The problem is the quality of Internet search, which is much lower than the state of the art in the information retrieval literature. The barrier to industrial adoption of this research has been the high expense in human expertise and computational power which the best techniques require. We intend to achieve scalability by computing the singular value decomposition (SVD) in near-linear time and space. SVD can be used with appropriate data to create vector spaces whose operations simulate the human judgments that create the data. These simulations can also be used to perform information indexing by mapping data into the judgment space (j-space), and to perform information retrieval by mapping queries into the j-space and retrieving data vectors in order of similarity to the query vector. The conceptual information matching performed by J-space systems, made newly scalable, when combined with existing, already-scalable techniques like web crawling, keyword matching and PageRank, promises to give state-of-the-art performance and the ability to scale with the growth of the Internet.

Competitive technologies either cannot scale with the Internet or cannot achieve the best possible quality of results. Decades of research and industrial experience demonstrate a quality/quantity tradeoff that can be overcome only with expensive human expertise or computational costs that grow quadraticly with the size of the Internet. We plan to create an indexing process that scales, and apply this process to create an index of the Internet that provides both textual pattern matching and conceptual information matching, thus overcoming the traditional 50/50 trade-off between precision and recall. This index will provide an intrinsically valuable service. We plan to validate this index empirically and to publish and publicize the results, thus contributing to the advance of information science and enhancing the credibility of our technology. We also plan to create an open-source community to maintain evolve our technology, and thereby to find allies to support our ongoing mutual success. An Internet search service can be used directly to generate advertising revenue from online users. A successful Internet service company can also be sold to its competitors.

TABLE OF CONTENTS

For font size and page formatting specifications, see GPG section II.B.2.

	Total No. of Pages	Page No.* (Optional)*
Cover Sheet for Proposal to the National Science Foundation		
Project Summary (not to exceed 1 page)	1	
Table of Contents	1	
Project Description (Including Results from Prior NSF Support) (not to exceed 15 pages) **(Exceed only if allowed by a specific program announcement/solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)**	9	
References Cited	3	
Biographical Sketches (Not to exceed 2 pages each)	8	
Budget (Plus up to 3 pages of budget justification)	3	
Current and Pending Support	2	
Facilities, Equipment and Other Resources	1	
Special Information/Supplementary Documentation	0	
Appendix (List below.) **(Include only if allowed by a specific program announcement/ solicitation or if approved in advance by the appropriate NSF Assistant Director or designee)**		
Appendix Items:		

*Proposers may select any numbering mechanism for the proposal. The entire proposal however, must be paginated. Complete both columns only if the proposal is numbered consecutively.

Internet services

The overall goal of this project is to create a profitable enterprise in the Internet services industry. We envision a company that fosters an alliance of for-profit and non-profit service providers and software developers, utilizing our technology as a core component in their own products and services.

Outside of highly specialized services such as the Bloomberg terminal the Internet services market is a market for advertising, which since 2002 has been increasingly pay-per-click. Google dominates this market with it's Adsense program, which auctions off keywords to the highest bidder, places the winning binners' ads on to pages that match the keywords, and charges the bid amount for every ad clicked on. 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. We estimate that Google averages over .20 cents for every search result page it displays..

Most of the essential components of an Internet search service like Google are now well understood (e.g. Brin & Page 1998, Kobayashi & Takeda 2000, Ghemawat et al. 2003, Mehta et al. 2007, Dean & Ghemawat 2008, Chang et al. 2008) and are available as open source from Hadoop Lucene, Nutch and related projects (Kimbal et al. 2008). But these systems share a limitation long known in the information retrieval literature: there is a trade off between quality and quantity such that the sum of precision (the percentage of the results that are relevant information) and recall (the percentage of all the relevant information that is tin he results) very rarely exceeds 100%, and is typically much worse (Salton 1975). For a study of the major search engines in 2004 the result for the best query on the best engine was 122%, and the overall result for Google was only 49% (Shafi and Rather, 2005). In practical terms: for the typical query many results are not relevant and many relevant results are missing, and to do better the user is forced to guess just which keywords were used in the desired documents.

