

7. LEAVES OF ABSENCE

University, Company or Organization at which Leave was taken	Type of Leave	Dates
PROOF Centre of Excellence	50% secondment	July 1, 2009– June 30, 2016
BC Cancer Research	Sabbatical	July 1, 2005– June 30, 2006
AT & T Research Center	Sabbatical	July 1, 1998–Feb 28, 1999
University of Munich	Sabbatical	April 1, 1999–June 30, 1999

8. TEACHING

(a) *Areas of special interest and accomplishments*

I. Curriculum Development: the Machine Learning Course

In July 2013, I was asked to teach CPSC 340 in September 2013. On reviewing the course content, and in consultation with several faculty members, I worked to turn 340 into more a hands-on laboratory oriented course than just a mathematics oriented course. In two months time, I added 10 new tutorials and 3 new assignments. The changes were generally well received by the students. Nonetheless, there is much work to be done.

Based on a student survey on the content, I got many ideas to improve the course. I obtained CWSEI funding to get one of my post-doctoral fellows, Yashar Mehdad, to design and implement a whole new set of tutorials and assignments, that would integrate better with the lectures. In particular, we designed 8 modules: with 3 modules on classification, 2 modules on regression, one on clustering, one on principal component analysis and one on mixture models. We also focused mainly on two text datasets, so that the students can see how different techniques can be applied to the same datasets. For the most part, the laboratory modules follow the lectures. Once relevant concepts and methods were introduced in the lectures, To complement the laboratory modules, we designed four new assignments: two on classification, one on clustering and the last one on ensembling. All the assignments and modules were done in R, which today is the most prominent data analytics language and software platform.

II. Curriculum Development: Database Courses

My area of specialization is database systems. Database systems is a rapidly changing/maturing field. However, for various reasons, UBC did not have a long tradition in this area before I joined. Consequently, in the past several years, I, together with George Tsiknis, have spent considerable effort to re-design and update the curriculum of CPSC 304 and 404, the sequence of undergraduate database courses in UBC. Specifically:

In 1995, we removed outdated topics from CPSC 304 (e.g., file management), and moved forward topics from CPSC 404 to 304 (e.g., introduction of the relational model). This allowed newer topics to be included in 404 (e.g., distributed databases).

In 1998, after making the observation that many of our co-op students worked with relational database systems in their work terms, we re-designed CPSC 304 so that we could teach all the essential relational database skills (e.g., formulation of SQL queries) as early as possible. This also forced a new curriculum to be developed for CPSC 404. The new curriculum includes some topics previously taught in the old version of CPSC 304 (e.g., query optimization), as well as emerging topics (e.g., data mining). The implementation of this curriculum change took 2 years to complete. I was the first to teach the new

CPSC 304 curriculum in 1999, and the first to teach the new 404 material in 2000.

In 2001, we developed a set of new tutorials for CPSC 304 to teach students JDBC, which allows the students to perform database operations within a Java host program. This is a significant job skill.

The changes that we did over the years appear to be very well received by the students. The two courses are now two of the most popular courses among our major students. Moreover, since the year 2000, CPSC 304 has become a required course for the software engineering option of the undergraduate program of the department of Electrical and Computer Engineering.

III. Curriculum Development: a Health Informatics Course

As I have interacted more and more people in the health care and medical sectors because of my research, it has been increasingly obvious that health professionals require a basic level of training in computer science. This belief is shared by many (not all) in the faculty of Medicine and the Medical School. Thus, in 2004, I put together a UBC Teaching and Enhancement Learning Fund (TLEF) project with Dr. Kendall Ho, associate Dean of Medicine, on designing a health informatics course that introduces computer science concepts important to health practice. This project was funded from April 2004 to March 2005; and it has just been renewed for another year. Through various means (including interviews with key informants), we identified a list of computer science topics considered essential to health professionals. The course IHHS 302 was first offered in 2006. As listed on the UBC Calendar, the title of the course is “Topics in Health Informatics for Health/Life Sciences Students”. The course was offered once in Point Grey campus and once in Okanagan campus. Both were very popular with the students. A bigger initiative on health informatics involves Continuing Studies. I have been a member of an executive working group exploring the possibility of UBC offering an applied Masters degree on health informatics. The working group consists of people from Faculty of Medicine, School of Nursing and Continuing Studies. While the discussions can be tricky at times, steady progress is being made.

IV. Curriculum Development: CPSC 111

Having taught CPSC 111, I observe that the existing course can be improved by recognizing that high school students have wildly varied training on information technology. Challenge exams help but are not adequate. Recently I explore with Paul Carter another possibility that we tentatively call CPSC 112. This course aims to provide a fast-track version of 111 for students with good background and to enrich the experience with a course project. We plan to submit a curriculum proposal soon, and hope to have the course offered in 2006.

V. Curriculum Development: CPSC 504 and 524

Laks Lakshmanan and I have intended to create a graduate level “general-purpose” database course for some time now. But with Rachel Pottinger’s arrival, this course became a reality and was created in 2004. The curriculum proposal passed through the Senate in 2005. The course provides an introduction to a few core topics in database systems research. The plan is to offer the course September 2005.

Before the aforementioned change, CPSC 504 was the data mining course. This data mining course is now given its proper title as CPSC 524. As part of the process, the course is updated with the latest topics and results. The course passed through the Senate in 2005.

VI. Honours Hour Experiment

Given the large class sizes of CPSC 304 and 404, there is a tendency that the best students find the pace of the courses much lower than they can handle. To allow them to learn more on the subject area, and to give them more challenges, in 2000 I experimented with a simple extension to CPSC 404. The so-called “honours hour” program added one hour per week to the regular classes. It covered interesting,

state-of-the-art research material not included in any exam. The students could freely drop in and out, and they came mainly for the knowledge. Throughout the entire term, there were at least 20 students who came voluntarily. I repeated the Honours Hour in 2001 and 2002. They were equally successful.

VII. Student Advising/Management

Since the fall of 2000, I have been a faculty advisor advising undergraduate and second degree students. I observed that certain parts of the advising can be automated with database technologies. Thus, starting from January 2001, I have been co-leading a project with George Tsiknis on developing student management tools. These tools help to manage student appointments, perform graduation and pre-requisite checks, and conduct course planning. Six undergraduate students developed a prototype version of these tools as CPSC 448 in the 00W session. During the summer of 2001, an undergraduate coop student was hired to further develop these tools. While my involvement was reduced by the participation of Patrice Belleville, the development continues. Recently the system is deployed to other departments of Science with great success.

VIII. Undergraduate Directed Studies

I have also been active in supervising undergraduate students either for their honour theses or directed studies. The following table provides a summary.

Session	Course Number	Number of students
S94	CPSC 448	3
W95	CPSC 448	2
S00	CPSC 448	1
F00	CPSC 449	2
W00	CPSC 448	7
W01	CPSC 448	2
S02	CPSC 448	1
S02	CPSC 449	1
W02	CPSC 448	2
W02	CPSC 449	1
W03	CPSC 449	1
W04	CPSC 449	1
W10	CPSC 448	2
W12	CPSC 448	1
W13	CPSC 448	1
W15	CPSC 448	1

(b) *Courses Taught at UBC*

Session	Course Number	Scheduled Hours	Class Size	Hours Taught			
				Lectures	Tutorials	Labs	Other
F92	CPSC 404	3/week	51	3/week			
W92	CPSC 504	3/week	13	3/week			
F93	CPSC 404	3/week	55	3/week			
W93	CPSC 504	3/week	7	3/week			
F94	CPSC 404	3/week	98	3/week			
W94	CPSC 504	3/week	12	3/week			
F95	CPSC 404	3/week	99	3/week			
W95	CPSC 504	3/week	14	3/week			
F96	CPSC 404	3/week	120	3/week			
W96	CPSC 504	3/week	14	3/week			
F97	CPSC 504	3/week	7	3/week			
W97	CPSC 304	3/week	180	3/week			
F99	CPSC 304	3/week	135	3/week			
F99	CPSC 504	3/week	12	3/week			
F00	CPSC 404	3/week	110	3/week			
W00	CPSC 504	3/week	43	3/week			
F01	CPSC 404	3/week	145	3/week			
W01	CPSC 304	3/week	135	3/week			
S02	CPSC 304	3/week	120	3/week			
F02	CPSC 504	3/week	16	3/week			
F03	CPSC 504	3/week	21	3/week			
W03	CPSC 111	3/week	86	3/week			
F04	CPSC 304	3/week	84	3/week			
F04	CPSC 111	3/week	77	3/week			
W07	CPSC 304	3/week	82	3/week			
W08	CPSC 304	3/week	90	3/week			
S10	CPSC 304	3/week	51	3/week			
W13	CPSC 340	3/week	93	3/week			
W14	CPSC 340	3/week	111	3/week			
W15	CPSC 304	3/week	155	3/week			

111 Introduction to Computation
304 Introduction to Relational Databases
340 Introduction to Machine Learning
404 Advanced Relational Databases
504 Database Design

