The Plan

Colvin & Company

Greg Colvin

PO Box 1418

Buena Vista, Colorado

81211

•

719-221-2575

greg@colvin.org

www.colvinco.net

Thursday, February 21, 2008

I. Executive Summary

Search less – find more
Advertise less – sell more
Deploy less – serve more

We are a team of seasoned entrepreneurs and industry veterans.

Our mission is to provide world-class Internet information services.

We will accomplish our mission by building a community of Internet service providers – providers of open-source infrastructure, proprietary applications, and advertisement.

We will support our providers by building an open, globally-distributed network of virtual servers for global-scale Internet applications.

We envision an open Internet-services company with
• virtual servers online in data centers worldwide,
• supporting a state-of-the-art distributed-computation infrastructure,
• running our superior Internet-search engine and a variety of other applications,
• while generating revenues via automated ad placement and direct charges for use of our resources by
• a community of providers of Internet services.

We will leverage the immanent commoditization of distributed computation for our always-on worldwide infrastructure.

We will build on our proprietary information-matching technology to support precise and responsive Internet-search and ad-placement services.

We will convince a world full of Internet application providers to program to our interfaces and deploy on our infrastructure.

We will focus our attention initially on professional researchers, which we will find on university campuses. From there we will ensure that the good word is spread.

We will capture our share of the global Internet advertising market, which is now about $40 billion in revenues, and has doubled over the last four years.

Research is underway; development awaits funding.

II. Business Idea

Our mission is to provide world-class Internet information services, beginning with Internet search and advertising.

- To the Internet user we say: "Spend less time searching and more time finding what you want."

- To the Internet advertiser we say: "Put your product in front of the customers who want it."

- To the Internet application provider we say, "Leverage the resources of our community to better serve your customers."

We will leverage the growth of utility computing, which provides virtualized, metered access to distributed computation and storage services, built on commodity hardware, cheap bandwidth, and open-source software.

We will provide the foundations for profitable information services:

- Proprietary Internet-search and automatic ad-placement services based on high-precision information matching that goes beyond matching words to matching concepts.

- An ongoing analysis of the content and structure of the Internet, capturing the relevance of links and language to the subjects they are about; which analysis supports our proprietary information-matching technology.

- Support for the community of infrastructure and application providers, advertisers, and users who will make our systems profitable.

We will make it easy for other Internet software providers to build on these foundations to deploy their applications – such as maps, phone books, news feeds, social networking, video sharing, online education, and more – in return for a share of the advertising revenues these services earn.

We will make it easy for advertisers to get their copy onto the right pages.

The Company is a startup seeking initial funding.

III. Industry Analysis

In the NAICS category of "Web Search Portals" and in Yahoo Financials' similar index of "Internet Information Providers," we can see some of our direct competition. As of June 2007 the Yahoo index had a market capitalization of $211 billion, of which well over $210 billion is represented below:

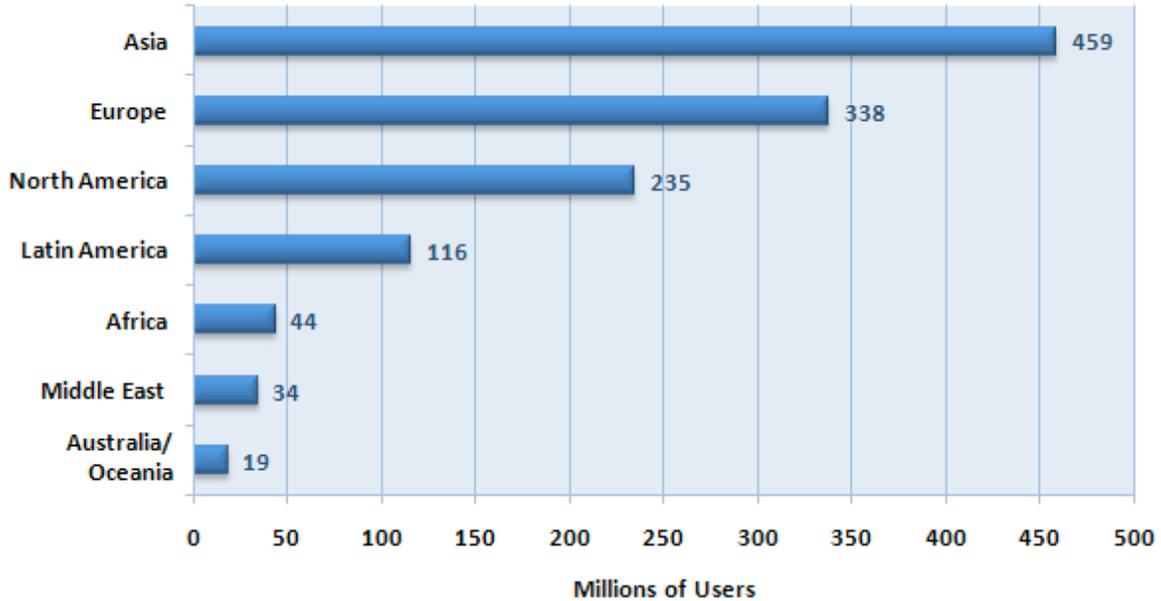| Top 13	Market Cap.	Revenue	Profit	Pages/Day
Google Inc.	150.64B	12.02B	27.35%	126M
Yahoo! Inc.	38.41B	6.53B	8.52%	48M
Expedia Inc.	7.41B	2.29B	6.32%	
Baidu.com, Inc.	4.41B	127.74M	31.02%	
Netease.com Inc.	2.18B	285.92M	56.53%	
CNET Networks Inc.	1.39B	395.82M	-9.90%	
Sohu.com Inc.	931.84M	136.91M	13.50%	
Global Sources Ltd.	798.91M	161.40M	18.68%	
Bankrate Inc.	792.41M	82.09M	24.17%	
InfoSpace Inc.	679.30M	368.11M	-0.62%	
Move, Inc.	649.36M	292.43M	1.96%	
The Knot, Inc.	582.75M	78.96M	7.77%	
Rediff.com India Ltd.	520.42M	28.68M	23.59%	

This sector is profitable, with the index averaging 19.9% profits. Although most of the profits remain in the US for now, five of the above companies are located in Asia and serve the Asian market. If recent growth trends continue for another seven years, Asia alone will have a billion Internet users, compared to 270 million in North America. This industry is, and will remain, a worldwide industry, growing in the service of growing worldwide demand.

Note the near-monopoly position of Google, at $151 billion accounting for 71% of the market capitalization of the index. Note also that Microsoft, although not listed in this index, has about 12% of the Internet search market, compared to Yahoo's 20% and Google's 54%.

