

Daniel Lemire, Ph.D.

Centre de recherche LICEF, TELUQ
Pavillon Saint-Urbain
100 Sherbrooke West
Montreal (Quebec)
H2X 3P2 Canada

lemire@gmail.com
<http://lemire.me/en/>

Canadian citizen

Experience

- 2004– . . .* Computer Science Professor
LICEF Research Center, TELUQ
Promoted to Full Professor in 2009
- 2002–2004* Research Officer in eBusiness
National Research Council of Canada (NRC)
Team Leader of the e-Health Research Group in 2002–2003
- 2001–2002* Assistant Professor
Acadia University
- 1999–2001* Entrepreneur
TechElements Inc. and Ondelette.com
- 1998–1999* Post-doctoral fellow
Institut de génie biomédial

Education

- 1995–1998* Ph.D.—“Iterative Schemas”
École Polytechnique de Montréal and **Université de Montréal**
Supervisors: Gilles Deslauriers and Serge Dubuc
- 1994–1995* M.Sc.—“A Priori Approximations for Nonlinear Systems”
University of Toronto
Supervisor: Catherine Sulem
- 1990–1994* B.Sc. with “High Distinction”
University of Toronto

Competitive scholarships and prizes as a student

- FCAR scholarship (Ph.D.) and NSERC scholarship (M.Sc. and Ph.D.);
- C. D. Howe Memorial for 4 years (≈\$50,000);
- NSERC Canada Scholarship for 4 years (\$4,000);
- St.Micheal’s College Dean’s medal (class of 1994);
- 3T0 Scholarship (UofT).

Secondary affiliations

- I am an adjunct professor at the University of New Brunswick where I supervise graduate students.

Recent Teaching Experience

Undergraduate (in French):

- INF 1250 - Introduction aux bases de données;
- INF 6460 - Recherche et filtrage d'informations;
- INF 6450 - Gestion de l'information avec XML;
- INF 9002 - Évaluation et analyse des systèmes d'information;
- INF 9004 - Informatique des entrepôts de données.

Graduate (in French):

- INF 6104 - Recherche d'informations et web;
- INF 6107 - Le Web social;
- INF 6408 - Informatique de l'analyse multidimensionnelle.

Moreover, I have chaired and cochaired several programs: two certificates, one bachelor degree, and four graduate programs including the master in Information Technology.

Professional Membership

- Association for Computing Machinery (ACM).

Research program

Individual Research Funding (external agencies)

- NSERC Discovery Grant 2012–2017 (\$140,000);
- NSERC Discovery Grant 2007–2012 (\$75,000);
- NSERC Discovery Grant 2003–2007 (\$48,000);
- FQRNT New Researcher Grant 2006–2008 (\$54,249).

Team Research Funding (external agencies)

- CFI Grant 2008–2009 (\$999,618) with G. Paquette and P. Valtchev;
- New Brunswick Innovation Fund 2007–2008 (\$10,000) with Owen Kaser;

Services (external)

Blog

- I run a blog at <http://lemire.me/>. I have over 4 000 subscribers.

Program committees (incomplete list)

- ACM Conference on Recommender Systems (RecSys) 2009–2012;
- ACM Conference on Information and Knowledge Management (CIKM) 2012;
- ACM Conference on Web Search and Data Mining (WSDM) 2013;
- Business Information Systems (BIS) 2011–2012;
- Web Intelligence, Mining and Semantics (WIMS) 2011–2012;
- Enterprise Information Systems (ICEIS) 2009–2012;
- Database and Expert Systems Applications (DEXA) 2009–2012;
- Web Information Systems and Technologies (WEBIST) 2007–2012;
- ACM/IEEE Joint Conference on Digital Libraries (JCDL) 2011–2012;
- Logistics, Informatics and Service Science (LISS) 2011–2012;
- Association for the Advancement of Artificial Intelligence (AAAI) 2008.

Editorial Board

- Atlantic Electronic Journal of Mathematics;
- Journal of Computers;
- International Journal on Advances in Intelligent Systems;
- Journal of Emerging Technologies in Web Intelligence.

