

Curriculum Vitæ
Diego G. Arroyuelo Billiardi

Información de Contacto

Académico Jornada Completa

Departamento de Informática,
Universidad Técnica Federico Santa María.
E-mail: darroyue@inf.utfsm.cl

Información Personal

- **Nombre:** Diego Gastón Arroyuelo Billiardi.
- **RUT:** 21.608.440-3.
- **Título:** Doctor en Ciencias, Mención Computación, Universidad de Chile.

Educación

Estudios Realizados

- (Julio 2008 – Marzo 2009) Estadía postdoctoral en la David R. Cheriton School of Computer Science, University of Waterloo, Canadá.
Profesores de contacto: Ian Munro y Alejandro López-Ortíz.
- (Agosto 2003 – Marzo 2009) Doctor en Ciencias, Mención Computación, Departamento de Ciencias de la Computación, Universidad de Chile.
Tesis de Doctorado: Lempel-Ziv Compressed Full-Text Self-Indexes.
Profesor Guía: Gonzalo Navarro (Departamento de Ciencias de la Computación, Universidad de Chile)
- (1997–2003) Licenciatura en Ciencias de la Computación (5 años de cursos más tesis), Universidad Nacional de San Luis, San Luis, Argentina.

Tesis: Árbol de Aproximación Espacial Dinámico con Pivotes Agregados.

Profesores Guía: Gonzalo Navarro (Universidad de Chile) y Nora Reyes (Universidad Nacional de San Luis).

Becas

- (Septiembre de 2007 – Agosto de 2008): Beca de *Yahoo! Research Latin America* para finalizar el doctorado en la Universidad de Chile.
- (Marzo de 2004 – Julio de 2007): Beca *CONICYT* para estudiantes de doctorado, para realizar el doctorado en la Universidad de Chile.
- (2006) Beca *CONICYT* para la asistencia a congresos en el extranjero, para asistir a la conferencia internacional *Combinatorial Pattern Matching (CPM'06)*.
- (2003): Beca del *Núcleo Milenio para la Investigación de la Web*, Universidad de Chile, para realizar el primer semestre del doctorado en la Universidad de Chile.
- (2003) Beca de arancel otorgada por la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile para el primer semestre del doctorado.

Investigación

Experiencia en Investigación

- (1 Abril 2009 – actualidad) Investigador en el laboratorio *Yahoo! Research Latin America*, Santiago, Chile.
- (1 Julio 2008 – 31 Marzo 2009) Visiting Scholar en el *Algorithms and Complexity Group*, *David R. Cheriton School of Computer Science*, *University of Waterloo*, Waterloo, Canadá. [Estadía postdoctoral]
Profesores: J. Ian Munro y Alejandro López Ortiz.

Proyectos de Investigación Asignados

Proyectos Nacionales

- Investigador Principal en proyecto FONDECYT de Iniciación en la Investigación 2012.
Título: Improved Data Compression for Text Search Engines
Número de Proyecto: 11121556
Duración: 2 años.
- Co-investigador en proyecto FONDECYT Regular 2011.
Título: Compact Data Structures for Information Retrieval
Investigador principal del proyecto: Gonzalo Navarro
Duración del Proyecto: 3 años.

Proyectos Internos Universidad Técnica Federico Santa María

- Investigador Principal en proyecto interno DGIP de Iniciación 2012.
Título: Representaciones Prácticas de Estructuras de Datos Sucintas y Comprimidas
Duración: 2 años

Publicaciones

Journals Internacionales Indexados por ISI

1. D. Arroyuelo, F. Claude, S. Maneth, V. Mäkinen, G. Navarro, K. Nguyen, J. Sirén, and N. Välimäki. Fast In-Memory XPath Search using Compressed Indexes. To appear in Software: Practice and Experience. *Wiley*, 2013.
2. D. Arroyuelo, G. Navarro y K. Sadakane. Stronger Lempel-Ziv Based Compressed Text Indexing. *Algorithmica* 62(1):54-101. *Springer*, 2012.
3. D. Arroyuelo, V. Gil-Costa, S. González, M. Marín y M. Oyarzún. Distributed Search based on Self-Indexed Compressed Text. *Information Processing and Management* 48:819–827. *Elsevier*, 2012.
4. D. Arroyuelo, F. Claude, R. Dorrigiv, S. Durocher, M. He, A. López-Ortiz, J. I. Munro, P. K. Nicholson, A. Salinger y M. Skala. Untangled Monotonic Chains and Adaptive Range Search. *Theoretical Computer Science* 412(32):4200–4211. *Elsevier*, 2011.
5. D. Arroyuelo y G. Navarro. Space-Efficient Construction of Lempel-Ziv Compressed Text Indexes. *Information and Computation*, 209(7):1070-1102. *Elsevier*, 2011.