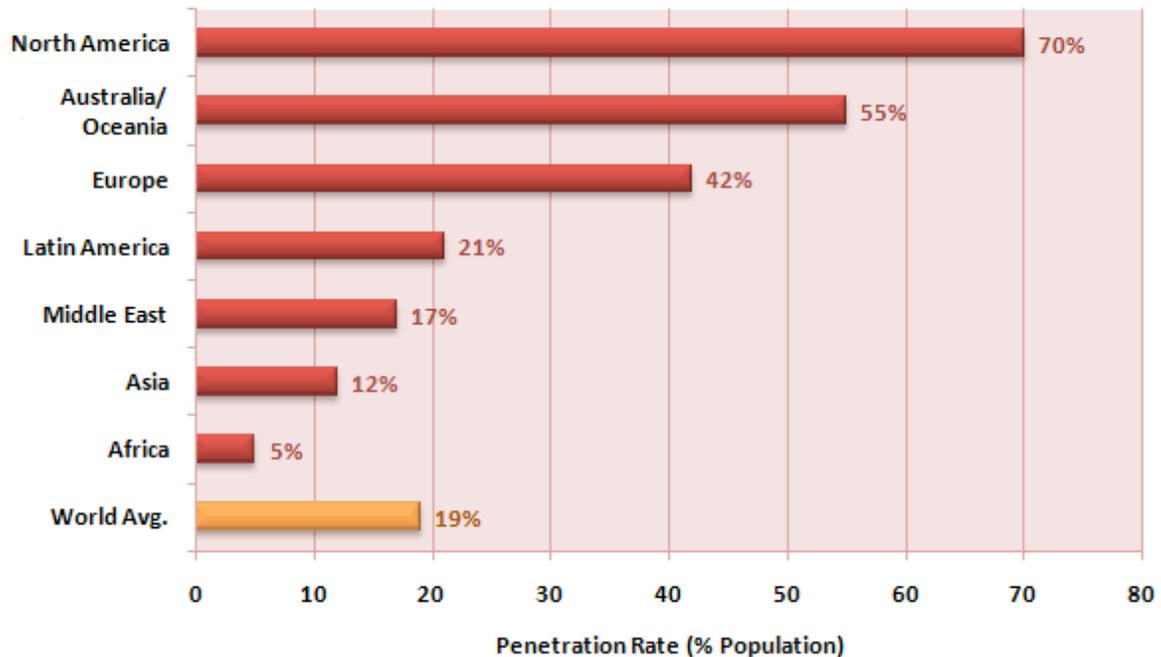## IV. Target Market

The possible consumers for our services are the soon-to-be seven billion people on the planet, of which so far over a billion are online <internetworldstats.com>. That's 17% of the planet online, 244% as many as just seven years ago. At that rate it will be just fourteen more years before the whole planet is as connected as North America is.

Internet Usage by World Region

Region	Millions of Users
Asia	459
Europe	338
North America	235
Latin America	116
Africa	44
Middle East	34
Australia/Oceania	19

Millions of Users

Internet Penetration by World Region

Region	Penetration Rate (% Population)
North America	70%
Australia/Oceania	55%
Europe	42%
Latin America	21%
Middle East	17%
Asia	12%
Africa	5%
World Avg.	19%

Penetration Rate (% Population)

Ultimately, the market for free information services is a market for advertising. The entire United States advertising market is about $300 billion in revenues, and the world market is about $600 billion. We are targeting the worldwide Internet advertising market, which saw about $40 billion in 2007 revenues, and has been doubling every few years. That revenue was obtained by placing ads on a trillion Web pages, displayed to a billion Internet users.

The Internet-advertising business has two sides. First, ads can be placed on the Company's own Web pages, especially on search-results pages. Second, ads can be placed on outside Web pages. Advertisers pay on a per-click or other basis, at prices set by auction, and in the case of outside placements the payments are shared with the outside owner. The key to success is to match the ads well to their readers, based on what queries they make and what text they are reading.

The synergy between advertising and Internet search is high. Most Web sessions begin with a search page, and when the right ads are placed on search-results pages they are likely to be just what the user wanted. So we can get far more ad views and clicks from our search pages than from externally placed ads, and we don't have to share the revenue from ads on our own pages.

V. Competitive Environment

Competitors in this market are in the business of turning commodity bandwidth, computation, and storage into Internet services, most visibly and profitably Internet search and automatic ad-placement services. Search is itself a commodity – you type in some words and you get back a list of presumably-relevant pages, along with some ads to click. So the power of suppliers is low, as other suppliers and downward integration by their customers threaten. The bargaining power of buyers remains high, as there are plenty of search portals competing for their business, and the cost of switching between these very similar products is low. There are plenty of new entrants like us joining the fray, and there are better technologies like ours waiting in the wings. And for those in the ring, the rivalry is fierce, with Alta Vista, Yahoo, and Google each having risen to prominence so far.

At least three competitive strategies make sense for us. None are easy:
1. Segment the market, so as to target chosen segments and attract loyal users.
2. Differentiate our products.
3. Provide the highest possible quality for the lowest possible price, at the least possible cost. Simply put, be the best.

Our proprietary technologies will provide some differentiation, but to compete as a new entrant in this industry we seek first to become and profitably remain the best Internet information service for at least one market segment, namely, professional researchers.

A barrier for us is the extensive statistical and linguistic analysis that we need to support our information-matching technology. But this same work becomes a barrier to our competitors, who will need to create better technology to compete with ours, or else buy ours to remain competitive.

Another barrier is the need to attract and nurture a community of providers of infrastructure, applications, and advertisement. We believe our years in industry, standards, academe, and open source give us the contacts and compassion we will need to finesse that barrier.