(c) *Graduate Students Supervised*

Student Name	Program Type	Year		Principal Supervisor	Co-Supervisor	Current Employer
		Start	Finish			
Jinhai Yang	M.Sc.	Jul. '93	Aug. '94	R. Ng		unknown
Xiaomei Tian	M.Sc.	Jul. '93	Sep. '94	R. Ng		Glenayre
Ed Knorr	M.Sc.	May '94	Aug. '95	R. Ng		UBC
Andishe Sedighian	M.Sc.	Jul. '94	Dec. '95	R. Ng		unknown
Dominic Tam	M.Sc.	May '95	Apr. '96	R. Ng		ALI
YuQing Yu	M.Sc.	May '95	Jun. '96	R. Ng		Glenayre
Dwi Faulus	M.Sc.	May '95	Aug. '96	R. Ng		MDA
Paul Shum	M.Sc.	May '95	Nov. '96	R. Ng		ALI
Carson Leung	M.Sc.	May '96	Aug. '97	R. Ng		UBC
Yaping Shi	M.Sc.	May '96	Aug. '97	R. Ng		Glenayre
Alex Pang	M.Sc.	Apr. '97	May '98	R. Ng		Lehman Brothers
Rita Dilek	M.Sc.	Oct. '94	Aug. '98	R. Ng		unknown
Malcolm Steenburgh	M.Sc.	Sep. '97	Aug. '98	R. Ng		Point Grey Research
Ivy Kwok	M.Sc.	Sep. '97	Aug. '99	R. Ng		Motorola
Jason Madar	M.Sc.	Sep. '97	Aug. '99	R. Ng		IBM
Teresa Mah	M.Sc.	Sep. '97	Aug. '99	R. Ng		Microsoft
Xiaoding Yi	M.Sc.	Sep. '97	Apr. '00	R. Ng	G. Tsiknis	unknown
Maja Dimitrijevic	M.Sc.	Jan. '00	Aug. '02	R. Ng		
Min Phan	M.Sc.	May '00	May '02	R. Ng		Crystal Decisions
Ruiyao Yang	M.Sc.	May '01	Dec '02	R. Ng		
Xiaodong Zhou	M.Sc.	May '01	Dec '02	R. Ng		UBC
Andrew Kwon	M.Sc.	May '01	Apr '03	H. Hoos	R. Ng	CMMT
Yuhan Cai	M.Sc.	May '03	Apr '04	R. Ng		U. Washington
Shofeng Bu	M.Sc.	May '03	Dec '04	R. Ng		UBC
Timothy Chan	M.Sc.	Jan. '04	Dec '05	R. Ng		
Zsuzanna Hollander	M.Sc.	Jan. '04	Sep '05	R. Ng		UBC
Qiang Kong	M.Sc.	Apr. '04	Mar '06	R. Ng		
Mingyue Tan	M.Sc.	Apr. '04	Mar '06	R. Ng		Microsoft
Jan Ulrich	M.Sc.	Sep. '07	Aug. '08	G. Carenini	R. Ng	
Maryam Tavafi	M.Sc.	Jul. '12	Aug. '13	G. Carenini	R. Ng	Google
Jill Slind	M.Sc.	Jul. '12	Jun '14	R. Ng		
Mandeep Takhar	M.Sc.	Jan. '13	Aug '14	R. Ng		
Tatsuro Oya	M.Sc.	Aug. '13	Aug '14	G. Carenini	R. Ng	
Weicong Liao	M.Sc.	May '14	Aug '15	G. Carenini	R. Ng	Microsoft
Bitan Nejat	M.Sc.	Jan '15		G. Carenini	R. Ng	
Kailang Jiang	M.Sc.	Jan '15		R. Ng	G. Carenini	
Halldor Thorhallsson	M.Sc.	Jan '16		R. Ng		
Louie Dinh	M.Sc.	Sep '15		R. Ng	S. Mostafavi	
Jordan Johnson	M.Sc.	Jan '16		G. Carenini	R. Ng	
Vaden Masrani	M.Sc.	Jan '16		G. Carenini	R. Ng	
Ed Knorr	Ph.D.	May '96	Apr. '02	R. Ng		UBC
Carson Leung	Ph.D.	Jan. '98	Dec. '02	R. Ng		Univ. Manitoba
Fatemah Alqallaf	Ph.D. (Stat.)	May '01	Apr. '03	Ruben Zamar	R. Ng	Kuwait Univ.
Xiaodong Zhou	Ph.D.	Jan '03	May '08	R. Ng		AOL
Sohrab Shah	Ph.D.	Apr '04	Sep '08	R. Ng		UBC
Shaofeng Bu	Ph.D.	Jan '05	Oct '10	R. Ng	Microsoft	
Hongrae Lee	Ph.D.	Sep. '06	Aug. '10	R. Ng		Google

Student Name	Program Type	Year		Principal Supervisor	Co-Supervisor	Current Employer
		Start	Finish			
Jafiq Joty	Ph.D	Sep. '08	Aug. '13	G. Carenini	R. Ng	Qatar Research
Shu Yang	Ph.D	Jan '15		R. Ng		
Joerg Sanders	Postdoctoral	Aug. '01	Dec. '02	R. Ng		Univ. Alberta
Ganesh Ramesh	Postdoctoral	Sep. '05	Dec. '07	R. Ng	L. Lakshmanan	Microsoft
Gabriel Murray	Postdoctoral	Nov. '07	Jun. '12	G. Carenini	R. Ng	Fraser Valley Univ.
Yashar Mehdad	Postdoctoral	Jul. '12	Dec. '14	G. Carenini	R. Ng	Yahoo Research
Shima Gerani	Postdoctoral	Aug. '13	Dec. '15	G. Carenini	R. Ng	
Feng Xu	Postdoctoral	Mar. '16		R. Ng		

(d) *Continuing Education Activities*

(e) *Visiting Lecturer (indicate university/organization and dates)*

(f) *Other: Undergraduate Students Supervised*

Student Name	Nature of Supervision	Year	
		Start	Finish
Chris Law	CPSC 448	May '94	Aug. '94
Yau Soon Loo	CPSC 448	May '94	Aug. '94
Bobby Pak Soo	CPSC 448	May '94	Aug. '94
Adeline Pang	NSERC intern	May '95	Aug. '95
Kin-Bong Fung	CPSC 448	Jan. '96	Apr. '96
Doris Lee	CPSC 448	Jan. '96	Apr. '96
Yuhan Cai	NSERC intern	May '00	Aug. '00
Monica Sleumer	CPSC 448	May '00	Aug. '00
Kelly Hon	CPSC 449	Sep. '00	Dec. '00
Monica Sleumer	CPSC 449	Sep. '00	Dec. '00
Mandy Choi	CPSC 448	Jan. '01	Apr. '01
Brian Hunter	CPSC 448	Jan. '01	Apr. '01
Joanne Leung	CPSC 448	Jan. '01	Apr. '01
Tom Shum	CPSC 448	Jan. '01	Apr. '01
Michelle Tang	CPSC 448	Jan. '01	Apr. '01
Fiona Wong	CPSC 448	Jan. '01	Apr. '01
Cynthia Zheng	CPSC 448	Jan. '01	Apr. '01
Bryon Kuo	CPSC 448	Jan. '02	Apr. '02
Timothy Chan	CPSC 448	Jan. '02	Apr. '02
Judy Chu	CPSC 449	May '02	Aug. '02
Paul Yan	CPSC 448	May '02	Aug. '02
Karen Brennan	CPSC 449	Jan. '03	Apr. '03
Zsusanna Hollander	CPSC 448	Jan. '03	Apr. '03
Ladan Mahabadi	CPSC 448	Jan. '03	Apr. '03
Grace Zhang	CPSC 449	Sep. '03	Apr. '04
Matt Ingham	CPSC 449	Sep. '04	Apr. '05
Bitu Nejat	CPSC 448	Jan. '13	Apr. '13
Shirley Yin	CPSC 448	Sep. '13	Apr. '14
Frances Russell	CPSC 448	Sep. '14	Apr. '15
Tom Jin	CPSC 448	Sep. '15	Apr. '16

9. SCHOLARLY AND PROFESSIONAL ACTIVITIES(a) *Areas of special interest and accomplishments*

My general area of research is database systems. In recent years, my research program revolves around three main areas: data mining, bioinformatics and image database management. Image database management can be viewed as a sub-area of multimedia systems, and I collaborate with colleagues in computer vision.

Data mining is a joint area between databases, statistics and machine learning, and I collaborate with researchers from the latter two areas. My collaborators include Ruben Zamar from Statistics, Laks Lakshmanan, Alan Wagner and Nando de Freitas.

I collaborate with two groups of medical researchers. The first group consists of Victor Ling, Wan Lam, Stephen Lam and Calum Macaulay from the BC Cancer Research Centre. The main focus of our work is on genomic analysis for lung cancer. I am an affiliated member of the Research Centre.

The second group of my collaborators consists of Bruce McManus and David Granville from the iCapture Centre at St Paul, and Rob McMaster from Medical Genetics at UBC. Our focuses are on transplanted organs and heart valve diseases. I am a principal investigator of the Centre.

