M. Elizabeth Day (SBN 177125)
eday@bdiplaw.com
BUNSOW DE MORY LLP
701 El Camino Real
Redwood City, CA 94063
Telephone: (650) 351-7248

Jason S. McManis (*pro hac vice* to be filed)
jmcmanis@azalaw.com
Weining Bai (*pro hac vice* to be filed)
wbai@azalaw.com
AHMAD, ZAVITSANOS & MENSING, PLLC
1221 McKinney Street, Suite 2500
Houston, Texas 77010
Telephone: (713) 655-1101

Attorneys For Plaintiff,
LookSmart Group, Inc.

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

LOOKSMART GROUP, INC.,

Plaintiff,

v.

GOOGLE, LLC,

Defendant.

Case No. 5:24-cv-7147

ORIGINAL COMPLAINT FOR PATENT INFRINGEMENT

JURY TRIAL DEMANDED

Plaintiff LookSmart Group, Inc. ("LookSmart") files this suit against Defendant Google, LLC ("Google" or "Defendant") for infringement of United States Patent No. 7,356,530 (the "'530 Patent" or "Asserted Patent") and alleges, with personal knowledge as to its own actions and on information and belief as to the actions of others, as follows:

THE PARTIES

1. Plaintiff LookSmart is a Nevada corporation with its principal place of business at 2 N. Central Ave., Ste. 1800, Phoenix, Arizona 85004.

2. Defendant Google is a Delaware corporation with its principal place of business at 1600 Amphitheatre Parkway, Mountain View, California 94043. Google can be served with process by serving its registered agent CSC-Lawyers Incorporating Services, 2710 Gateway Oaks Drive, Suite 150N, Sacramento, California, 95833.

3. Google sells and offers to sell products and services throughout the United States, including in this judicial district, and introduces products and services into the stream of commerce that incorporate infringing technology knowing that they would be sold in this judicial district and elsewhere in the United States.

JURISDICTION AND VENUE

4. This Court has subject matter jurisdiction over this patent infringement action pursuant to 28 U.S.C. §§ 1331 and 1338(a).

5. This Court has personal jurisdiction over Google because Google has its corporate headquarters in this District, has committed the infringement complained of in this District and throughout the state of California, and regularly conducts business and/or solicits business in this District including selling, using, and offering to sale products and services that infringe LookSmart's Asserted Patent. This Court also has personal jurisdiction because Google has placed infringing products and services into the stream of commerce, with the expectation they will be purchased and used by customers in California and in this District, such that said customers have purchased and used Google's infringing products and services, which has allowed Google to derive substantial benefits from infringing acts in this District and in California.

6. Venue is proper in this District pursuant to 28 U.S.C. §§1391 and 1400(b). Venue is also proper in this District because Google maintains a regular and established place of business in this District.

FACTUAL BACKGROUND

LookSmart's Patented Innovations

Background.

7. The World Wide Web ("Web") is "a universally accessible hypertext platform for sharing information over the Internet." U.S. Patent No. 7,356,530 (the "'530 Patent") at 1:37-40 (attached hereto as Ex. A). The Web consists of numerous servers that store information, which can be accessed by the public through means such as Uniform Resource Locators ("URLs") or web addresses. *See id*. at 1:46-2:5. Each URL corresponds to a specific web page that can be viewed in web browsers such as Internet Explorer, Google Chrome, or Microsoft Edge. *See id*. at 2:22-24. Additionally, web pages may include links (also known as hyperlinks or hypertext links) to other web pages. *See id*. at 2:26-29. Users can click on a link to open the web page associated with the URL to which the link directs. *See id*. at 2:29-31.

8. In 2001, estimates suggested that the Web contained over two billion publicly accessible web pages. *Id*. at 1:14-16. At that time, the Web was expanding by approximately seven million pages each day. *Id*. at 1:12-14. Since then, the Web has experienced exponential growth. Google, one of the most widely used search engines, reports that the Web now contains hundreds of billions of individual web pages.

