A Comparison of Two Desktop Search Engines: Google Desktop Search (Beta) vs. Windows XP Search Companion

Paula A. Farina

ABSTRACT

The beta version of "Google Desktop Search" (GDS) made its public debut in October 2004, touted by the folks at Google as being "how our brains would work if we had photographic memories" [3]. Similar to the way Google searches the web, GDS searches your computer's personal index and digs up files you may have forgotten you even had. This paper reveals the strengths and weaknesses of GDS and compares its performance to Windows XP's built-in "Search Companion" (WSC). Though GDS performs searches with lightning-fast speed, it only does so with a limited selection of file types due to the fact that it is still in its beta stage. WSC, on the other hand, can search for any file type, but it tends to be slower due to the fact that it searches in more locations than GDS does. This paper discusses the differences in performance between the two desktop search engines.

Keywords

Search engine, desktop search, precision, indexing service, Google Desktop Search, Windows Search Companion.

1. INTRODUCTION

For years, I have been what I call a "virtual pack rat" – a computer owner with a habit of hoarding files on my PC in random folders and on whatever drive space I can find. Saving a file is far less worrisome than emptying the recycling bin. I am a frequent user of Windows XP's "search" function – a.k.a., Search Companion – which I find slow and cumbersome. For the longest time, I have wished that I could just "google" my hard drive for instantaneous results. I guess I was not the only one wishing for this, as Google released the beta of its "Google Desktop" search engine on October 14, 2004 [4].

Hard drives are getting bigger and bigger, enabling those of us who already exhibit pack rat tendencies to indulge in this habit even more. Thus, there is an urgently increasing need for a better search tool to scour our computers and find that obscure file or email that we might otherwise never find. Over the years, a few companies have met this call and released an assortment of desktop search products, some available for free download and others available for a fee. Google's online search engine's popularity will no doubt catapult its desktop search engine ahead of the other free desktop engines on the market. People loyal to Googling the web will inevitably want to try Google Desktop. Though in its beta stage, Google Desktop marks an important point in the evolution of the relationship between man and computer, promising to make man the victor in the struggle against misplaced electronic files.

Google Desktop Search (Beta) is available for free download on Google's website [3]. Other popular web search engines have come out of the woodwork in recent months as well with free desktop search products, including Yahoo Desktop Search (Beta) and MSN Desktop Search [1], and a few have been around for some time now, including HotBot [1] and Copernic [4]. Google stands to be the biggest contender due to the overwhelming popularity of its online search engine. Google has legions of users who will find the look of its desktop search browser (almost identical to their web search browser) familiar and easy-to-use [4].

While each of these products promises to be better than any other out there, few have compared any of them to the desktop search function that has been a part of the Windows operating system for years, right under users' noses. Indeed, Microsoft has included a search function in the start menu of each of its operating system releases since Windows 95. Windows XP's Search Companion (WSC) has a "Windows Indexing Service" that may be switched on or off. Without indexing, WSC grinds away slowly through each file, searching word by word for the intended target. With indexing switched on, searches are notably faster. Narrow down the folders that are searched, excluding hidden files and folders and system folders, and the search is that much faster.

This paper aims to determine which is the overall superior desktop search mechanism in terms of performance: Google Desktop Search or Windows XP's Search Companion. Due to limited time and resources, no other products will be examined.

For the purposes of this study, performance is defined as an amalgamation of the following characteristics:

- Search time – the lower, the better;
- Number of total hits – the more, the better;
- Hit rate – the higher, the better;
- Precision of hits – the higher, the better.

Thus, this paper really looks at the following series of null hypotheses:

1) There is no statistically significant difference in search time between GDS and WSC when conducting a single-word search on file contents.
2) There is no statistically significant difference in total number of hits between GDS and WSC when conducting a single-word search on file contents.
3) There is no statistically significant difference in hit rate between GDS and WSC when conducting a single-word search on file contents.
4) There is no statistically significant difference in precision of search results between GDS and WSC when conducting a single-word search on file contents.