VI. Marketing Strategy

Search less — find more
Advertise less — sell more
Deploy less — serve more

We will begin with the submarket of professional researchers. In this we are playing to our strengths, as to date our information-matching technology has proven most useful to people doing in-depth research in a large set of documents, where finding all and only the relevant documents is the challenge.

Researchers, and the colleges and universities they often frequent, are a small market by Internet standards – with just four million PhDs, a million scientists, a million professors, and fifteen million college students counted by the US Census in 2000. But this is also a demographically attractive and influential market, with many discerning adopters, users, promoters, researchers and creators of search technology. From there, we can work to enchant people in other markets.

Much of our effort will be "stealth" or "viral" marketing – rather than buying advertising we will be cultivating allies, publishing papers and essays, and otherwise spreading the good word via the academic and trade literature, news media, and the Web. In particular, the open source projects and open services that we sponsor will bring us good words and goodwill, as will our early adopters in the research community. We will also do targeted advertising and direct sales, especially among potential advertisers.

VII. Operations

Internet services are worldwide. We will have people working from wherever they need to be, rather than gather everyone together. This saves money, and makes it easier to hire the best. We will maintain offices wherever we need to.

Internet services are hardware-intensive. We will need lots of servers with lots of disk storage and Internet bandwidth, distributed, eventually, all over the world. The 30 billion accessible Web pages total 300 terabytes, comprising about 100 terabytes of indexable text. Those numbers are doubling every four years. Initially we will buy a few servers so that we have them close by for development. Later we will rent virtual servers from Amazon and its competitors.

It is at this end of the business that we leverage commodity computing resources, as competition under Moore's Imperative keeps the cost of computing going down. Just as imperatively, the bandwidth of the Internet and the number of users continue to rise, creating more demand for services, thus using up all those newly affordable computing resources.

VIII. Information Technology

The Company's operations and assets are predominantly in the information technology domain. Among the Company's IT assets will be open-source software such as Linux operating systems, Gnu cluster management, Hadoop distributed-filesystem and distributed-computation software, Apache HTTP servers, and more. These are shared assets, yet vital nonetheless. We will sponsor and contribute to the open-source projects we are building on. Where necessary we will start open-source projects to create the software we need.

Also among the Company's IT assets will be proprietary software, including our search-engine and ad-placement services. We will build these services on our information-matching technology, which is based on judgment simulation – a psychometric technique invented in the 60's by our late colleague Peter Ossorio as a way to give computers the ability to make judgments like those made by people. Its first application was to simulate a library and librarian by giving a computer the ability to scan text and judge what it is about. Other applications have included search engines for NASA databases, robotics for a Mars rover, and the hierarchical classification of text for litigation support. When applied to information retrieval this approach reliably finds all and only the most relevant information, selecting and ranking documents not by mere word matching and popularity, but by their conceptual relevance to the query.

Developing and maintaining our information-matching technology will require an ongoing cycle of using standard and proprietary statistical tools to analyze well-organized portions of the Internet for linguistic patterns that can serve to organize and index the rest of the Internet. As part of this cycle, we also aggregate data from user behavior to help to identify and correct any weaknesses in the search results and ad placement. Continual improvement is the goal.

IX. Management & Exit Plan

The founders are a small group of seasoned entrepreneurs who have worked together off and on for decades. We are mostly veteran engineers who have what it takes to architect and implement Internet services. Our professional culture is that of Internet software development, with a premium on constant communication and relentless analysis in pursuit of the best available solutions, and a steady attitude of service towards the customers for and users of those solutions.

We intend to incorporate, and to operate our business as a partnership of principles. We plan to allocate equity so as to ensure that all are motivated by the Company's best interest, and provide fair salaries, benefits, and working conditions that make it possible for our people to work as hard as they can in those interests.

Our exit plan has five big milestones:
1. Financing, which will kick off development.
2. Beta testing, during which we first put our services online.
3. Public release of our online services.
4. Profitability, the prerequisite to a successful exit.
5. Exit via initial public offering or acquisition.

Exact timing will depend on market conditions.

