

# Daniel Lemire

Université du Québec (TELUQ)  
5800, rue Saint-Denis  
Bureau 1105  
Montréal (Québec)  
H2S 3L5 Canada

lemire@gmail.com  
<http://lemire.me/fr/>  
<https://github.com/lemire/>  
Citoyen canadien  
bilingue (français et anglais)

## Expérience

---

- 2004–... Professeur d'informatique  
Université du Québec (TELUQ)  
Titularisé en 2009  
Directeur de département (2017–2020)
- 2002–2004 Chercheur régulier en affaires électroniques  
Conseil national de recherches du Canada (CNRC)  
Chef de l'équipe de recherches en santé électronique (2002–2003)
- 2001–2002 Professeur adjoint  
Acadia University
- 1999–2001 Entrepreneur  
TechElements Inc. et Ondelette.com
- 1998–1999 Stagiaire postdoctoral  
Institut de génie biomédical

## 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 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 et à l'Université du Québec à Montréal où je dirige des étudiants aux cycles supérieurs.

## **Activités d'enseignement récentes**

---

Premier cycle:

- 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.

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.

J'ai dirigé des dizaines d'étudiants à la maîtrise. De plus, plusieurs de mes étudiants ont récemment obtenu leur doctorat:

- Erick Aokou Koffi (2018);
- Badis Merdaoui (2017);
- Jing Li, informatique (2016);
- Samy Chambi, informatique (2016);
- Hazel Webb, informatique (2010).

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

---

- Association for Computing Machinery (ACM) depuis 2002.

## **Programme de recherche**

---

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

Ma demande de subvention à la découverte au CRSNG en 2017 a obtenu les évaluations *Outstanding* (Excellence du chercheur), *Very Strong* (Mérite de la proposition) et *Very Strong* (Formation de personnel hautement qualifié). En 2017, j'ai reçu une subvention d'accélération ce qui signifie que j'ai été jugé comme ayant un programme de recherche de qualité supérieure.

- Subventions d'engagement partenarial du CRNSG (2018, 25 000\$);
- Subvention d'accélération à la découverte du CRSNG (2017–2020, 120 000\$);
- Subvention à la découverte du CRSNG (2017–2022, 210 000\$);
- Subvention à la découverte du CRSNG (2012–2017, 140 000\$);
- Subvention à la découverte du CRSNG (2007–2012, 75 000\$);
- Subvention à la découverte du CRSNG (2003–2007, 48 000\$);
- Subvention pour l'établissement de nouveaux chercheurs du FQRNT (2006–2008, 54 249\$).

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

- Fonds des leaders de la FCI (2016–2017, 797,481\$) avec N. Bélanger (responsable) et E. Filotas;
- Fondation CAA (2013–2014, 189 994\$) avec E. Vallières (responsable);

- Fonds des leaders de la FCI (2008–2009, 999 618\$) avec G. Paquette (responsable) et P. Valtchev;
- Fondation de l’innovation du Nouveau-Brunswick (2007–2008, 10 000\$) avec O. Kaser (responsable).

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

---

#### **Blogue**

- J’ai un blogue à l’adresse <http://lemire.me>. Il compte plus de 25 000 visiteurs uniques par mois. Il fait partie des blogues recommandés par l’ACM au sein de leur blogrollé

#### **Comités de lecture (liste partielle)**

- ACM Conference on Information Retrieval (SIGIR) 2015;
- ACM Conference on Recommender Systems (RecSys) 2009–2014, 2017;
- ACM Conference on Information and Knowledge Management (CIKM) 2012–2017;
- ACM Conference on Web Search and Data Mining (WSDM) 2013–2015;
- World Wide Web Conference (WWW) 2017, 2018;
- ACM/IEEE Joint Conference on Digital Libraries (JCDL) 2011–2017.

#### **Organismes subventionnaires**

- Au CRSNG, je suis membre du comité d’évaluation des subventions à la découverte (EG 1507) en Sciences informatiques (2018-2020).
- Au FQRNT, j’ai été membre du comité d’évaluation 03F (informatique théorique) depuis 2007.
- Toujours au FQRNT, j’ai été membre du comité d’évaluation 309 (subvention d’équipe en informatique) en 2006–2007, en 2013–2014, 2014–2015 et 2016–2017.
- Au CRSNG, j’ai un membre du comité d’évaluation du programme de subventions d’outils et d’instruments de recherche dans les sciences informatiques en 2012–2013, 2013–2014 et 2014–2015.