9. The growth of the Web can be attributed, in part, to the fact that any user with access to a web server can publish information online. *See* '530 Patent at 1:46-50. This accessibility removed many traditional obstacles to publishing, such as finding a publisher and covering printing costs, enabling worldwide access to information. Unlike traditional information repositories, such as public libraries, the Web lacks a central index. See *id.* at 2:15-17 ("The Web has no central index to the pages, such as that contained in a public library."). Instead, the Web consists of a "loosely linked set of pages." *Id.* at 2:4. Due to this lack of structured organization, it can be challenging to find relevant information on the Web. *See id.* at 2:18-20 ("Thus, the Web provides little structure to

support retrieval of specific information."). The rapid expansion of the Web only exacerbates this issue, making the retrieval of specific information even more difficult. *See id.* at 1:17-19 ("However, because of the Web's rapid growth and lack of central organization, millions of people cannot find specific information in an efficient manner.").

10. To access information on the Web, users have several options. First, if the user knows the URL of the desired page, they can navigate directly to that page. *See* '530 Patent at 2:34-35. Second, if the user knows the website hosting the desired page, they can visit the site and search within it for the relevant page. *See id.* at 2:35-37. Third, the user can browse through a web directory that organizes and categorizes websites. *See id.* at 2:39-60 (describing examples of web directories). Finally, users can employ a search engine to locate relevant websites by entering a keyword query. *See* '530 Patent at 2:61-3:12 (exemplarily describing web search engines).

11. However, both directory-based and keyword-based search methods presented challenges when it comes to efficiently finding information of interest at the time of invention. *See id.* at 1:11-19.

12. Given the vast size of the Web, a user's search query may return hundreds or thousands of relevant pages. However, the average user is neither inclined nor able to review all the documents that match their query. Instead, most users will typically only review one page of search results, which generally consists of around ten to twenty results. *See id.* at 3:8-10. "Therefore it is important to present the most relevant pages to the searchers at the top of the list, say in the first twenty results." *Id.* at 3:10-12.

13. Therefore, around the time of invention in the early 2000s, further innovation was needed to help users manage the overwhelming number of search results and allow them to see the most relevant results on the first page of the search results.

'530 Patent.

14. The '530 Patent is foundational in the field of internet search technology, particularly in the methods used to rank and index web pages. The innovations of the '530 Patent offer an advanced approach to improving internet search engine performance in terms of both the quality and speed of delivering search results. These improvements derive from the work of the LookSmart

team of inventors who focused on overcoming the limitations of early search algorithms by providing a more effective and reliable method for ranking web pages.

15. The '530 Patent describes and claims "methods for retrieving relevant information from a large collection of information such as that on the Internet and in particular the World Wide Web." *Id*. at 1:4-7. These methods provide technological solutions to the Internet-centric problem of locating relevant web pages in the Web, improving the ranking and reducing the latency of results generated in response to an Internet user's query. The analysis also enables a more sophisticated and accurate evaluation of a page's relevance, providing users with higher-quality search results faster than earlier systems.

16. The solutions recited by the claimed invention of the '530 Patent provide improved methods for determining a given page's overall rank, combining an intrinsic content score for the page at issue and the page weight of that page with an extrinsic ranking; improved methods for ranking pages based on their linking pages, anchor text on the linking pages, and the page weights of the linking pages, including improving the extrinsic ranking by examining linking pages for specific content and adjusting the anchor weight by page weight; and improved methods that reduce the time required to rank results, such as reducing the time required to rank results by associating ranking factors with keywords of each webpage indexed in a searchable data structure.

17. These are technological solutions necessarily rooted in computerized technology which address problems specific to computer networks and the Internet, including the specific problems of locating relevant information in a massive, loosely linked network of information content from countless authors and publishers.

18. When internet search engines first emerged, their primary focus was on analyzing keywords within a page's content to determine its relevance to a user's search query. However, these systems were vulnerable to spamming, where webmasters and page owners could manipulate the presence of keywords to elevate their pages in search rankings without providing valuable content. This issue resulted in suboptimal user experiences, as irrelevant or low-quality pages would often appear at or near the top of search results.

19. The inventors of the ’530 Patent addressed this internet search issue by going beyond simple keyword analysis. They devised a method that integrated both the internal content of a web page and the broader context of the Web’s hyperlink structure. This approach allows for the weighting of a web page based on both its intrinsic content and its connections to other Web pages, making it significantly harder for low-quality pages to dominate search rankings through manipulative tactics.

