

# Short Curriculum Vitae

## 1 Personal Information

Name	Ricardo Baeza-Yates
E-mail	rbaeza@acm.org
Personal website	<a href="http://www.baeza.cl/">http://www.baeza.cl/</a>
Languages	Spanish (native), English (fluent), Portuguese (good), and Catalan (reading and listening)

## 2 Current Activities

- WASP Professor (part-time) at KTH Royal Institute of Technology, Stockholm, Sweden, since 2025.
- Consultant for startups, companies, and venture capital in Brazil, Chile, Mexico, Spain, and USA; as well as for international organizations like United Nations and IADB (BID).
- Full Professor (part-time) at the Dept. of Engineering, Universitat Pompeu Fabra, Barcelona, Catalonia, Spain, since 2004.
- Professor (part-time) at the Dept. of Computing Sciences, Universidad de Chile, Santiago, Chile, since 1984 (full-professor since 1994, the youngest ever at that time).
- Member of the Steering Committee of the Global Partnership of AI at OECD, ACM's Policy Technology Committee (co-chair of the AI subcommittee), IEEE's AI Policy Committee, Catalonia's AI Observatory Advisory Board, and Safe Systems & Technologies Working Group at the World Economic Forum.
- Member of the Academia Europaea since 2023.
- Founding member of the Chilean Academy of Engineering since 2010.
- Member of the Chilean Academy of Sciences since 2023 (corresponding member during 2002-2022).
- Corresponding member of the Brazilian Academy of Sciences since 2019.
- Member of the advisory board of the National Portuguese Research Lab in CS, INESC, since 2012.
- Member of the advisory boards of the CSE Department of UC Riverside as well as the Department of Informatics Engineering, University of Porto, Portugal, since 2020.

## 3 Studies

- Ph.D. in Computer Science, University of Waterloo, Ontario, Canada, May 1989. Supervised by Gaston Gonnet, and working on searching algorithms for the New Oxford English Dictionary Project.
- M.Eng. in Electrical Engineering (Digital Systems and Automatic Control), Univ. 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 Awards, Distinctions, Fellowships and Scholarships

- Honorary Ph.D. from Univ. of Goteborg, Sweden, 2025.
- First Merit Award in Computer Science, Chilean Computer Science Society, 2025.
- Recognition as Co-Founder of the International Computer Programming Contest in Latin America and First Director of the Competition in Chile, Salvador, Brazil, 2025.
- **National Award for Applied and Technological Sciences** (given every other year), Chile, 2024.
- Person of the Year Award given by the Association of Technological Companies, CHILETEC, Chile, 2024.
- Rectoral Medal of the University of Chile, 2024.
- Named Full Member of the Chilean Academy of Sciences, 2023 (corresponding member since 2003).
- Fellowship, IWM - Institute for Human Sciences, Vienna, Austria, June 2023.
- Named Corresponding Member of the Brazilian Academy of Sciences, 2019.
- Award Salvà i Campillo to the Person of the Year, Catalan Association of Telecommunications Engineers with the cooperation of the Association of Computing Engineers of Catalonia, 2019.
- **National Research Award Ángela Ruiz Robles** for applied research and technology transfer, Scientific Computing Societies of Spain & BBVA Foundation, Spain, 2018.
- Best research award at Techpulse 2014, the main annual technical event of Yahoo. This award was obtained with Di Jiang, Fabrizio Silvestri and Beverly Harrison for a new algorithm that using deep learning allows to predict the next mobile app that a person will use with more than 90% accuracy.
- Computer World Spain Award to Yahoo Labs Barcelona for contributions to big data in 2014.
- Master inventor award, Yahoo, 2012.
- 2012 ASIS&T Best Information Science Book Award for the second edition of Modern Information Retrieval.
- Named IEEE Fellow for contributions to the development of computer science, 2011.
- National Award of the Chilean Association of Engineers, 2010.
- Founding member of the Chilean Academy of Engineering, 2010.
- Named ACM Fellow for contributions to information retrieval algorithms and techniques, 2009.
- CLEI Distinction to Achievements in Informatics, given by the Latin American association of CS departments to people that has contributed to the development of CS in the region, 2009.
- J. W. Graham Medal in Computing & Innovation given to distinguished alumni of the University of Waterloo, Canada, 2007.
- Chilean Fulbright Commission Fellowship in Information Technology, 2003.
- Manuel Montt award to “Modern Information Retrieval” (co-authored with Berthier Ribeiro-Neto) as the best technical book published by Chilean authors during 1997-2001, 2002.
- Recognition from Informatics Magazine (IDG Chile) for the Development of IT in Chile, 2001.
- COMPAQ Award to the best Brazilian research article in CS published in 1996 (with Eduardo Barbosa and Nivio Ziviani), 1997.
- Award “Ramón Salas Edwards” of the Institute of Engineers of Chile to the best research on engineering on the last 4 years in 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 Exact Sciences, 1993.
- Award from PC Software magazine to the best Chilean software, SearchCity, designed together with Pablo Palma in 1992. This Windows 3.1 tool predated Google Desktop but was too ahead in time to succeed.
- Five PhD scholarships from Ontario, Univ. of Waterloo, ICR and ITRC, May 1986-April 1989.
- Award “Marcos Orrego Puelma” of the Institute of Engineers of Chile to the best engineering student graduated in Chile during 1985.

## 5 Academic and Research Experience

- Director of Research, Institute for Experiential AI, Northeastern University at Silicon Valley, USA, 2021-2025.
- Director for Graduate Programs in Data Science (part-time), Northeastern University at Silicon Valley, USA, 2017-2020.
- ICREA Research Professor, Dept. of Information and Communication Technologies (now Dept. of Engineering), Universitat Pompeu Fabra, Barcelona, Spain, Dec 2004-Dec 2005. Part-time Professor since 2006.
- Faculty member at the Department of Computer Science, University of Chile, Santiago, Chile, 1985-2005 (Professor since April 1995). Part-time Professor since 2006.
- Founder and first director of the Center for Web Research ([www.cwr.cl](http://www.cwr.cl)), Santiago, Chile, 2002-2005.
- Chair of the Computer Science Dept. (Univ. of Chile) from June 1993 to September 1995, and also from November 2002 to November 2004.

## 6 Industry Experience

- CEO of NTENT Hispania, a R&D subsidiary of NTENT USA, Barcelona, Spain (April 2017 - December 2020).
- CTO of NTENT, a semantic search technology company, Carlsbad, USA (June 2016 - June 2020).
- Chief Research Scientist of Yahoo Labs from October 2015 to February 2016 in Sunnyvale, USA.
- VP of Research at Yahoo Inc. from August 2014 to February 2016 in Sunnyvale, USA.
- VP of Research for EMEA and Latin America at Yahoo Inc. from January 2006 to September 2015, where I started and lead the labs in Europe (Barcelona, Spain) and Latin America (Santiago, Chile).
- Member of the Scientific Advisory board of EURECAT, Barcelona (2015-2017).
- Member of the Executive Board of Barcelona Media Foundation (2010-2015).
- Supervised Yahoo Labs London from its inception in 2012 until 2015 and Yahoo Labs Haifa from its inception in 2008 until 2012.

## 7 Research Projects

At Yahoo Labs I was involved in many internal applied research projects in several areas, including web search, computational advertising (sponsored search and native), web usage data mining, and social media analytics. They included many aspects, in particular efficiency and scalability. For confidentiality reasons more details cannot be given.

I have lead or co-lead 28 research projects with funding from the European Union Research Programs, Ibero-American Research Projects (CYTED, ALFA), bilateral projects (Chile-Brazil, Chile-France, Chile-Spain), and national USA (NSF), Chilean (Millennium, CONICYT, FONDEF) and Spanish (CENIT, MICINN, AGAUR) Research Projects. The total funding obtained is of the order of ten million Euros.

## 8 Citations

According to Google Scholar there are more than 58,600 citations to more than 800 publications and other works (September 2025), with an h-index of 92. Citations below are mainly from this source (see my personal entry). I am the most cited researcher in algorithms and data structures, 2nd most cited in responsible AI and web search, the 3rd most cited in web mining, and 15th most cited in information retrieval.

