

# Daniel Lemire

Centre de recherche LICEF, TÉLUQ      lemire@acm.org  
Pavillon Saint-Urbain                      <http://www.daniel-lemire.com/fr/>  
Université du Québec à Montréal (UQAM)  
100, rue Sherbrooke Ouest  
Montréal (Québec)  
H2X 3P2 Canada

## Expérience

---

- 2004-...      Professeur en informatique  
Titularisé en 2009  
**Université du Québec à Montréal (UQAM)**  
<http://www.professeurs.uqam.ca/pages/lemire.daniel.htm>
- 2002-2004      Chercheur régulier en affaires électroniques  
Institut de technologie de l'information (affaires électroniques)  
**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  
Traitement du signal numérique (imagerie médicale et géophysique)  
**TechElements Inc. et Ondelette.com**
- 1998-1999      Stagiaire postdoctoral  
**Institut de génie biomédical** (Montréal, Canada)

## Formation

---

- 1995-1998      Doctorat (Ph.D.) en mathématiques de l'ingénieur  
« Schémas itératifs et ondelettes »  
**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 (≈50 000 \$);
- Bourse Canada du CRSNG (4 000 \$);
- Médaille du doyen de St.Michael's (1994);
- Bourse 3T0 (Université de Toronto) (1 000 \$).

### **Autres affiliations**

---

- Je suis professeur associé à l’Université du Nouveau-Brunswick au département d’informatique et de statistiques appliquées – avec habilitation à codiriger des thèses et mémoires.

### **Activités d’enseignement récentes**

---

- INF 9004 - Informatique des entrepôts de données;
- INF 6104 - Recherche d’informations et web;
- INF 6460 - Recherche et filtrage d’informations;
- INF 6450 - Gestion de l’information avec XML.

### **Participation au sein d’associations professionnelles**

---

- Association for Computing Machinery (ACM).

### **Supervision d’étudiants et de stagiaires**

---

Nom	Période	Programme
Nacima Hassaine	2010–...	M.Sc.
Idris Laabassi	2010	Stagiaire master (Lyon 2)
Loukmane Arzim	2010–...	Ph.D.
Eduardo Gutarra	2010–...	M.Sc.
Badis Merdaoui	2008–...	Ph.D.
Kamel Aouiche	2006–2008	postdoct.
Hazel Webb	2005–2010...	Ph.D.
Steve Keith	2004–2006	M.Sc.

### **Programme de recherche**

---

#### **Subventions individuelles de recherche (organismes externes)**

- Subvention du CRSNG (2007-2012, 75 000\$);
- Subvention du FQRNT (2006-2008, 54 249\$);
- Subvention du CRSNG (2003-2007, 48 000\$).

#### **Subventions de recherche en équipe**

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

#### **Services à la communauté (externe)**

---

##### **Blogue**

- J’ai un blogue à l’adresse <http://www.daniel-lemire.com>. Il compte plus de 2 400 lecteurs réguliers.

##### **Comités de lecture**

- ACM Conference on Recommender Systems (RecSys) 2009, 2010;

- International Conference on Enterprise Information Systems (ICEIS) 2009, 2010;
- International Conference on Database and Expert Systems Applications (DEXA) 2009, 2010;
- Conference on Web Information Systems and Technologies (WEBIST) 2007, 2008, 2009, 2010;
- Association for the Advancement of Artificial Intelligence (AAAI) 2008 – Web Track;

#### **Organismes subventionnaires**

- **Fonds Québécois de la recherche sur la Nature et les Technologies (FQRNT)**
  - De 2007 à 2010, j’ai été membre du comité d’évaluation 03F (informatique théorique).
  - En 2006–2007, j’étais membre du comité d’évaluation 309 dans le cadre du programme Projet de recherche en équipe.

#### **Soumis pour publication**

---

- [1] Daniel Lemire, The universality of iterated hashing over variable-length strings.
- [2] Hazel Webb, Owen Kaser, Daniel Lemire, Evaluating Multidimensional Queries by Diamond Dicing.
- [3] Daniel Lemire et Owen Kaser, Reordering Columns for Smaller Indexes.