X. Critical Risks and Contingencies

The Company can guard against recession and other negative economic and market conditions by raising sufficient funds and spending them wisely – being prepared to wait out bad times on the path to profitability and public value. Our strategy of focusing first on submarkets reduces the risk of not finding a market. We patent our inventions and maintain our trade secrets. Finally, we may be presented with friendly or hostile buyout offers ahead of plan. If they are good offers we should take them.

XI. Financials

This analysis is based on ten assumptions we think reasonable, in light of the industry leader's numbers. Analysis of alternatives indicates that weaker assumptions are untenable. So this analysis assumes that we will:

1. Develop and beta test for at least 2 years before booking revenue.
2. Serve 250,000 Web pages per day the first year after beta tests – that's 0.001% of the world market, or 0.625% of the university submarket, and half of Google's rate at that stage.
3. Grow the number of users by 20% a year – the growth rate of Internet users is 16%, and Google grew 50% in 2007.
4. Make $0.10 of revenue per page served – Google makes double that.
5. Spend 10% of revenue on advertising – Google spends 6%.
6. Spend 10% of revenue on marketing, including publishing, patents, and support for open source projects.
7. Spend $200,000 per employee per year – salary, benefits and expenses.
8. Spend $4,000 per server per year – e.g. 2 1.6-Ghz CPUs, 4 GB RAM, 0.5 TB disk, and 40 GB/mo Internet costs $266+/month at lpdedicated.com.
9. Serve 9,000 pages per day per employee – Google does the same, and served 500,000 a day with just 9 people in 1999.
10. Serve 2,000 pages per day per server – Google uses half as many servers, but our algorithms are more intensive. And before that, we need 200 servers, and growing, to index the growing Web.

The most critical assumptions are pages-per-day and revenue-per-page, which produce revenue, and pages-per-day-per-employee, which dominates costs. Therefore we estimate pages-per-day and revenue-per-page at only half of the industry leader's numbers.

	Year 1	Year 2	Year 3	Year 4	Year 5	Year 6
Revenue			$9,125,000	$10,950,000	$13,140,000	$15,768,000
Expenses	$616,000	$1,332,000	$7,500,000	$8,960,000	$10,472,000	$12,446,400
Profit (Loss)	($616,000)	($1,332,000)	$1,625,000	$1,990,000	$2,668,000	$3,321,600
Profit Margin			21.67%	22.21%	25.48%	26.69%
Employees	3	6	31	37	43	51
@ cost	$600,000	$1,200,000	$6,200,000	$7,400,000	$8,600,000	$10,200,000
Servers	4	33	325	390	468	562
@ cost	$16,000	$132,000	$1,300,000	$1,560,000	$1,872,000	$2,246,400
Ads & Marketing			$1,825,000	$2,190,000	$2,628,000	$3,153,600

XII. Key People

Greg Colvin started out over a year ago to recruit people, research technology, prototype software, and find investors for this venture, which will be his fourth startup. He has been working towards this vision, directly and indirectly, for over thirty years.

Greg has mastered many kinds of systems programming – simulations, databases, windowing systems, device drivers, virtual machines… – always in the service of better access to information and better tools for programming, on whatever platforms required. He can architect systems that don't collapse of their own complexity and empower teams of people to build those systems. Greg is a contributor to the ISO C++ Standard, a founding member of the Boost open-source C++ library project, and a member of the Advisory Board for the C++ Source online journal. Highlights of Greg's three-decade career include both technical and management positions:

• Oracle Corporation, Principal Member of Technical Staff
• Information Management Research, Senior Scientist
• Information Handling Services, Manager of Systems Development
• Reference Technology, Member of Technical Staff
• Information Access Systems, Vice President for Systems Development
• and short-term projects from embedded databases and multitasking kernels to voting-machine security

• Cornell University, PhD in Personality and Mathematical Psychology
• University of Colorado, Boulder, BA in Psychology and Computer Science

Some of Greg's presentations and publications include:

• Descriptive Metaphysics: On Science, Religion, and Wisdom. Society for Descriptive Psychology Conference, Estes Park, Colorado, October 2003
• In the Spirit of C. ACCU Conference Keynote, Oxford, England, April 2003
• Smart Pointers for C++ Garbage Collection. C/C++ Users Journal, 13:12, December 1995.
• Object Oriented Programming in C and C++. C Users Journal, 11:7, July 1993
• Exception Handling in ANSI C. C Users Journal, 9:8. August 1991
• Multitasking with Lightweight Threads. C Users Journal, 8:3. March 1990
• Synapsys: A Neural Network. C Users Journal, 7:3. April 1989
• Database indexing and retrieval. CD-ROM: Optical Publishing, Microsoft Press, 1987
• The current state of text retrieval. CD-ROM: The New Papyrus, Microsoft Press, 1986

Paul Zeiger is advising this venture, which will be his third startup. He plans to serve as chairman of our board of directors as part of an active retirement that includes consulting for H5 Technologies and teaching advanced Anasura yoga. Paul is a natural diplomat, and long-practiced at working across cultures, professional and personal.

Paul's four-decade career includes technical and management positions in both academia and industry, with up to 102 reports and a $9 million budget:

- US West Communications, Data Architect
- US West Communications, Director
- US West Communications, Member of Technical Staff
- Management Support Technology, Vice President for Product Development
- University of Colorado, Boulder, Chairman of Department of Computer Science
- University of British Columbia, Assistant Professor of Electrical Engineering

- Massachusetts Institute of Technology, BS, MS, and PhD in Electrical Engineering

Some of Paul's publications include:

- Toward a Rapprochement of Religion and Science. Advances in Descriptive Psychology Vol. 8, Descriptive Psychology Press, Ann Arbor, 2006
- Human Systems Issues in Software Engineering. Advances in Descriptive Psychology Vol. 5, Descriptive Psychology Press, Ann Arbor, 1990
- LDS/UCC: Intelligent control of the loan documentation process (with H. Joel Jeffrey, T. Schmidt, A.O. Putnam.). IEA/AIE, 2, 1989
- Very Special Languages and Representations of Recursively Enumerable Languages via Computation Histories (with David Haussler). Information and Control, 47(3), 1980
- Cascade Synthesis of Finite-State Machines. Information and Control, 10(4), 1967

Joe Jeffrey is also advising this venture, which will be his third startup. He will be serving with Paul Zeiger on our board of directors as he continues his teaching and research at NIU. Joe invented, and continues to invent, much of our information-matching technology. He combines uncompromising intellectual rigor with an utterly pragmatic approach to solving problems.

Highlights of Joe's three-decade career include:

- Northern Illinois University, Professor
- H5 Technologies, Founder and Chief Scientist
- Management Support Technology, Member of Technical Staff
- Bell Telephone Laboratories, Member of Technical Staff
- Vanderbilt University, Assistant Professor

- University of Colorado, Boulder, PhD in Computer Science
- California Institute of Technology, BS in Mathematics

Some of Joe's presentations and publications include:

- Adventures in San Francisco – Commercializing a Scientific Revolution. The Software Practitioner, May-June 2003
- Scalable Judgment Spaces. Presidential Address to the Society for Descriptive Psychology, Estes Park, Colorado, October 2002
- Wide Spectrum Information Search Engine. U.S. Patent 6,493,711, 2002.
- Managing Professionals: The Joy of Herding Cats. Society for Descriptive Psychology, Estes Park, Colorado, 1997
- Addressing the Essential Difficulties of Software Engineering, Journal of Systems and Software, 32, 1996
- A Logical Foundation for a Science of Consciousness: A Parametric Formulation. Conference on the Science of Consciousness, Tucson, Arizona, April 1996
- Relationship Definition and Management: Tools for Requirements Analysis (with A.O. Putman), Journal of Systems and Software, 24, 1994
- Judgment-Simulation Vector Spaces, Ch. 13, Advances in Computer Methods for Systematic Biology: Artificial Intelligence, Database, ComputerVision, Johns Hopkins University Press, 1993
- Chaos game visualization of sequences. Computers & Graphics 16(1), 1992
- Expert Document Retrieval Via Semantic Measurement, Expert Systems With Applications, 2(4), 1991
- Human systems analysis in the software engineering curriculum. Journal of Systems and Software 14(3), 1991
- LDS/UCC: Intelligent control of the loan documentation process (with H. Paul Zeiger, T. Schmidt and A.O. Putman). IEA/AIE, 2, 1989.
- A new type of information retrieval system. Proceedings of the 14th ACM Southeast Regional Conference. Birmingham, Alabama, 1976.