Journals Internacionales Indexados por Scopus

1. D. Arroyuelo y G. Navarro. Practical Approaches to Reduce the Space Requirement of Lempel-Ziv-Based Compressed Text Indices. *ACM Journal on Experimental Algorithmics* (JEA), Volumen 15, artículo 1.5. Diciembre de 2010. 54 páginas.

Conferencias Internacionales con Referato, con Proceedings Publicado por ACM, IEEE y SIAM

1. Diego Arroyuelo, Senén González, Mauricio Oyarzún y Victor Sepulveda. Run-Length Compressed Inverted Indexes with Higher Compression Ratio and Faster Query Processing. En *Proc. of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'13)*. Por aparecer.
2. Diego Arroyuelo, Senén González, Mauricio Marin, Mauricio Oyarzún y Torsten Suel. To Index or not to Index: Time-Space Trade-offs in Search Engines with Positional Ranking Functions. En *Proc. of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'12)*, páginas 255–264.

3. D. Arroyuelo, R. Cánovas, G. Navarro y K. Sadakane. Succinct Trees in Practice. En *Proc. 11th Workshop on Algorithm Engineering and Experiments (ALENEX'10)*, páginas 84–97. SIAM Press.
4. D. Arroyuelo, F. Claude, S. Maneth, V. Mäkinen, G. Navarro, K. Nguyen, J. Sirén y N. Välimäki. Fast In-Memory XPath Search Using Compressed Indexes. En *Proc. 26th IEEE International Conference on Data Engineering (ICDE'10)*, páginas 417–428. IEEE.

Conferencias Internacionales con Referato, con Proceedings Publicado por Springer

1. D. Arroyuelo, S. González y M. Oyarzún. Compressed Self-Indices Supporting Conjunctive Queries on Document Collections. *String Processing and Information Retrieval*, Lecture Notes in Computer Science, 6393:43–54. Springer, 2010.
2. D. Arroyuelo, F. Claude, R. Dorigiv, S. Durocher, M. He, A. López-Ortiz, J. I. Munro, P. K. Nicholson, A. Salinger y M. Skala. Untangled Monotonic Chains and Adaptive Range Search. *Algorithms and Computation*, Lecture Notes in Computer Science, 5878:203–212. Springer, 2009.
3. D. Arroyuelo. An Improved Succinct Representation for Dynamic k -ary Trees. *Combinatorial Pattern Matching*, Lecture Notes in Computer Science, 5029:277–289. Springer, 2008.
4. D. Arroyuelo y G. Navarro. A Lempel-Ziv Text Index on Secondary Storage. *Combinatorial Pattern Matching*, Lecture Notes in Computer Science, 4580:83–94. Springer, 2007.
5. D. Arroyuelo, G. Navarro y K. Sadakane. Reducing the Space Requirement of LZ-index. *Combinatorial Pattern Matching*, Lecture Notes in Computer Science, 4009:319–330. Springer, 2006. [ISI]
6. D. Arroyuelo y G. Navarro. Space-efficient Construction of LZ-index. *Algorithms and Computation*, Lecture Notes in Computer Science, 3827:1143–1152. Springer, 2005. [ISI]
7. D. Arroyuelo, F. Muñoz, G. Navarro y N. Reyes. Memory-adaptative dynamic Spatial Approximation Trees. *String Processing and Information Retrieval*, Lecture Notes in Computer Science, 2857:360–368. Springer, 2003. [ISI]

Conferencias Internacionales con Referato, con Proceedings Publicado por otros Editores

1. D. Arroyuelo y G. Navarro. Smaller and Faster LZ-indices. En *Proc. 18th International Workshop on Combinatorial Algorithms (IWOCA'07)*, páginas 11–20. College Publications, UK.

Publicaciones en otros Medios

- Opinión en la nota “La evolución de los buscadores: son más precisos, personalizados y enriquecen los datos”. Diario el Mercurio, Domingo 14 de Julio de 2013.
- Indexación y Compresión para Motores de Búsqueda. Por aparecer en “Bits de Ciencia”. 2013.