5) There is no statistically significant difference in search time between GDS and WSC when conducting a multiple-word (phrase) search on file contents.
6) There is no statistically significant difference in total number of hits between GDS and WSC when conducting a multiple-word (phrase) search on file contents.
7) There is no statistically significant difference in hit rate between GDS and WSC when conducting a multiple-word (phrase) search on file contents.
8) There is no statistically significant difference in precision of search results between GDS and WSC when conducting a multiple-word (phrase) search on file contents.
9) There is no statistically significant difference in search time between GDS and WSC when conducting a search by file name and type.
10) There is no statistically significant difference in total number of hits between GDS and WSC when conducting a search by file name and type.
11) There is no statistically significant difference in hit rate between GDS and WSC when conducting a search by file name and type.
12) There is no statistically significant difference in precision of search results between GDS and WSC when conducting a search by file name and type.

In addition, this paper discusses the differences in interface usability between GDS and WSC, which on a subjective level might affect performance.

2. PROCEDURE

Difficulties presented themselves almost immediately in trying to compare GDS and WCS. Because GDS is in its beta stage, it lacks scope. The current version can search for the following file types: MS Word documents (.doc), MS Excel documents (.xls), MS PowerPoint documents (.ppt), Outlook / Outlook Express emails, Web sites (Web history), text documents (.txt), and AOL chats. GDS only looks in folders where a user typically stores documents, i.e., it does not look in system folders [3]. WSC, on the other hand, is too broad in scope and tends to search every single folder on the hard drive, unless told otherwise. Nothing is off-limits to WSC (though it does not seem to search through emails), and this lack of focus is what slows it down, even with the indexing service enabled. While much can be said about the benefits of a comprehensive search, one that is too comprehensive will dig up obscure system files that no average user would understand, much less find useful.

In light of these differences and to make the comparison fair and ensure its validity, both search engines' preferences were adjusted to level the playing field, more or less. This paper focuses on the retrieval of documents including Word, Excel, PowerPoint, text, and the like; retrieval of emails, Web sites, and chats is outside the realm of this study. As already mentioned, GDS does not have the ability to search system folders – this omission was intentional. According to the folks at Google, GDS indexes and searches "the files and folders on your hard drive (the ones you actually look at, not the system files only your computer uses) [3]." Moreover, GDS preferences were adjusted to exclude the searching of the following folders: c:\I386, c:\DRIVERS, and c:\DELL. They were also adjusted such that the following items were not included in searches: Outlook emails, Outlook Express emails, Web history, and AOL chats. This was accomplished fairly effortlessly by single-clicking on the GDS "swirl" icon in the Windows system tray and selecting "Preferences" -- which in turn brought up the GDS browser (a regular Internet Explorer browser) -- making the appropriate changes, and clicking on the "save preferences" button.

Already mentioned was the fact that WSC can look for just about anything under the sun, so setting up its preferences required a little more work. The folders designated not-to-be-searched in GDS were also designated as such in WSC. This meant going to each undesirable folder and marking it "hidden" before even conducting a search, and then within the Search Companion window, making sure the boxes next to "search system folders" and "search hidden files and folders" were both deselected. Searching "all files and folders" was appropriate for the filename/type searches but too broad for the single-word and multiple-word searches; instead, these searches were performed on "documents (word processing, spreadsheets, etc.)".

Three experiments were conducted in the course of this study. One examined each search engine's performance in searching the contents of files for a single word, another examined each engine's performance in searching file contents for a two-or-three-word phrase, and the third examined each engine's performance in searching for specific files by name and extension (type). The experiments were performed on a Dell Inspiron 9200 laptop computer possessing 512MB of RAM. While each search was conducted, no other windows or applications were open. GDS features a built-in search timer. WSC, however, lacks this tool, so for WSC, a basic stopwatch (featuring 0.01-second precision) was used to measure search time.