Funding Agencies

- At the FQRNT, I have been a member of the review committee 03F (Theoretical Computer Science) since 2007. I was member of the review committee 309 (Computer Science team projects) in 2006–2007.

External examiner/referee

- External examiner for Ph.D. thesis:
Ahmad Taleb from Concordia University (2011);
- External referee for habilitation:
Sabine Loudcher Rabaseda from Lyon 2 University (2011);
- External referee for promotion:
Ken Pu from the Ontario Institute of Technology (2011).

Under Review

- [1] Owen Kaser and Daniel Lemire, Strongly universal string hashing is fast.
- [2] Zoltán Prekopcsák and Daniel Lemire, Time Series classification by class-based Mahalanobis distances.
- [3] Hazel Webb, Daniel Lemire, Owen Kaser, Diamond Dicing.

Peer-Reviewed Journal Articles

Impact factors of some journals: *Information Sciences* (2.8), *Pattern Recognition* (2.6), *Data & Knowledge Engineering* (1.7), *Computer Speech and Language* (1.4), and *Information Retrieval* (1.3).

- [4] Daniel Lemire, The universality of iterated hashing over variable-length strings, *Discrete Applied Mathematics* **160** (4-5), 2012.
- [5] Antonio Badia and Daniel Lemire, A Call to Arms: Revisiting Database Design, *SIGMOD Record* **40** 3, 2011.
- [6] Daniel Lemire and Owen Kaser, Reordering columns for smaller indexes, *Information Sciences* **181** 12, 2011.
- [7] Daniel Lemire and Owen Kaser, Recursive n-gram hashing is pairwise independent, at best, *Computer Speech and Language* **24** 4, 2010.
- [8] Daniel Lemire, Owen Kaser, Kamel Aouiche, Sorting improves word-aligned bitmap indexes, *Data & Knowledge Engineering* **69** 1, 2010.

- [9] Daniel Lemire, Faster retrieval with a two-pass Dynamic-Time-Warping lower bound, *Pattern Recognition* **42** 9, 2009.
- [10] Daniel Lemire, Martin Brooks, Yuhong Yan, An optimal linear time algorithm for quasi-monotonic segmentation, *International Journal of Computer Mathematics* **86** 7, 2009.
- [11] Kamel Aouiche, Daniel Lemire and Robert Godin, Web 2.0 OLAP: From data cubes to tag clouds, *Lecture Notes in Business Information Processing* **18**, 2009.
- [12] Daniel Lemire and Owen Kaser, Hierarchical Bin Buffering: Online local moments for dynamic external memory arrays, *ACM Transactions on Algorithms* **4** 1, 2008.
- [13] Daniel Lemire, Streaming maximum-minimum filter using no more than three comparisons per element, *Nordic Journal of Computing* **13** 4, 2006.
- [14] Owen Kaser and Daniel Lemire, Attribute value reordering for efficient hybrid OLAP, *Information Sciences* **176** 16, 2006.
- [15] Daniel Lemire, Harold Boley, Sean McGrath, Marcel Ball, Collaborative filtering and inference rules for context-aware learning object recommendation, *International Journal of Interactive Technology & Smart Education* **2** 3, 2005.
- [16] Daniel Lemire, Scale and Translation Invariant Collaborative Filtering Systems, *Information Retrieval* **8** 1, 2005.
- [17] Serge Dubuc, Daniel Lemire, Jean-Louis Merrien, Fourier analysis of 2-point Hermite interpolatory subdivision schemes, *Journal of Fourier Analysis and Applications* **7** 5, 2001.
- [18] Daniel Lemire, Chantal Pharand, Jean-Claude Rajaonah, Bruno Dubé, A.-Robert LeBlanc, Wavelet time entropy, T wave morphology and myocardial ischemia, *IEEE Transactions in Biomedical Engineering* **47** 7, 2000.
- [19] Gilles Deslauriers, Serge Dubuc, and Daniel Lemire, Une famille d'ondelettes biorthogonales sur l'intervalle obtenue par un schéma d'interpolation itérative, *Annales des Sciences Mathématiques du Québec* **23** 1, 1999.

In Peer-Reviewed Proceedings

According to Google Scholar, “Slope one predictors for online rating-based collaborative filtering” has **147 citations** whereas “Tag-cloud drawing: Algorithms for cloud visualization” received **122 citations**.