In September 2025, the ACM Digital Library had more than 10,200 citations to 356 publications, with more than 145,000 overall downloads. In February 2025, Semantic Scholar had more than 22,700 citations to 663 publications while DBLP had 510 recorded publications. At the same time, Semantic Scholar had more than 23,100 citations (three versions of my name) to 669 publications while DBLP had 526 recorded publications.

In March 2020, Microsoft Academic Search had 487 of my publications with almost 44,000 citations, appearing in 16th place in information retrieval and 42nd place in web related research (they do not longer publish statistics).

In May 2012, CiteSeerX found more than 5,380 independent citations and ranked me the 528th most cited author in CS from a universe of more than 800 thousand authors (since then has not been updated). In their statistics my information retrieval book is the third most cited work of 1999 and the 86th most cited overall (2015).

## 9 Selected Publications

Only recent and widely cited works are considered among more than 500 publications (9 books, 64 edited special journal issues and proceedings, 37 chapters in books, 30 invited papers, 126 international journal articles, 190 international conferences, 65 workshops and regional conferences, etc.).

We highlight papers with at least 100 citations in Google Scholar and 1,000 downloads in the ACM Digital Library.

### 9.1 Books

- [1] Omar Alonso and Ricardo Baeza-Yates (editors). *Information Retrieval: Advanced Topics and Techniques* (1st. ed.). Vol. 60. ACM Books, New York, NY, USA, 2025.
- [2] B.B. Cambazoglu and R. Baeza-Yates. *Scalability Challenges in Web Search Engines*, in *Synthesis Lectures on Information Concepts, Retrieval, and Services* (Gary Marchionini, editor). Morgan & Claypool Publishers, USA, 122 pages, 2015. More than **100** citations.
- [3] R. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval: the concepts and technology behind search*, second edition. Addison-Wesley, England, 913 pages, 2011. More than **22,100** citations.
- [4] 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). More than **970** citations.
- [5] W. Frakes and R. Baeza-Yates, editors. *Information Retrieval: Data Structures and Algorithms*. Prentice-Hall, Englewood Cliffs, NJ, USA, 504 pages, 1992. More than **3,970** citations.

## 9.2 Chapters in Books

- [1] R. Baeza-Yates, L. Murgai. Bias and the Web. In *Introduction to Digital Humanism*: 435-462. H. Werthner, C. Ghezzi, J. Kramer, J. Nida-Rümelin, B. Nuseibeh, E. Prem, A. Stanger. Springer, 2024.
- [2] P. Mika, R. Baeza-Yates. The Impact of the Web on Information Retrieval. In *Linking the World's Information: Essays on Tim Berners-Lee's Invention of the World Wide Web*. O. Seneviratne, J. Hendler (eds). ACM Books, New York, NY, USA, pp. 105–114, Sep 2023.
- [3] R. Baeza-Yates, U.M. Fayyad. The Attention Economy and the Impact of Artificial Intelligence. In *Perspectives on Digital Humanism*: 123-134. H. Werthner, E. Prem, E.A. Lee, C. Ghezzi (eds). Springer 2022.
- [4] L. Devillers, F. Fogelman-Soulié, R. Baeza-Yates. AI and Human Values: Inequalities, Biases, Fairness, Nudge and Feedback Loops. In *Reflections on AI for Humanity*. B. Braunschweig, M. Ghallab (eds). Springer LNAI 12600, pp. 66–89, 2021.
- [5] M. Rauschenberger, R. Baeza-Yates, and L. Rello. Technologies for Dyslexia (Chapter 31). In *Web Accessibility*, 2nd edition. S. Harper and Y. Yezilada (eds). Springer, pp. 603–627, 2019.
- [6] R. Baeza-Yates, R. Blanco, M. Castellanos. Web Text Mining. In *The Oxford Handbook of Computational Linguistics*, 2nd edition, R. Mitkow (editor). Oxford Press, 2017.
- [7] R. Baeza-Yates, A. Broder and Y. Maarek. The new frontier of Web Search Technology: Seven challenges. In *Search Computing-Trends and Developments*, S. Ceri and M. Brambilla (eds). Springer LNCS 6585, Berlin, 2011.
- [8] R. Baeza-Yates, P. Boldi, A. Bozzon, M. Brambilla, S. Ceri, G. Pasi. Trends in Search Interaction. In *Search Computing – Trends and Developments*, S. Ceri and M. Brambilla (eds). Springer LNCS 6585, Berlin, 2011.
- [9] R. Baeza-Yates. Privacy in Query Logs. In *Web Intelligence and Security*, M. Last and A. Kandel (eds). IOS Press, pp. 189-193, 2010.
- [10] R. Baeza-Yates, R. Jones, B. Poblete, M. Spiliopoulou. Issues with Privacy Preservation in Query Log Mining. In *Privacy-Aware Knowledge Discovery: Novel Applications and New Techniques*, F. Bonchi and E. Ferrari (eds). Chapman and Hall/CRC Press, pp. 353–368, 2010.
- [11] R. Baeza-Yates, P. Raghavan. Next Generation Web Search. In *Search Computing - Challenges and Directions*, S. Ceri and M. Brambilla (eds), LNCS 5950, Springer, pp. 11-23, 2010.
- [12] R. Baeza-Yates and C. Castillo. Web Searching. In *Encyclopedia of Language and Linguistics*, 2nd edition. G. Hirst (editor), Vol. 13, Elsevier: Oxford, 2005, 527-537.
- [13] R. Baeza-Yates. Query Usage Mining in Search Engines. In *Web Mining: Applications and Techniques*, A. Scime (editor). Idea Group, 2004, 307–321.
- [14] R. Baeza-Yates, B. Bustos, E. Chavez, N. Herrera, and G. Navarro. Clustering in Metric Spaces and Its Application to Information Retrieval. In *Clustering and Information Retrieval*, W. Wu and H. Xiong (eds), Kluwer Academic Publishers, 2003, pp. 1–34.
- [15] R. Baeza-Yates, A. Moffat and G. Navarro. Searching Large Text Collections. In *Handbook of Massive Data Sets*, J. Abello, P. Pardalos and M. Resende (eds), Kluwer Academic Publishers, 2002, pp. 195-244.
- [16] R. Baeza-Yates and P. Poblete. Searching. In *Handbook on Algorithms and Theory of Computation*, M. Atallah (editor), CRC Press, 1999.
- [17] R. Baeza-Yates. Searching: An algorithmic tour. In *Encyclopedia of Computer Science and Technology*, A. Kent and J.G. Williams (eds), vol. 37, pp. 331–359. Marcel Dekker Inc., 1997.

