Daniel Lemire, Ph.D.

Université TELUQ
Université du Québec
5800 Saint-Denis
Office 1105
Montreal (Quebec)
H2S 3L5 Canada

daniel.lemire@teluq.ca
blog: https://lemire.me/blog/
https://github.com/lemire/
Canadian citizen
bilingual (English and French)

Experience

2004–. . .	Computer Science Professor Université TELUQ Université du Québec – Ranked among the top 2% of scientists in the world (2025) according to the Stanford University/Elsevier list; – Promoted to Full Professor in 2009; – Department Chair (2017–2023); – Circle of Excellence (2019 and 2024), Université du Québec.
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 biomédical

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**

Ph.D. Students

- Fatma Miladi, Agents conversationnels pédagogiques basés sur l'intelligence artificielle générative pour améliorer l'apprentissage en ligne (Ph.D., 2025);
- François Lewis, Pertinence de l'ajout d'une composante sérieuse personnalisée à un jeu vidéo d'action dédié à la rééducation en lecture d'élèves ayant des difficultés d'apprentissage (Ph.D., 2024);
- Pierre Marie Ntang, Visualiser pour comprendre : application à la programmation vectorielle (Ph.D., 2023);
- Gary Germeil, La crédibilité du comportement d'un personnage non-joueur dans un jeu de type cartes à collectionner (Ph.D., 2022);
- Tarek Khei, Établir les correspondances entre les artéfacts d'un système d'information (Ph.D., 2020);
- Xueping Dai, Caractérisation du bruit électromagnétique dans les tunnels miniers profonds et application de la méthode de débruitage (Ph.D., 2019);
- Erick Aokou Koffi, Approche systémique de la conscience de la situation pour la réduction des erreurs de décision en environnement complexe : cas des projets informatiques (Ph.D., 2018);
- Badis Merdaoui, Simulation probabilistique d'acquisition des langues secondes (Ph.D., 2018);
- Samy Chambi, Les index bitmap compressés (Ph.D., 2017);
- Jing Li, Full Solution Indexing and Efficient Compressed Graph Representation for Web Service Composition (Ph.D., 2016);
- Hazel Webb, Properties and Applications of Diamond Cubes (Ph.D., 2010).

Professional Membership

- Association for Computing Machinery (ACM);
- National Academy of Scholars (NAS).

Research program

Individual Research Funding (external agencies, competitive grants)

- NSERC Discovery Grant 2024–2030 ($145,000) — Data Processing at Gigabytes Per Second;
- NSERC Accelerator Supplement 2017–2024 ($120,000);
- NSERC Discovery Grant 2017–2024 ($294,000) — Faster Compressed Indexes On Next-Generation Hardware;
- NSERC Engage (2018, $25,000);
- NSERC Discovery Grant 2012–2017 ($140,000);
- NSERC Discovery Grant 2007–2012 ($75,000);
- NSERC Discovery Grant 2003–2007 ($48,000);
- FRQNT New Researcher Grant 2006–2008 ($54,249).

Team-Based Research Funding (external agencies, competitive grants)

- AUF Grant (2024–2025, €20,000) with Valéry Psyché;
- CFI Grant (2020–2024, $720,000) with Hafedh Mili and Kim L Lavoie (PIs)
- NovaScience grant for Teaching AI (2020–2021, $140,000) with Valéry Psyché (PI);

- QPNR strategic grant on artificial intelligence (2020–2021, $45,000) with Isabelle Savard (lead);
- FRQNT team grant (2018–2021, $162,000) with Li Zhen Cheng (PI);
- CFI Grant 2016–2017 ($797,481) with N. Bélanger (PI) and E. Filotas;
- CAA Foundation (2013–2014, $189,994) with E. Vallières (PI);
- CFI Grant 2008–2009 ($999,618) with G. Paquette (PI) and P. Valtchev;
- New Brunswick Innovation Fund 2007–2008 ($10,000) with O. Kaser (PI).