#### **Évaluations externes**

- Évaluateur externe pour thèse de doctorat:
  - Kareem El Gebaly de l’Université Waterloo, Canada (2018) — dirigé par Jimmy Lin, Lukasz Golab and Ashraf Abounaga;
  - Mohammed Shaaban de l’Université Pierre et Marie Curie, France (2017) — dirigé par Patrick Garda;
  - Mehdi Boukhechba de l’UQAC, Canada (2016) — dirigé par Abdenour Bouzouane and Charles Gouin-Vallerand;
  - Hicham Assoudi de l’UQAM, Canada (2016) — dirigé par Hakim Lounis;
  - Khaled Dehdouh de Lyon 2, France (2015) — dirigé par Omar Boussaid;
  - Martin Leginus de l’Université Aalborg, Danemark (2015) — dirigé par Peter Dolog;

- Ahmad Taleb de l'Université Concordia, Canada (2011) — dirigé par Todd Eavis.
- Rapporteur pour dossier d'habilitation: Sabine Loudcher Rabaseda de l'Université Lyon 2, France (2011).
- Évaluatrice externe pour dossier de promotion:
  - Jason Sawin de l'Université of St. Thomas;
  - Amer Nizar AbuAli de la Philadelphia University;
  - Ken Pu du Ontario Institute of Technology;
  - Jinan Fiaidhi de Lakehead University.

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

Facteurs d'incidence de certaines revues : *Information Sciences* (2.8), *Pattern Recognition* (2.6), *Journal of the Association for Information Science and Technology* (2.0), *Data & Knowledge Engineering* (1.7), *ACM Transactions on Database Systems* (1.4), *Computer Speech and Language* (1.4), *Information Retrieval* (1.3), *ACM Transactions on Information Systems* (1.3) et *Information Systems* (1.2).

- [1] Daniel Lemire, Melissa E. O'Neill, Xorshift1024\*, Xorshift1024+, Xorshift128+ and Xoroshiro128+ Fail Statistical Tests for Linearity, *ACM Transactions on Modeling and Computer Simulation* (to appear)
- [2] Daniel Lemire, Fast Random Integer Generation in an Interval, *ACM Transactions on Modeling and Computer Simulation* (à paraître)
- [3] Wojciech Mula, Daniel Lemire, Faster Base64 Encoding and Decoding Using AVX2 Instructions, *ACM Transactions on the Web* (à paraître)
- [4] Daniel Lemire, Owen Kaser, Nathan Kurz, Luca Deri, Chris O'Hara, François Saint-Jacques, Gregory Ssi-Yan-Kai, Roaring Bitmaps: Implementation of an Optimized Software Library, *Software: Practice and Experience* **48** 4, 2018. <https://doi.org/10.1002/spe.2560>
- [5] Wojciech Mula, Nathan Kurz, Daniel Lemire, Faster Population Counts Using AVX2 Instructions, *Computer Journal* **61** 1, 2018. <https://doi.org/10.1093/comjnl/bxx046>
- [6] Antonio Badia et Daniel Lemire, On Desirable Semantics of Functional Dependencies over Databases with Incomplete Information, *Fundamenta Informaticae* **158** 4, 2018. <https://doi.org/10.3233/FI-2018-1651>
- [7] Daniel Lemire, Nathan Kurz, Christoph Rupp, Stream VByte: Faster Byte-Oriented Integer Compression. *Information Processing Letters* **130**, 2018. <https://doi.org/10.1016/j.ipl.2017.09.011>