Our unique value-added is an information-matching technology called J-Space. This proposal focuses on text retrieval, where we have proven able to beat the quality/quantity tradeoff, achieving combined precision and recall exceeding 160%. We have also been exploring the use of J-Space to improve the effectiveness and profitability of add placement.

J-Space technology

Our technology is based mathematically on the now-common approach of representing the content of text as vectors in a high-dimensional information space, and retrieving information based on the similarity of the vectors (Salton 1968). In Salton's vector model the dimensions of the space represent individual keywords, so that texts are retrieved based on keyword matching. In a J-Space the dimensions represent concepts, so that texts are retrieved based on conceptual similarity.

Information spaces for text retrieval can be created using a psychometric technique known as Classification Space Analysis or Judgment Space Analysis (Ossorio 1964, 1966). This technique is founded on a social-psychological analysis of the relation of subject matters, communities, and language. Put briefly, a subject matter is the world of a community, and communities use distinctive locutions to that correspond to the conceptual distinctions that they need to make in that world (Putman 1990). For example, for the subject matters of Astronomy or Baseball we have their corresponding worlds and communities, with concepts like "planetary system" and "solar mass" for Astronomy and "base hit" and "home run" for Baseball. Ossorio demonstrated that the locutions used by different communities are sufficiently distinctive that the subject matter of a text can most often be distinguished simply by observing these locutions.

Judgment Space Analysis involves identifying the concepts to be indexed, identifying the terms (words or phrases) that correspond to the concepts, collecting expert judgments of the relevance of each term to each concept, and

using factor analysis of the quantified judgments to build a vector space, known as a Judgment Space or J-Space, in which each concept is represented as an orthogonal dimension and each term is represented as a vector of loadings on the dimensions.

A document or query texts can be placed in a J-Space by combining the vectors of its constituent terms, and the relevance of a document to a query can then be estimated as vector similarity, where vector similarity can be indicated by such functions as Euclidean distance, cosine, or dot product. Ossorio called this operation Judgment Simulation, because it simulates the human behavior of making relevance judgments. An effective retrieval service can be implemented by simply returning an ordered list of the documents most similar to a query.

Over the years Ossorio and his colleagues (Jeffrey 1976, 1991, 1993, Ossorio & Kurtz 1989, Kurtz 1992, Kurtz 1993) have used judgment simulation to build several successful, and sometimes profitable, information-retrieval systems that have typically exceed 80% precision and 80% recall In the 1960s results this good were simply unprecedented, and purely statistical approaches like Latent Semantic Indexing (Deerwester et al. 1990) are only recently (Hoffman 1999) getting past that tradeoff.

Scaling Up

The problem with applying judgment space analysis at large scale has always been the time and expense of obtaining the requisite expert judgments. Jeffrey (2002) invented a means of drastically reducing the work by organizing the subject matters hierarchy, along the lines of traditional bibliographic systems like Dewey and Library of Congress. We estimate that an adequate set of judgments to replicate the scope and precision of these systems would cost circa $5 million. We also estimate that the scope of the Internet is at least an order of magnitude greater.

Yang & Chute (1992, 1994) have published an economical alternative, which appears to be an independent reinvention of the J-Space concept and a novel approach to building the space. Rather than work with direct judgments of the relevance of terms to texts, they begin with a corpus of texts which have been categorized by human experts. The problem then becomes simulating the term relevance judgments so as to reproduce the observed text categorization judgments, for which they derive a least-squares fit to a linear model that can be solved via singular value decomposition (SVD). In tests of this technique on a sample of categorized patient records they achieved simultaneous 84% precision and and 84% recall.

How can we obtain a sufficient sample of text judgments to build Internet-scale J-Spaces? By mining the Internet itself, which contains vast quantities of text and forms an almost incomprehensibly complex network of links, many of them representing direct human judgments.

In our current phase of research we are focusing on the Wikipedia (2009) category system, in which over 4 million articles are assigned to one or more of over 400 thousand categories. By observing the distribution of the words and phrases in Wikipedia text across the Wikipedia categories we can get a position for each word and phrase in the category space. The statistical distribution of words and phrases is at best an indirect indication of their relevance to the various categories, but work by Kurtz (2007, private communication) indicates that such statistics can be effective proxies for direct relevance judgments, provided the categories are well chosen and the text is well categorized. Unfortunately, Wikipedia does not always meet those criteria.