- La amenaza de los virus indestructibles. La Segunda Online, 30 de agosto de 2011.
<http://www.lasegunda.com/Noticias/CienciaTecnologia/2011/08/676828/La-amenaza-de-los-virus-indestructibles>

Manuscritos en Preparación

- D. Arroyuelo, C. Bonacic, V. Gil-Costa, M. Marin y G. Navarro. Distributed Text Search using Suffix Arrays. Enviado a *Parallel Computing* (ISI).
- D. Arroyuelo, P. Davoodi y S. S. Rao. Succinct Dynamic Cardinal Trees. En preparación. Por enviarse a *Algorithmica* (ISI).
- D. Arroyuelo. Dynamic Pivoting Data Structures based on Spatial Approximation. En preparación. Por enviarse a *Information Processing and Management* (ISI).

Estadías de Investigación

1. Estadía corta de investigación en *Polytechnic Institute, New York University*, Brooklyn. 30 de Abril al 25 de Mayo de 2012. New York, Estados Unidos.
Profesor: Torsten Suel.
2. Estadía corta de investigación en *NICTA*, Sydney, 17 al 28 de Noviembre de 2008. Sydney, Australia.
Profesor: Sebastian Maneth.
3. Estadía corta de investigación en *Department of Computer Science and Software Engineering, University of Melbourne*, 12 al 29 de Noviembre de 2007. Melbourne, Australia.
Profesor: Alistair Moffat.
4. Estadía corta de investigación en Algorithms and Complexity Group, *David R. Cheriton School of Computer Science, University of Waterloo*, 12 al 26 de Julio de 2007. Waterloo, Canadá.
Profesor: J. Ian Munro.
5. Estadía corta de investigación en *Dipartimento di Informatica, Università di Pisa*, 16 al 23 de Junio de 2006. Pisa, Italia.
Profesor: Paolo Ferragina.
6. Estadía corta de investigación en el *Department of Computer Science and Communication Engineering, Kyushu University*, 6 al 13 de Diciembre de 2005. Fukuoka, Japón.
Profesor: Kunihiko Sadakane.

Chair de Comités de Programa

Conferencias Internacionales

- PC Chair de la *32nd International Conference of the Chilean Computer Science Society (SCCC 2013)*, Temuco, Chile, 13–15 Noviembre de 2012.

Participación en Comités de Programa

Conferencias Internacionales

- *23rd Annual Symposium on Combinatorial Pattern Matching (CPM 2012)*, Helsinki, Finlandia, Julio de 2012.
PC Chairs: Juha Kärkkäinen y Jens Stoye.
- *XXXI International Conference of the Chilean Computer Science Society (SCCC 2012)*, Valparaíso, Chile, Noviembre de 2012.
PC Chairs: Rosa Alarcón y Pablo Barceló.
- *9th Latin American Theoretical Informatics Symposium (LATIN 2010)*, Oaxaca, México, Abril de 2010.
PC Chair: Alejandro López Ortíz.

Conferencias Nacionales y Regionales

- Lo mejor de lo nuestro 2013 (*LMLN 2013*).
- Lo mejor de lo nuestro 2012 (*LMLN 2012*).
- *XXXVIII Conferencia Latinoamericana en Informática (CLEI 2012)*, Medellín, Colombia, Octubre de 2012.
- *Encuentro de Tesistas (ET 2011)* de la Sociedad Chilena de Ciencias de la Computación. Curicó, Chile, Noviembre de 2011.

Referatos

Journals Internacionales

- Revisor para *Information Processing Letters* (Elsevier), *Algorithmica* (Springer), *The Computer Journal* (Oxford Journals), *Information Processing and Management* (Elsevier), *Journal of Experimental Algorithmics* (ACM), *Software: Practice and Experience Computer Science and Information Systems*, *Applied Computational Intelligence and Soft Computing*, *Journal of Discrete Algorithms* (JDA), *ACM Transactions on the Web* (TWEB), *Discrete Mathematics and Theoretical Computer Science* (DMTCS).
- Sub revisor para *ACM Transactions on Algorithms*.

Conferencias Internacionales

Sub revisor para: *Annual European Symposium on Algorithms (ESA'13)*, *Symposium on Combinatorial Pattern Matching (CPM'13)*, *Data Compression Conference (DCC'13)*, *ESA'12*, *Annual ACM SIGIR Conference (SIGIR'12)*, *SIGIR'11*, *ESA'11*, *Symposium on Principles of Database Systems (PODS'11)*, *DCC'11*, *International Conference on Database and Expert Systems Applications (DEXA'11)*, *Symposium on String Processing and Information Retrieval (SPIRE'11)*, *DCC'10*, *Symposium on Foundations of Information and Knowledge Systems (FoIKS'10)*, *Symposium on Theoretical Aspects of Computer Science (STACS'10)*, *SPIRE'10*, *Symposium on*

Discrete Algorithms (SODA'09), SPIRE'09, Workshop on Similarity Search and Applications (SISAP'09), Symposium on Experimental Algorithms (SEA'09), SPIRE'07, Workshop on Algorithm Engineering and Experiments (ALENEX'07), Conference on Information and Knowledge Management (ACM CIKM'06), SPIRE'06, Workshop on Experimental Algorithms (WEA'06).