#### **Articles scientifiques dans des revues avec comité de lecture**

---

- [CSL2010] Daniel Lemire and Owen Kaser, Recursive n-gram hashing is pairwise independent, at best, *Computer Speech and Language* **24** 4, 2010.
- [DKE2010] Daniel Lemire, Owen Kaser, Kamel Aouiche, Sorting improves word-aligned bitmap indexes, *Data & Knowledge Engineering* **69** 1, 2010.
- [PR2009] Daniel Lemire, Faster Retrieval with a Two-Pass Dynamic-Time-Warping Lower Bound, *Pattern Recognition* **42** 9, 2009.
- [IJCM2009] Daniel Lemire, Martin Brooks, Yuhong Yan, An Optimal Linear Time Algorithm for Quasi-Monotonic Segmentation, *International Journal of Computer Mathematics* **86** 7, 2009.
- [LNBIP2009] 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.
- [TALG2008] Daniel Lemire and Owen Kaser, Hierarchical Bin Buffering: Online Local Moments for Dynamic External Memory Arrays, *ACM Transactions on Algorithms* **4** 1, 2008.
- [NJC2006] Daniel Lemire, Streaming Maximum-Minimum Filter Using No More than Three Comparisons per Element, *Nordic Journal of Computing* **13** 4, 2006.
- [IS2006] Owen Kaser and Daniel Lemire, Attribute Value Reordering for Efficient Hybrid OLAP, *Information Sciences* **176** 16, 2006.

- [LL2005] 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.
- [IR2005] Daniel Lemire, Scale and Translation Invariant Collaborative Filtering Systems, *Information Retrieval* **8** 1, 2005. (Cit  au moins 25 fois selon Google Scholar.)
- [JFAA2001] 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.
- [ITBE2000] Daniel Lemire, Chantal Pharand, Jean-Claude Rajaonah, Bruno Dub , A.-Robert LeBlanc, Wavelet time entropy, T wave morphology and myocardial ischemia, *IEEE Transactions in Biomedical Engineering* **47** 7, 2000. (Cit  au moins 34 fois selon Google Scholar.)
- [ASMQ1999] Gilles Deslauriers, Serge Dubuc, and Daniel Lemire, Une famille d’ondelettes biorthogonales sur l’intervalle obtenue par un sch ma d’interpolation it rative, *Ann. Sci. Math. Qu bec* **23** 1, 1999.

#### **Communications avec arbitrage**

---

- [DOLAP08] Owen Kaser, Daniel Lemire, Kamel Aouiche, Histogram-Aware Sorting for Enhanced Word-Aligned Compression in Bitmap Indexes, DOLAP 2008, 2008.
- [BDA08] Kamel Aouiche, Daniel Lemire and Owen Kaser, Tri de la table de faits et compression des index bitmaps avec alignement sur les mots, BDA’08, 2008.
- [IDEAS08] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning Attributes From Data Cubes with Diamond Dicing, IDEAS’08, 2008.
- [WEBIST08] Kamel Aouiche, Daniel Lemire, Robert Godin, Collaborative OLAP with Tag Clouds: Web 2.0 OLAP Formalism and Experimental Evaluation, WEBIST 2008, 2008.
- [DOLAP07] Kamel Aouiche et Daniel Lemire, A Comparison of Five Probabilistic View-Size Estimation Techniques in OLAP, DOLAP 2007, pp. 17–24, 2007.
- [CASCON07] Owen Kaser et Daniel Lemire, Removing Manually-Generated Boilerplate from Electronic Texts: Experiments with Project Gutenberg e-Books. CASCON 2007, pp. 272–275, 2007.
- [WWW07] Owen Kaser et Daniel Lemire, Tag-Cloud Drawing: Algorithms for Cloud Visualization, Tagging and Metadata for Social Information Organization (WWW 2007), 2007. (Cit e au moins 59 fois selon Google Scholar.)
- [ICEIS07] Kamel Aouiche et Daniel Lemire, Unassuming View-Size Estimation Techniques in OLAP, An Experimental Comparison, ICEIS-2007, pp. 145–150, 2007.
- [SDM07] Daniel Lemire, A Better Alternative to Piecewise Linear Time Series Segmentation, SIAM Data Mining 2007, 2007.
- [CS06] Dan Kucerovsky et Daniel Lemire, Monotonicity Analysis over Chains and Curves. Curves and Surfaces 2006, pages 180–190, 2007.