We have recently built three prototype J-Spaces (Colvin, 2009) for three subsets of the Wikipedia: 4,982-Algebra articles , 89,031 Mathematics and Computer Science articles and 58,628 Psychology, Psychiatry, and Neurology articles. The most successful space is Algebra, which was small enough for careful hand-pruning of the included categories and articles. The larger, less-well-pruned

subsets don't do nearly so well. Statistically speaking, the Wikipedia is noisy data.

Fortunately,Yang (1995) shows that a truncated SVD approximation to her linear model can serve to filter the noise from categorization data. It appears that the important concepts are well represented by the largest singular values, and the noise of mistaken or idiosyncratic judgments is discarded with the smaller ones. An additional benefit of SVD is careful selectio and rotation of the singular vectors ab interpretable, visualizable structure can be found.

Until recently the best algorithms for SVD had polynomial complexity, such that matrice. But recent work by Achlioptas & Mcsherry(2007) and Frieze et. al (2004) show that it should be possible to approximate the SVD in near-linear time, rather than the impractically polynomial time of previous methods.

Project plan

The proposed project will utilize industry best practices for rapidly-iterated, open-source research and development.

Project information and computational resources are securely online, all electronic communication is securely archived, project workflow is automated with secure online systems, and key decisions are discussed and documented via email among appropriate stakeholders. Source code, documentation, and test cases will be maintained on a public version-control server, which will track all changes made to the project. The system software will be automatically built and tested daily, and tested labels published for release.

Wikis and Web pages will provide online documentation, blogs will support ongoing public announcements and discussion of the project, and archived email lists will be used for discussions among project contributors. Working prototypes of the technology will be kept running on our Web site, with mechanisms in place to monitor and log performance. The intent is to build

relations with potential allies from the start, and to leverage their efforts into continual evolutionary improvement of our technology.

To keep the evolution on track we must continually track the quality of our simulations. The established measures of the quality of information retrieval are precision and recall. But reliable estimates of precision and recall for a database as large as Wikipedia will require large samples and expensive expertise to obtain, so the proposed project will not yet attempt a formal evaluation of these metrics. Instead we will use cross validation as our primary indicator of progress, combined with ongoing discussion and qualitative analysis of our growing collection of "interesting" queries and result sets. The cross validation will be accomplished by randomly splitting our Wikipedia extracts in half, building a J-Space for one half, using the space to categorize the text in the other half, and computing the correlation between simulated and actual judgments. Cross validations will be performed as part of testing, and significant reductions in the correlations will be reported as regressions.

Given the tools described above it becomes possible to iteratively modify the software while maintaining forward progress. Changes that cause test regressions should never be checked in, but will even then be automatically detected by the automatic testing, which will generate email to the offending engineer and flag the label as failed. Changes can be easily backed out, divergent paths of evolution can be merged back together, and no change checked in is ever lost.

The plan of development begins with these improvements to our current system:

1. implement a near-linear SVD algorithm,
2. use the algorithm to approximate solutions to Yang & Chute's linear model, and
3. use these solutions to create J-Spaces.

We expect to make progress on these tasks before the project starts, and to have them done by the end of the second month at latest.

Once we have testable J-Spaces we will be looking for improvements. Known possibilities for improvement include:

1. use more powerful statistics (e.g. reduced-rank decomposition of Term-by-Text-by-Category tensors, Kolda 2009) to capture more structural information the data,
2. extract better phrases, e.g. detection of lexicalized noun phrases (Godby 2002),
3. combine J-space similarity with Boolean keyword matching and SQL in the judgment simulation, and, longer term,
4. engage the Wikipedia community to improve the quality of their category hierarchy.

The first three tasks may also be underway by the time the project starts, and should not take more than month each. A useful start to the fourth task would be to devise tools that can automatically flag inconsistencies in the category hierarchy, thus making it possible to correct the many problems without inspecting the entire hierarchy.

At every step of our R&D process we will have a demonstrable prototype, and by six months it should be very stable and performing well.