- (b) *Research or equivalent grants (indicate under COMP whether grants were obtained competitively (C) or non-competitively (NC))*

Granting Agency	Subject	Comp	\$ Per Year	Year	Principal Investigator	Co-Investigators
CFI Cyber Infrastructure	MERIDIAN: Marine Environmental Research Infrastructure for Data Integration and Application Network	C	\$1,600,000	2016-2019	Stan Matwin	Ng, Pottinger, Carenini and 10 others
Yahoo Research	An Effective and Scalable Email Summarization System	C	\$40,000	2015-2016	G. Carenini	R. Ng
NSERC Discovery	Managing and Modeling Time in Genomics	C	\$62,000	2014-2019	R. Ng	
Genome Canada	Clinical Implementation and Outcomes Evaluation of Blood-based Biomarkers for COPD Management	C	\$1,800,000	2013-2017	R. Ng	D. Sin
Genome BC	Clinical Implementation of Diagnostic Biomarker Assays in Heart and Kidney Transplantation	C	\$1,100,000	2012-2014	B. McManus	Keown, McMaster Ng
NSERC Discovery	A Data Mining Framework for Genomics Biomarkers and Signature Identification	C	\$70,000	2009-2014	R. Ng	
NSERC Strategic Network	Business Intelligence Network	C	\$1,000,000	2009-2014	R. Miller	Ng et al. (13 others)
NSERC Strategic	Nanosilver Effects on Amphibian wildlife using novel molecular assays	C	\$250,000	2009-2012	C.Helbing	Borchers, Lesperance, Ng
Genome BC	Frog Sentinel Comparative Omics for the Environment	C	\$150,000	2009-2010	C.Helbing	Borchers, Lesperance, Ng
CIHR	Biological Networks Regulating Organ Transplantation Outcome	C	\$100,000	2009-2010	A.Mui	Ng et al. (3 others)
Google	Summarization of Evaluative Text and Emails (granted)	C	US\$34,000	2007-2008	R.Ng	G.Carenini
Genome Canada	Better Biomarkers of Acute and Chronic Allograft Rejection	C	\$3,033,000	2004-2007	B. McManus	R. Ng et al.
Genome Canada	Application of Pharmacogenomics for Rational Chemotherapy of Lung Cancer	C	\$2,300,000	2004-2007	V. Ling	R. Ng et al.

Granting Agency	Subject	Comp	\$ Per Year	Year	Principal Investigator	Co-Investigators
Genome Canada	Development and Validation of Comparative Genomic Hybridization Arrays for Clinical Use in Cancer	C	\$1,500,000	2004-2007	W. Lam	R. Ng et al.
NSERC	Robust Mining of Biomedical Data	C	\$54,700	2004-2009	R. Ng	
TLEF	Information Technology Fundamentals	C	\$40,000	2004-2005	R. Ng	K. Ho et al.
NSERC	Robust Tools for Biomedical Data	C	\$48,686	2004-2004	R. Ng	
IRIS	Acquisition, Querying and Prediction of Motion Trajectories	C	\$190,000	2002-2005	R. Ng	Little et al.
IRIS	New Frontiers in Data Mining	C	\$145,000	2002-2005	L. Lakshmanan	R. Ng et al.
IRIS	Intelligent Computational Methods for the Analysis of Gene Expression Profiles	C	\$193,000	2002-2005	I. Jurisica	J. Glasgow H. Hoos R. Ng
MITACS	Towards Interactive Data Mining	C	\$130,000	2002-2004	R. Ng	K. Sevcik et al.
CFI	Institute for Computing, Information and Cognitive Systems	C	\$22,137,000	2000-2000	R. Ward	R. Ng et al.
NSERC	Funding for IBM Netfinity 5500 Equipment Grant	C	\$55,725	2000-2000	R. Ng	A. Wagner
NSERC	Optimized Tools for Exploratory Mining of Large Databases	C	\$45,000	2000-2004	R. Ng	
NSERC	Funding for IBM Netfinity 5500 Equipment Grant	C	\$55,725	2000-2000	R. Ng	A. Wagner
IBM Canada	Data Mining on the Internet	C	\$25,000	99-2001	R. Ng	
NSERC	Visual information management and analysis	C	\$35,500	96-2000	R. Ng	

Granting Agency	Subject	Comp	\$ Per Year	Year	Principal Investigator	Co-Investigators
IRIS	Building, Querying, Analyzing and Mining Data Warehouses on the Internet	C	\$180,000	98-2002	J. Han	L. Lakshmanan A. Mendelzon R.Ng
IRIS	Tools for Visual Information Management in Geographical and Biomedical Applications	C	\$330,000	98-2002	J. Snoeyink	R. Ng et al.
IBM Canada	Data Mining on the Internet	NC	\$25,000	98-99	R. Ng	
Workers' Compensation Board	Evaluation of Data Mining Opportunities in WCB	NC	\$6,500	98-99	R. Ng	
BC ASI	Evaluation of Data Mining Opportunities in WCB	C	\$6,500	98-99	R. Ng	
NSERC	Visual information management and analysis (Equipment Grant)	C	\$37,749	96-97	R. Ng	
NSERC	Distributed Multimedia Systems and Highspeed Communication Protocols (Equipment grant)	C	\$138,224	95-96	S. Vuong	R. Ng et al.
IRIS	Content- based Analysis and Abstraction for Visual Information Management	C	\$109,876	94-98	B. Woodham	J. Mulder R. Ng J. Snoeyink
IRIS	DBLEARN - Knowledge Discovery in Large Databases	C	\$171,278	94-98	J. Han	R. Ng et al.
NSERC	Semantics, computational and optimization techniques for databases with uncertainty and rules	C	\$24,000	93-96	R. Ng	
CITR	Continuous-media File Systems	C	\$125,000	93-98	G. Neufeld	N. Hutchinson M. Ito R. Ng

Granting Agency	Subject	Comp	\$ Per Year	Year	Principal Investigator	Co-Investigators
NSERC	Logging, Annotation and Navigation for Hypermedia Video Analysis Tools	C	\$102,060	92-95	K. Booth	R. Ng et al.
CICSR	Research Startup Funds	C	\$100,000	92-93	R. Ng	
UBC NSERC	Semantics, computational and optimization techniques for databases with uncertainty and rules (Equipment grant)	C	\$22,244	92-93	R. Ng	

NSERC Natural Sciences and Engineering Research Council
 IRIS Institute for Robotics & Intelligent Systems
 CITR Canadian Institute for Telecommunications Research
 CICSR Centre for Integrated Computer Systems Research

(c) *Research or equivalent contracts (indicate under COMP whether grants were obtained competitively (C) or non-competitively (NC))*

(d) *Invited Presentations*

- “Biomarker Discovery in Personalized Medicine,” Plenary Invited Talk, inaugural Fraser Lecture Series, Bellingham, Washington, April 8, 2016.
- “Data Science Meets Personalized Medicine,” Plenary Invited Talk, Symposium on Health Informatics, UBC Centennial Series, Vancouver, April 5, 2016.
- “Mining and Summarizing Conversation Data,” Plenary Invited Talk, Big Data 2016, Bilbao, Spain, February 13, 2016.
- “Biomarker Discovery in Personalized Medicine,” Plenary Invited Talk, International Conference on Advanced Information Technology, Services and Systems 2015, Marrakech, Morocco, December 16, 2015.
- “Biomarker Discovery in Personalized Medicine,” Plenary Invited Talk, 7th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Nancy, France, September 16, 2014.
- “Discourse Analysis and Summarization of asynchronous conversations,” Invited talk, National Institute of Informatics, Tokyo Japan, December 2013.
- “Conversation Extraction and Summarization,” Invited talk, VLDB 2012 Workshop on Real-time Business Intelligence, August 2012.

- “Better Medicine and Cheap Cost: Opportunities for Biomarker Development,” Invited talk, International Conference on Data Mining Workshop on Biological Data Mining and its Applications to Health Care, December 2011.
- “Extraction and Summarization of Conversations,” Invited talk, Singapore National University, April 2011.
- “Combinatorial Biomarker Development for End-stage Organ Failures,” Invited talk, Singapore National University, April 2011.
- “Conversation Extraction and Summarization,” Invited talk, University of Waterloo, January 2010, Canada.
- “Biomarkers in Transplantation,” Invited lecture, SAMSUNG Research, November 2009, Korea.
- “Conversation Extraction and Summarization,” Invited plenary lecture, Seoul National University, November 2009, Korea.
- “Conversation Extraction and Summarization,” Invited lecture, SAP Business Object, October 2009, Vancouver.
- “Data Mining Genomic Applications”, Distinguished Lecture Series, University of Alberta, March 2005.
- “Tools for Gene Expression Analysis” IBM Toronto Laboratory, May 2001.
- “Exploratory Association Rule Mining with Constraints”, IBM T. J. Watson Research Center, October, 1998.
- “Exploratory Association Rule Mining with Constraints”, Bell Laboratories, October, 1998.
- “Exploratory Association Rule Mining with Constraints”, University of Memphis, November, 1998.
- “Exploratory Association Rule Mining with Constraints”, University of Toronto, December, 1998.
- “Spatial Data Mining: Progress and Challenges”, AT &T Research Center, July 1997.
- “Spatial Data Mining: Progress & Challenges”, University of Hong Kong, December 1995.
- “Visual Information Management: Retrieval & Search”, Hong Kong University of Science & Technology, December 1995.
- “Visual Information Management: Retrieval & Search”, University of Alberta, April 1995.
- “Visual Information Management: Retrieval & Search”, George Mason University, March 1995.
- “Efficient and Effective Clustering Methods for Spatial Data Mining”, Concordia University, June 1994.
- “Maximizing Buffer and Disk Utilizations for News On-Demand”, McGill University, June 1994.