Program committees (incomplete list)

- WSDM 2027: The 20th International Conference on Web Search and Data Mining (Senior Program Committee)
- CIKM 2026: 34th ACM International Conference on Information and Knowledge (Senior Program committee)
- ECMLPKDD 2026: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (Program committee)
- WSDM 2026: The Nineteenth International Conference on Web Search and Data Mining (Senior Program Committee)
- ECMLPKDD 2025: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (Program committee)
- CIKM 2025: 33rd ACM International Conference on Information and Knowledge (Senior Program committee)
- BIGDACI 2025: 10th International Conference on Big Data Analytics, Data Mining and Computational Intelligence (Program Committee)
- ACM SIGIR 2025: The 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (Program committee)
- WSDM 2025: The Eighteenth International Conference on Web Search and Data Mining (Senior Program Committee)
- ECMLPKDD 2024: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (Program committee)
- ACM SIGIR 2024: The 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (Program committee)
- ACM SIGKDD 2023: 29th SIGKDD Conference on Knowledge Discovery and Data Mining (Program committee)
- ECMLPKDD 2023: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (Program committee)
- ACM SIGIR 2023: The 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (Program committee)
- WSDM 2022 15th ACM International WSDM Conference (Program committee)
- EDA 2022: 18e journées EDA Business Intelligence and Big Data (Program committee)
- SIGIR 2022: The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (Program committee)
- DOLAP 2022: 24th International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data (Program committee)
- CIKM 2021: 30th ACM International Conference on Information and Knowledge (Program committee)

- EDA 2021: 17e journées EDA Business Intelligence and Big Data (Program committee)
- ECML/PKDD 21: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (Program committee)
- DOLAP 2021: 23nd International Workshop On Design, Optimization, Languages and Analytical Processing of Big Dat (Program committee)
- SIGIR 2021: 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (Program committee)
- ACM SIGKDD 2021: 27th International Conference on Knowledge Discovery and Data Mining (Program committee)
- WSDM 2021: 14th ACM International WSDM Conference (Program committee)
- ASD 2021: 13th edition of the Conference on Advances in the Science of Data (Program committee)
- CIKM 2020: 29th ACM International Conference on Information and Knowledge (Program committee)
- SIGIR 2020: 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (Program committee)
- ECML-PKDD 2020: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (Area Chair)
- BBIGAP'2020: Second International Workshop for Business Intelligence & Big Data Applications (Program committee)
- RecSys 2020: 14th ACM Recommender Systems Conference (Program committee)
- ADBIS 2020: 24th European Conference on Advances in Databases and Information System (Program committee)
- EDML20: Second Workshop on Evaluation and Experimental Design (Program committee)
- DOLAP 2020: 22nd International Workshop On Design, Optimization, Languages and Analytical Processing of Big Data (Program committee)
- DaWak 2020: 22nd International Conference on Big Data Analytics and Knowledge Discovery (Program committee)

Funding Agencies

- At NSERC, member of the Computer Science Evaluation Group (EG 1507) for the Discovery Grants Program (four years, 2018–2021). In 2019–2020, I was co-chair of the committee.
- At the FRQNT, member of the review committee 03F (Theoretical Computer Science) in 2007–2023.
- Still at the FRQNT, Imember of the review committee 309 (Computer Science team projects) in 2006–2007, in 2013–2014, 2014–2015 and 2016–2017.
- At NSERC, member of the Computer Science Research Tools and Instruments (RTI) Review Committee in 2012–2013, in 2013–2014 and in 2014–2015.

External examiner/referee

- Jury member for the University of Quebec prizes (2024);
- External examiner for Ph.D. thesis:

    - Azim Afroozeh from Vrije Universiteit Amsterdam (2026) – supervised by par Peter Boncz;
    - Lockman Saleh from UQAM (2025) – supervised by Hafedh Mili et Mounir Boukadoum;
    - Jaël Champagne Gareau from UQAM (2024) — supervised by Éric Beaudry;
    - Nathan Maurice from la Sorbonne, France (2024) — supervised by Lionel Lacassagne;
    - Nigel Medforth from Simon Fraser University, Canada (2022) — supervised by Robert Cameron;
    - Luca Versari from Pisa University, Italy (2021) — supervised by Roberto Grossi;
    - Kareem El Gebaly from Waterloo University, Canada (2018) — supervised by Jimmy Lin, Lukasz Golab and Ashraf Aboulnaga;
    - Mohammed Shaaban from Université Pierre et Marie Curie, France (2017) — supervised by Patrick Garda;
    - Mehdi Boukhechba from UQAC, Canada (2016) — supervised by Abdenour Bouzouane and Charles Gouin-Vallerand;
    - Hicham Assoudi from UQAM, Canada (2016) — supervised by Hakim Lounis;
    - Khaled Dehdouh from Lyon 2, France (2015) — supervised by Omar Boussaid;
    - Martin Leginus from Aalborg University, Denmark (2015) — supervised by Peter Dolog;
    - Ahmad Taleb from Concordia University, Canada (2011) — supervised by Todd Eavis.
- External examiner for M.Sc. thesis:
    - Benjamin Lapointe-Pinel from UQAR, Canada (2024) — supervised by Steven Pigeon.
- External referee for habilitation:
  Sabine Loudcher Rabaseda from Lyon 2 University, France (2011)
  Bérenger Bramas from Inria Nancy, France (2025).
- External referee for promotion:
    - Guadalupe Canahuate from the University of Iowa (2024);
    - Jason Sawin from the University of St. Thomas (2015);
    - Amer Nizar AbuAli from Philadelphia University (2013);
    - Ken Pu from the Ontario Institute of Technology (2011);
    - Jinan Fiaidhi from Lakehead University (2005).

In 2020, one of two external reviewers for the M.Sc. in computer science program assessment at UQAC.

Peer-Reviewed Journal Articles

[1] Robert Clausecker, Daniel Lemire, Fixing ill-formed UTF-16 strings with SIMD instructions, Software: Practice and Experience, 2026. `https://doi.org/10.1002/spe.70105`

[2] Jaël Champagne Gareau, Daniel Lemire, Converting an Integer to a Decimal String in Under Two Nanoseconds. Software: Practice and Experience **56** 8, 2026. http://doi.org/10.1002/spe.70079

[3] Jaël Champagne Gareau, Daniel Lemire, Converting binary floating-point numbers to shortest decimal strings: An experimental review, Software: Practice and Experience **56** 4, 2026. https://doi.org/10.1002/spe.70056

[4] Savard, Isabelle; Léveillée, Anna-Kim; Dellazizzo, Laura; Latulippe, Karine; Plante, Patrick; Angulo Mendoza, Gustavo Adolfo; Roy, Jean-Sébastien; Lemire, Daniel, Parcours d'apprentissage en santé : orchestrer les ressources en accès libre validées et les technologies d'intelligence artificielle au sein d'un environnement numérique SPÉCIAL. Médiations et médiatisations 23, 2026. https://doi.org/10.52358/mm.vi23.504

[5] Clausecker, Robert; Lemire, Daniel et Schintke, Florian, Faster Positional-Population Counts for AVX2, AVX-512, and ASIMD, Concurrency and Computation: Practice and Experience **37** 27–28, 2025.

[6] François Lewis, Patrick Plante, Daniel Lemire, Pertinence d'une composante sérieuse personnalisée dans un jeu vidéo d'action dédié à la rééducation en lecture de jeunes élèves présentant des symptômes dyslexiques, Médiations et médiatisations 21, 2025. https://doi.org/10.52358/mm.vi21.449

[7] Daniel Lemire, Scanning HTML at Tens of Gigabytes per Second on ARM Processors, Software: Practice and Experience **55** 7, 2025. https://doi.org/10.1002/spe.3420

[8] Jeroen Koekkoek, Daniel Lemire, Parsing Millions of DNS Records per Second, Software: Practice and Experience **55** 4, 2025. https://doi.org/10.1002/spe.3396

[9] Nevin Brackett-Rozinsky, Daniel Lemire, Batched Ranged Random Integer Generation, Software: Practice and Experience **55** 1, 2025. https://doi.org/10.1002/spe.3369

[10] Blandine Courcot, Daniel Lemire, Nicolas Bélanger, Dynamics of soil water potential as a function of stand types in a temperate forest: Emphasis on flash droughts. Geoderma Regional **38**, 2024. https://doi.org/10.1016/j.geodrs.2024.e00850

[11] John Keiser, Daniel Lemire, On-Demand JSON: A Better Way to Parse Documents?, Software: Practice and Experience **54** 6, 2024. https://doi.org/10.1002/spe.3313

[12] Yagiz Nizipli, Daniel Lemire, Parsing Millions of URLs per Second, Software: Practice and Experience **54** 5, 2024. http://doi.org/10.1002/spe.3296