Directions for Further Research

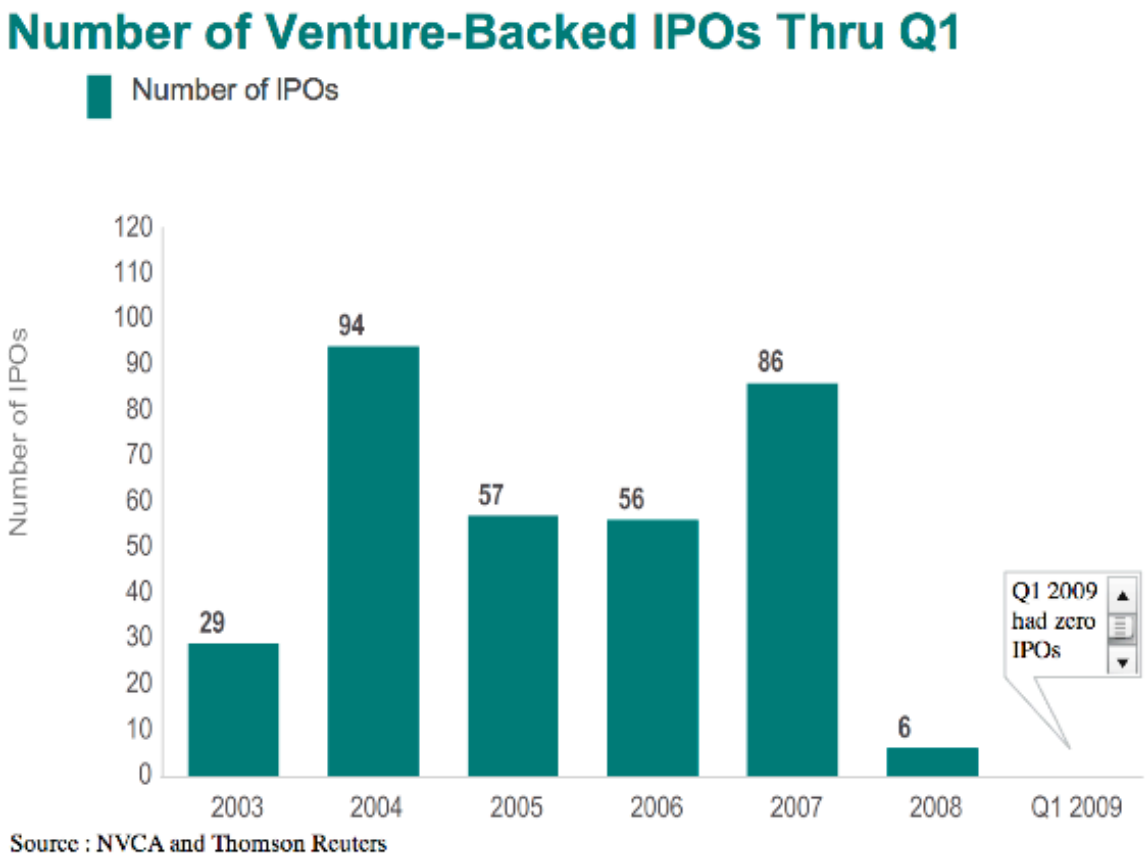
The success of this proposal would open many opportunities for ongoing research, including:

- Perform formal precision and recall assessments and publish the results.
- Use these assessments to identify weaknesses in our models that require conceptual or statistical improvements.
- Work with the Wikipedia community to improve and maintain their category system, and thus improving the quality of our system. This will be an application of Jeffrey's hierarchical judgment space model.
- Create tools for directly visualizing and navigating in information spaces.
- Start crawling the Web, working concentrically outward from Wikipedia and other important roots.

Commercialization

Our commercial goal is to create a business to profit from our technology by creating value for our customers and contributors.

In 2007 we drew up detailed plans for a well-financed venture that could provide a full range of internet services, but the market has changed substantially since then. The New York Times reported on April 1 that for there were no venture-backed initial public offerings in the first quarter of 2009, that the average number and size of acquisitions is down by half, and that more than half of those deals lose money. Google reports that their first quarter revenue is down 3% from the previous quarter. Yahoo reports that the average profitability of their Internet Information Providers index has slipped from circa 20% to circa 5%.