- “Efficient and Effective Clustering Methods for Spatial Data Mining”, University of Toronto, May 1994.
- “Parametric Query Optimization”, University of Aalborg, Denmark, July 1993.
- “Parametric Query Optimization”, Simon Fraser University, April 1993.
- “Flexible and Adaptable Buffer Management Techniques”, Oregon Graduate Institute, October 1992.

(e) *Other Presentations*

- “Data Warehousing and OLAP for Decision Support”, (Tutorial), 6th International Conference on Information & Knowledge Management, Las Vegas, November 1997.

(f) *Other*

The following table summarizes sabbatical or postdoctoral visitors that collaborated with me.

Visitor Name	Organization	Visit Type	Year	
			Start	Finish
Dr. Jiawei Han	SFU	Sabbatical	Jun. '97	Aug. '97
Dr. Jef Wijsen	Antwerp Univ., Belgium	Research	Dec. '97	Feb. '98
Dr. Jorg Sander	Munich Univ., Germany	Postdoctoral	Jul. '00	Aug. '01
Dr. Hideaki Kanai	Univ. Electro-communications, Japan	Research	Oct. '00	Aug. '01
Zunchao Li	Jiaotong Univ., China	Research	Oct. '00	Aug. '01

(g) *Conference Participation (Organizer, Keynote Speaker, etc.)*

I. Program Chair or Co-chair

- IEEE Conference on Data Engineering, June 2009.
- ACM SIGMOD Conference (General co-chair), June 2008.
- ACM SIGKDD Conference, July 2002.
- International Workshop on Data Mining and Discovery in Molecular Databases, January 1999.
- 1996 SIGMOD Workshop on Research Issues on Data Mining & Knowledge Discovery, Montreal, June 1996.

II. Keynote Presentations

- “Combinatorial Biomarker Development for End-stage Organ Failures,” Invited talk, plenary presentation, Second World DNA and Genome Day, April 2011, China.
- “A Comprehensive Combinatorial Biomarker Discovery Strategy,” Invited plenary presentation, BioIT World Expo, April 2010, Boston.
- “A Comprehensive Combinatorial Biomarker Discovery Strategy,” Invited plenary presentation, ADAPT Congress, September 2009, Washington DC.
- “Conversation Extraction and Summarization,” Invited plenary lecture in the 2009 joint International Conferences on Asia-Pacific Web Conference and Web-Age Information Management, April 3rd, 2009, China.

- “Incompleteness in Data Mining,” (with H. V. Jagadish) SIGMOD Workshop on Research Issues on Data Mining, May 2000.
- “Incompleteness in Data Mining,” CASCON Workshop on Data Mining, November 2000.
- “Discovering Outliers from Large Datasets”, 1998 Joint Statistical Meetings, August 1998, Dallas.
- “Are Spatial Data Special: from a Data Mining Perspective?”, Third International Conference on Knowledge Discovery and Data Mining, August 1997, Los Angeles.

III. Program Committee Membership

- International Conference on Very Large Databases (1997, 1998, 2001, 2004, 2007, 2009, 2011, 2014, 2015, 2016)
- ACM SIGMOD (1999, 2003, 2005, 2006, 2010, 2011, 2013)
- ACM SIGKDD (1996, 1997, 1998, 1999, 2000, 2001, 2003, 2004, 2005, 2006, 2013, 2014, 2015, 2016)
- International Conference on Data Engineering (2000, 2002, 2004, 2015, 2016)
- ACM Symposium on Principles of Database Systems (1997)
- SPIE Conference on Storage & Retrieval for Image & Video Databases (1997, 1998)
- International Workshop on Research Issues in Data Engineering: Continuous-Media Databases & Applications (1998)
- Workshop of Content-based Access of Image & Video Databases (1998)
- International Workshop on Multimedia Information Systems (1997)
- International Workshop on the Integration of Knowledge Discovery with Deductive and Object-Oriented Databases (1995)
- International Workshop on Uncertainty in Databases & Deductive Systems (1994)
- International Workshop on Interval Mathematics in Artificial Intelligence (1993)

(h) *Issued Patents*

- I. Method for using query templates in directory caches Inventors: Kapitskaia, Olga; Ng, Raymond; Srivastava, Divesh US patent # 6,904,433 Issue date: June 7, 2005

10. **SERVICE TO THE UNIVERSITY**

(a) *Memberships on committees, including offices held and dates*

I. Memberships Within Computer Science

- Chair, Colloquium Committee (1992-1996).
- Member of Ph.D. oral defense committee for Nevin Zhang (1994), Oliver Yu (1997).
- Member of Ph.D. proposal committees for Hing-Lung Lin (1994), Chris Healey (1994), Scott Flinn (1994), Alistair Veitch (1994), Jeff Beis (1995), Sanjoy Mukherjee (1997).

- Member, Executive Committee (1996-1997).
- Member, Graduate Admissions Committee (1997-1999).
- Member, Merit Evaluation Committee (1999-2000).
- Member, Undergraduate Affairs Committee (2000-2001).
- Member, Recruiting Committee (1994-1998, 1999-2001, 2014).
- Chair, CS Workload Sub-committee (2001).
- Chair, Undergraduate Affairs Committee (2001-2005).
- Vice Chair, Head Search Committee (2004-2004).
- Member, Faculty Affairs Committee (2006-2007).
- Associate Head on Faculty Affairs and Research (2008-2010).
- Chair, Faculty Affairs Committee (2008-2010).
- Member, Finance Committee (2008-2010).
- Acting co-head (2010-2011).
- Associate Head on Undergraduate Affairs (2011-2013).
- Chair of the Undergraduate Executive Committee (2011-2013).
- Member, Graduate Affairs Committee (2013-2014).
- Member, Recruiting Committee (2015-2016).

II. Memberships Outside of Computer Science

- Member, CICS R CFI Planning Committee and Co-investigator of the CICS R CFI Application (1999-2000).
- Member, Recruitment Committee of the Department of Statistics (2000-2001).
- Chair, CICS R Distinguished Lecture Series Committee (2000-2002).
- Chair, UBC ASI Exchange Coordination Committee (2000-2002).
- Member, University Campus Community Planning Committee (2003-2006).
- Member, the Dean's Advisory Committee On Promotion And Tenure (2006-2007, 2014).
- Member, ICICS Advisory Committee (2008-2010).
- Member, Faculty of Science Faculty Affairs Committee (2008-2010).
- Member, Recruiting Committee, CHiBi (2013-2014).
- Member, Faculty of Science Restricted Home Ownership Program Committee (2013-2016).
- Member, UBC Advanced Research Computing advisory group (2014-2016).

- Co-director, Master Of Data Science Program (2015-2016).
- Member, Faculty of Science Think Tank on Forming a Data Science Institute (2015-2016).
- Member, Faculty of Medicine Steering Committee on BC and Singapore health informatics program (2016).

(b) *Other service, including dates*

I. Within Computer Science

- Coop student advising - at least 1 student per term (1992-1997).
- Undergraduate student advising (2000-2001).
- Led the effort to set up a graduate mentoring program in Computer Science (2000-2001).
- Member in the Tri-mentoring Program (2002-2005).
- PhD supervisory committee for Mr Michael Lawrence (2009-2012).
- PhD supervisory committee for Ms Sarah Rastkar (2010-2013).
- PhD supervisory committee for Mr Samad Kardan (2011-2014).

II. Outside Computer Science

- Participation in the Science Mentor Program (1996).
- University Examiner for the Ph.D. oral defense of Belal Ali Moftah (Physics Dept., supervised by D. Measday) (1997).
- Ph.D. Defense Chair for Mr. Shahram Shirani (ECE) (March 2000).
- Participation in the MUG Program (September 2000).
- Helped the VP office on students to set up a tri-mentoring program for UBC students (2000-2001).
- Helped to establish a data mining lab under PIMS (2000-2001).
- Ph.D. Defense Chair for Mr. Robbie Nakatsu (Commerce) (July 2001).
- Ph.D. supervisory committee for Mr. Ofer Arazy (Commerce) (2001–2003).
- Ph.D. supervisory committee for Mr. Ryan Gandy (Interdisciplinary Studies Graduate Program) (2001–2003).
- Ph.D. supervisory committee for Ms. Anna Meredith (Pathology) (2010–2014).
- Ph.D. supervisory committee for Ms. Melanie Courtot (Medical Genetics) (2011-2014).
- Helped to develop a CFREF proposal for stage 1 and 2 of the competition (2014-2015).

11. SERVICE TO THE COMMUNITY

(a) *Memberships on scholarly societies, including offices held and dates*

(b) *Memberships on other societies, including offices held and dates*

(c) *Memberships on scholarly committees, including offices held and dates*

- Associate Director of the NSERC Strategic Network on Business Intelligence (2009-2012).
- Member of the Best Paper Award Committee of the 2009 ACM SIGKDD Conference on Data Mining (2009).
- Member of the NCE IRIS Research Management Committee (2001-2003).
- Member of the ACM SIGKDD Advisory Board (1999-2002).
- Member of the ACM Digital Review Board (1999-2000).

(d) *Memberships on other committees, including offices held and dates*

See section 9(g) for program committee participation.