Dave Abrahams is a master of C++ library development. He and his partners at Boost Consulting provide support, training, and development services to C++ projects. High-performance computing, distributed programming, and natural language and text processing are among their specialties. Dave is a contributor to the ISO C++ Standard, a founding member of the Boost project, and a member of the Advisory Board for the C++ Source online journal.

Dave's two-decade career includes:

- Boost Consulting, Principal and Founder
- C/C++ User's Journal, Editorial Board
- Altra Broadband, Inc., Software Architect
- Dragon Systems, Inc., Senior Engineer
- Mark of the Unicorn, Inc., Project Leader, Engineer, Researcher

- Berklee College of Music, Jazz Composition program
- Carnegie Mellon University, Computer Science PhD program
- University of Pennsylvania, BSE in Computer Science Engineering, Cum Laude

Dave's presentations and publications include:

- The Boost C++ Metaprogramming Library (with Aleksey Gurtovoy). ACCU Conference, Oxford, England 2003
- Error And Exception Handling. Boost.org, 2003
- Effects of Metaprogramming Style on Compilation Time (with Carlos Pinto Coelho). Boost.org, 2001
- Policy Adaptors and the Boost Iterator Adaptor Library (with Jeremy Siek). Accepted into the C++ Template Workshop at OOPSLA 2001
- Boost Coding Guidelines (with Nathan Meyers). Boost.org, 2001
- Generic Programming Techniques (with Jeremy Siek). Boost.org, 2001
- Exception Safety In Generic Components. Generic Programming, Proc. of a Dagstuhl Seminar, Lecture Notes on Computer Science 1766, for the Dagstuhl Conference on Generic Programming, Wadern, Germany, 1998

Mats Henricson is a master of Web-application programming, both client-side and server-side. He and his partners at Crisp AB in Stockholm, Sweden specialize in rapid application development, and will be key players in creating and supporting the applications that will drive our revenue. This will be Mats' fifth startup. Mats is a contributor to the ISO C++ Standard and the Java Community Process.

Highlights of Mat's two-decade career include:
* Crisp AB, Partner, Software Engineering Consultant
* Valtech AB, Software Engineering Consultant
* Avega AB, Software Engineering Consultant
* WebPutty Inc., Principal Software Engineer
* Enactex Inc., Senior Software Engineer
* Talaris Corporation, Senior Software Engineer
* Critical Path, Senior Software Engineer
* Henricson Technology AB, Software Engineering Consultant
* Contactor AB, Software Engineering Consultant
* Intact Media and Communications, Senior Software Engineer
* Swedish Postal Service, Project Leader
* Ellemtel Telecom Systems Labs, Team Manager

* Uppsala University, MS in Engineering Physics and Scientific Computing

Mats' publications include:

* State of the Art in Server Side Java, The Server Side, July 2005, http://www.theserverside.com/news/thread.tss?thread_id=35074
* Industrial Strength C++: Rules and Recommendations (with Erik Nyquist), Prentice Hall, 1996

João Abecasis is one of the developers of the software behind Sapo, the largest online advertising provider in Portugal. He is also one of the maintainers of the Portuguese distribution of Linux. He has been pursuing a PhD in Theoretical Biochemistry at the University of Lisbon, and is moving to the United States to accept a research internship at the Center for Computation & Technology, Louisiana State University. João is a contributor to Boost and other open-source projects, especially the Spirit C++ parsing library, and is a student of genetic algorithms, dynamical systems, and distributed computation.

João's still-short career includes:

- Associação para o Desenvolvimento das Telecomunicações e Técnicas de Informática, Software Developer
- Farmácia Valadas, Suc., Software Developer

- University of Lisbon, Licenciatura in Biochemistry