20. In particular, the ’530 Patent claims a system that crawls the Web to gather a collection of pages, then examines each page in relation to the likelihood that users will visit it, known as page weight. This process involves determining both intrinsic ranking factors (based on the content of the page) and extrinsic ranking factors (based on the links pointing to the page). These two factors are then integrated into a single score, which ranks pages for particular search terms. The precomputed rankings are stored in a database, allowing the search engine to quickly deliver ranked results when a user enters a query.

21. The ranking system at the core of the ’530 Patent fundamentally improves both the speed and quality of search results by incorporating the content of the subject Web page (such as by analyzing the actual text and metadata on the page to determine its relevance to specific search terms), the content of external links to the subject page from external pages (for example, by examining the anchor text and nearby content on the pages that link to the subject page, assessing how these links relate to the search terms), the page weight of the subject page (e.g., a probabilistic measure of how likely it is that a user will visit the page, based on the link structure of the Web), and the page weight of the linking pages (e.g., by evaluating the likelihood that users will visit the pages that link to the subject page, further refining the relevance score.).

22. The ’530 Patent’s combination of intrinsic and extrinsic ranking factors represents a substantial improvement over earlier, more rudimentary methods such as focusing solely on keyword density or simply counting the number of inbound links.

Improvements Over Prior Art

23. By introducing the concept of integrating intrinsic and extrinsic ranking factors, the invention described in the ’530 Patent allowed for a more nuanced and robust ranking system to

deliver the most relevant results to a search user and to do so quickly. *See, e.g.*, '530 Patent at 3:50-63, 13:9-47. The patent's innovations focused not only on counting the number of links but also on analyzing the content of both the subject page and the linking pages. This approach reduced the effectiveness of manipulative tactics like link spamming, as the quality and relevance of the links themselves became a critical part of the ranking process. *See, e.g.*, '530 Patent at 3:20-27.

24. Additionally, the '530 Patent introduced the concept of pre-indexing web pages by computing and storing these intrinsic and extrinsic scores before a user query was entered. *See, e.g.*, '530 Patent at 4:30-5:13. This method significantly reduced the time required to deliver quality search results, as the system could quickly retrieve pre-ranked pages from its index rather than calculating rankings on the fly. *See id.* (describing "fast access" and "fast retrieval"). This feature dramatically increases the speed at which search results are delivered because the engine does not need to perform complex calculations in real time. Instead, it can simply retrieve the pre-ranked pages from the index based on the user's query.

25. This pre-query ranking and indexing system represented a significant advancement over previous approaches, which often calculated rankings only after a user entered a query. By doing the heavy computational work in advance, the system ensured that users received their search results in a fraction of the time. This innovation also improved the overall user experience, as faster and more relevant results led to increased user satisfaction and engagement.

26. Finally, in this embodiment, the ranker combines the intrinsic and extrinsic ranking factors to generate a final rank for the given keyword-page pair. *See*, *e.g.*, '530 Patent at 5:4-34 (describing the ranker's operation). The query server utilizes the previously stored page weights and rank values for keyword-page pairs to efficiently return a sorted and ranked set of results in response to a keyword query. *See, e.g.*, '530 Patent at 5:29-34 (describing the query server's operation).

27. It is no surprise that, during prosecution of the '530 Patent, the Patent Office specifically remarked that the issued claims, which "produces a ranked, indexed database of words and related pages for producing ranked results in response to a search query" and reflects "building a searchable database indexed in accordance with selected words for producing a ranked set of search results in response for a query," constituted "a claimed result which is tangible, non-abstract,

real world result" as well as "a new and useful process or improvement thereof as required for patentability under 35 USC 101."

Google's Use of the '530 Patent's Innovations

28. Google, as the dominant player in the search engine market, has incorporated many of the techniques described in the '530 Patent into its own systems. Google's search results' performance demonstrate that Google employed pre-query indexing and weighted ranking factors, both of which are covered by the '530 Patent.

29. Reports indicate that Google began integrating aspects of the '530 Patent's innovations in the 2000s. This period coincides with several key updates to Google's search engine, including the "Florida" and "Austin" algorithm updates, which introduced more sophisticated ways of evaluating the relevance of web pages.

