Curriculum Vitae

1 Personal Information

Name	Ricardo Baeza-Yates
Address	Yahoo! Research Barcelona
	Ocata 1, 08003, Barcelona, SPAIN
Telephone	(34) 935421161
FAX	(34)935422896
E-mail	ricardo@baeza.cl, rbaeza@acm.org, rbaeza@ieee.org
URL	http://www.baeza.cl/

2 Current Activities

- Director of Yahoo! Research Barcelona, Spain, and Latin America, Santiago, Chile.

- Director of the Information Retrieval and Storage Group at the Barcelona Media Foundation, a technological center in Barcelona.

- Part-time Professor at Dept. of Technology, Pompeu Fabra University, Barcelona, Spain and at the Dept. of Computer Science, University of Chile, Santiago, Chile.

- Member of the Chilean Academy of Sciences from December 2002-present.

3 Studies

- Ph.D. in Computer Science, University of Waterloo, Ontario, Canada, May 1989. Supervised by Gaston Gonnet, and working in the New Oxford English Dictionary Project.

- M.Eng. in Electrical Engineering (Digital Systems and Automatic Control), Universidad de Chile, 1986.

- M.Sc. in Computer Science, Universidad de Chile, 1985.

- Electrical Engineer, specialization in Digital Systems, Universidad de Chile, 1985.

- Bachelor in Computer Science, Universidad de Chile, 1983.

4 Academic Experience

- President of CLEI (Latinamerican Center of Studies in Informatics, a Latinamerican association of more than 60 computer science departments in 14 countries) from 2000-2004.

- Member of the Board of Governors of the IEEE Computer Society (USA) from 2002-2004.

- International Coordinator of the Iberoamerican cooperation program in science and technology (CYTED) in the areas of Applied Electronics and Informatics, 2000-2004.

- Chair of the CS Dept. (Univ. of Chile) from June 1993-September 1995, and November 2002-November 2004.

- President of the Chilean Computer Science Society (SCCC) from October 1992 to November 1995 and from December 1996 to November 1998. Board member from December 1995 to November 1996 and from November 1998 to November 2002.

- Visiting Professor (sabbatical) at the LSI Dept., Polytechnic Univ. of Catalunya, Apr to Dec 1996 and the Mathematics Research Center (CRM), Barcelona, Jan to Mar 1996.

- Professor of Ph.D. programs at Univ. Pompeu Fabra, Barcelona, Spain; University Carlos III, Madrid, Spain; Univ. Rovira i Virgili, Tarragona, Spain; Polytechnic Univ. of Madrid, Spain; Polytechnic Univ. of Catalunya, Spain; and Catholic University, Santiago, Chile.

- Member of evaluation committee for FONDEF (technology transfer program), 1991-1995 and 2001-2002, and FONDECYT (basic research program), 1999-2000. National Research Council for Science and Technology, Chile.

5 Awards and Scholarships

- 1997 COMPAQ Award to the best Brazilian research article ([?] of Journals publications) in CS published in 1996 (with Eduardo Barbosa and Nivio Ziviani).

- Scholarship from the Education and Science Ministry of Spain to have a sabbatical year at the Polytechnic Univ. of Catalunya (1996).

- Award "Ramón Salas Edwards" of the Institute of Engineers of Chile to the best research on engineering on the last 4 years (1994).

- Recipient of a United States Information Service Important Visitor program (June 1994).

- Award of the Organization of American States (OAS) for young researchers in 1993 (Exact Sciences).

- Special mention, University of Chile Award for Young Researchers, 1992.

- Award "Marcos Orrego Puelma" of the Institute of Engineers of Chile to the best engineer student graduated in Chile during 1985.

- Ontario Graduate Scholarship, May 1988-April 1989.

- Institute for Computer Research scholarship for graduate studies, University of Waterloo, May 1986-August 1988.

6 Selected Research Projects

- Center for Web Research (www.cwr.cl), Millennium Nucleus, Mideplan, Chile, 2002-2005. Project leader.

- Web Mining, CONICYT (Chilean National Science Funding), 2002-2003. Project leader.

- Design and Analysis of Sequential and Parallel Algorithms with Applications, CONICYT (Chilean National Science Funding), 1999-2001. Project leader.

- Environment for Manipulation and Retrieval of Information in WWW (AMYRI), CYTED (Iberoamerican countries), 1997-1999. Project leader.