Given the market realities our current plans are to start slow, conserve cash, and stick to our core competencies. So over the next year our primary emphasis will be on our open-source project. This project promises to create a demonstrably excellent search engine for Wikipedia, which is a highly visible Web site. We plan to attract attention from the Wikipedia and broader "open" community, and publish demonstrations of our technology that attract attention and raise our credibility. And by operating in the open we plan to make many allies among our contributors and potential customers.

The next winning business model for making a profit from search technology has probably yet to be invented, but here are two reasonable scenarios that we can begin to execute now.

Scenario One. Crawl a lot of the Web. The entire Web would be best. The entire Wikipedia would a start. Everything reachable in six or seven degrees of separation from Wikipedia would be better. All of *.edu* and *.gov* might be useful. Find enough money to put the index online as a search service. Place ads on our results page, from the open market at the OpenX.org open-source project, from Doubleclick or AOL, or even from Google.

Scenario Two. Be the Firefox of search applications. Firefox is a powerful native executable for browsing the Web that competes credibly with Microsoft's dominant Internet Explorer. Firefox is maintained by the open-source Mozilla foundation and its for-profit subsidiary. One of its features is a search form that gets its results from Google, displays Google ads, and gets a share of the pay-for-click on those ads. We can build on the Firefox code to create a specialized browser that uses J-Space to formulate complex Boolean queries that tease all the relevant data out of Google, and we can display our results organizes by conceptual relevance. On this plan we don't need money to create and maintain an inverted index of the Web.

References

Achlioptas, D. & F. Mcsherry. Fast computation of low-rank matrix approximations. Journal of the ACM, v. 54:2, p. 611-617, 2007

Brin S., L. Page. The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems, v. 30:1-7, p. 107-117, 1998

Colvin, G. Wikos.org: An Experiment in Progress. http://wikos.org 2009

Chang F., J. Dean , S. Ghemawat, W. Hsieh, D. Wallach , M. Burrows, T. Chandra, A. Fikes & R. Gruber. Bigtable: A Distributed Storage System for Structured Data. ACM Transactions on Computer Systems, v. 26:2, p. 1-26, 2008

Deerwester, S., S. Dumais, T. Landauer, G. Furnas & L. Beck. Indexing by latent semantic analysis. Journal of the American Society for Information Science, v. 41, p. 36-40, 1990

Dean J. & S. Ghemawat. MapReduce: simplified data processing on large clusters. Communications of the ACM, v. 51:1, 2008

Frieze A., R. Kannan , S. Vempala. Fast monte-carlo algorithms for finding low-rank approximations. Journal of the ACM, v. 51:6, p. 1025-1041, 2004

Ghemawat S., H. Gobioff & S. Leung. The Google file system. Proceedings of the nineteenth ACM symposium on Operating systems principles, 2003

Giles, J. Nature, v. 438, p. 900-901, 2005

Godby, C. A Computational Study of Lexicalized Noun Phrases in English. Ohio State University, 2002

Jeffrey, J. A new type of information retrieval system. Proceedings of the 14th ACM Southeast Regional Conference, p. 271-277, 1976

Jeffrey, J. Expert Document Retrieval Via Semantic Measurement. Expert Systems With Applications, 2(4), 1991

Jeffrey, J. Judgment-Simulation Vector Spaces. Advances in Computer Methods for Systematic Biology: Artificial Intelligence, Database, ComputerVision. Johns Hopkins University Press, 1993

Jeffrey, J. Wide-spectrum information search engine. U.S. Patent 6,493,711, 2002

Kimball A., S. Michels-Slettvet & C. Bisciglia, Cluster computing for web-scale data processing, Proceedings of the 39th SIGCSE technical symposium on Computer science education, ACM, 2008

Kobayashi M., & K. Takeda. Information retrieval on the web. ACM Computing Surveys, v. 32:2, p.144-173, 2000

Kolda, T. & B. Bader. Tensor Decompositions and Applications. To appear in SIAM Review, v. 51:3, 2009