[13] Daniel Lemire, Exact Short Products From Truncated Multipliers, Computer Journal **67** 4, 2024. http://doi.org/10.1093/comjnl/bxad077

[14] Robert Clausecker, Daniel Lemire, Transcoding Unicode Characters with AVX-512 Instructions, Software: Practice and Experience **53** 12, 2023. https://doi.org/10.1002/spe.3261

[15] Noble Mushtak, Daniel Lemire, Fast Number Parsing Without Fallback, Software: Practice and Experience **53** 7, 2023. http://doi.org/10.1002/spe.3198

[16] Tom Humeau, Isabelle Savard, Daniel Lemire, Pierre-Olivier Dionne, Gustavo Adolfo Angulo Mendoza, Patrick Plante, Anne Marie Pinard, Jean-Sébastien Roy, FORCES 3 : Exploitation à des fins pédagogiques des données d'un portail d'apprentissage de l'autogestion de la douleur. Développement d'une architecture de collecte et d'analyse de données et d'un module de suivi du développement des compétences. Médiations et médiatisations **12**, 2022. https://doi.org/10.52358/mm.vi12.287

[17] Bahman Abbassi, Li-Zhen Cheng, Michel Jébrak, Daniel Lemire, 3D Geophysical Predictive Modeling by Spectral Feature Subset Selection in Mineral Exploration. Minerals **12** 1296, 2022. https://doi.org/10.3390/min12101296

[18] Thomas Mueller Graf, Daniel Lemire, Binary Fuse Filters: Fast and Smaller Than Xor Filters, Journal of Experimental Algorithmics **27**, 2022. https://doi.org/10.1145/3510449

[19] Daniel Lemire, Wojciech Muła, Transcoding Billions of Unicode Characters per Second with SIMD Instructions, Software: Practice and Experience **52** 2, 2022. https://doi.org/10.1002/spe.3036

[20] Marcus D. R. Klarqvist, Wojciech Muła, Daniel Lemire, Efficient Computation of Positional Population Counts Using SIMD Instructions, Concurrency and Computation: Practice and Experience **33** 17, 2021. https://doi.org/10.1002/cpe.6304

[21] Daniel Lemire, Colin Bartlett, Owen Kaser, Integer Division by Constants: Optimal Bounds, Heliyion **7** 6, 2021. https://doi.org/10.1016/j.heliyon.2021.e07442

[22] John Keiser, Daniel Lemire, Validating UTF-8 In Less Than One Instruction Per Byte, Software: Practice and Experience **51** 5, 2021. https://doi.org/10.1002/spe.2920

[23] Daniel Lemire, Number Parsing at a Gigabyte per Second, Software: Practice and Experience **51** 8, 2021. https://doi.org/10.1002/spe.2984

[24] François Lewis, Patrick Plante, Daniel Lemire, Pertinence, efficacité et principes pédagogiques de la réalité virtuelle et augmentée en contexte scolaire: une revue de littérature, Médiations et médiatisations **5**, 2021.

[25] Thomas Mueller Graf, Daniel Lemire, Xor Filters: Faster and Smaller Than Bloom and Cuckoo Filters, Journal of Experimental Algorithmics **25** 1, 2020. https://doi.org/10.1145/3376122

[26] Wojciech Muła, Daniel Lemire, Base64 encoding and decoding at almost the speed of a memory copy, Software: Practice and Experience **50** 2, 2020. https://doi.org/10.1002/spe.2777

[27] Xueping Dai, Li Zhen Cheng, Jean-Claude Mareschal, Daniel Lemire, Chong Liu, New method for denoising borehole transient electromagnetic data with discrete wavelet transform, Journal of Applied Geophysics **168**, 2019. https://doi.org/10.1016/j.jappgeo.2019.05.009

[28] Geoff Langdale, Daniel Lemire, Parsing Gigabytes of JSON per Second, VLDB Journal **28** 6, 2019. https://doi.org/10.1007/s00778-019-00578-5

[29] Daniel Lemire, Owen Kaser, Nathan Kurz, Faster Remainder by Direct Computation: Applications to Compilers and Software Libraries, Software: Practice and Experience **49** 6, 2019. https://doi.org/10.1002/spe.2689