30. Google's personalized search technology, which was introduced as a default feature for signed-in users in 2007, further demonstrates its reliance on pre-query indexing and content-weighted PageRank, two of the central innovations described in the '530 Patent. This personalized search feature allowed Google to deliver more relevant results by considering not only the content and link structure of web pages but also the user's individual preferences and search history.

31. The innovations described in the '530 Patent are not just theoretical improvements; they have had a profound impact on the commercial viability of search engines. Google, for example, generated over $150 billion in search-related revenue during the period in which it is believed to have implemented the patented techniques. The pre-query indexing and content-weighted ranking systems covered by the Asserted Patent enabled Google to deliver faster and more relevant search results, leading to increased user engagement and higher ad revenue.

32. The success of Google's search engine, which has culminated in its unparalleled dominance of the Internet search market, can be directly tied to its ability to deliver high-quality results quickly, which it accomplished by using the innovations claimed in the '530 Patent. By improving the speed and relevance of search results, the '530 Patent enabled Google to become the undisputed leader of the Internet search market, leaving competitors far behind.

33. Google knows, and has known, about the '530 Patent and that Google's search engine and algorithms practice LookSmart's claimed invention. Indeed, LooksSmart's '530 Patent has been cited as prior art to numerous of Google's own patents and patent applications for over a decade, including Google patents or patent applications related to search engine technology and algorithms to rank documents.

34. Given the integral nature of LooksSmart's invention to the foundation of Google's business (its search engine), the repeated citations to the '530 Patent in Google's own patent and patent applications relating to search technology and processes for ranking documents, Google clearly knew that it was practicing the claimed invention set forth in the '530 Patent but continued to do so without any authority or license.

CAUSES OF ACTION

Count I: Infringement of U.S. Patent No. 7,356,530

35. LookSmart incorporates and re-alleges, as though fully set forth herein, the factual allegations contained in the paragraphs above.

36. On April 8, 2008, United States Patent No. 7,356,530 was duly and legally issued for inventions entitled "Systems and Methods of Retrieving Relevant Information." LookSmart was assigned the '530 Patent and continues to hold all rights and interest in the '530 Patent. Attached hereto as Exhibit A is a true and correct copy of the '530 Patent.

37. Claim 1 recites a ranking method illustrative of the '530 Patent inventions:

> A computer-implemented method of ranking the relevancy of pages in a collection of pages including linking hypertext pages, comprising:
>
> crawling the World Wide Web to produce a collection of pages without limitation to topic;
>
> for each page in the collection of pages, examining a probability of visitors viewing a particular page to determine a page weight for said particular page;
>
> for each of a plurality of selected words, with regard to each of a plurality of selected pages in the collection of pages;

> determining an intrinsic ranking factor for use of a selected word on a selected page in the collection of pages by examining content related to the selected word on the selected page to determine a content score and adjusting the content score in accordance with the page weight of the selected page, and
>
> determining an extrinsic ranking factor for use of the selected word on the selected page by, for each linking page in the collection of pages containing an outbound hypertext link to the selected page, examining text associated with the outbound hypertext link on the linking page related to the selected word to determine an anchor weight for the linking page, adjusting the anchor weight in accordance with the page weight of the linking page and combining the adjusted anchor weights for all linking pages containing an outbound hypertext link to the selected page;
>
> ranking the selected page for the selected word by combining the intrinsic and extrinsic ranking factors related thereto; and then
>
> creating a database of the collection of pages indexed by the plurality of selected words, each indexed selected word in the database index associated with pages ranked for said each indexed selected word so that ranked search results are produced in response to a subsequent query which includes one or more of the selected words.