Kurtz, M. Advice from the Oracle: Really Intelligent Information Retrieval. Adding Intelligence to Information Retrieval: The Case of Astronomy and Related Space Sciences. Kluwer, 1992

Kurtz, M. Statistical Factor Spaces in the Astrophysics Data System. Intelligent Information Retrieval, Kluwer, 1993

Mehta A., A. Saberi , U. Vazirani & v. Vazirani. AdWords and generalized online matching, Journal of the ACM, v. 54:5, 2007

Ossorio, P. Classification Space Analysis. Defense Technical Information Center, Report AD0608034, 1964

Ossorio, P. Classification space: a multivariate procedure for automatic document indexing and retrieval. Multivariate Behavioral Research, v. 1, 1966

Ossorio, P. & M. Kurtz. Automated classification of resolved galaxies. Proceedings of the Third International Workshop on Data Analysis in Astronomy, Plenum Press, p. 121 , 1989

Putnam, A. Communities. Advances in Descriptive Psychology, v. 1, Descriptive Psychology Press, 1990

Salton, G. Automatic Information Organization and Retrieval. McGraw-Hill, 1968

Shafi S. & A. Rather. Precision and Recall of Five Search Engines for Retrieval of Scholarly Information in the Field of Biotechnology. Webology v. 2:2, 2005

Wikipedia. The English Wikipedia. Extracted by Freebase.com on 1/12/2009

Yang, Y. & C. Chute. A linear least squares fit mapping method for information retrieval from natural language texts. Proceedings of the 14th International Conference on Computational Linguistics, McGraw-Hill, pp. 447-453, 1992

Yang, Y. & C. Chute. An example-based mapping method for text categorization and retrieval. ACM Transactions on Information Systems, v. 12:3, pp. 252 - 277, 1994

Yang. Y. Noise reduction in a statistical approach to text categorization. Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval table of contents, pp. 256-263, 1995

Gregory Colvin

University of Colorado Boulder	Psychology	BA	1977
Cornell University	Psychology	PhD	1982

Metaweb Technologies	Software Engineer	2008-2009
	Entrepreneur	2006-2008
	Independent consultant	2006
Oracle Corporation	Principal Engineer	1998-2006
Information Management Research	Senior Scientist	1994-1998
XVT Software	Developer	1993-1994
Information Handling Services	Manager	1990-1993
	Independent consultant	1988-1990
Reference Technology	Technical Staff	1984-1988
Linguistic Research Institute	Architect	1984
Information Access Systems	Vice President	1983-1984
Cornell University	Instructor	1979-1982
University of Colorado, Boulder.	Research Assistant	1977-1978

Related publications:

- Colvin, G. Synapsys: A Neural Network. C Users Journal, v. 7:3, 1989
- Colvin, G. Database indexing and retrieval. CD-ROM: Optical Publishing. Microsoft Press, 1987
- Colvin, G. The current state of text retrieval. CD-ROM: The New Papyrus. Microsoft Press, 1986

Other publications:

- Colvin, G. How to Go Slow. The C++ Source, 2008 http://www.artima.com/cppsource/ how_to_go_slowP.html
- Colvin, G. Descriptive Metaphysics. Advances in Descriptive Psychology, v. 8, Descriptive Psychology Press, 2005
- Colvin, G. In the Spirit of C. The C++ Source, 2004 http://www.artima.com/cppsource/spiritofcP.html

- Colvin, G. Exception Safe Smart Pointers, ANSI/ISO 94-168/N0555, 1994.
- Colvin, G. Object Oriented Programming in C and C++. C Users Journal, v. 11:7, 1993

Synergistic Activities:

- Helped to found the C++ Source, a peer-reviewed, online journal for the C++ community.
- Helped to found the Boost.org open source project, which provides peer-reviewed portable C++ source libraries.
- Served on the ANSI C++ committee, where he championed the interests of working engineers and concentrated on mechanisms for enforcing program safety.
- Writes frequently for the trade press, providing techniques and tools to engineering community.
- Works for peace.

