sonera | medialab

Mobile Search Engines White Paper

Sonera MediaLab

www.medialab.sonera.fi
info@medialab.sonera.fi

November 15, 2002

Mobile Search Engines

1 INTRODUCTION

The popularity of mobile devices, such as personal digital assistants (PDAs), has grown rapidly over the last couple of years. The amount of Internet content targeted for these devices is constantly growing as well. For example, the number of subscribers of the most popular mobile data service in Japan, i-mode, has grown from 11 million in August 2000 to more than 34 million in August 2002. During the same period, the number of sites offering content specially formatted for i-mode phones has grown from 21,000 to over 57,000 [1]. Considering the amount of content available for mobile devices, it is obvious that finding information by following hyperlinks only is not efficient. Specialized search engines for mobile content are also needed.

Different mobile devices support different features of the HyperText Markup Language (HTML) or they use other markup languages; for example, most Wireless Application Protocol (WAP) phones use the Wireless Markup Language (WML) whereas i-mode phones utilize Compact HTML. In addition, the screen sizes of mobile devices vary a great deal. The screen size of a Pocket PC PDA is 320x240 pixels, while a cellular phone may have screen space for seven rows of 16 character wide text with graphics resolution around 100x45 pixels. These are the main reasons why traditional search engines available on the Web today are not directly suited for mobile devices. The popular Web search engines Google and Fast have search interfaces for some mobile devices, but many other Web search engines are still desktop oriented and not well suited for mobile devices, because:

- they cannot be accessed with a mobile device, because the search engines do not have a user interface for the markup language the mobile device uses.
- they find content that cannot be shown on the mobile device.
- the search query page or the result listing does not fit on the mobile device's screen or in its memory.
- they lack regional information; with mobile devices people often want to restrict the query to the location they currently are in (for example, find restaurants nearby).

In this white paper, we take a look at search engine characteristics (chapter 2) and the different file formats mobile search engines encounter (chapter 3). With a mobile search engine we mean search software or a service which can be used with a mobile phone, a communicator, a smart phone, a PDA or similar mobile device and returns documents suitable for the mobile device making the query. A mobile search engine can collect specific markup languages only, or it can collect various markup languages and make a markup conversion before sending the results to the mobile device. We will not discuss handheld PCs, tablet PCs or laptops, because the normal search engines can usually be used with them. Furthermore, we concentrate on textual information retrieval.

Currently, Open Directory [2] lists 16 wireless search engines in its Mobile Computing category, while Search Engine Watch [3] lists 10 search engines or services for wireless users. In chapter 4 we introduce five search engines targeted for mobile users.

2 THE ANATOMY OF A SEARCH ENGINE

In its basic form, a search engine consists of a crawler, a parser, an indexer and a query engine (Figure 2-1). The crawler – also known as a spider – fetches web pages which are files written in various markup languages, and gives the pages to the parser. The right parser module parses the markup, separates URLs to be fetched later, and forwards the text to the indexer.