- [8] Jing Li, Yuhong Yan, Daniel Lemire, Full Solution Indexing for top-K Web Service Composition, *IEEE Transactions on Services Computing* **11** 3, 2018. <http://dx.doi.org/10.1109/TSC.2016.2578924>
- [9] Dmytro Ivanchykhin, Sergey Ignatchenko, Daniel Lemire, Regular and almost universal hashing: an efficient implementation, *Software: Practice and Experience* **47** 10, 2017. <http://dx.doi.org/10.1002/spe.2461>
- [10] Daniel Lemire, Christoph Rupp, Upscaledb: Efficient Integer-Key Compression in a Key-Value Store using SIMD Instructions, *Information Systems* **66**, 2017. <http://dx.doi.org/10.1016/j.is.2017.01.002>
- [11] Samy Chambi, Daniel Lemire, Robert Godin, Vers de meilleures performances avec des Roaring bitmaps, *Technique et Science Informatiques* **35** 3, 2016.
- [12] Daniel Lemire, Gregory Ssi-Yan-Kai, Owen Kaser, Consistently faster and smaller compressed bitmaps with Roaring, *Software: Practice & Experience* **46** 11, 2016. <http://dx.doi.org/10.1002/spe.2402>
- [13] Daniel Lemire, Owen Kaser, Faster 64-bit universal hashing using carry-less multiplications, *Journal of Cryptographic Engineering* **6** 3, 2016. <http://dx.doi.org/10.1007/s13389-015-0110-5>
- [14] Daniel Lemire, Nathan Kurz, Leonid Boytsov, SIMD Compression and the Intersection of Sorted Integers, *Software: Practice & Experience* **46** 6, 2016. <http://dx.doi.org/10.1002/spe.2326>
- [15] Samy Chambi, Daniel Lemire, Owen Kaser, Robert Godin, Better bitmap performance with Roaring bitmaps, *Software: Practice & Experience* **46** 5, 2016. <http://dx.doi.org/10.1002/spe.2325>  
Cité au moins **58 fois**.
- [16] Owen Kaser and Daniel Lemire, Compressed bitmap indexes: beyond unions and intersections, *Software: Practice & Experience* **46** 2, 2016. <http://dx.doi.org/10.1002/spe.2289>
- [17] Adina Crainiceanu and Daniel Lemire, Multidimensional Bloom Filters, *Information Systems* **54**, 2015. <http://dx.doi.org/10.1016/j.is.2015.01.002>
- [18] Antonio Badia et Daniel Lemire, Functional dependencies with null markers, *Computer Journal* **58** 5, 2015. <http://dx.doi.org/10.1093/comjnl/bxu039>
- [19] Wayne Xin Zhao, Xudong Zhang, Daniel Lemire, Dongdong Shan, Jian-Yun Nie, Hongfei Yan, Ji-Rong Wen, A General

- SIMD-based Approach to Accelerating Compression Algorithms, *ACM Transactions on Information Systems* **45** 1, 2015. <http://dx.doi.org/10.1145/2735629>
- [20] Xiaodan Zhu, Peter Turney, Daniel Lemire, Andre Vellino, Measuring academic influence: Not all citations are equal, *Journal of the Association for Information Science and Technology* **66** 2, 2015. <http://dx.doi.org/10.1002/asi.23179>  
Cité au moins **56 fois**.
- [21] Daniel Lemire et Leonid Boytsov, Decoding billions of integers per second through vectorization, *Software: Practice & Experience* **45** 1, 2015. <http://dx.doi.org/10.1002/spe.2203>  
Cité au moins **114 fois**.
- [22] Owen Kaser et Daniel Lemire, Strongly universal string hashing is fast, *Computer Journal* **57** 11, 2014. <http://dx.doi.org/10.1093/comjnl/bxt070>
- [23] Hazel Webb, Daniel Lemire, Owen Kaser, Diamond Dicing, *Data & Knowledge Engineering* **86**, 2013. <http://arxiv.org/abs/1006.3726>
- [24] Zoltán Prekopcsák et Daniel Lemire, Time Series Classification by Class-Specific Mahalanobis Distances, *Advances in Data Analysis and Classification* **6** 3, 2012. <http://arxiv.org/abs/1010.1526>
- [25] Daniel Lemire, Owen Kaser, Eduardo Gutarra, Reordering Rows for Better Compression: Beyond the Lexicographic Order, *ACM Transactions on Database Systems* **37** 3, 2012. <http://arxiv.org/abs/1207.2189>
- [26] Daniel Lemire, The universality of iterated hashing over variable-length strings, *Discrete Applied Mathematics* **160** (4-5), 2012. <http://arxiv.org/abs/1008.1715>
- [27] Antonio Badia et Daniel Lemire, A Call to Arms: Revisiting Database Design, *SIGMOD Record* **40** 3, 2011. <http://arxiv.org/abs/1105.6001>
- [28] Daniel Lemire and Owen Kaser, Reordering columns for smaller indexes, *Information Sciences* **181** 12, 2011. <http://arxiv.org/abs/0909.1346>
- [29] Daniel Lemire and Owen Kaser, Recursive n-gram hashing is pairwise independent, at best, *Computer Speech and Language* **24** 4, 2010. <http://arxiv.org/abs/0705.4676>
- [30] Daniel Lemire, Owen Kaser, Kamel Aouiche, Sorting improves word-aligned bitmap indexes, *Data & Knowledge Engineering* **69** 1, 2010. <http://arxiv.org/abs/0901.3751>