38. Pursuant to 35 U.S.C. § 271(a), Google has directly infringed numerous claims of the '530 Patent, including at least claim 1 by having manufactured, used, sold, imported, and offered for sale Google Search's technology and services. For example, Google directly infringed at least claim 1 of the '530 Patent as discussed below[1]:

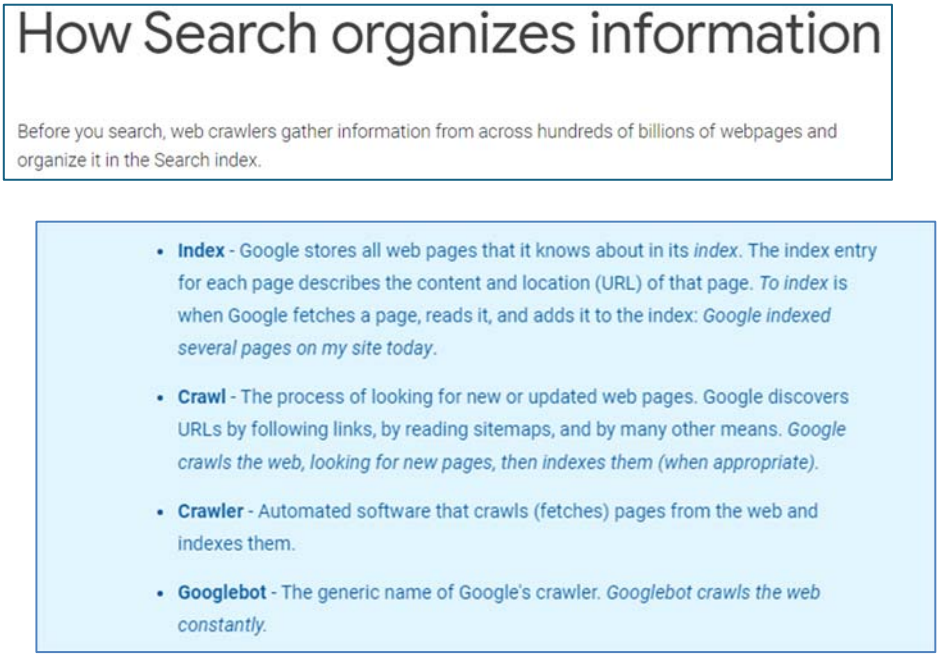
39. [1PRE] A computer-implemented method of ranking the relevancy of pages in a collection of pages including linking hypertext pages, comprising:

40. Insofar as the preamble is limiting, Google Search performed the steps below using software and/or hardware implemented in computers, such as servers, to rank the relevancy of web pages, including linking hypertext pages such as those on the World Wide Web containing links such as HTML links, for producing search results in response to user queries.

[1] The following exemplary citations are from sources available on or before April 15, 2021.

41. [1A] Crawling the World Wide Web to produce a collection of pages without limitation to topic:

42. Google Search crawled the World Wide Web to produce a collection of web pages. Google's crawling process, implemented by its Googlebot software, systematically scanned the web to discover and index new or updated web pages without any limitation on the topics of the pages crawled. These crawled pages were stored in Google's index for later retrieval when a user submits a search query.[2] Googlebot did not restrict its crawling to specific subject matters or topics.[3]



43. [1B] For each page in the collection of pages, examining a probability of visitors viewing a particular page to determine a page weight for said particular page:

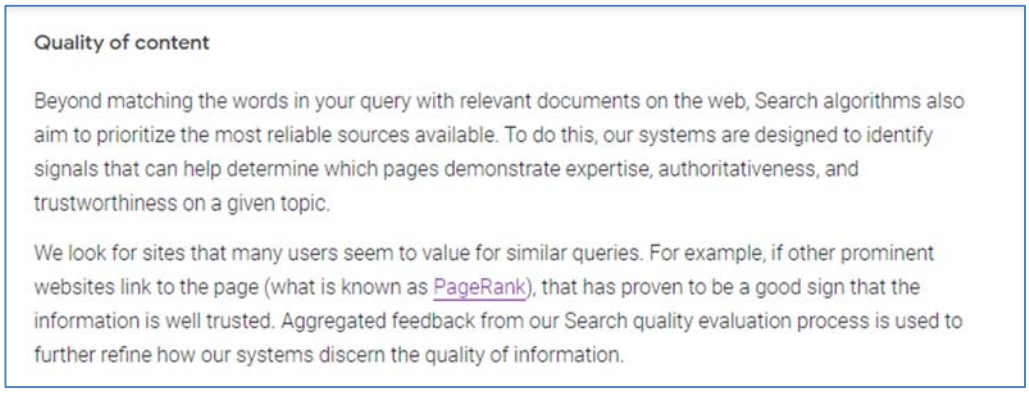
44. Google Search determined a page weight (*e.g.*, the PageRank) for each page in its index. PageRank evaluated the probability that users will visit a particular web page based on the number and quality of links to that page. The more important or trusted the pages linking to the subject page, the higher the PageRank score assigned to it. This system allowed Google to evaluate

---

[2] Google Developer Website – "How Search works".