[30] Daniel Lemire, Fast Random Integer Generation in an Interval, ACM Transactions on Modeling and Computer Simulation **29** 1, 2019. https://doi.org/10.1145/3230636

[31] Daniel Lemire, Melissa E. O'Neill, Xorshift1024*, Xorshift1024+, Xorshift128+ and Xoroshiro128+ Fail Statistical Tests for Linearity, ACM Transactions on Modeling and Computer Simulation **350**, 2019. https://doi.org/10.1016/j.cam.2018.10.019

[32] Wojciech Muła, Daniel Lemire, Faster Base64 Encoding and Decoding Using AVX2 Instructions, ACM Transactions on the Web **12** 3, 2018. https://doi.org/10.1145/3132709

[33] 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

[34] Wojciech Muła, Nathan Kurz, Daniel Lemire, Faster Population Counts Using AVX2 Instructions, Computer Journal **61** 1, 2018. https://doi.org/10.1093/comjnl/bxx046

[35] Antonio Badia and 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

[36] 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

[37] 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

[38] 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

[39] 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

[40] Samy Chambi, Daniel Lemire, Robert Godin, Vers de meilleures performances avec des Roaring bitmaps, Technique et Science Informatiques **35** 3, 2016.

[41] 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

[42] 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

[43] 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

[44] 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

[45] 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

[46] Adina Crainiceanu and Daniel Lemire, Bloofi: Multidimensional Bloom Filters, Information Systems **54**, 2015. http://dx.doi.org/10.1016/j.is.2015.01.002

[47] Antonio Badia and Daniel Lemire, Functional dependencies with null markers, Computer Journal **58** 5, 2015. http://dx.doi.org/10.1093/comjnl/bxu039

[48] 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

[49] 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

[50] Daniel Lemire and Leonid Boytsov, Decoding billions of integers per second through vectorization, Software: Practice & Experience **45** 1, 2015. http://dx.doi.org/10.1002/spe.2203

[51] Owen Kaser and Daniel Lemire, Strongly universal string hashing is fast, Computer Journal **57** 11, 2014. http://dx.doi.org/10.1093/comjnl/bxt070

[52] Hazel Webb, Daniel Lemire, Owen Kaser, Diamond Dicing, Data & Knowledge Engineering **86**, 2013. http://arxiv.org/abs/1006.3726

[53] Zoltán Prekopcsák and 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

[54] 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

[55] Daniel Lemire, The universality of iterated hashing over variable-length strings, *Discrete Applied Mathematics* **160** (4-5), 2012. http://arxiv.org/abs/1008.1715

[56] Antonio Badia and Daniel Lemire, A Call to Arms: Revisiting Database Design, *SIGMOD Record* **40** 3, 2011. http://arxiv.org/abs/1105.6001

[57] Daniel Lemire and Owen Kaser, Reordering columns for smaller indexes, *Information Sciences* **181** 12, 2011. http://arxiv.org/abs/0909.1346

[58] 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

[59] Daniel Lemire, Owen Kaser, Kamel Aouiche, Sorting improves word-aligned bitmap indexes, *Data & Knowledge Engineering* **69** 1, 2010. http://arxiv.org/abs/0901.3751

[60] Daniel Lemire, Faster retrieval with a two-pass Dynamic-Time-Warping lower bound, *Pattern Recognition* **42** 9, 2009. http://arxiv.org/abs/0811.3301

[61] 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

[62] 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

[63] Daniel Lemire and Owen Kaser, Hierarchical Bin Buffering: Online local moments for dynamic external memory arrays, *ACM Transactions on Algorithms* **4** 1, 2008.

[64] Daniel Lemire, Streaming maximum-minimum filter using no more than three comparisons per element, *Nordic Journal of Computing* **13** 4, 2006.

[65] Owen Kaser and Daniel Lemire, Attribute value reordering for efficient hybrid OLAP, *Information Sciences* **176** 16, 2006.

[66] 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.

[67] Daniel Lemire, Scale and Translation Invariant Collaborative Filtering Systems, *Information Retrieval* **8** 1, 2005.

[68] 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.

[69] 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.

[70] 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

[71] Valéry Psyché, Awa Diattara, Nour El Mawas and Daniel Lemire, Learning with RAG-Enhanced Language Models: A Mixed-Methods Study, Communications in Computer and Information Science, 2026 (to appear).