### 9.3 International Journals

- [1] M. Rauschenberger, R. Baeza-Yates, L. Rello. Screening Dyslexia Using Visual Auditory Computer Games and Machine Learning. *IEEE Access* 13, pp. 29541-29553, Feb 2025.
- [2] D. Ramírez-Cifuentes, R. Baeza-Yates, M. Lozano, A. Freire. Enhancing contact recommendation in social platforms through mental health awareness: Exploring Anorexia Nervosa as a case study. *PLOS One*, Feb 2025.
- [3] R. Ramachandranpillai, R. Baeza-Yates and F. Heintz. FairXAI - A Taxonomy and Framework for Fairness and Explainability Synergy in Machine Learning. *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1-18, Jan 2025.
- [4] D. Pedreschi, L. Pappalardo, E. Ferragina, R. Baeza-Yates, A-L. Barabási, F. Dignum, V. Dignum, T. Eliassi-Rad, F. Giannotti, J. Kertész, A. Knott, Y. Ioannidis, P. Lukowicz, A. Passarella, A.S. Pentland, J. Shawe-Taylor, A. Vespignani. Human-AI Coevolution. *Artificial Intelligence* 339, Feb. 2025.
- [5] A. Knott, D. Pedreschi, T. Jitsuzumi, S. Leavy, D. Eyers, T. Chakraborti, A. Trotman, S. Sundareswaran, R. Baeza-Yates, P. Biecek, A. Weller, P.D. Teal, S. Basu, M. Haklidi, V. Morini, S. Russell, Y. Bengio. AI content detection in the emerging information ecosystem: new obligations for media and tech companies. *Ethics and Information Technology* 26(4):63, 2024.
- [6] M. Andorra, A. Freire, I. Zubizarreta, N. Kerlero de Rosbo, S.D. Bos, M. Rinas, E.A. Høgestøl, S.A. de Rodez Benavent, T. Berge, S. Brune-Ingebretse, F. Ivaldi, M. Cellerino, M. Pardini, G. Vila, I. Pulido-Valdeolivas, E.H. Martinez-Lapiscina, S. Llufríu, A. Saiz, Y. Blanco, E. Martinez-Heras, E. Solana, P. Bäcker-Koduah, J. Behrens, J. Kuchling, S. Asseyer, M. Scheel, C. Chien, H. Zimmermann, S. Motamedi, J. Kauer-Bonin, A. Brandt, J. Saez-Rodriguez, L. G. Alexopoulos, F. Paul, H.F. Harbo, H. Shams, J. Oksenberg, A. Uccelli, R. Baeza-Yates, P. Villoslada. Predicting disease severity in multiple sclerosis using multimodal data and machine learning. *Journal of Neurology* 271(3), 1133-1149, March 2024.
- [7] R. Baeza-Yates, U.M. Fayyad. Responsible AI: An Urgent Mandate. *IEEE Intell. Systems* 39(1), pp. 12-17, 2024.
- [8] A. Knott, D. Pedreschi, R. Chatila, T. Chakraborti, S. Leavy, R. Baeza-Yates, D. Eyers, A. Trotman, P.D. Teal, P. Biecek, S. Russell, Y. Bengio. Generative AI models should include detection mechanisms as a condition for public release. *Ethics and Information Technology* 25, 55, 2023.
- [9] R. Baeza-Yates. An Introduction to Responsible AI. *European Review* 31(4):406-421, 2023.
- [10] E-M. Schön, M. Neumann, C. Hofmann-Stölting, R. Baeza-Yates and M. Rauschenberger. How are AI Assistants Changing Higher Education? *Frontiers in Computer Science* 5, 2023.
- [11] M. Zehlike, T. Sür, R. Baeza-Yates, F. Bonchi, C. Castillo, S. Hajian. Fair Top- $k$  ranking with multiple protected groups. *Information Processing & Management* 59(1), 2022.
- [12] D. Ramírez-Cifuentes, C. Llargeron, J. Tissier, R. Baeza-Yates, A. Freire. Enhanced Word Embedding Variations for the Detection of Substance Abuse and Mental Health Issues on Social Media Writings. *IEEE Access* 9, 2021.
- [13] D. Ramírez-Cifuentes, A. Freire, R. Baeza-Yates, N. Sanz Lamora, A. Álvarez, A. González-Rodríguez, M. Lozano Rochel, R. Llobet Vives, D.A. Velazquez, J.M. Gonfaus, J. González. Characterization of Anorexia Nervosa on Social Media: Textual, Visual, Relational, Behavioral, and Demographical Analysis. *Journal of Medical Internet Research* 23(7), 2021.
- [14] C. Ebell, R. Baeza-Yates, R. Benjamins, H. Cai, M. Coeckelbergh, T. Duarte, M. Hickok, A. Jacquet, A. Kim, J. Krijger, J. MacIntyre, P.B. Madhamshettiwar, L. Maffeo, J. Matthews, L.R. Medsker, P. Smith, S. Thais. Towards intellectual freedom in an AI Ethics Global Community. *AI Ethics* 1(2), pp. 131–140, 2021.
- [15] L. Rello, R. Baeza-Yates, A. Ali, J.P. Bigham, M. Serra. Predicting risk of dyslexia with an online gamified test. *Plos One*, Dec 2, 2020.

- [16] P. Villoslada, R. Baeza-Yates, J.C. Masdeu. Reclassifying Neurodegenerative Diseases. *Nature Biomedical Engineering*, Aug 2020.
- [17] D. Ramírez-Cifuentes, A. Freire, R. Baeza-Yates, J. Puntí Vidal, P. Medina-Bravo, J. Gonzalez, D.A. Velazquez, J.M. Gonfaus. Detection of Suicidal Ideation on Social Media: Multimodal, Relational, and Behavioral Analysis. *Journal of Medical Internet Research*, Jul 2020. More than **110** citations.
- [18] I. Tsalouchidou, R. Baeza-Yates, F. Bonchi, K. Liao, T. Sellis. Temporal Betweenness Centrality in Dynamic Graphs. *International Journal of Data Science and Analytics* 9, 257–272, 2020.
- [19] I. Tsalouchidou, F. Bonchi, G. de Francisi Morales, R. Baeza-Yates. Scalable Dynamic Graph Summarization. *IEEE TKDE* 32(2), 360-373, 2020.
- [20] R. Baeza-Yates. Bias on the Web. *CACM* 61(6), pp. 54–61, 2018. More than **620** citations and **23,700** downloads.
- [21] J. Lehmann, C. Castillo, M. Lalmas, R. Baeza-Yates. Story-focused Reading in Online News and its Potential for User Engagement. *JASIST* 68(4), pp. 869–883, 2017.
- [22] L. Rello, R. Baeza-Yates. How to Present more Readable Text for People with Dyslexia. *Universal Access in the Information Society* 16(1), pp. 29–49, Springer, 2017. More than **130** citations.
- [23] L. Rello, R. Baeza-Yates. The Effect of Font Type on Screen Readability by People with Dyslexia. *ACM Transactions on Accessibility* 8(4), pp. 15–33, 2016. More than **100** citations and **3,750** downloads.
- [24] D. Martin-Borregon, L. Maria Aiello, P. Grabowicz, A. Jaimes, R. Baeza-Yates. Characterization of online groups along space, time, and social dimensions. *EPJ Data Science* 3(1): 1–37, 2014.
- [25] S. Alvarez-García, R. Baeza-Yates, N. Brisaboa, J-L. Larriba-Pey, O. Pedreira. Automatic Multi-Partite Graph Generation from Arbitrary Data. *Journal of Systems and Software* 94: 72–86, 2014.
- [26] A. Francisco, R. Baeza-Yates, A. Oliveira. Mining query log graphs towards a query folksonomy. *Concurrency and Computation: Practice and Experience* 24(17): 2179–2192, 2012.
- [27] V. Pascual-Cid, R. Baeza-Yates, J.C. Dürsteler. Visual Web Mining for Website Evaluation. *Journal of Web Engineering* 9(4), 347–368, 2010.
- [28] B. Poblete, M. Spiliopolou, R. Baeza-Yates. Privacy-Preserving Query Log Mining for Business Confidentiality Protection, *ACM Transactions on the Web* 4 (3), 2010.
- [29] M. Marín, V. Gil-Costa, C. Bonacic, R. Baeza-Yates, I.D. Scherson. Sync/Async Parallel Search for Efficient Web Search Engines, *Parallel Computing* 36(4): 153-168, 2010.
- [30] D. Puppín, R. Perego, F. Silvestri, R. Baeza-Yates. Tuning the Capacity of Search Engines: Load-driven Routing and Incremental Caching to Reduce and Balance the Load, *ACM Trans. on Inform. Systems* 28 (2), Apr 2010.
- [31] R. Nilo, C. Saffie, K. Lilley, R. Baeza-Yates, V. Cambiazo, R. Campos-Vargas, M. Gonzalez, L.A. Meisel, J. Retamales, H. Silva, A. Orellana. Proteomic analysis of peach fruit mesocarp softening and chilling injury using difference gel electrophoresis (DIGE). *BMC Genomics*, 11–43, 2010. Almost **150** citations.
- [32] B. Jansen, T.J. Flaherty, R. Baeza-Yates, L. Hunter, B. Kitts, and J. Murphy. The Components and Impact of Sponsored Search, *IEEE Computer* 42 (5), May 2009, 98–101.
- [33] D. Nettleton and R. Baeza-Yates. Web Retrieval: Techniques for the Aggregation and Selection of Queries and Answers, *International Journal of Intelligent Systems* 23(12), 1223-1234, 2008.
- [34] R. Baeza-Yates, A. Gionis, F. Junqueira, V. Murdock, V. Plachouras, F. Silvestri. Design Trade-Offs for Search Engine Caching, *ACM Trans. on the Web*, 2(4), 2008. More than **130** citations and **1,010** downloads.
- [35] L. Becchetti, C. Castillo, Debora Donato, R. Baeza-Yates, S. Leonardi. Link Analysis for Web Spam Detection, *ACM Transactions on the Web* 2(1), Article 2, 2008. More than **170** citations and almost **2,000** downloads.