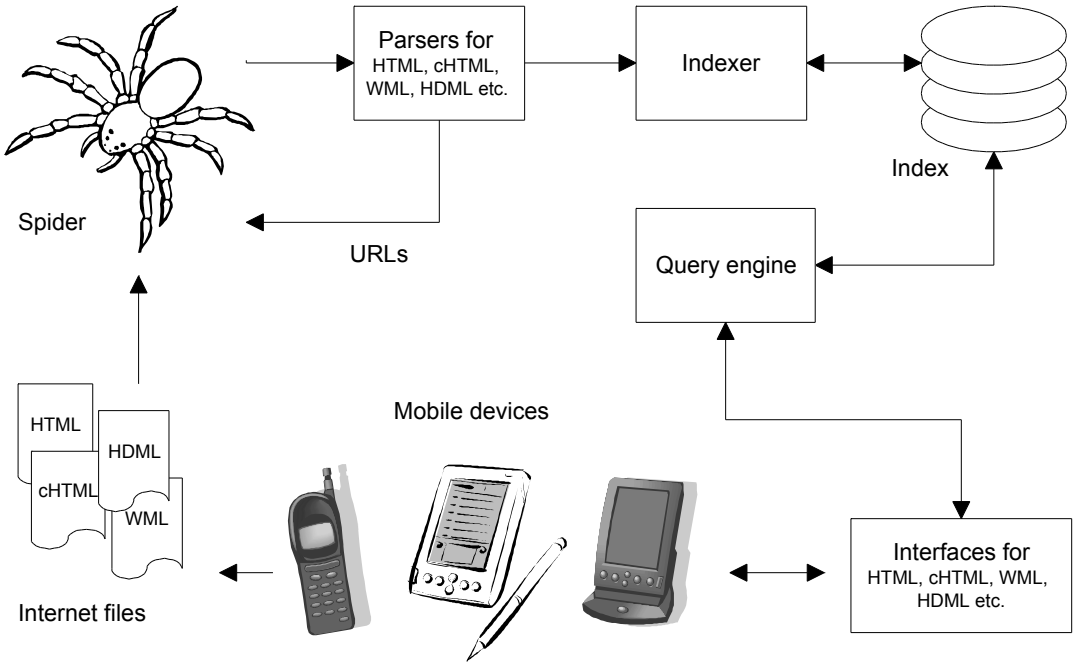

Figure 2-1. Search engine architecture.

The indexer saves the words and metadata of the page in the index database. The query engine receives search queries from the user, queries the index, and sends the search results back to the user. The results are returned in a format suitable for the device making the query. Only pages appropriate for that device are returned. The device can be identified from the headers it sends with the HyperText Transfer Protocol (HTTP) request. The architecture can also contain a conversion tool which can be used to convert markup languages not supported by the querying device into a supported one.

The index is updated when the crawler revisits the pages periodically to locate new, removed or changed pages. The ranking algorithms of the query engines are proprietary to the search engine companies, but most search engines most likely use the classical vector space model, where each document and a user query is represented as a multi-dimensional vector for ranking the search results. The similarity of these vectors can be determined mathematically by the cosine of the angle between these two vectors. The retrieved documents are sorted in a decreasing order of this degree of similarity.

Many guidelines for robot programmers are available on the Web, for an example, see [4]. Tim Bray, the writer of two large robots, points out the following technical issues for robot developers [5]:

- consider achieving high parallelism by flinging money at computers and RAM.
- learn the Robot Exclusion Protocol and take it seriously.
- if you're doing a robot that hits lots of sites whose operators you don't know, include a real human's email address in the HTTP request headers, and be responsive.
- consider doing domain name lookups as close as possible to the robot machines.
- assume that robot processes are going to freeze, lock up, go into loops, etc, from time to time, and build auto-recovery before you need it.
- to fetch pages, use a program which can be used as "get a maximum of X bytes from this URL and take a maximum of Y seconds to do it and come back when you're finished and tell me what happened".
- consider never reading more than X bytes from a URL, where X is, say, 10 K.

- consider crawling all the pages from a server, rather than going to a random server for each new request, but be really careful to avoid ever hitting any server with a lot of requests faster than it can deliver them. Track server response times to achieve this.
- ideally, auto-detect the language encoding being used and index non-English non-ASCII content in an intelligent way. As a fallback, be sure that you don't react to these pages by blowing up or indexing them incorrectly.
- read and use media type headers, but verify them.
- think a lot what you're going to do with redirects.
- consider just dumping URLs that break basic syntax rules.
- consider using a real relational database to store the data, rather than file system tricks.
- write highly compressed and maximally thorough logs, and be aggressive about watching for file system overload

Interface issues, such as small displays, tiny keypads and error-prone handwriting recognition still pose problems for mobile computing. To ease the searching for information, researchers have developed systems that ease text input with auto-completion, keyword suggestions, spoken queries and text summarization. For example, abbreviations and automatic removal of insignificant words can be used to fit the text on small screens.

