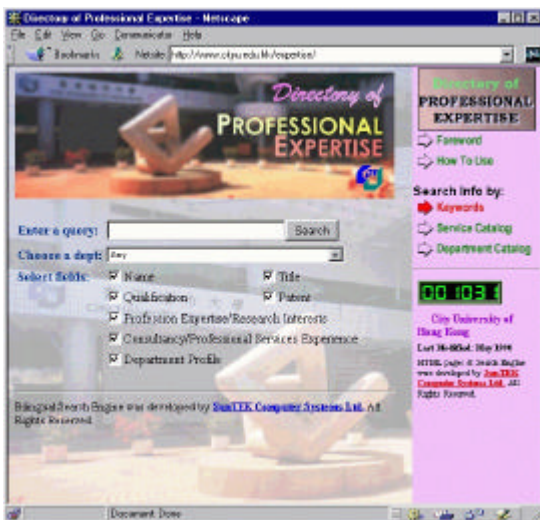


### The Hong Kong Government Information Center

The site run by the Hong Kong Government Information Center (<http://www.info.gov.hk/>) is one of the largest, if not the largest, portal in Hong Kong. Suntek Search Engine provides search on several hundred thousand web pages hosted on hundreds of government departments and semi-government organizations.

The speed and scalability of Suntek Search Engine are demonstrated by handling tens of thousands of queries each day.

The site supports multiple departmental search interfaces built on the same index database.



### City University of Hong Kong

The site at <http://www.cityu.edu.hk/expertise/> is a directory listing of the names, titles, qualifications, expertises, etc. of the professors in the University. Users can search the fields or full-texts of the faculty profiles. It is an excellent example for putting your company's employee profiles and product catalogs, etc., on the web.

The site at <http://www.cityu.edu.hk/cityutoday/> is a newsletter published by the University. It supports full-text and metatag search.

The site at <http://www.cityu.edu.hk/prospectus/> contains complete program and course information of the University. Users can search by course titles and/or course description.



### Hong Kong Institute of Education

Search at HKIED may look typical. However, a close look will first reveal that it supports search on both BIG5 and GB web pages. Suntek's code detection technology alleviates webmasters from explicitly separating and identifying the character set of each page. Pages using different codes can be mixed in the same folder, and Suntek will separate them.

Suntek allows GB queries to match against BIG5 pages and displays BIG5 pages in GB code – the code that the query uses – and vice versa.

Hidden from the interface is the authentication system that prohibits protected pages from being viewed by unauthorized users – not even the title will be shown in the result.

For more reference sites, please visit <http://www.suntek.com.hk/>

Suntek Computer Systems Limited Tel: 2784-7415 Fax: 2784-7413 [suntek@suntek.com.hk](mailto:suntek@suntek.com.hk) <http://www.suntek.com.hk/>