(e) *Editorships (list journal and dates)*

- Member of the editorial board of the IEEE Transactions on Big Data (2014-2017).
- Member of the editorial board of the VLDB conference Journal Track (2009-2011, 2014-2015).
- Member of the editorial board of the VLDB Journal (2001–2007).
- Member of the editorial board of the IEEE Transactions on Knowledge and Data Engineering (2002–2008).
- Guest Editor, Special Issue on Health Informatics, the Journal of SIGKDD Explorations (2007).
- Member of the editorial board of the International Journal of Intelligent Data Analysis (1996-2001).
- (Co-Guest Editor) Special Issue on Uncertainty Management in Deductive Databases, Theoretical Computer Science, 1997.
- (Co-Guest Editor) Special Issue on Knowledge Discovery and Data Mining, International Journal on Cooperative and Intelligent Systems, 1997.
- (Co-Guest Editor) Special Issue on 1997 Sigmod Data Mining Workshop, Data Mining and Knowledge Discovery Journal, 1998.

(f) *Reviewer (journal, agency, etc., including dates)*

- Canadian Cancer Society, Biomarker grant program (2014, 2015, 2016)
- Prostate Cancer Canada, Movember grant program (2014, 2015, 2016)
- IEEE Transactions on Knowledge and Data Engineering (April 1995, September 1997, June 2000)
- ACM Transactions on Database Systems (April 1996, June 1997)
- Journal of the ACM (August, 1998)
- Journal of Theoretical Computer Science (June 1997)
- VLDB Journal (July 1998, August 1999, August 2001)
- The International Journal of Information Systems (June 1996)
- The International Journal of Data Mining and Knowledge Discovery (December 1999, September 2000)
- See section 9(g) for program committee participation.

(g) *External examiner (indicate university and dates)*

- Wei Lu, Ph.D. April 1994, Simon Fraser University
- Yin Jiang, Ph.D. August 1994, James Cook University, Australia
- Ranabir Gupta, Ph.D. 1995, Simon Fraser University
- Joerg Sander, Ph.D. December 1998, University of Munich, Germany
- Markus Breunig, Ph.D. January 2001, University of Munich, Germany
- Wu Lee, Ph.D. May 2002, Carneige Mellon University, USA
- Themis Palpanas, Ph.D. March 2003, University of Toronto
- Toon Calders, Ph.D. May 2003, University of Antwerp, Belgium
- Rosa Karimi Adl, Ph.D. April 2013, University of Calgary

(h) *Consultant (indicate organization and dates)*

- ACD Systems, Victoria, BC (August 2000 to April 2001)

(i) *Other service to the community***12. AWARDS AND DISTINCTIONS**

(a) *Awards for Teaching (indicate name of award, awarding organizations and date)*

- Honourable mention as an incredible instructor in the Dept. of Computer Science, February 1996.
- Honourable mention as an incredible instructor in the Dept. of Computer Science, Term 2, Winter, 2000.
- Honourable mention as an incredible instructor in the Dept. of Computer Science, Term 2, Winter, 2001.

(b) *Awards for Scholarship (indicate name of award, awarding organizations and date)*

- In 1998, the paper entitled “Algorithms for Mining Distance-based Outliers from Large Datasets” (co-authored with Ed Knorr) got the highest referee rating of all the submitted papers. This paper appeared in the Special issue on the Best Papers of VLDB '98 of the VLDB Journal, 8, 3, Feb. 2000. The VLDB conference is one of the two most prestigious database conferences worldwide. The typical number of submission is 350 papers, and the typical acceptance rate is 1 out of 7.
- In 1999, the paper entitled “One-dimensional and Multi-dimensional Substring Selectivity Estimation” (co-authored with H. V. Jagadish, O. Kapitskaia and D. Srivastava) got the highest referee rating of all the submitted papers. This paper appeared in the Special issue on the Best Papers of VLDB '99 of the VLDB Journal, 9, 3, 2001.
- In 2001, the paper entitled “Robust Space Transformation for Distance-based Operations” (co-authored with Ed Knorr and Ruben Zamar) got the **Best Paper Award** for the 2001 ACM SIGKDD conference. This is the most prestigious conference worldwide for data mining research. The acceptance ratio was 20 out of 203.
- In 2004, the paper entitled “Indexing Spatio-temporal Trajectories with Chebyshev Polynomials” (co-authored with Yuhan Cai) got the **Best Paper Award** for the 2004 ACM SIGMOD conference. This is one the two most prestigious database conferences worldwide. The acceptance ratio was 69 out of 431.

(c) *Awards for Service (indicate name of award, awarding organizations and date)*

(d) *Other Awards*

13. OTHER RELEVANT INFORMATION (Maximum One Page)

(1998-Present)

My collaboration with Jiawei Han (Simon Fraser), Laks Lakshmanan (Concordia) and Alberto Mendelzon (Toronto) on knowledge discovery in databases has resulted in the award of an IRIS-3 grant, entitled “Querying and Mining the Internet”, and the publication of several conference papers.

(1998-Present)

My collaboration with Suzanne Fortier (Queen’s), Janice Glasgow (Queen’s), Chris Gold (Laval), Jim Little (UBC), David Lowe (UBC), Michael Mcallister (Dal-Tech), Tamer Ozsü (Alberta) and Bob Woodham (UBC) on content-based analysis and abstraction of visual information management has resulted in the award of an IRIS-3 grant, entitled “Tools for visual information management in geographical and biomedical applications.” Since the year 2000, I have been the project leader of the grant. I have also been leading and participating in two grant applications for IRIS funding from 2002 to 2005.

(1999-present)

I have been collaborating closely with Ruben Zamar (UBC, Statistics). Our first joint paper, co-authored with Ed Knorr, my PhD student, has been awarded the best paper award for the 2001 ACM SIGKDD conference. I am also co-supervising two PhD students from Statistics with Zamar.

(1993-present)

My collaboration with my former students resulted in the publication of many conference and journal papers.

THE UNIVERSITY OF BRITISH COLUMBIA
Publication Record

Date: May 1, 2016

Initials:

SURNAME: Ng

FIRST NAME: Raymond

MIDDLE NAME: Tak-yan

PUBLICATION SUMMARY:

Category	1a	1b	1c	2	3	4	5
Career Total	76	101	0	1	3	0	0
Last 5 Year Total	23	16	0	0	3	0	0

1. REFEREED PUBLICATIONS(a) *Journals*

(+ indicates publications with UBC postdoctoral fellows, graduate and/or undergraduate students)
(The default ordering of authors is alphabetical. For papers where the ordering is not alphabetical, the authors are listed in decreasing order of perceived contribution.)

1. Tkacova, R., Dai, D., Ng, R. et al. "Airway Hyperresponsiveness in COPD: A Marker of Asthma-COPD Overlap Syndrome?" *Journal of Allergy and Clinical Immunology*, (in press, accepted March 2016).
2. Meredith, A., Ng, R. et al. "Circulating Biomarker Responses to Medical Management Versus Mechanical Circulatory Support in Severe Inotrope-dependent Acute Heart Failure," *European Journal of Heart Failure*, (in press, accepted December 2015).
3. Sin, D., Ng, R., et al. "Biomarker Development for Chronic Obstructive Pulmonary Disease, from Discovery to Clinical Implementation", *American Journal of Respiratory and Critical Care Medicine*, 192, 10, November 2015, pp. 1162-1170. doi: 10.1164/rccm.201505-0871PP.
4. Obeidat, M., Ng, R. et al. "The Effect of Statins on Blood Gene Expression in COPD" *PLOS ONE*, October 2015. DOI:10.1371/journal.pone.0140022.
5. Khakban, A., Ng, R. et al. "10-Year Trends in Direct Costs of COPD: A Population Based Study," *CHEST*, 148, 3, September 2015:640-6. doi: 10.1378/chest.15-0721.
6. Joty, S., Carenini, G., Ng, R. "CODRA: A Novel Discriminative Framework for Rhetorical Analysis," *Journal of Computational Linguistics*, 41, 3, September 2015, pp. 385-435.
7. Quon, B., Ng, R. et al. "Discovery of Novel Plasma Protein Biomarkers to predict imminent Cystic Fibrosis Pulmonary Exacerbations using Multiple Reaction Monitoring Mass Spectrometry," *Journal of Thorax*, 71(3), March 2016:216-22. doi: 10.1136/thoraxjnl-2014-206710.
8. Min, J., Ng, R., Shim, K. "Aggregate query processing in the presence of duplicates in wireless sensor networks," *Information Sciences*, 297, pp. 1-20, 2015.