Both GDS and WSC had to perform an initial sweep of the hard drive to form an index. Though it was unclear how long the initial WSC indexing process took, the GDS process took approximately 10 minutes of idle CPU time. (According to Google, this process could take several hours, depending on how much material is on the user's hard drive [3].) Once each index was established, it was updated on a continual basis as the computer was put to everyday use.

Sample generation for the experiment consisted of coming up with three random lists: a list of 30 words, a list of 30 phrases, and a list of 30 filenames (with extensions). A sample size of $n = 30$ was selected as a decent sample size (again, given time and resource constraints) for a population ranging from 46 (total file names) to 1928 (total words) in size. The population from which the samples were drawn was generated by a thorough examination of each file, located "out in the open" on the laptop, that fell into one of the following filetype categories: Word documents, text files, Word Perfect documents, PowerPoint files, PDF files, and Excel spreadsheets. An item "out in the open" meant that it was on the desktop or in an unhidden subfolder of "Documents and Settings". This population, a master list of words, was run through the random number generator available at www.random.org [6]. For each search performed, search time, total number of hits, and precision were recorded. The precision of a given hit was gauged by whether the file contained the precise term upon which the search was performed. If not, the hit was deemed imprecise. As such, precision was measured and recorded as a proportion of precise hits to total hits [7].

3. RESULTS

Descriptive and inferential statistics were used to analyze the raw data. Because the sample was the same between the GDS and the WSC search runs, the correlated t-Test was the statistical analysis tool of choice for the data in this study. It was determined *a priori* to data collection that a significance level of $\alpha = 0.05$ would be used to avoid Type II errors and that the testing would be two-tailed in nature to avoid Type I errors, since the null hypotheses were non-directional. These standards are generally recommended by the scientific research community in order to distinguish between chance and a statistically significant effect [2]. In the following tables, "mean" refers to the average of the raw data, "var" refers to the variance between data points, "n" refers to the sample size (or number of observations), "df" stands for degrees of freedom, "t_{obs}" is the observed t value (or the t statistic), "$p(t_{obs}<=t_{crit})$ 2-tail" is the two-tailed probability that the t statistic calculated for the data is lower than or equal to the critical t-value, and "t_{crit} 2-tail" is the critical value of t, given a two-tailed distribution. The "Data Analysis" tool in Microsoft Excel was used to perform the following statistical analysis.

3.1 Single-Word Search

A correlated t-Test (n = 30) was conducted to evaluate whether search time for a single word differed significantly between GDS and WSC. As illustrated in Table 1, the absolute value of the t statistic calculated from the data ($t_{obs} = -159.8992923$) is much greater than the critical t value ($t_{crit} = 2.045230758$), and the two-tailed probability is much less than the α-level ($p = 2.83995E-44 < 0.05$). Thus, the mean search time for a single word differed significantly between GDS and WSC, with the GDS mean search time being significantly less than the WSC mean search time.

Table 1. Correlated t-Test for significance – Single-Word Search – Search Time

Statistic	Time GDS (sec)	Time WSC (sec)
mean	0.010333333	15.621
var	3.33333E-06	0.286754138
st dev	0.001825742	0.535494293
n	30	30
df	29	
t_{obs}	-159.8992923	
$p(t_{obs}<=t_{crit})$ 2-tail	2.83995E-44	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the total number of hits from a single-word search differed significantly between GDS and WSC. Table 2 shows that the absolute value of the t statistic calculated from the data ($t_{obs} = -3.466102336$) is greater than the critical t value ($t_{crit} = 2.045230758$), and the two-tailed probability is less than the α-level ($p = 0.001666294 < 0.05$). Thus, the mean total number of hits from a single-word search differed significantly between GDS and WSC, with the GDS mean total number of hits being significantly less than the WSC mean total number of hits.