- Cité au moins **109 fois**.
- [31] Daniel Lemire, Faster retrieval with a two-pass Dynamic-Time-Warping lower bound, *Pattern Recognition* **42** 9, 2009. <http://arxiv.org/abs/0811.3301>  
Cité au moins **122 fois**.
- [32] Daniel Lemire, Martin Brooks, Yuhong Yan, An optimal linear time algorithm for quasi-monotonic segmentation, *International Journal of Computer Mathematics* **86** 7, 2009. <http://arxiv.org/abs/0709.1166>
- [33] 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. <http://arxiv.org/abs/0905.2657>
- [34] Daniel Lemire and Owen Kaser, Hierarchical Bin Buffering: Online local moments for dynamic external memory arrays, *ACM Transactions on Algorithms* **4** 1, 2008.
- [35] Daniel Lemire, Streaming maximum-minimum filter using no more than three comparisons per element, *Nordic Journal of Computing* **13** 4, 2006.  
Cité au moins **60 fois**.
- [36] Owen Kaser and Daniel Lemire, Attribute value reordering for efficient hybrid OLAP, *Information Sciences* **176** 16, 2006.  
Cité au moins **33 fois**.
- [37] 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.  
Cité au moins **83 fois**.
- [38] Daniel Lemire, Scale and Translation Invariant Collaborative Filtering Systems, *Information Retrieval* **8** 1, 2005.  
Cité au moins **107 fois**.
- [39] 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.
- [40] 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.  
Cité au moins **83 fois**.
- [41] 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

---

- [42] Edmon Begoli, Jesús Camacho-Rodríguez, Julian Hyde, Michael J. Mior, Daniel Lemire, Apache Calcite: A Foundational Framework for Optimized Query Processing Over Heterogeneous Data Sources, ACM SIGMOD 2018, 2018.
- [43] Dara Aghamirkarimi, Daniel Lemire, Discovering the Smart Forests with Virtual Reality, ACM WIMS 2017, 2017.
- [44] Samy Chambi, Daniel Lemire, Robert Godin, Kamel Boukhalfa, Charles Allen, Fangjin Yang, Optimizing Druid with Roaring bitmaps, IDEAS 2016, 2016.
- [45] Jing Li, Yuhong Yan, Daniel Lemire, Scaling up Web Service Composition with the Skyline Operator, IEEE ICWS 2016, 2016.
- [46] Samy Chambi, Daniel Lemire, Robert Godin, Nouveaux modèles d'index bitmap compressés à 64 bits, EDA 2016, 2016.
- [47] Perrine Ruer, Charles Gouin-Vallerand, Le Zhang, Daniel Lemire, and Evelyne F. Vallières, An analysis tool for the contextual information from field experiments on driving fatigue, CONTEXT 2015, 2015.
- [48] Jing Li, Yuhong Yan, Daniel Lemire, A web service composition method based on compact K2-trees, IEEE SCC 2015, 2015.
- [49] Jeff Plaisance, Nathan Kurz, Daniel Lemire, Vectorized VByte Decoding, International Symposium on Web Algorithms 2015, 2015.
- [50] Samy Chambi, Daniel Lemire, Robert Godin, Owen Kaser, Roaring bitmap : nouveau modèle de compression bitmap, EDA 2014, 2014.
- [51] Jing Li, Yuhong Yan, Daniel Lemire, Full Solution Indexing Using Database for QoS-aware Web Service Composition, IEEE SCC 2014, 2014. **(Best Student Paper Award)**
- [52] Andre Vellino and Daniel Lemire, Extracting, Transforming and Archiving Scientific Data, VLDL 2011, 2011.
- [53] Owen Kaser, Daniel Lemire, Kamel Aouiche, Histogram-aware sorting for Enhanced Word-Aligned Compression in bitmap indexes, DOLAP 2008, 2008.
- [54] 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.
- [55] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning attributes from data cubes with Diamond Dicing, IDEAS 2008, 2008.
- [56] Kamel Aouiche, Daniel Lemire, Robert Godin, Collaborative OLAP with tag clouds: Web 2.0 OLAP formalism and experimental evaluation, WEBIST 2008, 2008.