Collaborator:s	Joe Jeffrey (NIU), Paul Zeiger, Pat Aucoin
Graduate Advisor:	Daryl Bem (Cornell University)

H. Joel Jeffrey

California Institute of Technology	Mathematics	BS	1969
University of Colorado Boulder	Computer Science	PhD	1974

Northern Illinois University	Professor	1992-present
H5 Technologies	Founder &Chief Scientist	1999-2002
Northern Illinois University	Associate Prof.	1983-1992
Management Support Technology	Technical Staff	1984-1987
Bell Telephone Laboratories	Technical Staff	1977-1983
Vanderbilt University	Assistant Prof.	1974-1977

Related publications:

- Jeffrey, J. Wide-spectrum information search engine. U.S. Patent 6,493,711, 2002
- Jeffrey, J. Judgment-Simulation Vector Spaces. Advances in Computer Methods for Systematic Biology: Artificial Intelligence, Database, ComputerVision. Johns Hopkins University Press, 1993
- Jeffrey, J. Expert Document Retrieval Via Semantic Measurement. Expert Systems With Applications, v. 2:4, p. 345-352, 1991
- Jeffrey, J. Theory and Practice of Knowledge Engineering. Advances in Descriptive Psychology, v. 5, Descriptive Psychology Press, 1990
- Jeffrey, J. A new type of information retrieval system. Proceedings of the 14th ACM Southeast Regional Conference. ACM, pp. 271-277, 1976

Other publications:

- Mathematical Description of Biological Structures, Mechanisms, and States. International Journal on Advances in Life Sciences, v. 1:1, to appear, 2009.
- Expanding the Range of Behavioral Factors in Economic Simulations, The Journal of Behavioral Finance 2006, Vol. 7:2, 97–106.

- Jeffrey, J. Human Systems Analysis in the Software Engineering Curriculum. Journal of Systems and Software, v . 4, p. 147-153, 1991
- Jeffrey, J. Chaos Game Representation of Genetic Sequences", Nucleic Acids Research, Vol. 18, No. 8, March 1990, pp. 2163-2170.
- A Unifying Framework for Software Technology Transfer", Proceedings of the Workshop on Transfer of Software Engineering Tool Technology, Santa Barbara, California, IEEE Computer Science Press, p. 82-85, 1988

Synergistic Activities:

- Research focus has been producing novel systems and techniques that apply to full-scale, real-world problems, as contrasted with theoretical or academic-scale problems
- Created one of the two of the first industrial expert systems, which applied a psychological analysis of social practices to build a question-answering expert system that helped Bell System engineers get their work done.
- Focus of teaching for the past 20 years has been to apply his psychological training and work experience to prepare his students for the real world of software engineering, and the focus of his writing to communicate that understanding to those already in the trenches.
- Founded a successful business applying judgment simulation to the problem of litigation support.
- Worked in his research, publication and presentations to make results from computer science and descriptive psychology useful to working biologists.

Collaborator:s	Gregory Colvin, Paul Zeiger, Pat Aucoin
Graduate Advisor:	Paul Zeiger (U S West Communication, retired)

Paschal J. Aucoin, Jr.

Rice University	Physics	BA	1954
University of Texas	Physics	MA	1956
University of California-Los Angeles	Mathematics	PhD	1970

Lockheed Martin -	Systems Engineer	1987-2002
	Independent Consultant	1986-1987
O'Connor Research	Vice President	1980-1986
Lockheed	Technical Lead	1977-1980
ARAMCO	Engineer	1974-1976
Ford Aerospace	Engineer	1974-1974
Digital Resources, Inc.	Consulting Scientist	1972-1974
Technology, Inc.	Biomedical Engineer	1971-1972
CCI-Marquardt	Supervisor, Analyst	1962-1970
Humble Oil	Engineer	1956-1962

Related publications:

- Erickson, J., P. Aucoin & P. Ossorio. Integrated Approach for Delineating Robotic Capabilities Needed for Future Space Exploration Tasks. Telemanipulator Technology and Space Telerobotics. SPIE, v. 2057, p.370-384, 1993
- Erickson, J., P. Aucoin & ,J. Dragg. Tools for Simulating Space Exploration Initiative (SEI) Operations as a Means of Obtaining Intelligent Systems Requirements. Cooperative Intelligent Robotics in Space III. SPIE, v. 1829, p. 482-494, 1992
- Erickson, J., P. Aucoin & P. Ossorio. Person-Like Intelligent Systems Architectures for Robotics Shared Control and Automated Operations. Cooperative Intelligent Robotics in Space III. SPIE, v. 1829, p. 149-160, 1992
- Aucoin, P. An AI-Based Approach to Initial Requirements Specification. Proceedings of the ROBEXS. ISA, v. 89, 1989