[72] Faten Slama, Daniel Lemire, Revolutionizing Code Optimization: A Deep Dive into AI-Powered Tools for Software Development, The 2nd International Generative AI and Computational Language Modelling Conference (GACLM 2025), 2025.

[73] Faten Slama, Daniel Lemire, Enhancing Developer Productivity: Benchmarking LLM-Powered Tools like GitHub Copilot and TabNine in Real-Time Coding Environments, The 11th IEEE International Conference on Intelligent Data and Security (IEEE IDS 2025), 2025.

[74] Fatma Miladi, Valéry Psyché, Awa Diattara, Nour El Mawas, Daniel Lemire, Evaluating a GPT-4 and Retrieval-Augmented Generation-Based Conversational Agent to Enhance Learning Experience in a MOOC, The 17th International Conference on Computer Supported Education (CSEDU 2025), 2025.

[75] Faten Slama, Daniel Lemire, Comparative Analysis of Loop-Free Function Evaluation Using ChatGPT and Copilot with C Bounded Model Checking, The International Conference on Foundation and Large Language Models (FLLM2024), 2024.

[76] Valéry Psyché, Jocelyne Kiss, Thomas Forissier, Jonathan Proulx-Guimond, Valérie Payen Jean-Baptiste, Josiane Koumenda, Fatma Miladi and Daniel Lemire, CO3PEDIA - An AI-Assisted Context-Aware Authoring Tool for Collaborative Blended Learning, AIED 2024 - 25th International Conference on Artificial Intelligence in Education, 2024.

[77] Fatma Miladi, Valéry Psyché and Daniel Lemire, Comparative Performance of GPT-4, RAG-Augmented GPT-4, and Students in MOOCs, Workshop on Breaking Barriers with Generative Intelligence (BBGI) held at the ITS 2024 - 20th International Conference on Intelligent Tutoring Systems, 2024.

[78] Fatma Miladi, Valéry Psyché and Daniel Lemire, Leveraging GPT-4 for Accuracy in Education: A Comparative Study on Retrieval-Augmented Generation in MOOCs, AIED 2024 - 25th International Conference on Artificial Intelligence in Education (LBR Track), 2024.

[79] Fatma Miladi, Daniel Lemire and Valéry Psyché, Learning Engagement and Peer Learning in MOOC: A Selective Systematic Review, ITS 2023 - 19th International Conference on Intelligent Tutoring Systems, 2023.

[80] Daniel Lemire, Unicode at Gigabytes per Second, SPIRE 2021, 2021.

[81] Pierre Marie Ntang, Daniel Lemire, SIMDGiraffe: Visualizing SIMD Functions, VISIGRAPP 2021, 2021.

[82] 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. `https://doi.org/10.1145/3183713.3190662`

[83] Dara Aghamirkarimi, Daniel Lemire, Discovering the Smart Forests with Virtual Reality, ACM WIMS 2017, 2017.

[84] Samy Chambi, Daniel Lemire, Robert Godin, Kamel Boukhalfa, Charles Allen, Fangjin Yang, Optimizing Druid with Roaring bitmaps, IDEAS 2016, 2016.

[85] Jing Li, Yuhong Yan, Daniel Lemire, Scaling up Web Service Composition with the Skyline Operator, IEEE ICWS 2016, 2016.

[86] Samy Chambi, Daniel Lemire, Robert Godin, Nouveaux modèles d'index bitmap compressés à 64 bits, EDA 2016, 2016.

[87] 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.

[88] Jing Li, Yuhong Yan, Daniel Lemire, A web service composition method based on compact K2-trees, IEEE SCC 2015, 2015.

[89] Jeff Plaisance, Nathan Kurz, Daniel Lemire, Vectorized VByte Decoding, International Symposium on Web Algorithms 2015, 2015.

[90] Samy Chambi, Daniel Lemire, Robert Godin, Owen Kaser, Roaring bitmap : nouveau modèle de compression bitmap, EDA 2014, 2014.

[91] 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)**

[92] Andre Vellino and Daniel Lemire, Extracting, Transforming and Archiving Scientific Data, VLDL 2011, 2011.