- [36] R. Baeza-Yates, P. Mika, and H. Zaragoza. Search, Web 2.0, and the Semantic Web. *IEEE Intelligent Systems* 23 (1), pp. 80–82, Jan-Feb 2008.
- [37] O. Alonso, M. Gertz and R. Baeza-Yates. On the Value of Temporal Information in Information Retrieval, *ACM SIGIR Forum* 41(2), 35–41, December 2007. More than **250** citations and **1,130** downloads.
- [38] R. Baeza-Yates, C. Hurtado and M. Mendoza. Improving Search Engines by Query Clustering, *JASIST* 58(12), 1793–1804, October 2007.
- [39] A. Moffat, W. Webber, J. Zobel and R. Baeza-Yates. A Pipelined Architecture for Distributed Text Query Evaluation. *Information Retrieval* 10(3):205-231, June 2007. More than **140** citations.
- [40] R. Baeza-Yates, C. Castillo and E.N. Efthimiadis: Characterization of National Web Domains. *ACM Transactions on Internet Technology* 7, Issue 2, 2007. More than **170** citations and **1,080** downloads.
- [41] R. Baeza-Yates and C. Castillo. Crawling the Infinite Web: Five Levels are Enough. *J. of Web Engineering* 6, pp. 49-72, 2007.
- [42] C. Badue, R. Baeza-Yates, B. Ribeiro-Neto, A. Ziviani, and N. Ziviani. Analyzing Imbalance among Homogeneous Index Servers in a Web Search System. *Information Processing & Management* 43, 592-608, 2007.
- [43] R. Baeza-Yates, P. Boldi and C. Castillo. Generic Damping Functions for Propagating Importance in Link-Based Ranking, *Journal of Internet Mathematics* 3(4), 445-478, 2006.
- [44] L. Meisel, B. Fonseca, S. Gonzalez, R. Baeza-Yates, V. Cambiazo, R. Campos, M. Gonzalez, A. Orellana, J. Retamales, H. Silva. A Rapid and Efficient Method for Purifying High Quality Total RNA from Peaches (*Prunus persica*) for Functional Genomics Analyses. *Biological Research* 38:83-88, 2005. More than **270** citations.
- [45] A. Jaimes, Javier Ruiz-del-Solar, R. Verschae, R. Baeza-Yates, C. Castillo, D. Yaksic, and E. Davis. On the Image Content of a Web Segment: Chile as a Case Study, *Journal of Web Engineering* 3 (2), 153–168, 2004.
- [46] R. Baeza-Yates. Information Retrieval in the Web: beyond current search engines, *International Journal on Approximated Reasoning* 34 (2-3), 97–104, 2003. More than **100** citations.
- [47] G. Navarro, R. Baeza-Yates, and J.M.A. Arcoverde. *Matchsimile*: A Flexible Approximate Matching Tool for Searching Proper Names. *JASIST* 54(1): 3-15, 2003.
- [48] R. Baeza-Yates and G. Navarro. XQL and Proximal Nodes (special issue on XML and Information Retrieval). *JASIST* 53(6): 504–514, 2002.
- [49] G. Navarro, R. Baeza-Yates, E. Sutinen and J. Tarhio. Indexing Methods for Approximate String Matching. *IEEE Data Engineering Bulletin* 24(4): 19–27, 2001. Almost **350** citations.
- [50] E. Chávez, G. Navarro, R. Baeza-Yates and J.L. Marroquín. Searching in Metric Spaces. *ACM Computing Surveys* 33(3):273–321, 2001. More than **1,960** citations and **7,330** downloads.
- [51] N. Ziviani, E.S. de Moura, G. Navarro and R. Baeza-Yates. Compression: A Key for Next-Generation Text Retrieval Systems. *IEEE Computer* 33(11):37-44 (cover feature), Nov 2000. More than **190** citations.
- [52] R. Baeza-Yates and G. Navarro. A Hybrid Indexing Method for Approximate String Matching *Journal of Discrete Algorithms (JDA)* 1(1), 205-239, 2000. More than **140** citations.
- [53] G. Navarro, E.S. de Moura, M.S. Neubert, N. Ziviani, R. Baeza-Yates. Adding Compression to Block Addressing Inverted Indexes. *Information Retrieval* 3(1): 49-77, 2000. More than **150** citations.
- [54] E.S. de Moura, G. Navarro, N. Ziviani and R. Baeza-Yates. Fast and flexible word searching on compressed text. *ACM Trans. on Information Systems* 18(2): 113-139, 2000. More than **340 citations** and **2,120** downloads.
- [55] G. Navarro, E. Barbosa, R. Baeza-Yates, W. Cunto, and N. Ziviani, Binary Searching with Non-uniform Costs and Its Application to Text Retrieval. *Algorithmica* 27 (2): 145–169, 2000.

- [56] 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 **150** citations.
- [57] R. Baeza-Yates and G. Navarro. Faster approximate string matching. *Algorithmica*, 23 (2):127–158, 1999. Almost **260** citations.
- [58] 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 **240** citations and **1,450** downloads.
- [59] 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 **200** citations and **2,650** downloads.
- [60] R. Baeza-Yates, E. Barbosa, and N. Ziviani. Hierarchies of indices for text searching. *Information Systems*, 21(6):497–514, 1996. **COMPAQ Award to the best Brazilian CS paper for that year.**
- [61] R. Baeza-Yates, J. Culberson, and G. Rawlins. Searching in the plane. *Information and Computation*, 106(2):234–252, Oct 1993. More than **610** citations.
- [62] R. Baeza-Yates and G.H. Gonnet. A new approach to text searching. *Communications of the ACM*, 35:74–82, Oct 1992. More than **1050** citations and **4,600** downloads.
- [63] R. Baeza-Yates. Searching Subsequences. *Theoretical Comp. Sci.*, 78:363–376, 1991. More than **130** citations.
- [64] R. Baeza-Yates. Improved string searching. *Software-Practice and Experience*, 19(3):257–271, 1989. Almost **130** citations.
- [65] R. Baeza-Yates. Expected behaviour of  $B^+$ -trees under random insertions. *Acta Informatica*, 26(5):439–472, 1989.

## 9.4 International Conferences

- [1] A. Sharif Rohani, R. Baeza-Yates. Measuring Bias. In *IEEE Big Data 2023*, Sorrento, Italy, Dec 2023.
- [2] S. Yfantidou, P. Sermpezis, A. Vakali, R. Baeza-Yates. Uncovering Bias in Personal Informatics. In *ACM Interact. Mob. Wearable Ubiquitous Technol.* 7(3): 139:1–139:30, Cancún, Mexico, Oct 2023. **Distinguished Paper Award** (8 of 205) and more than **1,790** downloads.
- [3] M. Trokhymovych, M. Aslam, A-J. Chou, R. Baeza-Yates, D. Sáez-Trumper. Fair Multilingual Vandalism Detection System for Wikipedia. In *KDD 2023*: 4981-4990, Long Beach, USA, Aug 2023.
- [4] B. Scarone, R. Baeza-Yates, E. Bernhardson. Understanding Search Behavior Bias in Wikipedia. In *BIAS 2023*: 134-146, Dublin, Ireland, Apr 2023.
- [5] R. Baeza-Yates, G. Delnevo. Exploration Trade-offs in Web Recommender Systems. In *IEEE Big Data 2022*, Osaka, Japan, Dec 2022.
- [6] R. Baeza-Yates, P. Villoslada. Human vs. Artificial Intelligence. In *IEEE CogMI 2022*: 40-48, Atlanta, USA, Dec 2023.
- [7] E. Graells-Garrido, R. Baeza-Yates. Bots don't Vote, but They Surely Bother! A Study of Anomalous Accounts in a National Referendum. In *ACM Web Science 2020*, Barcelona, Spain, Jun 2022.
- [8] I. Tsalouchidou, F. Bonchi, R. Baeza-Yates. Adaptive Community Search in Dynamic Networks. In *IEEE Big-Data 2020*, 987-995, 2020.
- [9] E. Graells-Garrido, R. Baeza-Yates, M. Lalmas. Every Colour You Are: Stance Prediction and Turnaround in Controversial Issues, In *ACM Web Science 2020*, Southampton, UK, Jun 2020.