9. Gunther, O., Ng, R., Le-Cao, K. et al. "Novel Multivariate Methods for Integration of Genomics and Proteomics Data: Applications in a Kidney Transplant Rejection Study," *OMICS*, 18(11), pp. 682695, November 2014.
10. Shannon, C., Tebbutt, S., Ng, R. et al. "Two-Stage, In Silico Deconvolution of the Lymphocyte Compartment of the Peripheral Whole Blood Transcriptome in the Context of Acute Kidney Allograft Rejection," *PLOS ONE*, April 2014;9(4):e95224. DOI: 10.1371/journal.pone.0095224.
11. Shin, H., Shannon, C., Tebbutt, S., Ng et al. "Variation in RNA-Seq transcriptome profiles of peripheral whole blood from healthy individuals with and without globin depletion," *PLOS ONE*, March 2014. DOI: 10.1371/journal.pone.009104.
12. Hollander Z, Lazarova M, McManus, B, Ng, R. et al. "Proteomic Biomarkers of Recovered Heart Function," *European Journal Heart Failure*, Feb 2014. DOI: 10.1002/ejhf.65.
13. Shin, H., Gunther, O., Ng,R. et al. "Longitudinal Analysis of Whole Blood Transcriptomes to Explore Molecular Signatures Associated With Acute Renal Allograft Rejection," *Bioinformatics and Biological Insights*, January 2014. DOI: 10.4137/BBI.S13376.
14. (+) Lin, D., Cohen-Freue, G., Ng, R. et al. "Plasma protein biosignatures for detection of cardiac allograft vasculopathy," *Journal of Heart and Lung Transplant*, July 2013;32(7):723-733. DOI: 10.1016/j.healun.2013.04.011.
15. Scherer, A., Gunther, O., Keown, P., Ng, R. et al. "Alteration of human blood cell transcriptome in uremia," *BMC Medical Genomics*, July 2013, 6:23. DOI: 10.1186/1755-8794-6-23.
16. (+) Joty, S., Carenini, G., Ng, R., "Topic Segmentation and Labeling in Asynchronous Conversations," *Journal of Artificial Intelligence Research*, 47, pp. 521-573, July 2013.
17. (+) Freue-Cohen, G., Meredith, A., Ng, R. et al. "Computational biomarker pipeline from discovery to clinical implementation: plasma proteomic biomarkers for cardiac transplantation," *PLOS Computational Biology*, April 2013. DOI: 10.1371/journal.pcbi.1002963.
18. (+) Hollander, Z., Chen, V., Ng, R. et al. "Predicting Acute Cardiac Rejection from Donor Heart and Pre-Transplant Recipient Blood Gene Expression," *Journal of Heart and Lung Transplantation*, 32, 2, pp. 259-265, February 2013.
19. Gunther, O., Chen, V., Ng, R. et al. "A computational pipeline for the development of multi-marker bio-signature panels and ensemble classifiers," *BMC Bioinformatics*, 2012, 13:326 doi:10.1186/1471-2105-13-326.
20. Shannon, C., Ng, R. et al. White Blood Cell Differentials Enrich Whole Blood Expression Data in the Context of Acute Cardiac Allograft Rejection, *Bioinformatics and Biology Insights*, 6, pp. 49-61, 2012.
21. (+) Lin, D., Hollander, Z, Ng, R. et al. "Molecular signatures of end-stage heart failure," *Journal of Cardiac Failure*, 17, 10, pp. 867-874, October 2011.
22. Gibb, E, Enfield, Ng, R. et al. Long non-coding RNAs are expressed in oral mucosa and altered in oral premalignant lesions, *Oral Oncology*, 47, 11, pp. 55-61, August 2011.
23. (+) Gunther, O., Ng, R. et al. "Effects of sample timing and treatment on gene expression in early acute renal allograft rejection," *Journal of Transplantation*, 91, 3, pp. 323-329, 2011.

24. (+) Freue-Cohen, G., Ng, R. et al. "Proteomic signatures in plasma during early acute renal allograft rejection", *Molecular and Cellular Proteomics*, 1, 9, pp. 1954–1967, September 2010.
25. Garmaroudi, F., McManus, B., Ng, R. et al. "Pairwise network mechanisms in the host signaling response to coxsackievirus B3 infection," *Proceedings of the National Academy of Science*, 107, 39, pp. 17053-17058, September 2010.
26. (+) Chari, R., Ng, R. et al. "A sequence-based approach to identify reference genes for gene expression analysis", *BMC Medical Genomics*, 3, 32, pp. 1-11, 2010.
27. Dong, X., Ng, R. et al. "Patient-Derived First Generation Xenografts of NonSmall Cell Lung Cancers: Promising Tools for Predicting Drug Responses for Personalized Chemotherapy," *Clinical Cancer Research*, 16, 5, pp. 1442-1451, February 2010.
28. Lonergan, K., Ng, R. et al. "Transcriptome Profiles of Carcinoma-in-Situ and Invasive Non-Small Cell Lung Cancer as Revealed by SAGE," *PLoS ONE*, 5, 2, e9162. doi:10.1371/journal.pone.0009162, February 2010.
29. (+) Gunther, O., Ng, R. et al. "Functional Genomic Analysis of Peripheral Blood During Early Acute Renal Allograft Rejection," *Journal of Transplantation*, 88, 7, pp. 942-951, doi: 10.1097/TP.0b013e3181b7ccc6, October 2009.
30. (+) Lee, H., Ng, R. and Shim, K. "Power-Law Based Estimation of Set Similarity Join Size," *Proceedings of the VLDB Endowment*, 2, 1, pp. 658-669, August 2009.
31. Garnis, C., Ng, R. et al. "Genomic imbalances in precancerous tissues signal oral cancer risk", *Molecular Cancer*, 8, 50, doi:10.1186/1476-4598-8-50, August 2009.
32. Zhang, Q., Hughes-Oliver, J. and Ng, R. "A Model-Based Ensembling Approach for Developing QSARs," *Journal of Chemical Information and Modeling*, 49, 8, pp. 1857–1865, 2009.
33. (+) Shah, S., Cheung, K.P., Ng, R., Murphy, K. et al. "Model based clustering of array CGH data," *Bioinformatics Journal*, 25, 12, i30-i38, 2009.
34. (+) Lin, D., Hollander, Z., Ng, R. et al. "Whole Blood Genomic Biomarkers of Acute Cardiac Allograft Rejection," *Journal of Heart and Lung Transplantation*, 28, 9, pp. 927–935, 2009.
35. (+) Tsui, I., Rosin, M., Zhang, L., Ng, R. and Lam, W. "Multiple aberrations of chromosome 3p detected in oral premalignant lesions," *Cancer Prevention Research*, 1, 424, doi: 10.1158/1940-6207, November 2008.
36. (+) Lakshmanan, L., Ng, R. and Ramesh, G. "On Disclosure Risk Analysis of Anonymized Itemsets in the Presence of Prior Knowledge", *ACM Transactions on Knowledge Discovery from Data*, 2, 3, October 2008.
37. (+) Cheung, K.J., Shah, S., Ng, R. et al. "Genome-wide profiling of follicular lymphoma by array comparative genomic hybridization reveals prognostically significant DNA copy number imbalances," *Blood Journal*, doi 10.1182/blood-2008-02-140616, August 2008.
38. (+) Chari, R., Coe, B., Ng, R. et al. "SIGMA2: A system for the integrative genomic multi-dimensional analysis of cancer genomes, epigenomes, and transcriptomes," *BMC Bioinformatics*, 9, 422, doi: 10.1186/1471-2105-9-422, October 2008.

39. (+) Lee, E., Chari, R., Lam, A., Ng, R., Yee, J., English, J., Evans, K., MacAulay, C., Lam, S. and Lam, L. Disruption of the non-canonical Wnt pathway in lung squamous cell carcinoma. *Clinical Medicine: Oncology*, 2, pp. 169–179, 2008.
40. (+) Chari, R., Lonergan, K., Ng, R., MacAulay, C., Lam, W. and Lam, S., “Effect of Active Smoking on the Human Bronchial Epithelium Transcriptome,” *BMC Genomics*, 8:297, pp. 1-13, August 2007.
41. (+) Shah, S., Lam, W., Ng, R. and Murphy, K., “Modelling Recurrent DNA Copy Number Alterations in array CGH Data,” *Bioinformatics Journal*, 23, 13, pp. 450–458, August 2007.
42. (+) Cohen-Freue, G., Hollander, Z., Ng, R. et al. “MDQC: a new quality assessment method for microarrays based on quality control reports,” *Bioinformatics Journal*, 23, 23, pp. 3162–3139, 2007.
43. Kim, C., Shim, K. and Ng, R., “SQUIRE: Sequential Pattern Mining with Quantities,” *International Journal of Systems and Software*, 80, 10, pp. 1726–1745, 2007.
44. Wong, K., Ng, R., Lam, W. et al. “A comprehensive analysis of common copy number variations in the human genome,” *American Journal of Human Genetics*, vol. 80, pp. 1-14, January 2007.
45. Malossini, A., Blanzieri, E. and Ng, R., “Detecting potential labeling errors in microarrays by data perturbation,” *Bioinformatics Journal*, 22, 17, pp. 2214-2121, July 2006.
46. Calders, T., Lakshmanan, L., Ng, R., “Expressive Power of an Algebra for Data Mining,” *ACM Transactions on Database Systems*, 31, pp. 1169–1214, Dec 2006.
47. Lonergan, K., Lam, W., Ng, R. et al., “Identification of novel lung genes in bronchial epithelium by serial analysis of gene expression,” *American Journal of Respiratory Cell and Molecular Biology*, 35, 6, pp. 651–661, Dec 2006.
48. (+) Chilson, J., Ng, R., Wagner, A. and Zamar, R., “Parallel Computation of High Dimensional Robust Correlation and Covariance Matrices,” *Journal of Algorithmica*, Special Issue on Coarse Grained Parallel Algorithms For Scientific Applications, 45, 3, pp. 403–431, July 2006.
49. (+) Shah, S., Xuan, X., DeLeeuw, R., Khojasteh, M., Lam, W., Ng, R. and Murphy, K., “Integrating copy number polymorphisms into array CGH analysis using a robust HMM,” *Bioinformatics Journal*, 22, 14, pp. 431–439, May 2006.
50. (+) Sander, J., Ng, R., Sleumer, M., Yuen, M., and Jones, S., “A Methodology for Analyzing SAGE Libraries For Cancer Profiling”, Special Issue on Genomic Information Retrieval, *ACM Transactions on Information Systems*, 23, 1, pp. 35–60, 2005.
51. (+) Ying, A., Murphy, G., Ng, R., and Chu-Carroll, M., “Predicting Source Code Changes by Mining Change History”, *IEEE Transactions on Software Engineering*, 30, 9, pp. 574–586, September, 2004.
52. (+) Lakshmanan, L., Leung, C. and Ng, R., “Efficient Dynamic Mining of Constrained Frequent Sets,” *ACM Transactions on Database Systems*, 28, 4, pp. 337-389, 2003.
53. (+) Kwon, A., Hoos, H., and Ng, R. “Inference of Transcriptional Regulation Relationships from Gene Expression Data,” *the Bioinformatics Journal*, 19, 8, pp. 905-912, May 2003.
54. Calders, T., Ng, R. and Wijssen, J., “Searching for Dependencies at Multiple Abstraction Levels”, *ACM Transactions on Database Systems*, 27, 3, pp. 229-260, 2002.