3 MARKUP LANGUAGES FOR MOBILE DEVICES

In this chapter, markup languages used with mobile devices are discussed in more detail. Figure 3-1 outlines the evolution of markup languages during the past ten years. Arrows in the picture mean that the language is based on the ancestor specification, or the language borrows ideas from its ancestor. Traditionally WAP enabled mobile phones have used WML, while other mobile devices have used variants of HTML. Fortunately the WML and HTML camps have finally joined forces as the new WAP 2.0 specification uses XHTML Mobile Profile as the markup language for WAP content. In all probability all new mobile devices will support XHTML Mobile Profile in the future. Unfortunately, a mobile search engine should still know how to handle other mobile markup languages, if it tries to index older document types or if it tries to serve to older mobile devices as well.

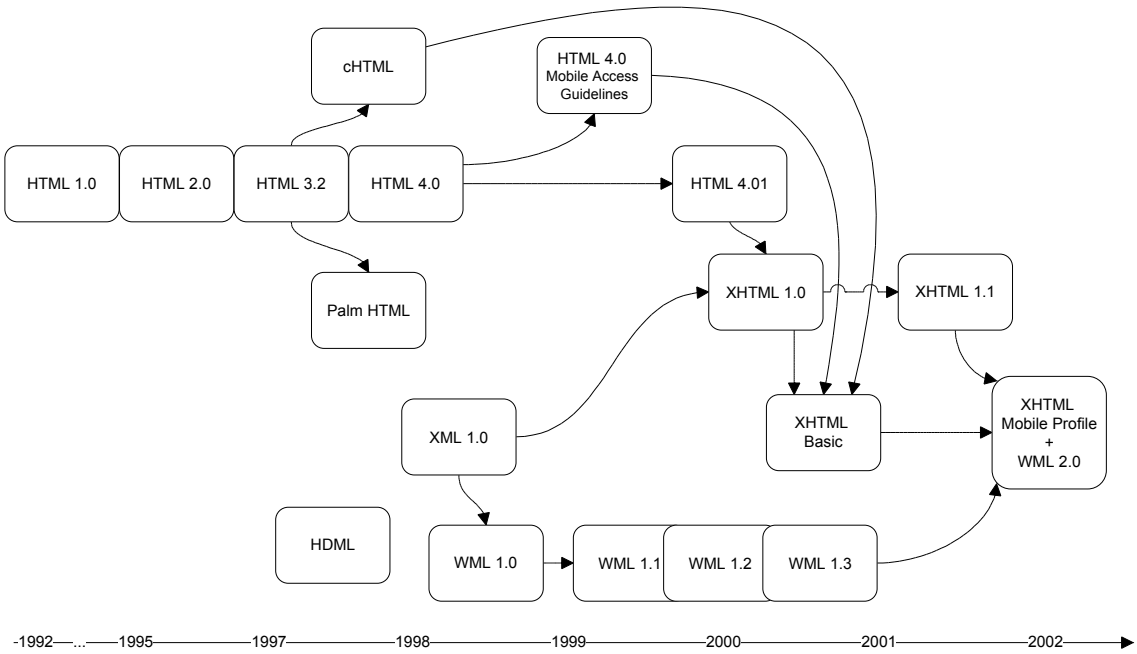

Figure 3-1. The evolution of markup languages.

3.1 HyperText Markup Language

HyperText Markup Language (HTML) is the *de facto* standard used on the Web. The current version, 4.01, is a successor to versions 4.0, 3.2, 2.0 and 1.0. The language was designed for desktop devices, but new mobile devices have included support for different versions of HTML as well. Because of the limitations of mobile devices (small display space, limited display capabilities, input methods etc.) all HTML features are not appropriate. A Pocket PC 2000 powered PDA can handle HTML version 3.2. The new Pocket PC 2002 understands more advanced HTML features. The Benefon Q mobile phone can also display HTML pages to some extent, as it includes Microsoft's Mobile Explorer. The Microsoft Windows powered Orange SPV smart phone has a 16-bit colour display with a resolution of 176x220 pixels and it can handle HTML version 3.2. The Nokia 9210i Communicator has HTML 4.01 support and it understands frames and cascading style sheets. HTML browsers for mobile Java enabled devices are also available. For example, the Reqwireless WebViewer HTML 4.01 browser can be used on the Nokia 7650 mobile phone which has a color display of 176x208 pixels. Also, Opera Software has announced a technology which reformats web sites to fit inside the mobile device's screen width, eliminating the need for horizontal scrolling.

