

Daniel Lemire

Centre de recherche LICEF, TÉLUQ
Pavillon Saint-Urbain
Université du Québec à Montréal (UQAM)
100, rue Sherbrooke Ouest
Montréal (Québec)
H2X 3P2 Canada

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

Citoyen canadien

Expérience

- 2004-... Professeur, UER Science et technologie
Titularisé en 2009
TÉLUQ, Université du Québec à Montréal (UQAM)
- 2002-2004 Chercheur régulier en affaires électroniques
Institut de technologie de l'information
Conseil national de recherches du Canada (CNRC)
- 2002-2003 Chef de l'équipe de recherches en santé électronique
Institut de technologie de l'information
Conseil national de recherches du Canada (CNRC)
- 2001-2002 Professeur adjoint
Acadia University (Wolfville, Canada)
- 1999-2001 Entrepreneur
TechElements Inc. et **Ondelette.com**
- 1998-1999 Stagiaire postdoctoral
Institut de génie biomédical (Montréal, Canada)

Formation

- 1995-1998 Doctorat (Ph.D.) – *Schémas itératifs*
École Polytechnique de Montréal et Université de Montréal
Directeurs : Prof. Gilles Deslauriers et Serge Dubuc
- 1994-1995 Maîtrise – *Approximation a priori dans les systèmes non linéaires*
University of Toronto
Directrice : Catherine Sulem
- 1990-1994 Baccalauréat avec mention *High Distinction*
University of Toronto

Bourses et prix d'excellence à titre d'étudiant

- Bourse d'étude du FCAR (doctorat) et CRSNG (maîtrise et doctorat);
- Bourse *C. D. Howe Memorial* ($\approx 50\,000$ \$);
- Bourse Canada du CRSNG (4 000 \$);
- Médaille du doyen de St.Michael's (1994);
- Bourse 3T0 (Université de Toronto).

Autres affiliations

- Je suis professeur associé à l'Université du Nouveau-Brunswick où je dirige des étudiants aux cycles supérieurs.

Activités d'enseignement récentes

Premier cycle:

- 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.

Cycles supérieurs:

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

De plus, j'ai dirigé et codirigé plusieurs programmes incluant deux certificats, un baccalauréat, et quatre programmes aux cycles supérieurs dont la maîtrise en technologie de l'information.

Participation au sein d'associations professionnelles

- Association for Computing Machinery (ACM).

Supervision d'étudiants et de stagiaires

- Charbel Matni (M.Sc., 2011-...);
- Djamel Garar (M.Sc., 2011-...);
- Sahbi Khtatfa (M.Sc., 2011-...);
- Minfang Tao (Master, 2011);
- Ludovic Bocken (Ph.D., 2011-...);
- Toufik Lazri (M.Sc., 2010-...);
- Grégory Hallé Petiot (M.Sc., 2010-...);
- Eduardo Gutarra (M.Sc., 2010-...);
- Julien Marais (M.Sc., 2011-...);
- Ikhtear Sharif (M.Sc., 2011-...);
- Badis Merdaoui (Ph.D., 2008-...).

Diplômés récents :

- Hamid Douzi (D.E.S.S., 2010–2011) est développeur en intelligence d'affaire chez IMBD (Montréal, Canada);
- Hazel Webb (Ph.D., 2005–2010) est professeure adjointe à l'Université du Nouveau-Brunswick (Saint-Jean, Canada);
- Idris Laabassi (Master, 2010) est gestionnaire des ventes et des technologies de l'information chez OTIS (Libye);
- Kamel Aouiche (Postdoct., 2006–2008) est consultant chez Objet direct (Lyon, France);
- Steve Keith (M.Sc., 2004–2006) est architecte de base de données chez Innovatia (Saint-Jean, Canada).

Programme de recherche

Subventions individuelles de recherche (organismes externes)

- Subvention à la découverte du CRSNG (2007-2012, 75 000\$);
- Subvention pour l'établissement de nouveaux chercheurs du FQRNT (2006-2008, 54 249\$);
- Subvention à la découverte du CRSNG (2003-2007, 48 000\$).

Subventions de recherche en équipe

- Fonds des leaders de la FCI (2008-2009, 999 618\$) avec G. Paquette et P. Valtchev;
- Fondation de l'innovation du Nouveau-Brunswick (2007-2008, 10 000\$) avec O. Kaser;
- Programme de collaboration Nouveau-Brunswick/Québec (2007-2008, 1 825\$) avec O. Kaser.

Services à la communauté (externe)

Blogue

- J'ai un blogue à l'adresse <http://lemire.me>. Il compte plus de 3 500 lecteurs réguliers. En 2010, mon site web a été visité 500 000 fois.