[3] Google Developer Website – "SEO Starter Guide".

and rank the pages in its index based on their likelihood of being visited by users, as required by the claim.[4]



Quality of content

Beyond matching the words in your query with relevant documents on the web, Search algorithms also aim to prioritize the most reliable sources available. To do this, our systems are designed to identify signals that can help determine which pages demonstrate expertise, authoritativeness, and trustworthiness on a given topic.

We look for sites that many users seem to value for similar queries. For example, if other prominent websites link to the page (what is known as PageRank), that has proven to be a good sign that the information is well trusted. Aggregated feedback from our Search quality evaluation process is used to further refine how our systems discern the quality of information.

45. [1C] For each of a plurality of selected words, with regard to each of a plurality of selected pages in the collection of pages:

46. Google Search processed a plurality of selected words or keywords in its index, associating these words with a plurality of selected web pages. When a user enters a search query, the search engine retrieved and ranked pages from its index that are relevant to the selected keywords to help determine the ranking of the pages and how the content relates to the query.[5]

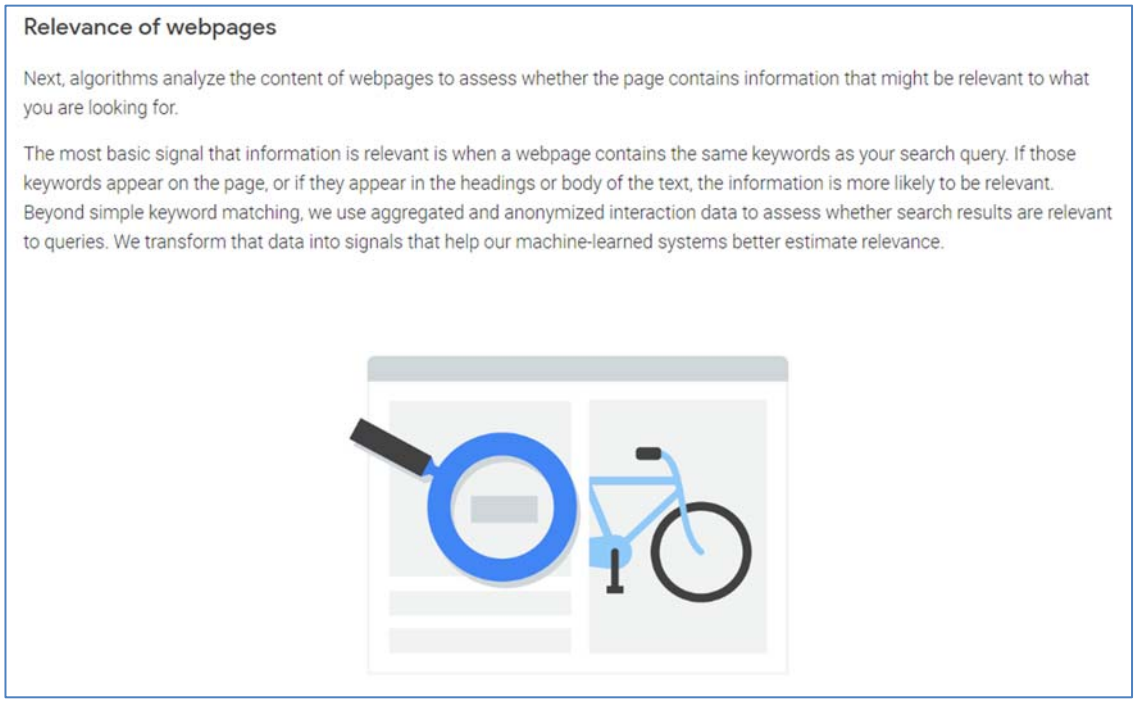
47. [1D] Determining an intrinsic ranking factor for use of a selected word on a selected page in the collection of pages by examining content related to the selected word on the selected page to determine a content score and adjusting the content score in accordance with the page weight of the selected page:

48. Google Search determined an intrinsic ranking factor by analyzing the content on each page related to the selected word. Google's algorithms evaluated whether the content on a page is relevant to the user's search query by looking for the presence of keywords and other relevant information. Based on this analysis, Google assigned a content score to the page, which is then

[4] Google Developer Website – "How Search Works"; G. Hotchkiss, *Gord Interviews Marissa Mayer on Personalization*, WebProNews (Feb. 23, 2007).