3.2 HTML 4.01 Strict

The World Wide Web Consortium advises to use HTML 4.01 Strict [7] which excludes the presentation attributes and elements, as a baseline for documents targeted for mobile devices. An even better option is to use XHTML (discussed later), as it is newer and more scalable. The media type of an HTML document is `text/html` and to inform that the document conforms to the strict recommendation, the first line of an HTML document should contain:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01//EN"
"http://www.w3.org/TR/html4/strict.dtd">
```

3.3 Handheld Device Markup Language

The Handheld Device Markup Language (HDML) is Openwave's (formerly known as phone.com) proprietary language which can only be viewed on mobile phones that use the Openwave browser. It was created before WML, it uses the Handheld Device Transport Protocol (HDTP) instead of WAP and it is mainly used in North America. Newer versions of Openwave browsers also interpret WML. HDML is the only language that older phones understand, but it may be on its way to obsolescence since all new phones support fully XML-compliant WML or (X)HTML.

The media type of an HDML page is `text/x-hdml`. So, in order to serve HDML pages, web server administrators must add the correct media type to the server configuration files.

3.4 Compact HTML

Compact HTML [8], or cHTML for short, is a subset of HTML targeted for small appliances such as smart phones, communicators and mobile PDAs. i-mode is the most popular service that uses cHTML. The language is defined so that all basic operations can be performed using a combination of four buttons. Scrolling is not supported, because it is assumed that properly designed pages will fit on a single screen. cHTML includes support for GIF images, but JPEG images, tables, image maps, multiple character fonts and styles, background color and images, frames and style sheets are excluded, because mobile devices have display constraints, and limited power, memory and storage resources. The cHTML media type is `text/html` and the document type is defined as

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD Compact HTML 1.0 Draft//EN">.
```


3.5 Palm HTML

Palm HTML [9], the markup language used on Palm PDAs, is based on the HTML 3.2 specification, but the following 3.2 features are not supported: cascading style sheets, frames, image maps, animated GIFs, Java applets, JavaScript, layers, nested tables, multiple font faces, and the <vspace>, <sub>, <sup>, <link>, <isindex> tags. On the other hand, Palm HTML includes several new tags, variables and meta tags not found in the HTML specification. For example, the *PalmComputingPlatform* meta tag identifies that the HTML page is Palm friendly. The media type for Palm HTML is `text/html` and the document type is declared with

```
<!DOCTYPE HTML PUBLIC "-//POS//DTD Palm OS HTML 3.2 Final//EN">.
```

3.6 XHTML Basic

The Extensible HyperText Markup Language (XHTML) is a family of document types and modules in XML format. The first recommendation, XHTML 1.0, is a successor of HTML 4.01, and basically it is HTML 4.01 in XML format. The second recommendation, XHTML Basic [6], is designed for Web clients that do not support the full set of XHTML features, for example, mobile phones, PDAs, and set-top boxes. It also borrows ideas from cHTML and HTML 4.0 Guidelines for Mobile Access note. Among other things, XHTML Basic does not include style sheets, frames, scripts and events, but it does contain:

- basic text including headings, paragraphs, and lists,
- hyperlinks and links to related documents,
- basic forms and basic tables,
- images, and meta information.