Comités de lecture (liste partielle)

- ACM Conference on Recommender Systems (RecSys) 2009, 2010, 2011;
- Business Information Systems (BIS) 2011;
- International Conference on Enterprise Information Systems (ICEIS) 2009, 2010, 2011;
- International Conference on Database and Expert Systems Applications (DEXA) 2009, 2010, 2011;
- Conference on Web Information Systems and Technologies (WEBIST) 2007, 2008, 2009, 2010, 2011;
- ACM/IEEE Joint Conference on Digital Libraries (JCDL) 2011;
- International Conference on Web Intelligence, Mining and Semantics (WIMS) 2011;
- International Conference on Logistics, Informatics and Service Science (LISS) 2011;
- Association for the Advancement of Artificial Intelligence (AAAI) 2008 – Web Track.

Comités éditoriaux au sein de revues scientifiques

- Open Research Computation;
- Atlantic Electronic Journal of Mathematics;
- Journal of Computers;
- Journal of Systems and Management Sciences;
- Journal of Emerging Technologies in Web Intelligence.

Organismes subventionnaires

- Au FQRNT, je suis membre du comité d'évaluation 03F depuis 2007 et j'ai été membre du comité d'évaluation 309 en 2006-2007.

Évaluations externes

- Évaluateur externe pour thèse de doctorat: Ahmad Taleb de l'Université Concordia (2011);
- Rapporteur pour dossier d'habilitation: Sabine Loudcher Rabaseda de l'Université Lyon 2 (2011);
- Évaluateur externe pour dossier de promotion: Ken Pu de l'Ontario Institute of Technology (2011).

Soumis pour publication

- [1] Zoltán Prekopcsák et Daniel Lemire, Time series classification by class-based Mahalanobis distances.
- [2] Hazel Webb, Daniel Lemire, Owen Kaser, Diamond Dicing.

Articles scientifiques dans des revues avec comité de lecture

Facteurs d'incidence de certaines revues : *Information Sciences* (2.8), *Pattern Recognition* (2.6), *Data & Knowledge Engineering* (1.7), *Computer Speech and Language* (1.4) et *Information Retrieval* (1.3).

- [3] Daniel Lemire, The universality of iterated hashing over variable-length strings, *Discrete Applied Mathematics* **160** (4-5), 2012.
- [4] Antonio Badia et Daniel Lemire, A Call to Arms: Revisiting Database Design, *SIGMOD Record* **40** 3, 2011.
- [5] Daniel Lemire and Owen Kaser, Reordering columns for smaller indexes, *Information Sciences* **181** 12, 2011.
- [6] Daniel Lemire and Owen Kaser, Recursive n-gram hashing is pairwise independent, at best, *Computer Speech and Language* **24** 4, 2010.
- [7] Daniel Lemire, Owen Kaser, Kamel Aouiche, Sorting improves word-aligned bitmap indexes, *Data & Knowledge Engineering* **69** 1, 2010.
- [8] Daniel Lemire, Faster retrieval with a two-pass Dynamic-Time-Warping lower bound, *Pattern Recognition* **42** 9, 2009.
- [9] Daniel Lemire, Martin Brooks, Yuhong Yan, An optimal linear time algorithm for quasi-monotonic segmentation, *International Journal of Computer Mathematics* **86** 7, 2009.
- [10] 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.

- [11] Daniel Lemire and Owen Kaser, Hierarchical Bin Buffering: Online local moments for dynamic external memory arrays, *ACM Transactions on Algorithms* **4** 1, 2008.
- [12] Daniel Lemire, Streaming maximum-minimum filter using no more than three comparisons per element, *Nordic Journal of Computing* **13** 4, 2006.
- [13] Owen Kaser and Daniel Lemire, Attribute value reordering for efficient hybrid OLAP, *Information Sciences* **176** 16, 2006.
- [14] 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.
- [15] Daniel Lemire, Scale and Translation Invariant Collaborative Filtering Systems, *Information Retrieval* **8** 1, 2005.
- [16] 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.
- [17] 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.
- [18] 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.

Communications avec arbitrage

Selon Google Scholar, ma communication “Slope one predictors for online rating-based collaborative filtering” a été citée plus de **167 fois** alors que ma communication “Tag-cloud drawing: Algorithms for cloud visualization” a reçu plus de **115 citations**.

- [19] Andre Vellino and Daniel Lemire, Extracting, Transforming and Archiving Scientific Data, VLDL 2011, 2011.
- [20] Owen Kaser, Daniel Lemire, Kamel Aouiche, Histogram-aware sorting for Enhanced Word-Aligned Compression in bitmap indexes, DOLAP 2008, 2008.
- [21] 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.
- [22] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning attributes from data cubes with Diamond Dicing, IDEAS 2008, 2008.
- [23] Kamel Aouiche, Daniel Lemire, Robert Godin, Collaborative OLAP with tag clouds: Web 2.0 OLAP formalism and experimental evaluation, WEBIST 2008, 2008.