[5] G. Hotchkiss, *Gord Interviews Marissa Mayer on Personalization*, WebProNews (Feb. 23, 2007).

adjusted based on the PageRank or page weight of the page. The more authoritative or popular the page (as determined by PageRank), the higher the adjusted content score.[6]



49. [1E] Determining an extrinsic ranking factor for use of the selected word on the selected page by, for each linking page in the collection of pages containing an outbound hypertext link to the selected page, examining text associated with the outbound hypertext link on the linking page related to the selected word to determine an anchor weight for the linking page, adjusting the anchor weight in accordance with the page weight of the linking page and combining the adjusted anchor weights for all linking pages containing an outbound hypertext link to the selected page:

50. Google Search determined an extrinsic ranking factor by examining the anchor text and surrounding content on pages that link to the subject page. This anchor text helped Google understand the context of the link and how it relates to the selected search terms. Google assigned an anchor weight to each link, which is adjusted based on the PageRank (*e.g.*, the page weight) of

[6] Google Developer Website – "How Search Works"; *see also* U.S. Patent No. 7,260,573 (Jeh et al.) at 7:36-52. The personalized search team's algorithm adjusted (using, for example, function $F1$) the content score for each page in the collection of pages in accordance with the page's page weight (e.g., page importance score *PS*, such as PageRank); *id*. at 9:42-58.

the linking page. If a high-ranking, trusted page linked to the subject page, the anchor weight of that link was increased, further boosting the subject page's rank. Google combined the adjusted anchor weights from all linking pages to determine the extrinsic ranking factor for the subject page, as claimed in the '530 Patent.[7]

> **Write good link text**
>
> Link text is the visible text inside a link. This text tells users and Google something about the page you're linking to. Links on your page may be internal—pointing to other pages on your site—or external—leading to content on other sites. In either of these cases, the better your anchor text is, the easier it is for users to navigate and for Google to understand what the page you're linking to is about.

51. [1F] Ranking the selected page for the selected word by combining the intrinsic and extrinsic ranking factors related thereto:

52. Google Search ranked each page for a given search query by combining the intrinsic ranking factors (such as based on the content of the page) with the extrinsic ranking factors (*e.g.*, based on the links and anchor text from other pages). This combined score determined the overall relevance of the page for the selected search terms and influences the page's placement in the search results. The integration of intrinsic and extrinsic factors was a direct implementation of the claimed ranking process in the '530 Patent.[8]

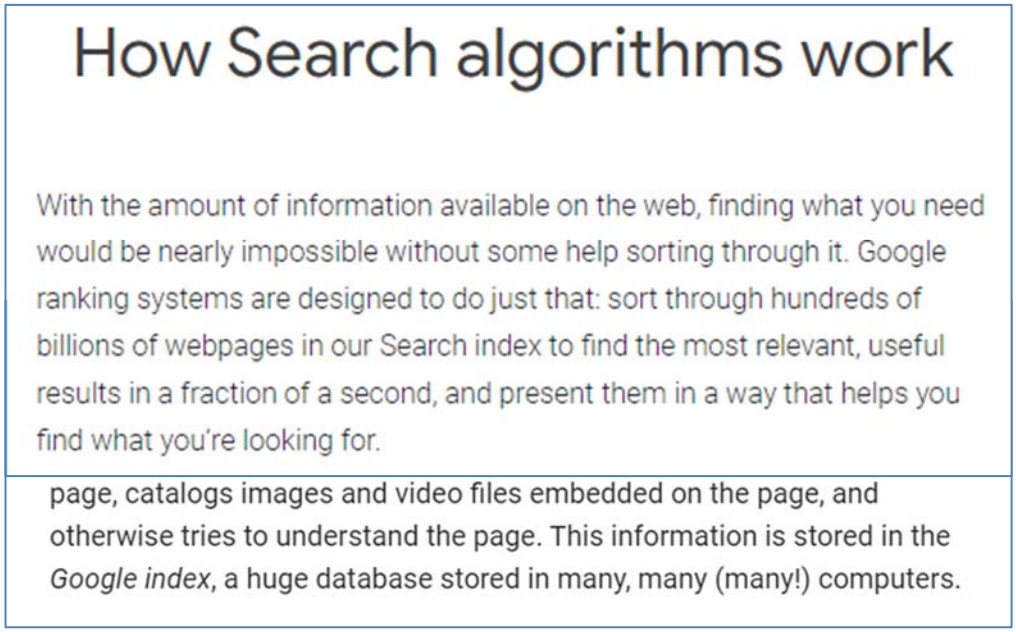
> These ranking systems are made up of not one, but a whole series of algorithms. To give you the most useful information, Search algorithms look at many factors, including the words of your query, relevance and usability of pages, expertise of sources, and your location and settings. The weight applied to each factor varies depending on the nature of your query—for example, the freshness of the content plays a bigger role in answering queries about current news topics than it does about dictionary definitions.