55. Ng, R. and Han, J., "CLARANS: A Method for Clustering Objects for Spatial Data Mining", *IEEE Trans. Knowledge & Data Engineering*, 14, 5, pp. 1003-1016, 2002.
56. Jagadish, H.V., Kapitskaia, O., Ng, R., and Srivastava, D., "One-Dimensional and Multi-Dimensional Substring Selectivity Estimation", Special issue on the Best Papers of VLDB '99, *VLDB Journal*, 9, 3, pp. 214-230, 2001.
57. (+) Knorr, E., Ng, R., and Tucakov, V., "Distance-based Outliers: Algorithms and Applications", Special issue on the Best Papers of VLDB '98, *VLDB Journal*, 8, 3, pp. 237-253, Feb. 2000.
58. (+) Ng, R. and Tam, D., "An Analysis of Multi-level Filtering for Image Queries", *IEEE Trans. Knowledge & Data Engineering*, 11, 6, Dec. 1999.
59. Han, J., Lakshmanan, L. and Ng, R., "Human-centred Multi-dimensional Data Analysis and Mining - The Constraints Way", *IEEE Computer*, Special Issue on Data Mining, 32, 8, pp. 46-50, 1999.
60. (+) Ng, R. and Shum, P., "Optimal Clip Ordering for News-on-Demand Systems", *VLDB Journal*, 7, 4, pp. 239-252, Dec. 1998.
61. Ng, R., "Semantics, Consistency & Query Processing of Empirical Deductive Databases", *IEEE Trans. on Knowledge & Data Engineering*, 9, 1, pp. 32-49, Feb. 1997.
62. Ng, R., "Reasoning with Uncertainty in Deductive Databases and Logic Programs", *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems*, 2, 3, pp. 261-316, 1997.
63. (+) Ng, R. and Faulus, D., "An Expressive Language & Interface for Image Querying", Special Issue on Image & Video Databases, *Journal of Machine Vision and Applications*, 10, pp. 74-85, 1997.
64. (+) Ng, R. and Tian, X., "Incremental Algorithms for Optimizing Model Computation Based on Partial Instantiation", *Journal of Logic Programming*, 30, 3, pp. 179-204, March 1997.
65. Ioannidis, Y., Ng, R., Sellis, T. and Shim, K., "Parametric Query Optimization", *VLDB Journal*, 6, 2, pp. 132-151, May 1997.
66. (+) Ng, R. and Yang, J., "An Analysis of Buffer Sharing and Prefetching Techniques for Multimedia Systems", *ACM-Springer Journal of Multimedia Systems*, 4, 2, pp. 55-69, 1996.
67. (+) Knorr, E. and Ng, R., "Finding Aggregate Proximity Relationships and Commonalities in Spatial Data Mining", *IEEE Trans. on Knowledge & Data Engineering*, 8, 6, pp. 884-497, Dec. 1996.
68. Bell, C., Nerode, A., Ng, R. and Subrahmanian, V.S., "Implementing Deductive Databases by Mixed Integer Programming", *ACM Transactions on Database Systems*, 21, 2, pp. 238-269, 1996.
69. (+) Makaroff, D. and Ng, R., "Schemes for Implementing Buffer Sharing in Continuous-Media Systems", *Information Systems*, 20, 6, pp. 445-464, 1995.
70. Faloutsos, C., Ng, R. and Sellis, T., "Adaptive and Flexible Buffer Allocation", *IEEE Transactions on Computers*, 44, 4, pp. 546-560, 1995.
71. Nerode, A., Ng, R. and Subrahmanian, V.S., "Computing Circumscriptive Databases, Part I: Theory and Algorithms", *Information and Computation*, 116, 1, pp. 58-90, 1995.
72. Ng, R. and Subrahmanian, V.S., "Stable Semantics for Probabilistic Deductive Databases", *Information and Computation*, 110, 1, pp. 42-83, 1994.

73. Bell, C., Nerode, A., Ng, R. and Subrahmanian, V.S., “Mixed Integer Programming Methods for Computing Nonmonotonic Deductive Databases”, *Journal of ACM*, 41, 6, pp. 1178-1215, 1994.
74. Ng, R. and Subrahmanian, V.S., “A Semantical Framework for Supporting Subjective and Conditional Probabilities in Deductive Databases”, *Journal of Automated Reasoning* 10, 2, pp. 191-235, 1993.
75. Ng, R. and Subrahmanian, V.S., “Probabilistic Logic Programming”, *Information and Computation*, 101, 2, pp. 150-201, 1992.
76. Sellis, T., Roussopoulos, N. and Ng, R., “Efficient Compilation of Large Rule Bases using Logical Access Paths”, *Information Systems*, 15, 1, pp. 73-84, 1990.

(b) *Conference Proceedings*

(Acceptance rates are provided, whenever available.)

1. (+) Allen, K., Carenini, G., Ng, R. “Detecting Disagreement in Conversations using Pseudo-Monologic Rhetorical Structure,” *Conference on Empirical Methods in NLP and the Conference on Natural Language Learning*, October, 2014.
2. (+) Gerani, S., Mehdad, Y., Carenini, G., Ng, R. “Abstractive Summarization of Product Reviews Using Discourse Structure,” *Conference on Empirical Methods in NLP and the Conference on Natural Language Learning*, October, 2014.
3. (+) Tatsuuro, O., Mehdad, Y., Carenini G., Ng, R., “A Template-based Abstractive Meeting Summarization: Leveraging Summary and Source Text Relationships,” *International Conference on Natural Language Generation*, August 2014.
4. (+) Mehdad, Y., Carenini, G. and Ng, R. “Abstractive Summarization of Spoken and Written Conversations Based on Phrasal Queries,” *Annual meeting of the Association for Computational Linguistics*, August 2014.
5. Dang, X., Assent, I., Ng, R., Zimek, A. and Schubert, E. “ Discriminative Features for Identifying and Interpreting Outliers,” *The 30th IEEE International Conference on Data Engineering (ICDE 2014)*, Chicago USA, April 2014.
6. Micenkova, B., Ng, R., Dang, X. and Assent, I. “Explaining Outliers for Validation and Analysis,” *The IEEE International Conference on Data Mining (ICDM 2013)*, Dallas USA, December 2013.
7. Dang, X., Micenkova, B., Assent, I. and Ng, R. “Local Outlier Detection with Interpretation,” *European Conference on Machine Learning (ECML PKDD 2013)*, Prague, Czech Republic, September 2013.
8. (+) Tavafi, M., Mehdad, Y., Joty, S., Carenini, G. and Ng, R. “Dialogue Act Recognition in Synchronous and Asynchronous Conversations,” *The 14th Annual Meeting on Discourse and Dialogue (SIGDIAL)*, Metz, France, August 2013.
9. (+) Joty, S., Carenini, G., Ng, R., Mehdad, Y., “Combining Intra- and Multi-sentential Rhetorical Parsing for Document-level Discourse Analysis,” *Annual meeting of the Association for Computational Linguistics*, August 2013.
10. (+) Mehdad, Y., Carenini, G., Ng, R., Joty, S., “Towards Topic Labeling with Phrase Entailment and Aggregation,” *14th Annual Conference of the North American Chapter of the Association for Computational Linguistics*, June 2013.