[93] Owen Kaser, Daniel Lemire, Kamel Aouiche, Histogram-aware sorting for Enhanced Word-Aligned Compression in bitmap indexes, DOLAP 2008, 2008.

[94] 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.

[95] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning attributes from data cubes with Diamond Dicing, IDEAS 2008, 2008.

[96] Kamel Aouiche, Daniel Lemire, Robert Godin, Collaborative OLAP with tag clouds: Web 2.0 OLAP formalism and experimental evaluation, WEBIST 2008, 2008.

[97] Kamel Aouiche and Daniel Lemire, A comparison of five probabilistic view-size estimation techniques in OLAP, DOLAP 2007, 2007.

[98] Owen Kaser and Daniel Lemire, Removing manually-generated boilerplate from electronic texts: Experiments with project Gutenberg e-books, CASCON 2007, 2007.

[99] Owen Kaser and Daniel Lemire, Tag-Cloud Drawing: Algorithms for cloud visualization, Tagging and Metadata for Social Information Organization (WWW 2007), 2007.

[100] Kamel Aouiche and Daniel Lemire, Unassuming view-size estimation techniques in OLAP: An experimental comparison, ICEIS 2007, 2007.

[101] Daniel Lemire, A better alternative to piecewise linear time series segmentation, SIAM Data Mining 2007, 2007.

[102] Dan Kucerovsky and Daniel Lemire, Monotonicity analysis over chains and curves, Curves and Surfaces 2006, 2007.

[103] Owen Kaser, Daniel Lemire, Steven Keith, The LitOLAP Project: data warehousing with literature, CaSTA 2006, 2006.

[104] Daniel Lemire, Martin Brooks, Yuhong Yan, An optimal linear time algorithm for quasi-monotonic segmentation, ICDM 2005, 2005.

[105] Will Fitzgerald, Daniel Lemire, Martin Brooks, Quasi-monotonic segmentation of state variable behavior for reactive control, AAAI 2005, 2005.

[106] Yuhong Yan, Martin Brooks, Daniel Lemire, Scale-based monotonicity analysis in qualitative modelling with flat segments, IJCAI 2005, 2005.

[107] Daniel Lemire and Anna Maclachlan, Slope One predictors for online rating-based collaborative filtering. SIAM Data Mining 2005, 2005.

[108] Yuhong Yan, Daniel Lemire, Martin Brooks, Monotone pieces analysis for qualitative modeling, MONET 2004, 2004.

[109] 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.

[110] Owen Kaser and Daniel Lemire, Attribute value reordering for efficient hybrid OLAP. DOLAP 2003, 2003.

[111] Daniel Lemire, A family of 4-point dyadic high resolution subdivision schemes, Curves and Surfaces 2002, 2003.

[112] Daniel Lemire, Wavelet-based relative prefix sum methods for range sum queries in data cubes, CASCON 2002, 2002. **(Best Paper Award)**

Books

[113] Daniel Lemire, Faster Than You Think: Essays on Thinking Better and Building Faster, 163 pages, 2026. `https://www.amazon.com/dp/B0HDYG8FVK`

[114] Daniel Lemire, Maîtriser la programmation: Des tests à la performance en Go Paperback, 274 pages, 2025. `https://www.amazon.ca/-/fr/dp/B0FMPGWGFQ/`

[115] Daniel Lemire, Mastering Programming: From Testing to Performance in Go, 262 pages, 2025. `https://www.amazon.ca/-/fr/dp/B0FMPGSWR5/`

[116] Daniel Lemire, Neila Mezghani, Élodie Boissières, Robert Godin, Habib Louafi, Richmond Osei, Shadi Shuraida, Renée-Maria Schmitt, Dragos Vieru, La science des données: Théorie et applications avec R et Python, 300 pages, 2024. `https://www.amazon.ca/-/fr/dp/B0D53QXGKM/`

[117] Robert Godin, Daniel Lemire, Programmation avec Python: des jeux au Web, 322 pages, 2024. `https://www.amazon.ca/-/fr/Robert-Godin/dp/B0CVX9296P/`

[118] Robert Godin, Daniel Lemire, Java pas à pas, Introduction à la programmation et au langage Java, 272 pages, 2023. `https://www.amazon.ca/-/fr/Robert-Godin/dp/B0CR7RW87Y/`

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

Special Issues and editorials

[120] 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.