- [20] Andre Vellino and Daniel Lemire, Extracting, Transforming and Archiving Scientific Data, VLDL 2011, 2011.
- [21] Owen Kaser, Daniel Lemire, Kamel Aouiche, Histogram-aware sorting for Enhanced Word-Aligned Compression in bitmap indexes, DOLAP 2008, 2008.
- [22] Kamel Aouiche, Daniel Lemire, Owen Kaser, Tri de la table de faits et compression des index bitmaps avec alignement sur les mots, BDA 2008, 2008.
- [23] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning attributes from data cubes with Diamond Dicing, IDEAS 2008, 2008.
- [24] Kamel Aouiche, Daniel Lemire, Robert Godin, Collaborative OLAP with tag clouds: Web 2.0 OLAP formalism and experimental evaluation, WEBIST 2008, 2008.
- [25] Kamel Aouiche and Daniel Lemire, A comparison of five probabilistic view-size estimation techniques in OLAP, DOLAP 2007, 2007.
- [26] Owen Kaser and Daniel Lemire, Removing manually-generated boilerplate from electronic texts: Experiments with project Gutenberg e-books, CASCON 2007, 2007.
- [27] Owen Kaser and Daniel Lemire, Tag-Cloud Drawing: Algorithms for cloud visualization, Tagging and Metadata for Social Information Organization (WWW 2007), 2007.
- [28] Kamel Aouiche and Daniel Lemire, Unassuming view-size estimation techniques in OLAP: An experimental comparison, ICEIS 2007, 2007.
- [29] Daniel Lemire, A better alternative to piecewise linear time series segmentation, SIAM Data Mining 2007, 2007.
- [30] Dan Kucerovsky and Daniel Lemire, Monotonicity analysis over chains and curves, Curves and Surfaces 2006, 2007.
- [31] Owen Kaser, Daniel Lemire, Steven Keith, The LitOLAP Project: data warehousing with literature, CaSTA 2006, 2006.
- [32] Daniel Lemire, Martin Brooks, Yuhong Yan, An optimal linear time algorithm for quasi-monotonic segmentation, ICDM 2005, 2005.
- [33] Will Fitzgerald, Daniel Lemire, Martin Brooks, Quasi-monotonic segmentation of state variable behavior for reactive control, AAAI 2005, 2005.

- [34] Yuhong Yan, Martin Brooks, Daniel Lemire, Scale-based monotonicity analysis in qualitative modelling with flat segments, IJCAI 2005, 2005.
- [35] Daniel Lemire and Anna Maclachlan, Slope One predictors for online rating-based collaborative filtering. SIAM Data Mining 2005, 2005.
- [36] Yuhong Yan, Daniel Lemire, Martin Brooks, Monotone pieces analysis for qualitative modeling, MONET 2004, 2004.
- [37] Michelle Anderson, Marcel Ball, Harold Boley, Stephen Greene, Nancy Howse, Daniel Lemire, Sean McGrath, RACOFI: A rule-applying collaborative filtering system, IEEE/WIC COLA 2003, 2003.
- [38] Owen Kaser and Daniel Lemire, Attribute value reordering for efficient hybrid OLAP. DOLAP 2003, 2003.
- [39] Daniel Lemire, A family of 4-point dyadic high resolution subdivision schemes, Curves and Surfaces 2002, 2003.
- [40] Daniel Lemire, Wavelet-based relative prefix sum methods for range sum queries in data cubes, CASCON 2002, 2002. (**Best Paper Award**)

Books

- [41] Mamadou Tadiou Koné and Daniel Lemire (Eds.), Canadian Semantic Web, Springer, 2006.

Special Issues

- [42] Daniel Lemire and Richard Hotte (Eds.), Special issue on learning and the social web, Journal of Emerging Technologies in Web Intelligence **2** 1, 2010.
- [43] Mamadou Tadiou Koné and Daniel Lemire (Eds.), Special issue on Canadian Semantic Web, Computational Intelligence **23** 3, 2007.