- [57] Kamel Aouiche and Daniel Lemire, A comparison of five probabilistic view-size estimation techniques in OLAP, DOLAP 2007, 2007.
- [58] Owen Kaser and Daniel Lemire, Removing manually-generated boilerplate from electronic texts: Experiments with project Gutenberg e-books, CASCON 2007, 2007.
- [59] Owen Kaser and Daniel Lemire, Tag-Cloud Drawing: Algorithms for cloud visualization, Tagging and Metadata for Social Information Organization (WWW 2007), 2007.  
Cité au moins **257 fois**.
- [60] Kamel Aouiche and Daniel Lemire, Unassuming view-size estimation techniques in OLAP: An experimental comparison, ICEIS 2007, 2007.
- [61] Daniel Lemire, A better alternative to piecewise linear time series segmentation, SIAM Data Mining 2007, 2007.  
Cité au moins **71 fois**.
- [62] Dan Kucerovsky and Daniel Lemire, Monotonicity analysis over chains and curves, Curves and Surfaces 2006, 2007.
- [63] Owen Kaser, Daniel Lemire, Steven Keith, The LitOLAP Project: data warehousing with literature, CaSTA 2006, 2006.
- [64] Daniel Lemire, Martin Brooks, Yuhong Yan, An optimal linear time algorithm for quasi-monotonic segmentation, ICDM 2005, 2005.
- [65] Will Fitzgerald, Daniel Lemire, Martin Brooks, Quasi-monotonic segmentation of state variable behavior for reactive control, AAAI 2005, 2005.
- [66] Yuhong Yan, Martin Brooks, Daniel Lemire, Scale-based monotonicity analysis in qualitative modelling with flat segments, IJCAI 2005, 2005.
- [67] Daniel Lemire and Anna Maclachlan, Slope One predictors for online rating-based collaborative filtering. SIAM Data Mining 2005, 2005.  
Cité au moins **607 fois**.
- [68] Yuhong Yan, Daniel Lemire, Martin Brooks, Monotone pieces analysis for qualitative modeling, MONET 2004, 2004.
- [69] 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.  
Cité au moins **150 fois**.
- [70] Owen Kaser and Daniel Lemire, Attribute value reordering for efficient hybrid OLAP. DOLAP 2003, 2003.

- [71] Daniel Lemire, A family of 4-point dyadic high resolution subdivision schemes, *Curves and Surfaces 2002*, 2003.
- [72] Daniel Lemire, Wavelet-based relative prefix sum methods for range sum queries in data cubes, *CASCON 2002*, 2002. (**Best Paper Award**)

---

#### Livres

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

---

#### Numéros spéciaux et éditoriaux

- [74] Cameron Neylon, Jan Aerts, C. Titus Brown, Daniel Lemire, Jarrod Millman, Peter Murray-Rust, Fernando Perez, Neil Saunders, Arfon Smith, Gaël Varoquaux and Egon Willighagen, Changing computational research: The challenges ahead, *Source Code for Biology and Medicine* **7** (2), 2012.
- [75] Daniel Lemire and Richard Hotte (Eds.), Special issue on learning and the social web, *Journal of Emerging Technologies in Web Intelligence* **2** 1, 2010.
- [76] Mamadou Tadiou Koné and Daniel Lemire (Eds.), Special issue on Canadian Semantic Web, *Computational Intelligence* **23** 3, 2007.

---

#### Chapitres de livre

- [77] 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

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

---

#### Rapports techniques

- [79] Owen Kaser and Daniel Lemire, Threshold and Symmetric Functions over Bitmaps, UNBSJ CSAS Technical Report TR-14-001, 2014.
- [80] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning attributes from data cubes with Diamond Dicing, UNBSJ CSAS Technical Report TR-08-011, 2008.
- [81] 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.

- [82] Daniel Lemire and Owen Kaser, One-pass, one-hash n-gram statistics estimation, UNBSJ CSAS Technical Report TR-06-001, 2006.
- [83] Steven Keith, Owen Kaser, Daniel Lemire, Analyzing large collections of electronic text using OLAP, UNBSJ CSAS Technical Report TR-05-001, 2005.
- [84] 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.
- [85] Daniel Lemire, Local interpolation by high resolution subdivision schemes. Technical Report 200205-01, Acadia University, 2002.