- Data Structures: Design, Analysis and Applications, CONICYT (Chilean National Science Funding), 1995-98. Project leader.

- Textual Databases: Algorithms, Interfaces and Applications, CONICYT (Chilean National Science Funding), 1993-94. Project leader.

- Algorithms for Text Processing, Fundación Andes, Chile-Brazil cooperation program (with Univ. of Campinas and Univ. Fed. of Minas Gerais), 1993-94. Project leader.

- Program Visualization, CONICYT (Chilean National Science Funding), 1991-1992. Project leader.

- Integrated Software with Graphical Interfaces for Text Processing, Fundación Andes C-11001, 1990-1992. Project leader.

7   Citations

According to ISI's Web of Science, on September 2004 there were about 1,200 citations to my work. According to CiteSeer (www.researchindex.com), there were about 2,900 citations in Internet documents (which are in general different from ISI's journal citations). In July 2005 CiteSeer found 1,700 independent citations and ranked me among the most 650 cited authors in CS from a universe of more than 750 thousand authors. The main 5 references have the number of citations in boldface.

8   Selected Publications

We select the best and newest publications from more than 200 of them.

8.1   Books

R.Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval, ACM Press/Addison-Wesley, England, 513 pages, 1999. Part of the book is available in my personal WWW page. This book has been translated to Chinese and Korean and printed in special editions for China and India. More than 1,600 citations.

G.H. Gonnet and R.Baeza-Yates. Handbook of Algorithms and Data Structures - In Pascal and C. Addison-Wesley, Wokingham, UK, 424 pages, 1991. (second edition). Source code is available in my personal WWW page. More than 300 citations.

W.Frakes and R.Baeza-Yates, editors. Information Retrieval: Data Structures and Algorithms. Prentice-Hall, Englewood Cliffs, NJ, USA, 504 pages, 1992. Source code is available in my personal WWW page. More than 800 citations.

8.2   Editor of Proceedings

R. Baeza-Yates, E. Chvez, M. Crochemore (Eds.): Combinatorial Pattern Matching Proc. 14th Annual Symposium in Combinatorial Pattern Matching (CPM 2003), LNCS 2676, Springer, Morelia, Michoacn, Mexico, June 25-27, 2003.

R. Baeza-Yates, Ugo Montanari, and N. Santoro, editors. Foundations of Information Technology in the Era of Network and Mobile Computing (Proc. II IFIP Theoretical Computer Science Symposium), Kluwer, August 2002 (held in Montreal, Canada).

Micheline Beaulieu, Ricardo Baeza-Yates, Sung Hyon Myaeng, editors. Proc. 25th ACM SIGIR, ACM Press, August 2002 (held in Tampere, Finland).

8.3   Editor of Journal Issues

Ricardo Baeza-Yates and Maxime Crochemore, editors. Special issue on Combinatorial Pattern Matching, Journal of Discrete Algorithms 3 (2-4), 2005.

R.Baeza-Yates, D. Carmel, Y. Maarek and A. Sofer. Special Issue in XML and Information Retrieval, Journal of American Society on Information Systems and Technology 53(6): 2002.

R.Baeza-Yates and E. Goles. Special Issue in Latin American Theoretical Informatics, Theoretical Computer Science 181 (1), July 1997 (225 pages).

8.4 Chapters in Books

Ricardo Baeza-Yates, Alistair Moffat and Gonzalo Navarro. Searching Large Text Collections. In J. Abello, P. Pardalos and M. Resende (editors), Handbook of Massive Data Sets, Kluwer Academic Publishers, 2002, pages 195-244.

R.Baeza-Yates and P.Poblete. Searching. In Mikhail Atallah, editor, Handbook on Algorithms and Theory of Computation, chapter2. CRC Press, 1999.

R.Baeza-Yates. Searching: An algorithmic tour. In Allen Kent and JamesG. Williams, editors, Encyclopedia of Computer Science and Technology, volume37, pages 331-359. Marcel Dekker, Inc., 1997.

8.5 International Journals

Ricardo Baeza-Yates. Information Retrieval in the Web: beyond current search engines, International Journal on Approximated Reasoning 34 (2-3), 97-104, 2003.

Edgar Chvez, Gonzalo Navarro, Ricardo Baeza-Yates and Jos Luis Marroqun. Proximity Searching in Metric Spaces. ACM Computing Surveys 33(3):273-321, 2001. More than 150 citations.