[7] Google Developer Website – "SEO Starter Guide"; U.S. Patent No. 7,260,573 (Jeh et al.) at 9:4-41.

[8] Google Developer Website – "How Search Works"; (U.S. Patent No. 7,260,573 (Jeh et al.) at 9:42-58.); M. Cutts, Text links and PageRank, Sept. 1, 2005.

53. [1G] Creating a database of the collection of pages indexed by the plurality of selected words, each indexed selected word in the database index associated with pages ranked for said each indexed selected word so that ranked search results are produced in response to a subsequent query which includes one or more of the selected words:

54. Google Search maintained a pre-indexed database of web pages that are associated with specific keywords. This index allowed Google to quickly retrieve and rank pages based on the selected words in a user's search query. When a query was entered, Google's search engine accessed the pre-computed rankings stored in the index to produce a set of relevant, ranked search results. This process of creating and using an indexed database for search results practiced the process claimed in the '530 Patent.[9]



How Search algorithms work

With the amount of information available on the web, finding what you need would be nearly impossible without some help sorting through it. Google ranking systems are designed to do just that: sort through hundreds of billions of webpages in our Search index to find the most relevant, useful results in a fraction of a second, and present them in a way that helps you find what you're looking for.

page, catalogs images and video files embedded on the page, and otherwise tries to understand the page. This information is stored in the *Google index*, a huge database stored in many, many (many!) computers.

[9] Google Developer Website – "How Search Works" ; G. Hotchkiss, *Gord Interviews Marissa Mayer on Personalization*, WebProNews (Feb. 23, 2007).

JURY DEMAND

55. LookSmart hereby demands a trial by jury on all issues.

PRAYER

Wherefore, LookSmart prays for entry of judgment as follows:

56. A judgment in favor of LookSmart that Google has infringed, either literally and/or under the doctrine of equivalents, the Asserted Patent;

57. An award of damages in favor of LookSmart adequate to compensate LookSmart for Google's infringement of the Asserted Patent which shall in no event be less than a reasonable royalty, together with interest and cost as fixed by the Court pursuant to 35 U.S.C. § 284;

58. An award of attorneys' fees pursuant to 35 U.S.C. § 285 or as otherwise permitted by law in an amount deemed just and appropriate by the Court;

59. Pre- and post-judgment interest;

60. An award of costs and expenses as deemed appropriate by the Court; and

61. Any other legal or equitable relief to which LookSmart is justly entitled.

Dated: October 14, 2024

Respectfully submitted,

/s/ M. Elizabeth Day

M. Elizabeth Day (SBN 177125)
eday@bdiplaw.com
BUNSOW DE MORY LLP
701 El Camino Real
Redwood City, CA 94063
Telephone: (650) 351-7248

Jason S. McManis (pending pro hac vice)
jmcmanis@azalaw.com
Weining Bai (pending pro hac vice)
wbai@azalaw.com
AHMAD, ZAVITSANOS & MENSING, PLLC
1221 McKinney Street, Suite 2500
Houston, Texas 77010
(713) 655-1101

Attorneys for Plaintiff\
LookSmart Group, Inc.