
NEWS RELEASE

Infonortics Ltd

15 Market Place, Tetbury, GL8 8DD, England

Contact: Harry Collier

Phone: +44 (0)1666 505 772

Fax: +44 (0)1666 505 774

E-mail: hcollier@infonortics.com

Website: www.infonortics.com

Google, We Hardly Know You!

THE GOOGLE LEGACY

A New Book by Stephen E. Arnold

What kind of company is Google?

The world mostly knows this high-flying, publicly traded American West Coast company as the upstart that has revolutionised search.

Wrong, says the author of a new book. Google is much more. New, radical, and overlooked, Google is this era's transformational computing platform and could be about to unseat Microsoft from its throne.

Google is not just about search: search is merely one application you can load on its processor, Internet technology expert Stephen E. Arnold says in **The Google Legacy** (Infonortics; September 2005; \$180 / €145, PDF version). Although Google has been releasing a series of separate applications programs, the company is starting to assemble the mosaic pieces into a bigger picture. Its future will be about leveraging its innovative hardware/software infrastructure. In so doing, just as Microsoft replaced IBM, Google promises to replace Microsoft as network computing comes of age.

Written for business readers, especially senior executives of mid to large-sized, knowledge-based corporations, **The Google Legacy** places Google under a microscope, dissects Google's technology, evaluates its potential and determines that Google's future lies beyond search. The book's 11 chapter headings include "Google's First Principles," "Google Basics," "Google Technology," "Google Relevance Ranking and Search Engine Optimization," "Gmail and Google Maps," "Google Clustering: News and Enhanced Search," "Google Print and Scholar," "Google: A New Force in Enterprise Search," "Google APIs: Netting Developers," "Google Goes Personal," and "The Google Legacy." Three appendices provide lists of Google patents, publishers who have indicated some type of relationship with Google, and universities working with Google – information that, according to the author, Google has not made easy to locate or update.

“. . . At Google, from its inception, Google software and Google hardware have been tightly coupled," Arnold observes. "Google is not a software company nor is it a hardware company. Google is, like IBM, a company that owes its existence to both hardware and software. Unlike IBM, Google has a business model that is advertiser supported. Technically, Google is conceptually closer to IBM (at one time a hardware and software company) than it is to Microsoft (primarily a software company) or Yahoo! (an integrator of multiple software)."

Among the book's critical insights:

Google's computing platform – named the Googleplex by Arnold after the name given by the company to its Mountain View headquarters complex – is a better (faster, cheaper and simpler to operate) computer processor and operating system than systems now available from competitors. Its price advantage is five or six to one over other hardware. Massively parallelised and distributed, its processing capability can be expanded indefinitely. As a virtual system or network utility, the user simply faces no need for backup or setup or restore.

Google has re-coded Linux [the well-known open-source alternative to Microsoft's Windows operating system] to meet its needs. This recoding enables Google to deploy numerous current and future applications – 50 or more – without degrading performance.

Google products have the potential to be assembled into a version of MS Office – including word processing – and many other applications.

Such insights underpin Arnold's conclusions that, if Google can avoid or overcome certain pitfalls and hurdles, "Google is poised to become the heir to Microsoft." The author sets forth these legal, management and marketing obstacles.

The book also identifies and explains a series of incremental hardware and software innovations "not fully appreciated by Google's competitors, analysts or users" that have given Google its competitive edge. "The net of these advantages is that Google does not have a search system. Google has a supercomputer that delivers applications. Some of these applications are free for the user; for example, search. Other applications are for Google's 4,000 employees; for example, the programmers who craft applications for the Googleplex and employees who use the formidable number-crunching capabilities of the Googleplex to figure out what users are doing, how to maximize advertising revenue from billions of online clicks in real time, and improve the search experience."

About the Author

A sought-after consultant, popular lecturer and established author on technology, Stephen E. Arnold has six books and over 50 articles appearing in a wide range of media to his credit. Among his books are five on the Internet, including *Publishing on the Internet: A New Medium for a New Millennium* (1996) and *New Trajectories of the Internet* (2001). These books identify trends, impacts and new technologies. His last book, *The Enterprise Search Report*, was a comprehensive overview of 28 search solution providers and best practices. Arnold, who heads Arnold Information Technology, is based in the Louisville, Kentucky, suburb of Harrod's Creek.

The Google Legacy (Infonortics, Tetbury, England; www.infonortics.com; August 2005). Available in online PDF version only; 280 printed pages. \$185 / €165 per download at: <http://www.infonortics.com/publications/google/google-legacy.html>

QUOTES FROM

THE GOOGLE LEGACY

(Permission by Infonortics is given to use the following material with proper attribution)

From *The Google Legacy* by Stephen E. Arnold, to be published online by Infonortics, (www.infonortics.com), in September 2005:

Chapter One: Google's First Principles

. . . the argument advanced in this monograph—Google is poised to become the heir to Microsoft. p. 14

* * *

. . . Commodity hardware [low-priced PCs available from retail outlets] could run fast only if Google were able to make commodity hardware reliable. Google solved this problem by changing Linux to read files at blinding speed without the cost of fibre-optic drives and gigabit hardware. Google figured out how to keep certain overhead costs down by changing the way its Linux-based system recognized new racks of servers and eliminating traditional tape back up processes. Google absorbed the most useful notions from Beowulf, River, and dozens of other research computing initiatives and baked them into its operating system. The payoff was that huge computing resources could be plugged into the Google system (hereinafter, "Googleplex") and become available without much human fiddling. In effect, Google can add storage, computational resources and special purpose functions to the Googleplex that sprawls over more than 15 data centres and more than 100,000 computing devices in a simple way. Google looked at how a Mac or XP laptop recognizes a USB mouse. Google implemented the "plug and play" approach on a much larger scale. The sum of these incremental improvements is not fully appreciated by Google's competitors, analysts or users:

- Google can put more computing horsepower under its hood more cheaply than its competitors. The cost for 80 Google servers, memory and storage is at least one-fifth less than of comparable resources from IBM, Hewlett-Packard or Sun Microsystems. Even Dell Computers cost about 30 percent more than Google's commodity approach.
- Google has streamlined certain system administration tasks. The use of snap in servers, cables on the front of server racks, and automatic fail-over allows Google to cut costs. At the same time, Google's approach to data centre set up allows Google to get new resources online faster, with fewer personnel, and with less configuration time than hosting companies using traditional administrative processes. Google is more nimble than most of its rivals.
- Google's microcode engineering of Linux solves many of the bottleneck problems that plague various flavours of Unix, Solaris and Windows Advanced Server. The Googleplex performs certain types of calculations faster, delivers data transfers that may be order of magnitude faster than standard operating systems' performance, and scales quickly.

Other points of differentiation emerge from the discussion of specific Google services and technologies elsewhere in this monograph.

The net of these advantages is that Google does not have a search system; Google has a supercomputer that delivers applications. Some of these applications are free for the user; for example, search. Other applications are for Google's 4,000 employees; for example, the programmers who craft applications for the Googleplex and employees who use the formidable number-crunching capabilities of the Googleplex to figure out what users are

doing, how to maximize advertising revenue from billions of online clicks in real time, and improve the search experience.

Bottomline: Google is poised to be the next Microsoft. (pps. 16-17; footnotes omitted)

* * *

Virtual applications are the future of computing. A *virtual application* is, stripping away the techno-jargon, software that allows the user to perform a task on any device as long as an Internet connection is available. Some of the virtual application's code may reside on the user's computing device, which may be a BlackBerry, hand-held computer, or public Internet café PC. For Gmail, Google's ad-supported e-mail service, the software and the data reside "somewhere on the Internet." For Google addicts, Google eliminates the hassle of software installation and alleviates worries about losing e-mail to laptop crash. (pps. 18-19; footnote omitted)

* * *

A major drawback for organizations is that virtual applications may have a copy of data and fine-grained data about what users did what on the system. Organizations want to keep these data in their sandbox, not in a third-party provider's sandbox "somewhere out there on the Internet." (p. 19)

* * *

Google is a virtual application company. (p. 19)

* * *

Google's integrated Googleplex gives it an advantage selling advertisements to advertisers who want data about usage patterns. Yahoo has a team in India struggling to make Yahoo's data repositories talk to one another. Google just asks, and the Googleplex spits out data at the same mind-boggling speed Google Search works in dozens of languages. (p. 21)

* * *

. . . Google can introduce applications that can challenge Microsoft Office or weave together desktop search, the Google Appliance and the Googleplex to create a platform on which to build enterprise applications. (p. 30)

* * *

Google is the network computer for an increasing number of users each day. Google's operating system is the platform for a new generation of virtual applications. Google's commodity devices are Oracle's nCube supercomputer at an affordable price: free, because of advertiser support. . . .

A young programmer in Beijing or Bangkok is now influenced by Google. If that programmer wakes up one day and Google has disappeared, for what system will the programmer develop?

One possible answer is that these young programmers would first build a better Googleplex and then craft virtual applications for it. (p. 33)

CHAPTER THREE: GOOGLE TECHNOLOGY

. . . Users do not give much thought to what technology underpins a routine query or the 300 million queries Google handles each day. In a single second, Google's technology handles around 340 queries in dozens of languages from users worldwide. (p. 59)

* *

The software requires a suitable hardware and network infrastructure in which to operate. Without Google's hardware and software, there would be no Google. Hardware and software are inextricably linked at Google. With each new advance in software, Google's engineers must make correspondingly significant advances in hardware. And when hardware engineers come up with an advance, the software engineers greedily use that advance to up the functionality of their software. (p. 59)

* * *

. . . The synergy between software and hardware is perhaps one of Google's major accomplishments. (p. 59)

* * *

At Google, from its inception, Google software and Google hardware have been tightly coupled. Google is not a software company nor is it a hardware company. Google is, like IBM, a company that owes its existence to both hardware and software. Unlike IBM, Google has a business model that is advertiser supported. Technically, Google is conceptually closer to IBM (at one time a hardware and software company) than it is to Microsoft (primarily a software company) or Yahoo! (an integrator of multiple software). (p. 60)

* * *

. . . If these 2002 data can be accepted, Google spends one-third for more computing horsepower and disc space than companies spend using a traditional server architecture. (p. 72)

* * *

The biggest boost to Google's technical approach is that its competitors are following different, more expensive approaches. Yahoo is a fruit cake of hardware, operating systems, and applications coded at different times in different languages by different people. Microsoft uses its own operating systems but relies on other operating systems as well, including Solaris. Microsoft's must invest in hardware to squeeze performance out of its platforms. Yahoo wrestles with its many different platforms. Microsoft seems powerless to enhance the speed of its operating system. Both are digital ostriches burying their heads in their own marketing material.
(p. 78)