- [10] R. Baeza-Yates, Z. Liaghat. Quality-efficiency trade-offs in machine learning for text processing. In *IEEE Big-Data 2017*, Boston, USA, Dec 2017.
- [11] Meike Zehlike, F. Bonchi, C. Castillo, Sara Hajian, Mohamed Megahed, R. Baeza-Yates. FA\*IR: A Fair Top-k Ranking Algorithm. In *ACM CIKM 2017*, Singapore, Nov 2017. More than **600** citations and **2,120** downloads.
- [12] R. Baeza-Yates, L. Rello. The Impact of Misspellings on Reading Comprehension for People with Dyslexia. In *International Dyslexia Association Annual Conference*, Atlanta, Nov 2017.
- [13] L. Recalde, David Nettleton, R. Baeza-Yates, Ludovico Boratto. Detection of Trending Topic Communities: Bridging Content Creators and Distributors. In *Hypertext 2017*, Prague, Czech Republic, Jul 2017.
- [14] L. Li, H. Deng, A. Dong, Y. Chang, R. Baeza-Yates, H. Zha. Exploring Query Auto-Completion and Click Logs for Contextual-Aware Web Search & Query Suggestion. In *WWW 2017*, pp. 539–548, Perth, Australia, 2017.
- [15] I. Tsalouchidou, G. De Francisci Morales, F. Bonchi, R. Baeza-Yates. Scalable Dynamic Graph Summarization. In *IEEE Big Data 2016*, pp. 1032–1039, Washington D.C., USA, Dec 2016.
- [16] L. Aiello, I. Arapakis, R. Baeza-Yates, X. Bai, N. Barbieri, A. Mantrach, F. Silvestri. The Role of Relevance in Sponsored Search. In *CIKM 2016*, pp. 185–194, Indianapolis, USA, Nov 2016.
- [17] M. Grbovic, D. Nemanja, V. Radosavljevic, F. Silvestri, R. Baeza-Yates, A. Feng, E. Ordentlich, L. Yang, G. Owens. Scalable Semantic Matching of Search Queries to Ads in Sponsored Search Advertising. In *SIGIR 2016*, pp. 375–384, Pisa, Italy, July 2016.
- [18] A. Zhang, A. Goyal, R. Baeza-Yates, Y. Chang, J. Han, C.A. Gunter, H. Deng. Towards Mobile Query Auto-Completion: A Mobile Application-Aware Approach. In *WWW 2016*, Montreal, Canada, April 2016.
- [19] E. Graells-Garrido, M. Lalmas, R. Baeza-Yates. Encouraging Diversity and Representation Awareness in Geographically Centralized Content. In *ACM Intelligent Users Interfaces*, Sonoma, CA, USA, March 2016.
- [20] E. Graells-Garrido, M. Lalmas, R. Baeza-Yates. Data Portraits and Intermediary Topics: Encouraging Exploration of Politically Diverse. In *ACM Intelligent Users Interfaces*, Sonoma, CA, USA, March 2016.
- [21] L. Li, H. Deng, A. Dong, Y. Chang, H. Zha, R. Baeza-Yates. Analyzing User’s Sequential Behavior in Query Auto-Completion via Markov Processes. In *ACM SIGIR 2015*, Santiago, Chile, August 2015.
- [22] R. Baeza-Yates, L. Rello, J. Dembowski. CASSA: A Context-Aware Synonym Simplification Algorithm. In *NAACL-HLT 2015*, Denver, USA, June 2015.
- [23] R. Baeza-Yates, P. Boldi, F. Chierichetti. Essential Pages are Easy to Find. In *WWW 2015*, Florence, May 2015. **Nominated to the best paper award.**
- [24] R. Baeza-Yates, Di Jiang, F. Silvestri, B. Harrison. Predicting the Next App that You Are Going to Use. In *WSDM 2015*, Shanghai, Feb 2015. More than **210** citations and **1,800** downloads.
- [25] D. Saez-Trumper, Y. Liu, R. Baeza-Yates, B. Krishnamurthy, A. Mislove. Beyond CPM and CPC: determining the value of users on OSNs. In *COSN’14: Second ACM conference on Online Social Networks*. ACM, Dublin, Ireland, Oct 2014.
- [26] L. Rello, R. Baeza-Yates. Evaluation of Dyswebxia, A Reading App Designed for People with Dyslexia. W4A 2014, Seoul, South Korea, Apr 2014. **Best paper runner up.**
- [27] C. Aslay, N. Barbieri, F. Bonchi, R. Baeza-Yates. Online Topic-Aware Influence Maximization Queries. In *EDBT 2014*, Athens, Greece, 295–306, Mar 2014. More than **120** citations.
- [28] G. Frances, X. Bai, B.B. Cambazoglu, R. Baeza-Yates. Improving the Efficiency of Multi-site Web Search Engines. In *WSDM 2014*, New York, USA, 3–12, Feb 2014.

- [29] J. Lehmann, M. Lalmas, G. Dupret, R. Baeza-Yates. Online Multitasking - What Web Analytics think we do and what we really do. In *CIKM 2013*, Burlingame, Oct 2013.
- [30] E. Yom-Tov, M. Lalmas, R. Baeza-Yates, G. Dupret, J. Lehmann, P. Donmez. Measuring Inter-Site Engagement. In *IEEE International Conference on Big Data*, Santa Clara, USA, Oct 2013.
- [31] L. Rello, R. Baeza-Yates. Good fonts for dyslexia. In *ACM ASSETS 2013*, Bellevue, USA, Oct 2013. More than **4,530** downloads.
- [32] H. Vahabi, M. Ackerman, D. Loker, R. Baeza-Yates, A. López-Ortiz. Orthogonal query recommendation. In *ACM RecSys 2013*, Hong Kong, China, pp. 33–40, Oct 2013.
- [33] L. Rello, R. Baeza-Yates, L. Dempere-Marco, H. Saggion. Frequent Words Improve Readability and Short Words Improve Understandability for People with Dyslexia. In *INTERACT*, Cape Town, South Africa, pp. 203–219, 2013. More than **200** citations.
- [34] L. Rello, R. Baeza-Yates, S. Bott, H. Saggion. Simplify or help?: text simplification strategies for people with dyslexia. In *W4A*, Rio de Janeiro, Brazil, May 2013. **Best paper award**, **530** citations and **1,200** downloads.
- [35] D. Saez-Trumper, G. Comarela, V. Almeida, R. Baeza-Yates, F. Benevenuto. Finding Trendsetters in Information Networks. In *ACM KDD 2012*, Beijing, China, Aug 2012. More than **100** citations.
- [36] L. Rello, G. Kanvinde, R. Baeza-Yates. Layout Guidelines for Web Text and a Web Service to Improve Accessibility for Dyslexics. In *W4A 2012*, Lyon, France, Apr 2012. More than **170** citations and **1,700** downloads.
- [37] L. Rello, R. Baeza-Yates, R. Mitkov. Elliphant: Improved Automatic Detection of Zero Subjects and Impersonal Constructions in Spanish. In *EACL 2012*, Avignon, France, Apr 2012, 706–715.
- [38] R. Baeza-Yates, Simon Jonassen. Modeling Static Caching in Web Search Engines. In *ECIR 2012*, Barcelona, Spain, Springer LNCS 7224, Apr 2012, 436–446.
- [39] C. González-Caro, R. Baeza-Yates. A Multi-faceted Approach to Query Intent Classification. In *SPIRE 2011*, Pisa, Italy, Oct 2011, Springer LNCS 7024, 368–379.
- [40] V. Jethava, L. Calderón-Benavides, C. Bhattacharyya, D. Dubhashi, R. Baeza-Yates. Scalable multi-dimensional user intent identification using tree structured distributions. In *SIGIR 2011*, Beijing, China, Jul 2011.
- [41] O. Alonso, R. Baeza-Yates. Design and Implementation of Relevance Assessments using Crowdsourcing. In *ECIR 2011*, Dublin, Ireland, Springer LNCS 6611, Apr 2011, 153–164. More than **200** citations.
- [42] S. Ding, J. Attenberg, R. Baeza-Yates, T. Suel. Batch Query Processing for Web Search Engines. In *WSDM 2011*, ACM Press, Hong Kong, Feb 2011, 137–146.
- [43] B.B. Cambazoglu, H. Zaragoza, R. Baeza-Yates. Web Search Solved? All Result Rankings the Same? In *ACM CIKM 2010*, Toronto, Nov 2010, 529–538.
- [44] A.P. Francisco, R. Baeza-Yates, A.L. Oliveira. Mining Large Query Induced Graphs towards a Hierarchical Query Folksonomy. *SPIRE 2010*, Springer LNCS 6393, Los Cabos, Oct 2010, 237–242.
- [45] O. Alonso, M. Gertz, R. Baeza-Yates. Temporal Analysis of Document Collections: Framework and Applications. *SPIRE 2010*, Springer LNCS 6393, Los Cabos, Oct 2010, 290–296.
- [46] I. Bormino, D. Donato, R. Baeza-Yates. Coniunge et Impera: Multiple-graph mining for Query-log analysis. In *PKDD'2010*, Barcelona, Spain, Sept 2010, 168–183.
- [47] B.B. Cambazoglu, E. Varol, E. Kayaaslan, C. Aykanat, R. Baeza-Yates. Query Forwarding in Geographically Distributed Search Engines. In *SIGIR'2010*, Geneva, Switzerland, July 2010, 90–97.
- [48] R. Baeza-Yates, A. Gionis, F. Junqueira, V. Plachouras, Luca Telloli. On the Feasibility of Multi-Site Web Search Engines. In *ACM CIKM 2009*, Hong Kong, China, Nov 2009, 425–434. **Best paper award**.