Table 2. Correlated t-Test for significance – Single-Word Search – Total Hits

Statistic	Total Hits GDS	Total Hits WSC
mean	6.266666667	12.1
var	40.89195402	128.162069
st dev	6.394681698	11.32086874
n	30	30
df	29	
t_{obs}	-3.466102336	
$p(t_{obs}<=t_{crit})$ 2-tail	0.001666294	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the hit rate of a single-word search differed significantly between GDS and WSC. As shown in Table 3, the absolute value of the t statistic calculated from the data ($t_{obs} = 5.250827812$) is greater than the critical t value ($t_{crit} = 2.045230758$), and the two-tailed probability is much less than the α-level ($p = 1.26325E-05 < 0.05$). Thus, the mean hit rate of a single-word search differed significantly between GDS and WSC, with the GDS mean hit rate being significantly greater than the WSC mean hit rate.

Table 3. Correlated t-Test for significance – Single-Word Search – Hit Rate

Statistic	Hit Rate GDS	Hit Rate WSC
mean	610	0.78014712
var	404379.3103	0.54094924
st dev	635.9082562	0.735492515
n	30	30
df	29	
t_{obs}	5.250827812	
$p(t_{obs}<=t_{crit})$ 2-tail	1.26325e-05	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the precision of search results from a single-word search differed significantly between GDS and WSC. The absolute value of the t statistic calculated from the data ($t_{obs} = 3.353356199$) is greater than the critical t value ($t_{crit} = 2.045230758$), and the two-tailed probability is less than the α-level ($p = 0.00223587 < 0.05$), as listed in Table 4. Thus, the mean precision of search results from a single-word search differed significantly between GDS and WSC, with the GDS mean precision being significantly greater than the WSC mean precision.

Table 4. Correlated t-Test for significance – Single-Word Search – Precision

Statistic	Precision GDS	Precision WSC
mean	1	0.8701555
var	0	0.0449789
st dev	0	0.212082285
n	30	30
df	29	
t_{obs}	3.353356199	
$p(t_{obs}<=t_{crit})$ 2-tail	0.00223587	
t_{crit} 2-tail	2.045230758	

3.2 Multiple-Word Search

A correlated t-Test (n = 30) was conducted to evaluate whether search time for a phrase differed significantly between GDS and WSC. Table 5 shows that the absolute value of the t statistic calculated from the data (t_{obs} = -692.3642539) is much greater than the critical t value (t_{crit} = 2.045230758), and the two-tailed probability is much less than the α-level (p = 1.00391E-62 < 0.05). Thus, the mean search time for a phrase differed significantly between GDS and WSC, with the GDS mean search time being significantly less than the WSC mean search time.

Table 5. Correlated t-Test for significance – Phrase Search – Search Time

Statistic	Time GDS (sec)	Time WSC (sec)
mean	0.012333333	15.755
var	5.98851E-05	0.015115517
st dev	0.007738544	0.12294518
n	30	30
df	29	
t_{obs}	-692.3642539	
$p(t_{obs}<=t_{crit})$ 2-tail	1.00391E-62	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the total number of hits from a phrase search differed significantly between GDS and WSC. The absolute value of the t statistic calculated from the data (t_{obs} = -11.80262688) is greater than the critical t value (t_{crit} = 2.045230758), and the two-tailed probability is much less than the α-level (p = 1.35415E-12 < 0.05), as illustrated in Table 6. Thus, the mean total number of hits from a phrase search differed significantly between GDS and WSC, with the GDS mean total number of hits being significantly less than the WSC mean total number of hits.

Table 6. Correlated t-Test for significance – Phrase Search – Total Hits

Statistic	Total Hits GDS	Total Hits WSC
mean	2.066666667	5.033333333
var	7.650574713	5.964367816
st dev	2.765967229	2.442205523
n	30	30
Df	29	
t_{obs}	-11.80262688	
$p(t_{obs}<=t_{crit})$ 2-tail	1.35415E-12	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the hit rate of a phrase search differed significantly between GDS and WSC. Table 7 shows that the absolute value of the t statistic calculated from the data (t_{obs} = 5.611655452) is greater than the critical t value (t_{crit} = 2.045230758), and the two-tailed probability is much less than the α-level (p = 4.6444E-06 < 0.05). Thus, the mean hit rate of a phrase search differed significantly between GDS and WSC, with the GDS mean hit rate being significantly greater than the WSC mean hit rate.