#### Ouvrages d'enseignement

---

- [INF9004] Daniel Lemire, Informatique des entrepôts de données, 2012. Cours en ligne : <http://benhur.telug.ca/SPIP/inf9004/>.
- [INF6104] Daniel Lemire, Recherche d'informations et web, 2008. Cours en ligne : <http://benhur.telug.ca/SPIP/inf6104/>.
- [INF6460] Daniel Lemire, Recherche et filtrage d'informations, 2007. Cours en ligne : <http://benhur.telug.ca/SPIP/inf6460/>.
- [INF6450] Daniel Lemire, Gestion de l'information avec XML, ISBN 2762418747, 2005. (Révisé en 2006, 2007 et 2008.) En ligne : <http://www.telug.ca/inf6450/>.

#### Conférences publiques récentes

---

- “Engineering Fast Indexes for Big Data Applications” au Spark Summit East (Boston, 2017)
- “Algorithms: How content finds you” au Sommet de la découvrabilité (CRTC, Toronto, 2016)
- “Pour la perennité de nos contenus nationaux : l'enjeu de la visibilité” aux rencontres de l'ADISQ (Montréal, 2016)

#### Logiciel open source (sélection)

---

- Daniel Lemire et al., Roaring Bitmaps, 2014. (<http://roaringbitmap.org>)  
Je suis un des auteurs de plusieurs mises en œuvre de la structure de données Roaring (Java, C, C++, Go, Python, etc.). Cette librairie est utilisée par Apache Spark, Apache Kylin (le moteur OLAP de la société eBay), Metamarkets' Druid, Apache Hive, Apache Tez, Netflix Atlas, LinkedIn Pinot, Pilosa, Microsoft Visual Studio Team Services (VSTS), Jive Miru et ainsi de suite. Une mise en œuvre du format fait partie d'Apache Lucene est utilisée par des systèmes tels que Solr et Elastic. Lucene est le moteur de recherche utilisé par des sites tels que Wikipédia.

- Daniel Lemire et al., JavaFastPFOR: A simple integer compression library in Java, 2012. (<https://github.com/lemire/JavaFastPFOR>)  
Une classe dérivée de cette bibliothèque fait partie d'Apache Lucene. La bibliothèque a aussi été intégrée à Apache Parquet et Apache NiFi. Elle fait partie de LinkedIn Pinot. Une version modifiée de cette bibliothèque fait partie du moteur de recherche Terrier (<http://terrier.org/>).
- Daniel Lemire et al., The FastPFOR C++ library: Fast integer compression, 2012. (<https://github.com/lemire/FastPFor>)  
Cette bibliothèque a été régulièrement citée par des articles et communications scientifiques comme une référence et une base pour de nouvelles mises en œuvre.
- Daniel Lemire, FrameOfReference, 2015. (<https://github.com/lemire/FrameOfReference>)  
FrameOfReference a été adaptée par Apache Arrow et Apache Impala.
- Daniel Lemire et al., JavaEWAH, A compressed alternative to the Java BitSet class, 2009. (<https://github.com/lemire/javaewah>)  
Ce projet a reçu une reconnaissance du *Google's Open Source Programs* en 2012.  
JavaEWAH est utilisée par Apache Hive pour les index bitmap. Elle **fait partie des principales distributions Linux** (RedHat, Ubuntu). Elle est utilisée par JGit.
- Daniel Lemire, EWAHBoolArray: A compressed bitmap class in C++, 2008. (<https://github.com/lemire/EWAHBoolArray>)  
Utilisé par le moteur OLAP Hustle.  
Cette bibliothèque fut adaptée en C et intégrée au système de contrôle de version Git par les ingénieurs de GitHub. Ainsi donc, la plupart des utilisateurs de GitHub (et ils sont des millions) bénéficient de cette bibliothèque et des travaux de recherche l'ayant appuyée.

### **Passe-temps**

---

- Je passe plusieurs heures chaque semaine à jouer à des jeux vidéos. Je partage cette passion avec mes deux garçons.
- Mes auteurs de science-fiction préférés sont Charles Stross, Peter F. Hamilton et Alastair Reynolds.
- Je construis des voiliers radio-commandés.
- Depuis l'âge de douze ans, je programme des ordinateurs pour m'amuser.