- [49] O. Alonso, M. Gertz, R. Baeza-Yates. Clustering and Exploring Search Results using Timeline Constructions. In *ACM CIKM 2009*, Hong Kong, China, Nov 2009. More than **180** citations and **1,430** downloads.
- [50] R. Baeza-Yates, C. Middleton, C. Castillo. The Geographical Life of Search. In *ACM/IEEE/WIC Web Intelligence*, Milano, Italy, September 2009.
- [51] R. Baeza-Yates, V. Murdock, C. Hauff. Efficiency Trade-offs in Two-tier Web Search Systems. In *SIGIR 2009*, Boston, USA, July 2009, 163-170.
- [52] B.B. Cambazoglu, V. Plachouras, R. Baeza-Yates. Quantifying Performance and Quality Gains in Distributed Web Search Engines. In *SIGIR 2009*, Boston, USA, July 2009, 411-418.
- [53] Alvaro Pereira, R. Baeza-Yates, N. Ziviani, Jesús Bisbal. A Model for Fast Web Mining Prototyping. In *ACM WSDM 2009*, Barcelona, Spain, Feb 2009. More than **1,220** downloads.
- [54] Claudia Hauff, V. Murdock, R. Baeza-Yates. Improved Query Difficulty Prediction for the Web. In *CIKM 2008*, Napa Valley, USA, Oct 2008. More than **110** citations and **1,010** downloads.
- [55] G. Skobeltsyn, F. Junqueira, V. Plachouras, R. Baeza-Yates. ResIn: A Combination of Result Caching and Index Pruning for High-performance Web Search Engines. In *SIGIR 2008*, Singapore, Jul 2008.
- [56] R. Baeza-Yates, Álvaro R. Pereira Jr., N. Ziviani. Genealogical trees on the Web: a search engine user perspective. In *WWW 2008*, Beijing, China, Apr 2008, 367-376.
- [57] B. Poblete, R. Baeza-Yates. Query-sets: using implicit feedback and query patterns to organize Web documents. In *WWW 2008*, Beijing, China, Apr 2008, 41-50. More than **100** citations.
- [58] R. Baeza-Yates, F. Junqueira, V. Plachouras, H.F. Witschel. Admission Policies for Caches of Search Engine Results, In *SPIRE 2007*, Springer LNCS 4726, Santiago, Chile, Oct 2007, 74-85.
- [59] R. Baeza-Yates, A. Tiberi. Extracting Semantic Relations from Query Logs. In *ACM KDD 2007*, San Jose, California, USA, August 2007, 76–85. More than **410** citations and almost **3,400** downloads.
- [60] R. Baeza-Yates, A. Gionis, F. Junqueira, V. Murdock, V. Plachouras, F. Silvestri. The Impact of Caching on Search Engines. In *ACM SIGIR*, Amsterdam, Netherlands, July 2007, 183–190. More than **290** citations and **1,430** downloads.
- [61] R. Baeza-Yates, L. Calderón-Benavides, C. González-Caro. The Intention Behind Web Queries. *SPIRE 2006*, Springer LNCS 4209, Glasgow, Scotland, Oct 2006. Almost **280** citations.
- [62] R. Baeza-Yates, P. Boldi, C. Castillo. Generalizing PageRank: Damping Functions for Link-Based Ranking Algorithms. In *ACM SIGIR 2006*. Seattle, USA, Aug 2006. More than **140** citations and **1,450** downloads.
- [63] L. Becchetti, C. Castillo, D. Donato, S. Leonardi, R. Baeza-Yates. Link-Based Characterization and Detection of Web Spam. In *Adversarial Information Retrieval on the Web*. Seattle, USA, Aug 2006. More than **240** citations.
- [64] R. Baeza-Yates, C. Castillo, M. Marín, A. Rodríguez. Crawling a Country: Better Strategies than Breadth-First for Page Ordering. In *WWW 2005*, Chiba, Japan, May 2005. More than **190** citations and **1,320** downloads.
- [65] R. Baeza-Yates, J. Ruiz-del-Solar, R. Verschae, C. Hurtado, C. Castillo. Content-based Image Retrieval and Characterization on Specific Web Collections. In *CIVR'04:w*, Springer LNCS 3115, Dublin, Ireland, July 2004.
- [66] R. Baeza-Yates. A Fast Set Intersection Algorithm for Sorted Sequences, In *15th Combinatorial Pattern Matching 2004*, Springer LNCS 3109, Istanbul, Turkey, July 2004. More than **150** citations.
- [67] R. Baeza-Yates, C. Hurtado, M. Mendoza. Query Recommendation Using Query Logs in Search Engines. In *EDBT Workshop on Clustering Information over the Web*, Heraklion, Crete, Greece, March 14-18, 2004, Revised Selected Papers. Springer LNCS 3268, pp. 588-596. More than **820** citations.