Table 7. Correlated t-Test for significance – Phrase Search – Hit Rate

Statistic	Hit Rate GDS	Hit Rate WSC
mean	150	0.319663732
var	21379.31034	0.024182933
st dev	146.2166555	0.155508628
n	30	30
df	29	
t_{obs}	5.611655452	
$p(t_{obs}<=t_{crit})$ 2-tail	4.6444E-06	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the precision of search results for a phrase differed significantly between GDS and WSC. As Table 8 depicts, the mean precision for a phrase search between GDS and WSC was exactly the same at 1 (i.e., all hits were relevant in both cases). Thus, the mean precision of search results from a phrase search did not differ significantly between GDS and WSC.

Table 8. Correlated t-Test for significance – Phrase Search – Precision

Statistic	Precision GDS	Precision WSC
mean	1	1
var	0	0
st dev	0	0
n	30	30
df	29	
t_{obs}	undefined	
$p(t_{obs}<=t_{crit})$ 2-tail	undefined	
t_{crit} 2-tail	undefined	

3.3 Filename/type Search

A correlated t-Test (n = 30) was conducted to evaluate whether search time for a filename/type search differed significantly between GDS and WSC. The absolute value of the t statistic calculated from the data (t_{obs} = -31.99125278) is much greater than the critical t value (t_{crit} = 2.045230758), and the two-tailed probability is much less than the α-level (p = 3.59399E-24 < 0.05), as given in Table 9. Thus, the mean search time for a filename/type search differed significantly between GDS and WSC, with the GDS mean search time being significantly less than the WSC mean search time.

Table 9. Correlated t-Test for significance – Filename/type Search – Search Time

Statistic	Time GDS (sec)	Time WSC (sec)
mean	0.099666667	0.581
var	0.00303092	0.004478276
st dev	0.055053788	0.066919921
n	30	30
df	29	
t_{obs}	-31.99125278	
$p(t_{obs}<=t_{crit})$ 2-tail	3.59399E-24	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the total number of hits from a filename/type search differed significantly between GDS and WSC. Table 10 illustrates that the absolute value of the t statistic calculated from the data (t_{obs} = 0.921043264) is less than the critical t value (t_{crit} = 2.045230758), and the two-tailed probability is greater than the α-level ($p = 0.364626879 > 0.05$). Thus, the mean total number of hits from a filename/type search did not differ significantly between GDS and WSC. Though the mean number of hits garnered by GDS was greater than that of WSC, the degree to which it was greater was insignificant.

Table 10. Correlated t-Test for significance – Filename/type Search – Total Hits

Statistic	Total Hits GDS	Total Hits WSC
mean	1.4	1.1
var	4.248275862	0.162068966
st dev	2.061134606	0.4025779
n	30	30
df	29	
t_{obs}	0.921043264	
$p(t_{obs}<=t_{crit})$ 2-tail	0.364626879	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the hit rate of a filename/type search differed significantly between GDS and WSC. As shown in Table 11, the absolute value of the t statistic calculated from the data (t_{obs} = 5.601941931) is greater than the critical t value (t_{crit} = 2.045230758), and the two-tailed probability is much less than the α-level ($p = 4.7709E-06 < 0.05$). Thus, the mean hit rate of a filename/type search differed significantly between GDS and WSC, with the GDS mean hit rate being significantly greater than the WSC mean hit rate.