Book chapters

- [44] Sylvie Noël and Daniel Lemire, On the challenges of collaborative data processing, in Collaborative Information Behaviour: User Engagement and Communication Sharing (edited by Jonathan Foster), IGI Global, 2010.

Magazines

- [45] Daniel Lemire, Marketing your ideas: Don't sell yourself short, ACM XRDS: Crossroads **16** (4), 2010.

Technical Reports

- [46] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning attributes from data cubes with Diamond Dicing, UNBSJ CSAS Technical Report TR-08-011, 2008.
- [47] Owen Kaser and Daniel Lemire, Removing manually-generated boilerplate from electronic texts: Experiments with project Gutenberg e-books. UNBSJ CSAS Technical Report TR-07-001, 2007.
- [48] Daniel Lemire and Owen Kaser, One-pass, one-hash n-gram statistics estimation, UNBSJ CSAS Technical Report TR-06-001, 2006.
- [49] Steven Keith, Owen Kaser, Daniel Lemire, Analyzing large collections of electronic text using OLAP, UNBSJ CSAS Technical Report TR-05-001, 2005.
- [50] Jean-Michel Nonglaton, Franco Lenardon, and Daniel Lemire, Wavelet shrinkage of LINAC III and protons synchrotron booster transformers by the Haar Transform, CERN Technical Report Number AB-Note-2003-033, avril 2003, VNRC 45816.
- [51] Daniel Lemire, Local interpolation by high resolution subdivision schemes. Technical Report 200205-01, Acadia University, 2002.

Lecture Notes

- [INF9004] Daniel Lemire, Informatique des entrepôts de données (Data Warehousing and Computer Science), 2012. Online course: <http://benhur.teluq.ca/SPIP/inf9004/>.
- [INF6104] Daniel Lemire, Recherche d'informations et web (Information Retrieval and the Web), 2008. Online course: <http://benhur.teluq.ca/SPIP/inf6104/>.
- [INF6460] Daniel Lemire, Recherche et filtrage d'informations (Information Retrieval and Filtering), 2007. Online course: <http://benhur.teluq.ca/SPIP/inf6460/>.
- [INF6450] Daniel Lemire, Gestion de l'information avec XML (Information Management with XML), ISBN 2762418747, 2005. (Revised in 2006, 2007 and 2008.) Online course: <http://www.teluq.ca/inf6450/>.

Open Source Software

- Daniel Lemire, JavaEWAH, A compressed alternative to the Java BitSet class, 2009. (<http://code.google.com/p/javaewah/>, used by Apache Hive for bitmap indexing)

- Daniel Lemire, Rolling Hash C++ Library, 2009 (<http://code.google.com/p/ngramhashing/>)
- Daniel Lemire, Lemur Bitmap Index C++ Library, 2008. (<http://code.google.com/p/lemurbitmapindex/>)
- Daniel Lemire and Owen Kaser, Tag-Cloud Drawing Software, 2008. (<http://code.google.com/p/tagclouddrawing/>)
- Daniel Lemire, Fast Nearest-Neighbor Retrieval under the Dynamic Time Warping, 2008 (<http://code.google.com/p/lbimproved/>, over 1 400 downloads).
- Kamel Aouiche, Daniel Lemire, Unassuming materialized view-size estimation techniques in OLAP (C++), 2007 (<http://code.google.com/p/viewsizeestimation/>).
- Daniel Lemire, Fast maximum-minimum filters implemented in C++, 2007 (<http://code.google.com/p/maxminfilters/>).
- Daniel Lemire and Owen Kaser, Hierarchical Bin Buffering Library in C++, 2007 (<http://code.google.com/p/hierarchicalbinbuffering/>).
- Daniel Lemire, Cofi: A Java-Based Collaborative Filtering Library, 2003 (<http://www.nongnu.org/cofi/>).
- Daniel Lemire and Owen Kaser, Lemur OLAP, 2003 (<http://www.nongnu.org/lemur/>).
- Daniel Lemire and Raj Bhaskar, WebForum, 1999 (<http://webforum.sourceforge.net/>).

Hobbies

- My favorite science-fiction authors are Charles Stross, Peter F. Hamilton and Alastair Reynolds.
- I build radio-controlled sailboat models.