- [68] F. Ccheda, R. Baeza-Yates. An Optimistic Model for Searching Web Directories. In *26th European Conference on IR Research*. Springer LNCS 2997, pp 364-377, March 2004.
- [69] R. Baeza-Yates, F. Saint-Jean. A Three Level Search Engine Index based in Query Log Distribution. In *SPIRE 2003*, Springer LNCS 2857, Manaus, Brazil, Oct 2003. More than **110** citations.
- [70] R. Baeza-Yates, C. Castillo. Balancing Volume, Quality and Freshness in Web Crawling. In *Hybrid Intelligent Systems 2002*, IOS Press, Santiago, Dec 2002.
- [71] R. Baeza-Yates, F. Saint-Jean, C. Castillo. Web Dynamics, Age and Page Quality, In *SPIRE 2002*, Springer LNCS 2476, Lisbon, Portugal, September 2002. More than **140** citations.
- [72] R. Baeza-Yates, C. Castillo, Relating Web Characteristics with Link Based Web Page Ranking, In *SPIRE 2001*, IEEE CS Press, Laguna San Rafael, Chile, pp. 21-32, Nov 2001. More than **50** citations.
- [73] C. Badue, R. Baeza-Yates, B. Ribeiro-Neto, N. Ziviani. Distributed Query Processing Using Partitioned Inverted Files, In *SPIRE 2001*, IEEE CS Press, Laguna San Rafael, Chile, pp. 10-20, Nov 2001. More than **170** citations.
- [74] R. Baeza-Yates, G. Navarro. Fast approximate string matching in a dictionary. In *SPIRE'98*, Sta. Cruz de la Sierra, Bolivia, Sep 1998. IEEE CS Press, 14-22. More than **130** citations.
- [75] E. Moura, G. Navarro, N. Ziviani, R. Baeza-Yates. Fast searching on compressed text allowing errors. In *21th SIGIR*, Melbourne, Australia, 1998, 298-306. More than **60** citations.
- [76] R. Baeza-Yates, G. Navarro. Block-addressing indices for approximate text retrieval. In *Proc. CIKM'97*, pp. 1-8, Las Vegas, USA, 1997.
- [77] R. Baeza-Yates, G. Navarro. A faster algorithm for approximate string matching. In *Combinatorial Pattern Matching*, LNCS 1075, pp. 1-23, Irvine, CA, Jun 1996. More than **160** citations.
- [78] R. Baeza-Yates. Visualization of large answers in text databases. In *Advanced User Interfaces (AVI'96)*, pp. 101-107, Gubbio, Italy, May 1996. ACM Press. More than **210** citations.
- [79] E. Barbosa, G. Navarro, R. Baeza-Yates, Chris Perleberg, N. Ziviani. Optimized binary search and text retrieval. In *European Symposium on Algorithms*, Corfu, Greece, September 1995.
- [80] G. Navarro, R. Baeza-Yates. A language for queries on structure and contents of textual databases. In *18th ACM SIGIR*, pp. 93-101, Seattle, WA, USA, July 1995. More than **120** citations.
- [81] R. Baeza-Yates, W. Cunto, U. Manber, S. Wu. Proximity matching using fixed queries trees. In *5th Combinatorial Pattern Matching*, LNCS 807, pp. 198-212, Asilomar, CA, June 1994. Springer. Almost **320** citations.
- [82] R. Baeza-Yates, C.H. Perleberg. Fast and practical approximate pattern matching. In *Combinatorial Pattern Matching*, Springer LNCS 644, pp. 185-192, Tucson, AZ, April/May 1992. More than **180** citations.
- [83] R. Baeza-Yates, G.H. Gonnet, M. Régnier. Analysis of Boyer-Moore-type string searching algorithms. In *ACM-SIAM SODA'90*, pp. 328-343, San Francisco, Jan 1990. More than **1,700** downloads.
- [84] R. Baeza-Yates, G.H. Gonnet. A new approach to text searching. In *Proc. of 12th ACM SIGIR*, pp. 168-175, Cambridge, Mass., June 1989. More than **950** downloads.
- [85] R. Baeza-Yates, G.H. Gonnet. Efficient text searching of regular expressions. In *ICALP'89*, Springer LNCS 372, pp. 46-62, Stresa, Italy, July 1989.

## 10 Patents

1. Ricardo Baeza-Yates and Barbara Poblete. User Query Data Mining and Related Techniques, Yahoo! Inc. and Univ. of Chile, September 2006. USA Patent #7,617,208, granted Nov 10, 2009.
2. Ricardo Baeza-Yates. Techniques for Searching Future Events, Yahoo! Inc. and Univ. of Chile, August 2006. USA Patent #7,668,813, granted Feb 23, 2010.
3. Luciano Barbosa, Flavio Junqueira, Vassilis Plachouras, Ricardo Baeza-Yates. Method and System for Quantifying the Quality of Search Results based on Cohesion, Yahoo! Inc., December 2007. USA patent #7,720,870, granted May 18, 2010.
4. Ricardo Baeza-Yates, Aristides Gionis, Flavio Junqueira and Vassilis Plachouras. System and Method for Caching Posting Lists, Yahoo! Inc., October 2007. USA Patent #7,890,488, granted Feb 5, 2011.
5. Ricardo Baeza-Yates, Alessandro Tiberi. Extracting Semantic Relations from Query Logs, Yahoo! Inc., December 2007. USA Patent #7,895,235, granted Feb 22, 2011.
6. Ricardo Baeza-Yates, Aristidis Gionis, Flavio Junqueira, Vassilis Plachouras and Luca Telloli. System and Methodology for a Multi-Site Search Engine, Yahoo! Inc., October 2008. USA patent #8,095,545, granted January 10, 2012.
7. Barbara Poblete, Ricardo Baeza-Yates. Classifying Documents using Implicit Feedback and Query Patterns, Yahoo! Inc., July 2008. USA patent # 8,645,369, granted February 4, 2014.
8. Barbara Poblete, Ricardo Baeza-Yates. Methods for Web Site Analysis, Yahoo! Inc., April 2010. USA patent #8,751,632, granted June 10, 2014.
9. Fabrizio Silvestri, Ricardo Baeza-Yates, Beverly Harrison, Di Jiang. Predicting the Next Application that You are going to Use on Aviate, Yahoo! Inc., December 2014. USA Patent 9971972B2, granted May 15, 2018.
10. Lizi Zhang, Hongbo Deng, Amit Goyal, Yi Chang, Ricardo Baeza-Yates. Computerized System and Method for Search Query Auto-completion, Yahoo! Inc, September 2015. USA Patent #10049149B2, granted August 14, 2018. More than **100** citations.
11. Manjana Chandrasekharan, Keiko Horiguchi, Amanda Stent, Ricardo Baeza-Yates, Jeffrey Kuwano, Achint Thomas, Yi Chang. Audio Verification, Yahoo! Inc., September 2015. USA Patent #10277581B2, granted April 30, 2019.
12. Ricardo Baeza-Yates, Berkant Barla Cambazoglu, Darshan M. Shankaralingappa, Matteo Catena. Pipeline for Document Scoring, NTENT, September 2019. USA Patent #11436235, granted September 6, 2022.
13. Ricardo Baeza-Yates, Juan Garay, Sebastián Acuña-Keller, Andrés Villavicencio-Theodoluz, Sebastián Gálvez-Ortiz. Cryptographic systems and methods for providing services to authenticated users, Khosla Ventures, October 2023. USA Patent #US12137166B2, granted November 2024.
14. María José Margabit Sagredo, Ricardo Baeza-Yates, Mary Isabel Buenaventura Keane, Soffi Arrigoni Quintela, Manuel Ignacio Cartagena Herrera, Vladimir Giovanni Araujo Vásquez. Machine learning techniques for detecting, measuring, and mitigating bias within textual content. Theodora AI, December 2024. USA Patent #US12579364B1, granted March 17, 2026.

## 11 Mentoring

I have mentored dozens of postdoctoral students in Spain and USA and supervised 33 Ph.D. students in Brazil, Chile, Italy, México, Spain (66%), Switzerland, and USA, where 52% were women and 67% were from developing countries. Similarly, I have graduated 23 master students in Chile (65%), Spain and USA, where 30% were women and 70% were from developing countries. I have also supervised the final engineering projects of 20 students, most of them in Chile.

## 12 Teaching Record

I have taught the following undergraduate courses in Spanish at Universidad de Chile (1985-2004) and Universitat Pompeu Fabra, Spain (2005-2024) (most recent first, although many were taught more than once): Information Retrieval and Web Analytics, Information Retrieval (also taught in English at University of Canterbury, New Zealand, in 2004), Data Structures, Design and Analysis of Algorithms, Bioinformatics, Introduction to Computer Programming, Software Project, Object Oriented Programming, Software Workshop on Graphical Interfaces, Computer Architecture, Automata Theory and Formal Languages, Design and Construction of Compilers, Information Systems, Programming Languages, Introduction to Computer Graphics, Numerical Algorithms, Digital Systems, Operating Systems, and Software Tools.

I have taught the following graduate courses (university, language and years are indicated): Responsible AI (Technical University of Vienna, English, 2024), Algorithms (Northeastern University, English, 2017-19), Web Retrieval and Data Mining (Universitat Pompeu Fabra, English, 2010-14), Web Search (Univ. Carlos III, Spain, Spanish, 2005-07), Advanced Data Structures, Problem Solving Seminar, Visualization Seminar and Text Searching Seminar (Universidad de Chile, Spanish, 1990-95).

My article on teaching algorithms published in SIGACT News in 1995 has more than 1,300 downloads from the ACM Digital Library.