Table 11. Correlated t-Test for significance – Filename/type Search – Hit Rate

Statistic	Hit Rate GDS	Hit Rate WSC
mean	13.26388889	1.904639567
var	129.1625602	0.414991481
st dev	11.36497075	0.644198325
n	30	30
df	29	
t_{obs}	5.601941931	
$p(t_{obs}<=t_{crit})$ 2-tail	4.7709E-06	
t_{crit} 2-tail	2.045230758	

A correlated t-Test (n = 30) was conducted to evaluate whether the precision of search results from a filename/type search differed significantly between GDS and WSC. Table 12 shows that the absolute value of the t statistic calculated from the data (t_{obs} = -5.639825052) is greater than the critical t value (t_{crit} = 2.045230758), and the two-tailed probability is much less than the α-level ($p = 4.29629E-06 < 0.05$). Thus, the mean precision of search results from a filename/type search differed significantly between GDS and WSC, with the GDS mean precision being significantly less than the WSC mean precision.

Table 12. Correlated t-Test for significance – Filename/type Search – Precision

Statistic	Precision GDS	Precision WSC
mean	0.500555556	0.977777778
var	0.218400064	0.014814815
st dev	0.467332926	0.121716124
n	30	30
df	29	
t_{obs}	-5.639825052	
$p(t_{obs}<=t_{crit})$ 2-tail	4.29629E-06	
t_{crit} 2-tail	2.045230758	

4. DISCUSSION & CONCLUSION

Statistical analysis performed on the data indicates that there is a statistically significant difference between the performances of GDS and WSC, thus rejecting the null hypothesis that there is no statistically significant difference between the two engines' performances.

In terms of search time, GDS is significantly better than WSC. A GDS search brings up results instantaneously, in a fraction of a second, while WSC is not exactly instantaneous, taking a sustained amount of time (~15-20 seconds). Both are pretty quick, but GDS is what I like to call "lightning fast".

In terms of the total number of hits resulting from a search, WSC was the overall victor, pulling up significantly more hits on average than GDS in the single-word search and phrase search exercises. GDS seemed to lag in updating its index, so while WSC found recently created files, GDS frequently did not. GDS did not perform as wide a sweep as WSC, and GDS did not match partial filenames [1]. It is important to note, however, that many of the extra hits made by WSC over GDS were text-file cookies, and as such, were deemed irrelevant. When the search was done on a specific filename/type, the total number of hits was about even in the case of GDS and WSC.

The substantial speed advantage GDS has over WSC outweighed WSC's tendency to garner more hits, thus making GDS's hit rate on average significantly higher than that of WSC.

Precision was key to determining differences between the performances of GDS versus WSC. Regardless of whether the engine was faster or gathered up more hits, if it lacked relevant results, its performance would be diminished. The behavior of the engines varied widely depending on the type of search being performed. GDS performed better in the single-word experiment; GDS scoured files for precisely the word in question, whereas WSC even brought back instances in which the word in question was part of a larger word. For instance, performing a search for files containing the word "art" with WSC would not only bring back files containing "art", but also files containing words such as "part", "compartment", "cartilage", "smart", "start", and "article". This inherent imprecision, in many cases, is a considerable weakness exhibited by WSC. Then again, if the target file actually contains the word "arts", using the word "art" as a search term would result in a successful hit for WSC, but a failed hit for GDS. WSC performed better in the filename/type search, because for whatever reason, while WSC tended to bring up the exact file being searched for, GDS was not as consistently precise. In some cases, notably in file names containing spaces and/or symbols/punctuation, GDS failed to find the target file altogether. No doubt this is a weakness due to the engine's beta status. GDS and WSC performed equally well in terms of searches by phrase, pulling up only those files containing the phrases being sought.

In terms of usability of interface, GDS is the clear winner. To start a new search, the user simply double-clicks on the GDS icon in the bottom right of the XP taskbar to bring up the search browser, types in the search term, and clicks the "Search" button. In the blink of an eye, the results appear, listed in order of Google-judged precision or date, depending on how the user would rather have results ranked. Within the browser main page, the user can click on "Preferences" to adjust the search preferences. The search browser even displays the search time and lists a brief "abstract" of the items found in the search, with the search term in boldface.