[121] Daniel Lemire and Richard Hotte (Eds.), Special issue on learning and the social web, Journal of Emerging Technologies in Web Intelligence **2** 1, 2010.

[122] Mamadou Tadiou Koné and Daniel Lemire (Eds.), Special issue on Canadian Semantic Web, Computational Intelligence **23** 3, 2007.

Book chapters

[123] Daniel Lemire, Patrick Plante, Nicolas Bélanger, Isabelle Savard, Marc-André Carle, Dolorès Grossemy et Caroline Boilard, Des robots conversationnels en soutien à l'apprentissage à distance, dans L'éducation à l'épreuve de l'intelligence artificielle (édité par Ambroise Baillifard et Henrietta Carbonel), Épistémé, 2025.

[124] 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

[125] Daniel Lemire, Will AI Flood Us with Irrelevant Papers? Communications of the ACM **67** 9, 2024.

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

Technical Reports

[127] Daniel Lemire, Colin Bartlett, Owen Kaser, Integer Division by Constants: Optimal Bounds, UNBSJ CSAS Technical Report TR-20-001, 2020.

[128] Owen Kaser and Daniel Lemire, Threshold and Symmetric Functions over Bitmaps, UNBSJ CSAS Technical Report TR-14-001, 2014.

[129] Hazel Webb, Owen Kaser, Daniel Lemire, Pruning attributes from data cubes with Diamond Dicing, UNBSJ CSAS Technical Report TR-08-011, 2008.

[130] 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.

[131] Daniel Lemire and Owen Kaser, One-pass, one-hash n-gram statistics estimation, UNBSJ CSAS Technical Report TR-06-001, 2006.

[132] Steven Keith, Owen Kaser, Daniel Lemire, Analyzing large collections of electronic text using OLAP, UNBSJ CSAS Technical Report TR-05-001, 2005.

[133] 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.

[134] Daniel Lemire, Local interpolation by high resolution subdivision schemes. Technical Report 200205-01, Acadia University, 2002.

Lecture Notes

[INF1220] Robert Godin, Daniel Lemire, Java Pas à Pas, Introduction à la Programmation et au Langage Java, 228 pages, 2023 `https://github.com/RobertGodin/JavaPasAPas`.

[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 (sample)

- simdjson: Parsing gigabytes of JSON per second, C++ `https://simdjson.org`
- RoaringBitmap: A better compressed bitset in Java `https://github.com/RoaringBitmap/RoaringBitmap`
- fast_float: fast and exact implementation of the C++ from_chars functions for float and double types `https://fastfloat.github.io/fast_float/`
- simdutf: Unicode routines: billions of characters per second `https://simdutf.github.io/simdutf/`
- WHATWG-compliant and fast URL parser written in modern C++ `https://github.com/ada-url/ada`

Recent Public Appearances

Most of my talks are available online at `https://youtube.com/playlist?list=PLlCBCOY8ivBf_G4erndahVgOf2_5Jj8bI`.

- "Ada: Parsing Millions of URLs per Second" at NodeConf EU 2023 (Ireland, 2023)
- "Accurate and efficient software microbenchmarks" at BID (Montreal, 2023)
- "Unicode at gigabytes per second" at SPIRE (virtual, 2021)
- "Parsing numbers at a gigabyte per second" at MIT Fast Code Seminar (virtual, 2021)
- "Floating-point Number Parsing w/Perfect Accuracy at GB/sec" at Go Systems Conf (virtual, 2020)
- "Parsing JSON Really Quickly: Lessons Learned" at QCon San Francisco (San Francisco, 2019)
- "Next Generation Indexes For Big Data Engineering" at ODSC East (Boston, 2018)
- "Engineering Fast Indexes for Big Data Applications" at the Spark Summit East (Boston, 2017)
- "Algorithms: How content finds you" at the Discoverability Summit (CRTC, Toronto, 2016)
- "Pour la perennité de nos contenus nationaux : l'enjeu de la visibilité" at the "rencontres de l'ADISQ" (Montreal, 2016)

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.Michael's College Dean's medal (class of 1994);
- 3T0 Scholarship (UofT).

Prizes

- Award of Excellence for Achievement in Research (all fields), Université du Québec (2020)