Charlas Dictadas

Charlas Internacionales

- (Oct. 2010) “Practical Rank/Select Data Structures with Applications to Text Search Engines”, en el *5th Workshop on Compression, Text, and Algorithms*, Los Cabos, México. (<http://spire2010.natix.org/?f=workshop>).
- (Feb. 2009) “Succinct Cardinal Trees Supporting Subpath Queries”, en el Department of Computer Science, McMaster University, Hamilton, Canadá. Invitado por el Profesor William F. Smyth.
- (Nov. 2008) “Constructing Lempel-Ziv text indexes in reduced-memory scenarios”, en el *3rd Workshop on Compression, Text, and Algorithms*, University of Melbourne, Melbourne, Australia. (<http://spire2008.csse.unimelb.edu.au/workshop08.html>).
- (Nov. 2007) “Lempel-Ziv Compressed Indices”, en el *Information Retrieval Workshop*, Department of Computer Science and Software Engineering, University of Melbourne.
- (Nov. 2007) “An Improved Succinct Dynamic k -ary Tree Representation”, en el *2nd Workshop on Compression, Text, and Algorithms*, Departamento de Ciencias de la Computación, Universidad de Chile. (<http://www.ciw.cl/eventos/compression-workshop.html>).
- (Jul. 2007) “Searching Lempel-Ziv-Compressed Texts” en el *Algorithms and Complexity Seminar Series* (<http://algcomp.uwaterloo.ca/seminars.html>), David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Canadá. Abstract disponible en <http://algcomp.uwaterloo.ca/event.php?code=102>.
- (Dec. 2005) “Reducing the Space Requirement of LZ-index”, en el Department of Computer Science and Communication Engineering, Kyushu University, Fukuoka, Japón.
- (Nov. 2005) “Lempel-Ziv Compressed Full-Text Self-Indexes”, en el *1st Workshop on Compression, Text, and Algorithms*, Departamento de Ciencias de la Computación, Universidad de Chile (<http://www.ciw.cl/eventos/compression-workshop.html>).

Charlas Nacionales y Regionales

- (Nov. 2010) “Búsqueda en texto comprimido”, en el Ciclo de Charlas 2010, organizadas por el Departamento de Ingeniería en Informática, Universidad de Santiago de Chile.

- (Nov. 2009) “Representaciones Sucintas de Árboles, con Aplicación en Consultas XPath”, en el taller *Tópicos Emergentes en la Web*, organizado por el Centro de Investigación de la Web, Núcleo Científico Milenio.
- “A Lempel-Ziv Text Index on Secondary Storage” en las “Jornadas de Postgrado DCC”, Departamento de Ciencias de la Computación, Universidad de Chile, Santiago, Chile.
- (Feb. 2006) “Búsqueda en Texto: Reduciendo el Espacio de LZ-index” en el Departamento de Informática, Universidad Nacional de San Luis, San Luis, Argentina.

Participación en Conferencias y Workshops

Asistencia a Conferencias Internacionales

ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR 2013**), **SIGIR 2012**, String Processing and Information Retrieval (**SPIRE 2010**), **SPIRE 2008**, Combinatorial Pattern Matching (**CPM 2008**), Workshop on Combinatorial Algorithms (**IWOCA 2007**), **SPIRE 2007**, **CPM 2007**, **CPM 2006**, Algorithms and Computation (**ISAAC 2005**), **SPIRE 2005**, Latin American Web Congress (**LAWEB 2005**).

Participación en Workshops Internacionales

- *String Masters at McMaster Workshop (SAM'07)*, McMaster University, Hamilton, Canadá, 3 al 6 de Julio de 2007. Profesor organizador: William F. Smyth, McMaster University.
- *Workshop on Space-Conscious Algorithms*, Bertinoro, Italia, 10 al 15 Junio de 2006. La participación fue sólo por invitación de los organizadores.
<http://roquefort.di.unipi.it/~ferrax/Bertinoro06/>

Docencia

Cargos Docentes

- (Julio 2011 – actualidad) Académico Jornada Completa, Departamento de Informática, Universidad Técnica Federico Santa María.