WSC, on the other hand, generally needs to be started from the Start menu, unless the user knows the "magic" shortcut that consists of pressing the Windows key and F-key simultaneously. Either action brings up the WSC window and the "Search Companion", which can be chosen from a variety of annoying animated characters, including a dog that "fetches" items for you, known in some circles as "Fido the Time-Killing Windows Dog" [5]. (Fortunately, the user can opt to turn this feature off, and it was turned off for the purposes of this study.) At that point, the user must then determine whether to search for "pictures, music, or video", "documents (word processing, spreadsheet, etc.)", "all files and folders", or "computers or people". Upon selecting one of these options, the user must then type in the search term, with the option of narrowing the search by clicking on "more advanced options". In other words, there is no way to go to a convenient, centralized place to change the preferences in WSC. All the clicking around is tedious and time-consuming. WSC, moreover, does not give a preview of file contents the way GDS does, forcing the user to open each file to determine whether it is the intended target.

In light of the results of this study, it is difficult to make a summative statement that declares which search engine is best in terms of overall performance as defined earlier in this paper. Each has its own strengths and weaknesses. Though GDS is incredibly fast, it lacks the depth and precision of WSC, which in turn is slow and cumbersome in comparison to GDS. By and large, GDS has the potential to surpass WSC's performance in all facets when it matures beyond the beta stage. There are a few things that Google needs to work out prior to achieving this, but the current product puts them in a very good position and seems to be based on a highly efficient algorithm. Google gained its popularity as a web search engine, and undoubtedly its followers will want to try its desktop tool. For now, GDS and WSC seem to pick up for one another where the other leaves off, so perhaps until a "final release" is available, the GDS beta should be used as an adjunct tool to WSC.

5. FURTHER RESEARCH

Further research must be done to compare other desktop search tools to GDS and WSC. Since GDS was released, other players known for their web browsers have entered or will soon be entering the desktop search market, including Yahoo, MSN, HotBot, and Ask Jeeves [1]. America Online is also rumored to be working on a desktop search product. Besides these competitors, GDS is also up against smaller companies already offering desktop search products (some not available for free download), including X1, Vivisimo, and Copernic [4].

The community of PC users would also benefit from a repeat of this study once the full-blown (i.e., non-beta) version of GDS is released, though it is unclear when that will actually happen. Google claims that the full-blown version will be capable of searching more types of files.

In addition, future plans at Google include developing a Mac version of GDS [3]. When and if that happens, research should be done to compare GDS to the desktop search tools available to Mac users, most notably its upcoming Spotlight for the Tiger operating system, a desktop search tool that promises to "find anything on your computer as quickly as you type" [8].

6. REFERENCES

[1] P. Boutin, "Keeper Finders: Five new programs that let you search your hard drive without having a seizure," *Slate*, 31 Dec. 2004; http://slate.msn.com/id/2111643/.

[2] R.B. Burns, *Introduction to Research Methods*, Sage Publications, 2000.

[3] "Google Desktop Beta," 22 Oct. 2004; http://desktop.google.com/.

[4] S. Olsen, "Google Unveils Desktop Search," *CNET News.com*, 14 Oct. 2004; http://news.com.com/2100-1024_3-5408765.html.

[5] D. Pogue, "Google Takes On Your Desktop," *The New York Times On The Web*, 21 Oct. 2004; http://www.nytimes.com/2004/10/21/technology/circuits/21stat.html.

[6] "Random.org – True Random Number Service," Jan. 2005; http://random.org/nform.html.

[7] D. Sullivan, "Search Engine Glossary," *Search Engine Watch*, 17 Jan. 2005; http://searchenginewatch.com /facts/article.php/2156001.

[8] "Tiger Preview – Spotlight," 17 Jan. 2005; http://www.apple.com/macosx/tiger/spotlight.html.