XHTML Basic documents should be served as `application/xhtml+xml`. A document in XHTML Basic format should contain a document type declaration as follows:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML Basic 1.0//EN"
"http://www.w3.org/TR/xhtml-basic/xhtml-basic10.dtd">
```

3.7 XHTML Mobile Profile

Mobile Profile [16] extends XHTML Basic by bringing enhanced functionality to application authors, including more presentation elements and support for internal style sheets. The media type for XHTML Mobile Profile is `application/vnd.wap.xhtml+xml`. A properly formatted DOCTYPE declaration is specified as:

```
<!DOCTYPE html PUBLIC "-//WAPFORUM//DTD XHTML Mobile 1.0//EN"
"http://www.wapforum.org/DTD/xhtml-mobile10.dtd">
```

3.8 Wireless Markup Language

The Wireless Markup Language (WML) is the XML based markup language used with the Wireless Application Protocol (WAP). WML is designed for small narrowband devices and it includes four major functional areas:

- *text presentation and layout:* WML includes text and image support including formatting and layout commands.
- *deck/card organizational metaphor:* all information in WML is organized into a collection of cards and decks (documents).
- *navigation:* WML includes support for managing the navigation between cards and decks.
- *string parameterisation and state management:* all WML decks can be parameterised using a state model.

The new WML 2.0 document type is used only for backward compatibility and it is not intended for content authoring. WAP 2 content is created with XHTML Mobile Profile. The WML 2.0

document type consists of XHTML Mobile Profile plus WAP extensions which enable the delivery of WML 1.x content to WAP 2 clients, achieving backward compatibility in a transparent manner to the end user. The new WAP specification allows the device manufacturers to develop handsets that support applications with new features, such as color, multimedia messaging service (MMS), large file downloading, improved navigational functions, and user-friendly menus.

WML 2.0 documents are identified by the media type `application/wml+xml` while older WML documents are identified by `text/vnd.wap.wml`.

3.9 Distinguishing Different Markup Languages

The search engine should only return pages the mobile device can interpret. That is why the engine or its crawler should somehow distinguish between different page formats. It is easy to distinguish between HDML, WML and XHTML pages, because a properly configured Web server identifies these formats with particular media types (Table 3-1).

Table 3-1. Media types of mobile markup languages.

Markup Language	Media Type
HTML 4.01 Strict	text/html
HDML	text/x-hdml
cHTML	text/html
Palm HTML	text/html
XHTML Basic	application/xhtml+xml
XHTML Mobile Profile	application/vnd.wap.xhtml+xml
WML 1.x	text/vnd.wap.wml
WML 2.0	application/wml+xml

The situation with cHTML, Palm HTML and ordinary HTML pages can be tricky; they all have the media type `text/html`. In that case, file suffixes, HTML tags, document type declaration, page content, page sizse or URL patterns can be used to distinguish between pages suitable for a particular device. For example, if the page URL contains */i/* or */i-mode/* and the page does not contain frames, there is a great likelihood that the page is suitable for i-mode phone. Also, the page can be classified as a PDA page if it contains tags like:

```
<meta name="PalmPilotFriendly" content="True">
<meta name="HandheldFriendly" content="True">
<meta name="PalmComputingPlatform" content="True">
```


4 SOME MOBILE SEARCH ENGINES

In this chapter, we introduce five search engines targeted for mobile devices; WithAir concentrates to i-mode content only, IndexCell's service collects WML content only, whereas Fast, Pinpoint and Google collect several document types. Table 4-1 at the end of this chapter summarizes the features of these search engines.

4.1 FAST Mobile Search

FAST Mobile Search [10] has interfaces for WAP, HDML and i-mode phones, and also for other mobile devices like the iPAQ PDA at /wml/, /hdml/, /i/, and /pda/ on mobile.alltheweb.com server, respectively. The search returns content appropriate for the current device only. For example, WAP search finds files with the content-type `text/vnd.wap.wml` and HDML search finds files with the content-type `text/x-hdml`. The help page does not tell how FAST distinguishes between an ordinary HTML, i-mode or a PDA page. The index page claims the search service has indexed over 10 million documents.

4.2 WithAir

WithAir [11] allows the user to search over one million i-mode pages. It distinguishes an i-mode page from a web page by scoring structural characteristics like cHTML tags, URLs, and page size. For example, a page loses points if it contains tags unsupported by cHTML, such as <frame>. Higher scoring pages are more likely to be i-mode pages. WithAir reports its crawlers can distinguish between i-mode pages and Web pages with a precision of 99 %.

WithAir offers regional information searching by classifying i-mode pages according to the place names they contain, including addresses, zip codes, phone numbers, or train station names. A page that contains more place names related to a particular region is considered more likely to contain information related to that region. WithAir offers searching of 370,000 regional information pages extracted from over one million i-mode pages with a evaluated precision of 93 %.

Additionally, WithAir has two forms of navigation which simplify keyword input and search result output: keyword predictive navigation helps users input frequently used keywords taken from the query history and reduces the operation cost of entering search keywords. Purpose-predictive navigation reduces the trial-and-error in browsing search results.

4.3 Pinpoint Mobile Search Engine and Mobile Directory

Pinpoint Networks [12] offers a hosted and customizable search solution targeted for mobile operators. It enables the user to access thousands of wireless content services and web sites. Mobile Search detects and recognizes the type of the device making a query, and delivers results in the format best suited for that device. At least WML, HDML, cHTML and Palm pages are supported. The search allows the user to access pages in his or her preferred language, personalization and quick search help in the entering of client-defined search terms with fewer keystrokes.

Mobile Directory contains material reviewed and classified by Pinpoint's editorial staff. Directory categories and sub-categories can also be customized to meet each client's requirements. The directory can be integrated with Mobile Search. Pinpoint also has a tool which allows users to use a PC web browser to preview and navigate the wireless content before accessing it on a mobile device. For administrators there are usage reporting tools. Pinpoint's search customers include Terra Lycos, Verizon Wireless and Rogers AT&T Wireless.

4.4 Google Wireless Search

Figure 4-1. Google interfaces for a desktop computer, WAP enabled mobile phone and Palm PDA.

The famous web search engine Google also has wireless search solutions [13] for WAP, PDA, i-mode, and J-Sky devices at /wml, /palm, /imode and /jsky on the www.google.com server. Figure 4-1 shows search results for the query 'sonera' on a desktop computer, on a Nokia 6210 WAP phone and on a Palm m505 PDA.

Google Number Search for WAP devices found at http://www.466453.com/ simplifies the input of search queries and minimizes keypad strokes on a mobile phone. For example, the query 'sonera' can be entered as 766372 instead of 777766666337772.

Google says its wireless searches allow the searching for pages created for a particular mobile device plus the entire Web with Google's conversion system that translates Web pages into a format understandable by a mobile device. In this way users can access more than 3 billion pages. When tested, the conversion worked for WML pages, but Google's Palm optimized search didn't convert the pages found. The Palm search returned pages which were not meant to be shown on a PDA.

Like the Google Web search, Google's wireless searches use the PageRank [14] technology which examines the links between pages to reach accurate search results. It is possible to restrict the search to a given country or language. Google also sells customizable versions of its wireless searches.

4.5 IndexCell Wireless Search

IndexCell Communications [15] develops a UK specific wireless search engine which currently indexes WML content at the top level .uk domain only. WML pages can be submitted to the service via an "add URL" feature available on the IndexCell Web site. Only the main entry needs to be submitted as the IndexCell robot program will find the rest by following links. A WAP search interface is available at http://indexcell.com/.

IndexCell suggests that a card title be included in all WML documents (decks), the first card being the most important. Usually one or two words are enough and the total number of characters in the title should not exceed 18, as the output device's screen may be limited.

IndexCell also proposes the use of Wireless Meta Tags similar to the *Description* and *Keywords* Meta Tags used on HTML pages. The Description wireless meta tag is up to 60 characters long while the keywords meta tag should contain four to eight most important words. These meta tags are placed between the <head> and </head> tags of WML decks. These tags are indexed in addition to the ordinary text. With these suggestions in mind, a correctly formatted WML deck ready for indexing would look something like this:

```
<?xml version="1.0"?>
    <!DOCTYPE wml PUBLIC "-//WAPFORUM//DTD WML 1.1//EN"
    "http://www.WAPforum.org/DTD/wml_1.1.xml">
 <wml>
   <head>
   <meta name="description" content="We offer services from here to there
    and back again!"/>
   <meta name="keywords" content="business, services, production, me"/>
   </head>
   <card title="Info Title" id="home">
   </card>
 </wml>
```

For deck exclusion, the company suggests the use of a *robots.txt* file or Robots Meta tags similar to their HTML counterparts. The wireless version of the *noindex* Meta tag looks like

```
<meta name="robots" content="noindex"/>
```

and it should be placed at the top of the WML deck. If one wants to remove an entry from the search index, he or she needs to either delete the WML deck or use the *noindex* wireless meta tag and then re-submit the URL to IndexCell.

Table 4-1. Summary of mobile search engine features.

Feature / Engine	FAST Mobile	WithAir	Pinpoint Mobile	Google Wireless	IndexCell Wireless
WML support	Yes	No	Yes	Yes	Yes
HDML support	Yes	No	Yes	Yes	No
cHTML / i-mode support	Yes	Yes	Yes	Yes	No
PDA / Palm support	Yes	No	Yes	Yes	No
Language restriction	Yes	No	Yes	Yes	No
Domain / country restrict.	Yes	No	Not known	Yes	.uk only
Region restriction	No	Yes	No	No	No
Indexed documents	> 10 million	> 1 million	Thousands	> 3 billion*	Not known

* With on-the-fly conversion, otherwise 4-5 million.

5 MEDIALAB MOBILESEARCH

Figure 5-1. PDA and WAP interfaces for MobileSearch.

As an example of a search service for mobile devices, we developed our own search pilot – MobileSearch [17]. For queries we use third party text indexes which contain information on WML and HTML documents over the globe. We do not have control over what gets indexed and there is no mobile specific content index available. Fortunately, the query APIs of the third party indexes allow us to restrict the query by document features, file name endings or URL patters, among other things. Currently MobileSearch supports WAP phones, i-mode phones and PDA devices. Figure 5-1 illustrates the PDA and WAP interfaces of MobileSearch. The WAP interface locates WML pages by restricting the query to the WML file type. i-mode queries are restricted to HTML documents which do not contain unsupported HTML features, like frames, scripts or tables, and which do contain i-mode or its variants in the URL path. PDA documents are extracted in a similar way. In i-mode search, we also rank Japanese documents higher, because i-mode documents are usually written in Japanese. Our simple approach works quite well, but sometimes we get back result documents which are not suitable for an i-mode phone's or a PDA device's screen. This situation could be improved by accepting only documents which are smaller in size than a given limit value. Unfortunately, the query APIs do not allow restricting the query to documents regarding their size. To accommodate this, we sort the ranked result documents by size in ascending order with the smallest document appearing first before presenting the results to the end user. This way, there is a great likelihood that the pages occurring first in the result list are suitable for an i-mode or a PDA device.

6 SUMMARY

In this white paper we took a look at text search engines for mobile devices, how they operate in principle, and what kind of markup language types they collect, index and search. We also briefly discussed the limitations of mobile devices and lastly introduced five search engines that can be used with mobile devices.

With the amount of content targeted for mobile devices in many different markup languages, it is not an easy task to develop a good large-scale search engine. Many technical issues have to be taken into account if one wants to develop even a smaller search engine. Also, the interface issues with small displays place many restrictions. Researchers have used approaches such as text summarization, text auto-completion, keyword prediction and spoken queries to facilitate information retrieval on small devices.