Gonzalo Navarro and Ricardo Baeza-Yates. Improving an Algorithm for Approximate String Matching. Algorithmica 30(4):473-502, 2001.

Edleno de Moura, Gonzalo Navarro, Nivio Ziviani and Ricardo Baeza-Yates. Direct Pattern Matching on Compressed Text Allowing Errors. ACM Transactions on Information Systems (TOIS) 18(2): 113-139, 2000. More than 50 citations.

R. Baeza-Yates and G. Navarro. Block-Addressing Indices for Approximate Text Retrieval. J. of the American Society for Information Science (JASIS) 51(1):69-82, Jan 2000. More than 40 citations.

R.A. Baeza-Yates and G.Navarro. Faster approximate string matching. Algorithmica, 23 (2):127-158, 1999. More than 70 citations.

G.Navarro and R.Baeza-Yates. Proximal Nodes: a model to query document databases by content and structure. ACM TOIS, 15(4):401-435, Oct 1997. More than 100 citations.

R.A. Baeza-Yates and G.Navarro. Integrating contents and structure in text retrieval. ACM SIGMOD Record, 25(1):67-79, March 1996. More than 80 citations.

R.Baeza-Yates and G.H. Gonnet. Fast text searching for regular expressions or automaton searching on tries. Journal of the ACM, 43(6):915-936, Nov 1996. More than 40 citations.

R.A. Baeza-Yates and C.H. Perleberg. Fast and practical approximate pattern matching. Information Processing Letters, 59:21-27, 1996. More than 40 citations.

R.A. Baeza-Yates, J.Culberson, and G.Rawlins. Searching in the Plane. Information and Computation, 106(2):234-252, Oct 1993. More than 160 citations.

R.Baeza-Yates and G.H. Gonnet. A new approach to text searching. Communications of the ACM, 35:74-82, Oct 1992. More than 160 citations.

U.Manber and R.Baeza-Yates. An algorithm for string matching with a sequence of don't cares. Information Processing Letters, 37:133-136, February 1991. More than 50 citations.

R.Baeza-Yates. Searching subsequences. Theoretical Computer Science, 78:363-376, 1991. More than 50 citations.

R.A. Baeza-Yates and P-. Larson. Performance of B-trees with partial expansions. IEEE Trans. on Knowledge and Data Engineering, 1:248-257, June 1989. More than 30 citations.

R.Baeza-Yates. Improved string searching. Software-Practice and Experience, 19(3):257-271, 1989. More than 25 citations.

Ricardo Baeza-Yates, Liliana Caldern, Cristina Gonzlez. The Intention Behind Web Queries. SPIRE 2006, LNCS Springer, Glasgow, Scotland, October 2006.

Ricardo Baeza-Yates, Paolo Boldi, Carlos Castillo. Generalizing PageRank: Damping Functions for Link-Based Ranking Algorithms. In Proceedings of ACM SIGIR 2006. Seattle, Washington, USA, August 2006. ACM Press.

R. Baeza-Yates, C. Castillo, M. Marin, A. Rodriguez. Crawling a Country: Better Strategies than Breadth-First for Page Ordering. In WWW 2005, Industrial Track, ACM Press, Chiba, Japan, May 2005.

R. Baeza-Yates, A Fast Set Intersection Algorithm for Sorted Sequences, In 15th Combinatorial Pattern Matching 2004, LNCS, Springer, Istanbul, Turkey, July 2004.

Eduardo Barbosa, Gonzalo Navarro, Ricardo Baeza-Yates, Chris Perleberg, and Nivio Ziviani. Optimized binary search and text retrieval. In European Symposium on Algorithms, Corfu, Greece, September 1995.

Gonzalo Navarro and Ricardo Baeza-Yates. A language for queries on structure and contents of textual databases. In 18th ACM Conference on Research and Development in Information Retrieval (SIGIR'95), pages 93-101, Seattle, WA, USA, July 1995. More than 60 citations.

R.A. Baeza-Yates, W.Cunto, U.Manber, and S.Wu. Proximity matching using fixed queries trees. In M.Crochemore and D.Gusfield, editors, 5th Combinatorial Pattern Matching, LNCS 807, pages 198-212, Asilomar, CA, June 1994. Springer-Verlag. More than 50 citations.