## 13 Editorial Boards

I am editor of the following journals: Information Systems (Elsevier, Netherlands) since 2002, Computing (Springer, Germany) since 2013, Journal of Web Engineering (USA) since 2002, Theoretical Informatics and Applications - RAIRO (France) since 1998. Grammars (Netherlands) since 2004, Journal of the SBC (Brazil) since 2002, First Monday (USA) since 2001, and El Profesional de la Información (Spain) since 2003.

I was also member of the editorial boards of ACM Transactions on Information Systems (ACM, USA) from 2002 to 2014, Information Processing & Management (Elsevier, Netherlands) from 2005 to 2011 (since 2007 as Associated Editor), VLDB Journal (Springer, Germany) from 2001 to 2007, Computer Science Review (Elsevier, Netherlands) from 2013 to 2014, Information Technology for Development (Netherlands), International Journal of Web Information Systems (UK), Journal of Digital Management (India-South Korea) from 2003 to 2005, Computación y Sistemas (México), Novática (Spain), Revista de la SCCC (Chile) and Revista Colombiana de Informática (Colombia).

## 14 Selected Videos of Keynotes and Seminars

Here we give a sample of videos in the last 10 years from more than 100 talks, podcasts, interviews, etc.:

- “• The Limitations of Data, ML & Us”. IIIT Bangalore, India, Nov 2025. URL: <https://youtu.be/rVjcMFOOXKA?t=860>
- “The Dangers of Artificial Intelligence” (in Spanish with English subtitles). The AI Club, podcast with Jon Hernández, Aug 2025. YouTube: <https://www.youtube.com/watch?v=q3Vym6Mz7QE&pp=ygULYmFlmEteWF0ZXM%3D> More than **65 thousand** views.
- “From Semantic to Predictive Search: Evolution or Involution? SEO Week, New York, May 2025. URL: <https://ipullrank.com/seo-week-2025-ricardo-baeza-yates>
- “Responsible AI” (in Spanish). 50th Latin American Congress, Bahia Blanca, Argentina, Aug 2024. YouTube: <https://www.youtube.com/watch?v=VwF4GRPaZBA>. More than **1,700** views in two different versions.
- “The Shades of AI and its Challenges” (in Spanish), Congress of the Future, Santiago, Chile, Jan 2024. YouTube: <https://www.youtube.com/watch?v=ms3MmgP7Fhw>. More than 800 views.

- “Responsible AI: Challenges & Recommendations”, Digital Humanism Seminar, Vienna, Austria, Jun 2023. YouTube: <https://www.youtube.com/watch?v=UbgxJyGNsS>. More than **1,000** views.
- “The Dangers of Generative AI and Its Impact in Higher Education” (in Spanish), Master Class, Universidad de Chile, May 2023. YouTube: <https://www.youtube.com/watch?v=a2Ttj1thdYA>. More than **2,200** views.
- “Ethics in AI: A Challenging Task”, The GobLab Ethics in AI Global Perspective, Oct 2021. <https://www.youtube.com/watch?v=aQcf6TaE6Gg>. More than 500 views.
- “Bias on Search and Recommender Systems”, Keynote at RecSys 2020, Rio de Janeiro, Brazil, Jun 2020. YouTube: [https://www.youtube.com/watch?v=8zetbdx4\\_08&t=60s](https://www.youtube.com/watch?v=8zetbdx4_08&t=60s). More than **2,500** views.
- “Bias on the Web”, Mozilla’s Emerging Technologies Speaker Series, May 2020. YouTube: <https://www.youtube.com/watch?v=brr8cZHf2ec&t=69s>. More than 600 views.
- “Data and Algorithmic Bias in the Web”, San Francisco Bay ACM Professional Chapter, USA, Sep 2017. YouTube: [https://www.youtube.com/watch?v=W\\_mqrGAp3aU](https://www.youtube.com/watch?v=W_mqrGAp3aU). More than 850 views.
- Bias in the Web, Gödel Annual Seminar, Wien, Austria, Jun 2017. YouTube: <https://www.youtube.com/watch?v=CguWbbQeruM>. More than 350 views.
- “Who Owns the Data in this Digital Era?” (in Spanish), Congress of the Future, Santiago, Chile, Jan 2016. YouTube: <https://www.youtube.com/watch?v=fFlkOKUa2M0>. More than 600 views.
- “Does Big Data Empowers the Web?” (in Spanish), TEDx Barcelona, July 2015. YouTube: <https://www.youtube.com/watch?v=JvO1P8yGVqA> (more than **11,000** views).

During 2020 to 2022 I was host of two online programs in Chile, the first about understanding the data of the pandemia and the second about data and AI. These weekly programs were watched by more than **100 thousand** people.

## 15 Service at Conferences

- Member of the steering committees of SPIRE since 1998 (co-founder) and AMW since 2006 (co-founder). I was part of the steering committee of ACM WSDM from 2008 to 2017 (co-founder), ACM CIKM from 2014 to 2020, ICWSM from 2014 to 2018, OAIR from 2011 to 2015, ESSIR from 2009 to 2014, and LATIN from April 1995 to March 2006.
- Invited speaker or tutorial lecturer at many conferences including ACM SIGIR, ACM SIGMOD, ACM WWW, ACM WSDM, VLDB, ACM CIKM, ACM Recsys, SDM, ACM Hypertext, ACM Web Science, ECDL, ICDE, ECIR, ECML/PKDD, ESWC, UMAP, ICWE, EDBT, ISSI, SPIRE, CIAC, CLEF, NLDB, SAMT, SSDBM, SISAP, ICTIR, DSAI, ICWI, WISE, CRWIG, Simultools, SRDS, IDA, LA-WEB, etc.; as well as in many workshops.
- Conference co-chair of SPIRE 2025, London, United Kingdom, Sep 2025; Conference co-chair of IEEE Big-Data 2024, Washington DC, USA, Dec 2024; Conference co-chair of SPIRE 2024, Puerto Vallarta, Mexico, Sep 2024; Conference co-chair of ACM KDD 2024, Barcelona, Spain, Aug 2024; Co-chair of the ACM Fairness, Accountability and Transparency Conference (FAcT), Rio de Janeiro, Brazil, Jun 2024; Conference co-chair of DESIRES 2022, San Jose, USA, Aug 2022. Conference co-chair of ACM Web Science 2022, Barcelona, Spain, Jun 2022. Co-chair of The Web Conference 2019 (formerly WWW), San Francisco, USA, May 2019; Conference co-chair of IEEE Big Data 2017, Boston, USA, Dec 2017; Co-organizing chair of ESSIR 2017, Barcelona, Spain, Sep 2017; Conference co-chair of IEEE ICDM 2016, Barcelona, Spain, Dec 2016; Conference chair of ACM SIGIR 2015, Santiago, Chile, Aug 2015; Conference chair of LA-WEB 2014, Ouro

Preto, Brazil, Oct 2014; Conference chair of AMW 2014, Cartagena de Indias, Colombia, Jun 2014; Conference chair of LA-WEB 2012, Cartagena de Indias, Colombia, Oct 2012; Conference chair of ECIR 2012, Barcelona, Spain, Apr 2012; Co-organizing chair of ESSIR 2009, Padova, Italy, Sep 2009; Conference chair of ACM WSDM 2009, Barcelona, Spain, Feb 2009; Conference chair of LA-WEB 2008, Vila Velha, Brazil, Oct 2008; Conference chair of SPIRE 2007 and LA-WEB 2007 (collocated) Santiago, Chile, Oct 2007; Conference co-chair of ACM SIGIR 2015 in Bahia, Brazil, Aug 2005; Conference chair of the first Latin American Web conference (LA-WEB 2003) in Santiago, Chile, Nov 2003; Conference co-chair of IEEE PROMSMmNet'97, Santiago, Chile, Nov 1997; Conference co-chair of LATIN'95, Viña del Mar, Chile, Apr 1995; Conference chair of the 2nd South American Workshop on String Processing, Viña del Mar, Chile, Apr 1995; and Conference chair of the XI International Conference of the Chilean Computer Science Society, Santiago, Chile, Oct 1991.

- Other services such as program committee chair, senior member or member; workshop organizer; and seminars given in more than 50 countries, have been omitted.