- [CaSTA06] Owen Kaser, Daniel Lemire, Steven Keith, The LitOLAP Project: Data Warehousing with Literature, CaSTA 2006, Fredericton, 2006.
- [ICDM05] Daniel Lemire, Martin Brooks et Yuhong Yan, An Optimal Linear Time Algorithm for Quasi-Monotonic Segmentation, ICDM-05, pp. 709–712, novembre 2005.
- [AAAI05] Will Fitzgerald, Daniel Lemire et Martin Brooks, Quasi-monotonic segmentation of state variable behavior for reactive control, AAAI05, pp. 1145–1151, juillet 2005.
- [IJCAI05] Yuhong Yan, Martin Brooks, Daniel Lemire, Scale-Based Monotonicity Analysis in Qualitative Modelling with Flat Segments, IJCAI’05, pp. 400–405, juillet 2005.
- [SDM2005] Daniel Lemire et Anna Maclachlan, Slope One Predictors for Online Rating-Based Collaborative Filtering. SIAM Data Mining 2005, pp. 471–476, 2005. (Citée au moins 77 fois selon Google Scholar.)
- [MBR2004] Yuhong Yan, Daniel Lemire et Martin Brooks, Monotonicity Analysis for Constructing Qualitative Models, MBR’04, 2004.
- [Monet2004] Yuhong Yan, Daniel Lemire et Martin Brooks, Monotone Pieces Analysis for Qualitative Modeling, ECAI MONET 2004, 2004.
- [COLA2003] Michelle Anderson, Marcel Ball, Harold Boley, Stephen Greene, Nancy Howse, Daniel Lemire, Sean McGrath, RACOFI: A Rule-Appling Collaborative Filtering System, IEEE/WIC COLA’03, Halifax, Canada, octobre 2003. (Citée au moins 56 fois selon Google Scholar.)
- [DOLAP2003] Owen Kaser et Daniel Lemire, Attribute Value Reordering for Efficient Hybrid OLAP. DOLAP’03, New Orleans, Louisiane, 7 novembre 2003.
- [CS2003] Daniel Lemire, A Family of 4-point Dyadic High Resolution Subdivision Schemes. Curves and Surfaces 2002, Saint-Malo, France, 2003.
- [CASCON2002] Daniel Lemire, Wavelet-Based Relative Prefix Sum Methods for Range Sum Queries in Data Cubes. CASCON2002, Toronto, Canada, 2002. (**Best Paper Award**)

#### **Livres**

---

- [CSWWS2006] Mamadou Tadiou Koné, Daniel Lemire (éditeurs), Canadian Semantic Web, Semantic Web and Beyond: Computing for Human Experience, volume 4, Springer, 2006, ISBN 0-387-29815-0.

#### **Numéros spéciaux**

---

- [JETWI2010] Daniel Lemire and Richard Hotte (éditeurs), Special Issue on Learning and the Social Web, Journal of Emerging Technologies in Web Intelligence, volume 2 no 1, Academy Publisher, 2010, ISBN : 1798-0461.
- [CI2007] Mamadou Tadiou Koné, Daniel Lemire (éditeurs), Special Issue on Canadian Semantic Web, Computational Intelligence, Blackwell Publishing, volume 23 no 3, pages 299–392, août 2007.

## **Chapitres de livre**

---

[Collab2010] Sylvie Noël, Daniel Lemire, On the Challenges of Collaborative Data Processing, dans Collaborative Information Behaviour: User Engagement and Communication Sharing (édité par Jonathan Foster), IGI Global, avril 2010, ISBN-10: 161520797X.