Dirección de Tesis

- Mauricio Silva O., alumno del Doctorado en Ciencias de la Ingeniería, Mención Informática. Departamento de Ingeniería Informática, Universidad de Santiago de Chile.
Título de la Tesis: Índices Comprimidos Auto-Indexados para Motores de Búsqueda.
Profesor Co-Guía: Mauricio Marin.
- Senén González, alumno del Magister en Ciencias, Mención Computación. Departamento de Ciencias de la Computación, Universidad de Chile.
Título de la Tesis: Auto índices comprimidos para recuperación de posiciones en colecciones de documentos.
Profesor Co-Guía: Gonzalo Navarro.

Cursos Diseñados

- (2010) Rediseño del curso “Recuperación de la Información”, para la Ing. Civil en Ciencias de la Computación, Universidad de Chile.
- (2010) Participé en la creación del curso de postgrado “Compresión de Texto”, junto con el Prof. Gonzalo Navarro, Universidad de Chile.
- (2011) Creación del curso de postgrado “Compresión de Texto”, Departamento de Informática, Universidad Técnica Federico Santa María.

Cursos Regulares Dictados

Universidad Técnica Federico Santa María

- Estructuras de Datos (Pregrado): Primavera 2011, Otoño 2012, Primavera 2012, Otoño 2013.
- Fundamentos de Informática II (Pregrado): Primavera 2011, Primavera 2012.
- Compresión de Texto (Postgrado): Otoño 2012.

Otras Universidades

- (Otoño 2011) Compresión de Texto, Departamento de Ciencias de la Computación, Universidad de Chile. Junto con Gonzalo Navarro.
- (Otoño 2010 y 2011) Participación en el dictado del curso “Tecnologías de la Web”, en el Departamento de Ciencias de la Computación, Universidad de Santiago de Chile. Profesor a cargo del Curso: Mauricio Marín.
- (Primavera 2010) Estructuras de datos compactas, Departamento de Ciencias de la Computación, Universidad de Chile. Junto con Gonzalo Navarro.
- (Primavera 2010) Recuperación de la Información, Departamento de Ciencias de la Computación, Universidad de Chile. Junto con Barbara Poblete, Marcelo Mendoza y Mauricio Marín.
- (Primavera 2010) Bioinformática, Departamento de Ciencias de la Computación, Universidad de Chile. Junto con Gonzalo Navarro, Alvaro Olivera y Oriana Salazar.
- (Otoño 2010) Compresión de Texto, Departamento de Ciencias de la Computación, Universidad de Chile. Junto con Gonzalo Navarro.
- (Primavera 2009) Recuperación de la Información, Departamento de Ciencias de la Computación, Universidad de Chile. Junto con Marcelo Mendoza, Flavio Ferrarotti y Mauricio Marín.

Participación en Comités de Tesis de Grado y Postgrado

- (Septiembre 2013) Miembro del comité de la Tesis de Magister en Ciencias, Mención Computación de Felipe Bravo, Departamento de Ciencias de la Computación, Universidad de Chile.
- (Enero 2013) Miembro del comité de la Tesis de Magister en Ciencias, Mención Computación de Daniel Valenzuela, Departamento de Ciencias de la Computación, Universidad de Chile.
- (Diciembre 2012) Miembro del comité de la Tesis de Magister en Ciencias, Mención Computación de Eliana Providel, Departamento de Ciencias de la Computación, Universidad de Chile.
- (Abril 2012) Profesor Correferente de la Memoria de Ingeniería Civil Informática de Matias Henriquez, Departamento de Informática, Universidad Técnica Federico Santa María,
- (Noviembre 2010) Miembro del comité de la Tesis de Magister en Ciencias, Mención Computación de Sebastian Kreft, Departamento de Ciencias de la Computación, Universidad de Chile.

Cursos Intensivos Dictados

- (Mayo de 2010) Participación en el dictado del curso de postgrado “Técnicas de Computación para la Web y Escalabilidad”, en el Departamento de Informática, Universidad Nacional de La Plata, Argentina.
Profesor a cargo del Curso: Mauricio Marín.

Desarrollo de Software

- Participación en el desarrollo de software para motores de búsqueda que actualmente está en producción para Yahoo!
- Participación en el desarrollo de software para el sitio *Pizza&Chili*, a cargo de los profesores Gonzalo Navarro (U. Chile) y Paolo Ferragina (U. Pisa, Italia). Este sitio se está volviendo el estándar para la comparación práctica de índices para texto. Mi participación consistió en programar los índices comprimidos de texto basados en el algoritmo de compresión de Lempel-Ziv.
URL: <http://pizzachili.dcc.uchile.cl/indexes/LZ-index/>