Other publications:

- Newell, A. & P. Aucoin. Semidispersive Wave Systems, Journal of Fluid Mechanics, v. 49:3, p.593-609, 1971.
- Newell A., C. Lange & P. Aucoin. Random Convection. Journal of Fluid Mechanics, v. 40, p. 513-542, 1970.
- Stechman R., R. Reis & P. Aucoin. Prediction of Rocket Exhaust Flowfields. Journal of Spacecraft and Rockets, v. 7:2, p. 155-159, 1970

Synergistic Activities:

- Continuing development of enhancements in modeling and simulation systems for man-made missions (such as space missions) to be put into service next year.
- Lead role in designing and operating the NASA Internet Web site for Advance Life Support missions including the Fall 1997 90 Day mission.
- Participation in the Texas Historical Commission Archeological Stewardship Network - searching for new sites and investigating existing ones.
- Creation of a curriculum of learning for teachers of archeology based upon specific episodes of virtual archeological investigations and the provision of associated information in the issues involved. Reviewed and approved by the Texas Archeological Society's teaching administrator.

Collaborators: Gregory Colvin, Paul Zeiger, Joe Jeffrey (NIU)
Graduate Advisor: Alan C. Newell (University of Arizona)
.

H. Paul Zeiger

MIT	Electrical Engineering	BS	1958
MIT	Electrical Engineering	MS	1960
MIT	Electrical Engineering	PHD	1964
MIT	Electrical Engineering	Fellow	1965

Independent Consulting		1998-2005
U S West Communications	Data Architect	1997-1998
U S West Communications	Director	1994-1997
U S West Advanced Technologies	Technical Director	1993-1994
U S West Advanced Technologies	Technical Staff	1989-1993
Michael Smith & Associates	Manager	1987-1989
Management Support Technolgy	Vice President	1982-1987
Health Management Systems	Consultant	1981-1982
University of Colorado Boulder	Department Chair	1978-1981
University of Colorado Boulder	Associate Professor	1972-1978
Stanford University	Summer Associate	1967-1970
University of Colorado Boulder	Assistant Professor	1966-1972

Related publications:

- Zeiger, P. Ho's Algorithm, Commutative Diagrams and the Uniqueness of Minimal Linear Systems. Information and Control, v. 11:1-2, p. 71-79, 1967.
- Arbib M. & P. Zeiger. On the Relevance of Abstract Algebra to Control Theory. Automatica, v. 5, p. 589-606, 1969
- Zeiger, P. & A. McEwen. Approximate Linear Realizations of Given Dimension via Ho's Algorithm. IEEE Transactions on Automatic Control, v. AC-19:2, p. 153, 1974
- Zeiger, P. Descriptive Psychology in Computing. Descriptive Psychology Bulletin, v. 1, p. 6, 1979

- Ossorio, P. & P. Zeiger. Behavior Pattern Indexing for Modus Operandi Files. Proceedings of the Fifth Annual law Enforcement Data Processing Management Symposium, 1981.

Other publications:

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

Synergistic Activities:

- Worked as Ford Fellow to revamp MIT's foundational EE course.
- Representative to the American National Standards Committee for C++.
- President and Founding member of Society for Descriptive Psychology.
- Executive Vice President of the Arizona Senior Academy, a non-profit offering courses, lectures, concerts, and other programs to residents of the Rincon Valley.
- Teacher of yoga to and advocate for Parkinson's patients. Participated in the production of a book to raise funds for the American Parkinson's Disease Association. Lectured and participated in talk shows on the subject of living with PD and the use of exercise therapies for it.

Collaborators:	Gregory Colvin, Pat Aucoin, Joe Jeffrey (NIU)
Co-author:	Richard Singer (Webster University)
Graduate Advisor:	Trenchard More (IBM, retired)