R.Baeza-Yates. Text retrieval: Theory and practice. In J.van Leeuwen, editor, 12th IFIP World Computer Congress, Volume I, volume Algorithms, Software, Architecture, pages 465-476, Madrid, Spain, September 1992. Elsevier Science. More than 40 citations.

R.A. Baeza-Yates. An adaptive overflow technique for the B-tree. In F.Bancilhon, C.Thanos, and D.Tsichritzis, editors, Extending Data Base Technology Conference (EDBT 90), pages 16-28, Venice, March 1990. Springer Verlag Lecture Notes in Computer Science 416.

8.6   Invited Papers

Ricardo Baeza-Yates. Applications of Web Query Mining. In European Conference on Information Retrieval (ECIR'05), D. Losada, J. Fernndez-Luna (editors), Springer LNCS 3408, Santiago de Compostela, Spain, March 2005, 7-22.

Ricardo Baeza-Yates. Challenges in the Interaction of Natural Language Processing and Information Retrieval. CICLING 2004, Springer LNAI, Seoul, Korea, February 2004.

R. Baeza-Yates, T. Jones, and G. Rawlins. New Approaches to Information Management: Attribute-Centric Data Systems, 7th International Symposium on String Processing and Information Retrieval, A Corua, Spain, 27-29 September, 2000, IEEE Computer Science Press, 18-27.

9   Ph.D. Theses Supervised

1. Castillo, Carlos. Effective Web Crawling, Ph.D. in Computer Science, Univ. de Chile, November 2004.

2. Chávez, Edgar. Similarity Searching on Metric Spaces, Ph.D. in Computer Science, CIMAT, México, September 1999. Co-supervised with José Luis Marroquín.

3. Vegas, Jesús. An Information Retrieval System on Structure and Content, Ph.D. in Computer Science, Univ. de Valladolid, Spain, May 1999. Co-supervised with Pablo de la Fuente (in Spanish).

4. Navarro, Gonzalo. Approximate Text Searching, Ph.D. in Computer Science, Univ. de Chile, December 1998.

5. Barbosa, Eduardo. Efficient Text Retrieval Methods for Secondary Memory, Ph.D. in Computer Science, Univ. Fed. de Minas Gerais, Brazil, July 1995. Co-supervised with Nivio Ziviani.

We have omitted 13 M.Sc. theses (two of them won the award to the best Latinoamerican thesis) and 9 Final Engineering Projects.

10 Other Relevant Information

- Managing Editor of the CLEI Electronic Journal at `www.clei.cl` (1998-). Editorial committee member of the following journals: ACM Transactions on Information Systems (USA), VLDB Journal (USA), Information Systems (USA), Journal of Web Engineering (USA), Theoretical Informatics and Applications - RAIRO (France), Information and Software Technology (Netherlands), International Journal of Web Information Systems (UK), Grammars (Spain), Journal of Digital Management (South Korea), Journal of the SBC (Brazil), First Monday (USA), El Profesional de la Información (Spain), Revista de la SCCC (Chile) and Revista Colombiana de Informática.

- In the last 3 years have been: General co-chair for the ACM Conference in Information Retrieval (SIGIR) (Bahia, Brazil, 2005); Program committee co-chair for the Mexican Computer Science Conference (ENC 2004, Colima, Mexico, September 2004); Program committee co-chair for the First Latinamerican Web Conference (LA-WEB, Santiago, Chile, November 2003); Program committee co-chair for the Combinatorial Pattern Matching (CPM, Morelia, Mexico, June 2003); Program committee chair for the Americas of the ACM Conference in Information Retrieval (SIGIR, Tampere, Finland, Aug 2002); and Program committee co-chair of IFIP TCS'2002 (Montreal, Canada, Aug 2002).

- Program Committee member of more than a 100 congresses and referee for more than 30 journals. External reviewer for Ph.D. theses in Australia, Brazil, France, Italy, and Spain. I have evaluated grant proposals for the National Science Foundation (USA), NSERC (Canada), ISF (Israel), CONICYT (Chile), CNPq (Brazil), GACR (Check Republic), RCN (Norway), CONICET (Argentina), CONICYT (Uruguay), and Chilean/European Community projects.

- Invited speaker at more than 135 institutions in 29 countries.

- Member of the ACM, AMS, Chilean Computer Science Society (SCCC), EATCS, IEEE (Senior Member), and SIAM.

- My teaching and professional experience has been omitted.