## **Rapports techniques**

---

- [0807.1734] Daniel Lemire, Faster Sequential Search with a Two-Pass Dynamic-Time-Warping Lower Bound, 2008. (arXiv:0807.1734)
- [TRD08011] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning Attributes From Data Cubes with Diamond Dicing, UNBSJ CSAS Technical Report TR-08-011, 2008.
- [TRD07001] Owen Kaser et Daniel Lemire, Removing Manually-Generated Boilerplate from Electronic Texts: Experiments with Project Gutenberg e-Books. UNBSJ CSAS Technical Report TR-07-001, 2007.
- [TRD06001] Daniel Lemire et Owen Kaser, One-Pass, One-Hash n-Gram Statistics Estimation, Rapport technique UNBSJ CSAS TR-06-001, octobre 2006.
- [TRD05001] Steven Keith, Owen Kaser, Daniel Lemire, Analyzing Large Collections of Electronic Text Using OLAP, Rapport technique UNBSJ CSAS TR-05-001, juin 2005.
- [TRD01] Daniel Lemire, Sean McGrath, Implementing a Rating-Based Item-to-Item Recommender System in PHP/SQL, Rapport technique D-01, janvier 2005.
- [BELL2004-TR] Daniel Lemire, Sean McGrath et Bruce Spencer, The Design and Architecture of a Collaborative Filtering Web site — inDiscover.net, Rapport technique du CNRC présenté à Bell Canada, octobre 2004, 14 p.
- [DOLAP2003-TR] Owen Kaser et Daniel Lemire, Attribute Value Reordering for Efficient Hybrid OLAP, janvier 2004, 29 p., CNRC 46509.
- [CERN2003] 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.
- [CHIPP2003] Irina Kondratova, Scott Flinn, and Daniel Lemire, Teleoncology Technology Assessment Report for the Teleoncology Project. Soumis au projet de téléoncologie, CHIPP, Health Canada, Beauséjour Hospital Corporation. Moncton, New Brunswick, février 2003, 25 p., CNRC 45799.
- [Acadia200205-01] Daniel Lemire, Local Interpolation by High Resolution Subdivision Schemes. Technical Report 200205-01, Acadia University, avril 2002.
- [THEM2001-02] Daniel Lemire, Baseline asymmetry, Tau projection, B-field estimation and automatic half-cycle rejection, THEM Geophysics Inc., THEM2001-02, février 2001.
- [THEM2001-01] Daniel Lemire, Off-time denoising: is stacking the solution?, THEM Geophysics Inc., THEM2001-01, janvier 2001.

- [THEM1999-02] Daniel Lemire, Rapport sur un nouvel algorithme de correction de la ligne de base pour THEM Geophysics, Splines interpolation des moyennes et moindres carrés (SIM), THEM Geophysics Inc., THEM1999-02, 1999.
- [CIRA1999] Alain Béliveau et Daniel Lemire, White paper sur la compression d'images par la Fast Wavelet Transform, société Compression d'Images, Réseaux et Automatismes, mars 1999.
- [EPM/RT-97/28] Gilles Deslauriers, Serge Dubuc et Daniel Lemire, Dérivées de l'interpolation itérative de Lagrange et les ondelettes b-adiques de Cohen-Daubechies-Feauveau, Rapport technique EPM/RT-97/28, École Polytechnique de Montréal, Montréal, avril 1997.

### **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/>, plus de 400 téléchargements)
- Daniel Lemire, Rolling Hash C++ Library, 2009 (<http://code.google.com/p/ngramhashing/>, plus de 350 téléchargements)
- Daniel Lemire, Lemur Bitmap Index C++ Library, 2008. (<http://code.google.com/p/lemurbitmapindex/>, plus de 750 téléchargements)
- 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/>, plus de 1 400 téléchargements).
- 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/>, plus de 4 000 téléchargements).