- [24] Kamel Aouiche and Daniel Lemire, A comparison of five probabilistic view-size estimation techniques in OLAP, DOLAP 2007, 2007.
- [25] Owen Kaser and Daniel Lemire, Removing manually-generated boilerplate from electronic texts: Experiments with project Gutenberg e-books, CASCON 2007, 2007.
- [26] Owen Kaser and Daniel Lemire, Tag-Cloud Drawing: Algorithms for cloud visualization, Tagging and Metadata for Social Information Organization (WWW 2007), 2007.
- [27] Kamel Aouiche and Daniel Lemire, Unassuming view-size estimation techniques in OLAP: An experimental comparison, ICEIS 2007, 2007.
- [28] Daniel Lemire, A better alternative to piecewise linear time series segmentation, SIAM Data Mining 2007, 2007.
- [29] Dan Kucerovsky and Daniel Lemire, Monotonicity analysis over chains and curves, Curves and Surfaces 2006, 2007.
- [30] Owen Kaser, Daniel Lemire, Steven Keith, The LitOLAP Project: data warehousing with literature, CaSTA 2006, 2006.
- [31] Daniel Lemire, Martin Brooks, Yuhong Yan, An optimal linear time algorithm for quasi-monotonic segmentation, ICDM 2005, 2005.
- [32] Will Fitzgerald, Daniel Lemire, Martin Brooks, Quasi-monotonic segmentation of state variable behavior for reactive control, AAI 2005, 2005.
- [33] Yuhong Yan, Martin Brooks, Daniel Lemire, Scale-based monotonicity analysis in qualitative modelling with flat segments, IJCAI 2005, 2005.
- [34] Daniel Lemire and Anna Maclachlan, Slope One predictors for online rating-based collaborative filtering. SIAM Data Mining 2005, 2005.
- [35] Yuhong Yan, Daniel Lemire, Martin Brooks, Monotone pieces analysis for qualitative modeling, MONET 2004, 2004.
- [36] 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.
- [37] Owen Kaser and Daniel Lemire, Attribute value reordering for efficient hybrid OLAP. DOLAP 2003, 2003.
- [38] Daniel Lemire, A family of 4-point dyadic high resolution subdivision schemes, Curves and Surfaces 2002, 2003.
- [39] Daniel Lemire, Wavelet-based relative prefix sum methods for range sum queries in data cubes, CASCON 2002, 2002. (**Best Paper Award**)

Livres

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

Numéros spéciaux

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

Chapitres de livre

- [43] 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

- [ECAP2007] Daniel Lemire, Marketing your ideas: Don't sell yourself short, ACM XRDS: Crossroads **16** (4), 2010.

Rapports techniques

- [44] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning attributes from data cubes with Diamond Dicing, UNBSJ CSAS Technical Report TR-08-011, 2008.
- [45] 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.
- [46] Daniel Lemire and Owen Kaser, One-pass, one-hash n-gram statistics estimation, UNBSJ CSAS Technical Report TR-06-001, 2006.
- [47] Steven Keith, Owen Kaser, Daniel Lemire, Analyzing large collections of electronic text using OLAP, UNBSJ CSAS Technical Report TR-05-001, 2005.
- [48] 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.
- [49] Daniel Lemire, Local interpolation by high resolution subdivision schemes. Technical Report 200205-01, Acadia University, 2002.

Ouvrages d'enseignement

- [INF6104] Daniel Lemire, Recherche d'informations et web, Université du Québec à Montréal, 2008. Cours en ligne: <http://benhur.teluq.uqam.ca/SPIP/inf6104/>.
- [INF6460] Daniel Lemire, Recherche et filtrage d'informations, Université du Québec à Montréal, 2007. Cours en ligne: <http://benhur.teluq.uqam.ca/SPIP/inf6460/>.
- [INF6450] Daniel Lemire, Gestion de l'information avec XML, Université du Québec à Montréal, ISBN 2762418747, 2005. (Révisé en 2006, 2007 et 2008.) En ligne: <http://www.teluq.uqam.ca/inf6450/>.

Logiciel open source

- Daniel Lemire, JavaEWAH, A compressed alternative to the Java BitSet class, 2009. (<http://code.google.com/p/javaewah/>, used by Apache Hive for its bitmap indexes)
- 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, 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/>).
- Kamel Aouiche et 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/>, plus de 150 téléchargements).
- Daniel Lemire et 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 et Owen Kaser, Lemur OLAP Library, 2003 (<http://www.nongnu.org/lemur/>).
- Daniel Lemire et Raj Bhaskar, WebForum, 1999 (<http://webforum.sourceforge.net/>).

Passe-temps

- Mes auteurs de science-fiction préférés sont Charles Stross, Peter F. Hamilton et Alastair Reynolds.
- Je construis des voiliers radio-commandés.