Just like with ordinary web pages, the Robot Exclusion Protocol can be used with mobile content to exclude unwanted pages from the search index. IndexCell strongly encourages the use of this technique. Meta tags can be added to mobile content to help the engines to give more relevant search results and to classify pages for different devices.

It is not wise to serve pages a mobile device cannot display. To distinguish between different markup languages, page media type, document type declaration, file suffixes, page content, HTML tags, page size, or URL patterns can be used in addition to meta tags. Another approach is to use a converter which translates any page to a format suitable for the device. Google uses this method for its wireless searches.

With mobile devices, people often want to restrict the query to the location they currently are in. They would like to find a shop or a restaurant nearby, for example. Currently, mobile search engines do not have good support for regional information. Of the five introduced engines, only WithAir contains support for regional information. Fast and Google allow the search to be restricted to a country or language, which is certainly not enough in the mobile world. Of the presented mobile search engines, Fast, Google and Pinpoint index more than one markup language types, with Google offering the best search experience with a much bigger wireless index than its competitors. WithAir concentrates to i-mode content only, but it does the job well with a nice region restriction feature. IndexCell collects .uk specific WML pages only.

Without doubt browsers on mobile devices will come closer to their desktop counterparts in the future and they will have better support for the XHTML standard; what works on the desktop will also work on mobile device. As WAP now supports XHTML, it aims to reduce development costs, allowing developers to write applications for both PC and WAP clients concurrently. Hopefully a common markup language also helps search technology providers to develop search engines for both desktop computers and mobile devices.

REFERENCES

[1] NTT DoCoMo: Subscriber Growth, http://www.nttdocomo.com/html/subscriber_growth.html [Accessed October 7, 2002]

[2] Open Directory Project, http://dmoz.org/ [Accessed October 8, 2002]

[3] Search Engine Watch: Tips About Internet Search Engines & Search Engine Submission, http://www.searchenginewatch.com [Accessed October 8, 2002]

[4] Avi Rappoport, Checklist for Search Robot Crawling and Indexing, http://www.searchtools.com/robots/robot-checklist.html [Accessed August 27, 2002]

[5] Tim Bray, Re: Perl and LWP robots, Emails to the Internet robots and spiders discussion list, Mar 7, 2002.

[6] Mark Baker *et al.*, XHTML Basic, W3C Recommendation, 19 December 2000, http://www.w3.org/TR/xhtml-basic/ [Accessed October 8, 2002]

[7] Tomihisa Kamada *et al.*, HTML 4.0 Guidelines for Mobile Access, W3C Note, 15 March 1999, http://www.w3.org/TR/NOTE-html40-mobile/ [Accessed October 8, 2002]

[8] Tomihisa Kamada, Compact HTML for Small Information Appliances, Access Co. Ltd., W3C Note, February 9, 1998, http://www.w3.org/TR/1998/NOTE-compactHTML-19980209/ [Accessed August 28, 2002]

[9] Gary Hillerson, Web Clipping Developer's Guide, Palm, Inc., May 1, 2001, http://www.palmos.com/dev/support/docs/webclipping/PalmWebClippingFront.html [Accessed 28, 2002]

[10] FAST Mobile Search, http://mobile.alltheweb.com/ [Accessed August 28, 2002]

[11] Hideki Kawai, Susumu Akamine, Koji Kida, Katsushi Matsuda, Toshikazu Fukushima, Developing and Evaluation of the WithAir Mobile Search Engine, *Proc. 11th Int. Word-Wide Web Conf.*, 2002, http://www2002.org/CDROM/poster/102.pdf [Accessed August 28, 2002]

[12] Pinpoint Networks Web site, http://www.pinpoint.com/ [Accessed August 28, 2002]

[13] Google Wireless Services, http://www.google.com/wireless/ [Accessed August 28, 2002]

[14] Ian Rogers, The Google Page Rank Algorithm and How It Works, IPR Computing Ltd., May 16, 2002. http://www.iprcom.com/papers/pagerank/ [Accessed August 27, 2002]

[15] IndexCell Communications Web site, http://www.indexcell.com/ [Accessed August 28, 2002]

[16] WAP Forum, XHTML Mobile Profile, Oct 29, 2001. http://www.wapforum.org [Accessed November 6, 2002]

[17] Sonera MediaLab MobileSearch, http://www.medialab.sonera.fi/mobilesearch/ [Accessed November 11, 2002]