11. Dang, X., Micenkova, B., Assent, I., Ng, R., “Outlier Detection with Space Transformation and Spectral Analysis,” *SIAM International Conference on Data Mining*, May 2013.
12. (+) Joty, S., Carenini, G., Ng, R. “A Novel Discriminative Framework for Sentence-Level Discourse Analysis,” *Conference on Empirical Methods in NLP and the Conference on Natural Language Learning*, June 2012.
13. (+) Murray G., Carenini G., Ng, R., “Detecting abstractive sentence communities: a new summarization task,” *Fifth International Conference on Natural Language Processing*, November 2011.
14. (+) Murray G., Carenini G., Ng, R., “Generating Abstracts of Meeting Conversations: A User Study,” *International Conference on Natural Language Generation*, August 2010.
15. (+) Sandu, O, Carenini, G., Ng, R., “Domain Adaptation to Summarize Human Conversations”, *ACL 2010 Workshop on Domain Adaptation for Natural Language Processing*, July 2010.
16. (+) Murray G. , Carenini G., Ng, R., “Interpretation and Transformation for Abstracting Conversations,” *11th Annual Conference of the North American Chapter of the Association for Computational Linguistics*, June 2010.
17. (+) Joty S., Carenini G., Murray G., Ng, R., “Finding Topics in Emails: Is LDA Enough?” *NIPS '09 workshop on Applications for Topic Models: Text and Beyond*, December 2009.
18. (+) Cheung J., Carenini G, Ng, R., “Optimization-based Content Selection for Opinion Summarization,” *ACL-IJCNLP '09 Workshop on Language Generation and Summarization*, August 2009.
19. (+) Lee, H., Ng, R. and Shim, K. “ Approximate Substring Selectivity Estimation,” *Proc. of the 2009 International Conference on Extending Database Technology*, pp. 827–838, March 2009.
20. (+) Ulrich, J., Carenini, G., Murray, G., Ng. R. “Regression-Based Summarization of Email Conversations,” *Pro. of the Third International Conference on Weblogs and Social Media*, May 2009.
21. (+) Carenini, G., Ng, R. and Zhou, X. “Summarizing Emails with Conversational Cohesion and Subjectivity,” *Proc. of 2008 American Society of Computational Linguistics*, June 2008.
22. (+) Lee, H., Ng, R. and Shim, K., “Extending Q-grams to Estimate Selectivity of Strings with Low Edit Distance”, *Proc. of 33rd International Conference on Very Large Databases*, pp. 195–206, September 2007.
23. (+) Carenini, G., Ng, R. and Zhou, X., “Summarization of Email Conversations,” *Proc. World Wide Web Conference*, pp. 91-100, May 2007 (acceptance ratio about 10%).
24. Gokhale, C., Gupta, N., Kumar, P., Lakshmanan, L., Ng, R. and Prakash, A., “Complex Group-by Queries for XML,” *Proc. 2007 International Conference on Data Engineering* , pp. 646–655, April 2007 (acceptance ratio about 15%).
25. (+) Bu, S., Lakshamanan, L., Ng, R. and Ramesh, G., “Preservation of Patterns and Input-Output Privacy,” *Proc. 2007 International Conference on Data Engineering* , pp. 696–705, April 2007 (acceptance ratio about 15%).

26. (+) Carenini, G., Ng, R. and Pauls, A., “Multi-document summarization of evaluative text,” *Proc. 11th Conference of the European Chapter of the Association for Computational Linguistics*, April 2006.
27. (+) Chari, R., Lonergan, K., deLeeuw, R., Shadeo, A., Chi, B., Tsao, M., Jones, S., Marra, M., Ling, V., Ng, R., MacAulay, C., Lam, S. and Lam, W., “Identification of novel lung-specific genes in bronchial epithelium using SAGE,” *Proc. American Association of Cancer Research Annual Conference*, Paper #852, April 2006.
28. (+) Carenini, G., Ng, R. and Pauls, A., “Interactive Multimedia Summaries of evaluative text,” *Proc. 11th International Conference on Intelligent User Interfaces*, pp. 124–131, January 2006.
29. (+) Carenini, G., Ng, R. and Zwart, E., “Extracting Knowledge from Evaluative Text,” *Third International Conference on Knowledge Capture*, pp. 11-18, October 2005.
30. (+) Bu, S., Lakshmanan, L. and Ng, R., “MDL Summarization with Holes,” *Proceedings of 31st International Conference on Very Large Databases*, pp. 433–444, August 2005.
31. (+) Carenini, G., Ng, R. and Zhou, X., “Scalable Discovery of Hidden Emails from Large Folders,” *Proceedings of ACM SIGKDD Conference*, pp. 544–549, August 2005.
32. (+) Lakshmanan, L., Ng, R., and Ramesh, G., “To Do or Not To Do: The Dilemma of Disclosing Anonymized Data,” *Proceedings of ACM SIGMOD Conference*, pp. 61–72, 2005.
33. (+) Chan, T., Ng, R., McCaulay, C., and Lam, W., “Finding Biomarkers Specific for Early and Late Stages of Lung Cancer Using SAGE Data”, *11th World Conference on Lung Cancer. Lung Cancer*, Supplement 2, 49:S120, P-032, July 2005.
34. (+) Chari, R., Ng, R., McCaulay, C., Lam, S. and Lam, W., “Effect of Inhaled Budesonide in Smokers with Bronchial Dysplasia,” *11th World Conference on Lung Cancer. Lung Cancer*, Supplement 2, 49:S120, P-017, July 2005.
35. (+) Lam, S., Garnis, C., Lonergan, K., Chan, T., Chari, R., Ng, R., MacAulay, C. and Lam, W., “Genetic alternations in pre-invasive lung cancer,” *11th World Conference on Lung Cancer. Lung Cancer*, Supplement 2, 49:S63, E-40, July 2005.
36. (+) Carenini, G., Ng, R., Zhou, X. and Zwart, E., “Discovery and Regeneration of Hidden Emails”, *Proceedings of the 2005 ACM Symposium on Applied Computing*, pp. 503–510, March 2005.
37. (+) Chen, L. and Ng, R., “On The Marriage of Lp-norms and Edit Distance,” *Proceedings of 30th International Conference on Very Large Databases*, pp 792–803, August 2004 (Acceptance rate: 81/504).
38. (+) Chilson, J., Ng, R., Wagner, A. and Zamar, R., “Parallel computation of high dimensional robust correlation and covariance matrices”, *Proceedings of the ACM SIGKDD*, pp 533-538, August 2004.
39. (+) Ng, R. and Cai, Y. “Indexing Spatio-temporal Trajectories using Chebyshev Polynomials”, (**Best Paper Award**), *Proceedings of the ACM SIGMOD*, pp. 599–610, June 2004. (Acceptance rate: 69/431).
40. Jagadish, H., Ng, R., Ooi, B. and Tung, K., “ItCompress: an Iterative Semantic Compression Algorithm,” *the 20th International Conference on Data Engineering*, pp. 646–657, March 2004.

41. Kim, C., Lim, J., Ng, R. and Shim, K., "SQUIRE: Sequential Pattern Mining with Quantities," *the 20th International Conference on Data Engineering*, pp. 827, March 2004..
42. (+) Lonergan, K., MacAulay, C., Ng, R., Chan, T., Chari, R., Lam, S. and Lam, W., "Comparing gene expression profiles during development of squamous cell lung carcinoma," *Proc. 95th American Association for Cancer Research Annual Conference*, March 2004.
43. (+) Chan, T., Ng, R., McCaulay, C., and Lam, W. "Using the Permutation Test to Analyze Lung Cancer SAGE Libraries", *Proceedings of the 10th World Conference on Lung Cancer*, Supplement 2, 41:S120, P-035, August 2003.
44. (+) Chan, T., MacAulay, C, Lam, W., Lam, S., Lonergan, K., Chari, R., Jones, S., Marra, M. and Ng, R., "Evaluating Housekeeping Genes Used In Lung Microarray Analysis With Lung SAGE Data", *Proceedings of the 10th World Conference on Lung Cancer. Lung Cancer*, Supplement 2, 41:S120, P-117, August 2003.
45. (+) Kwon, A., Hoos, H., and Ng, R. "Inference of Transcriptional Regulation Relationships from Gene Expression Data," *Proceedings of the 2003 ACM Symposium on Applied Computing*, p. 135-140, 2003.
46. (+) Ng, R., Chan, T., and Kuo, B., "Analyzing Brain and Breast SAGE Libraries", *Proceedings of the 10th International Conference on Intelligent Systems for Molecular Biology*, August 2002.
47. (+) Ng, R., Sleumer, M., Yuen, M., Sander, J. and Jones, S., "Hierarchical Cluster Analysis and Classification of SAGE Data", *Proceedings of the 10th International Conference on Intelligent Systems for Molecular Biology*, August 2002.
48. (+) Lakshmanan, L., Ng, R., Wang, C., Zhou, X. and Johnson, J. "The Generalized MDL Approach for Summarization", *Proceedings of 28th International Conference on Very Large Databases*, pp 766-777, August 2002 (Acceptance rate: 59/431).
49. (+) Leung, K., Ng, R. and Mannila, H. "OSSM: A Segmentation Approach to Optimize Frequency Counting", *Proceedings of International Conference on data engineering*, pp 583-592, March 2002 (Acceptance rate: 54/287).
50. (+) Knorr, E., Ng, R. and Zamar, R. "Robust space transformation for distance-based operations", (**Best Paper Award**), *Proceedings of the ACM SIGKDD*, pp 126-135, August 2001. (Acceptance rate: 20/203).
51. (+) Ng, R., Sander, J. and Sleumer, M. "Hierarchical cluster analysis of SAGE data for cancer profiling", *Proceedings of BIODDD Workshop on Data Mining in Bioinformatics*, pp 65-72, August 2001.
52. Lin, S., Ozsu, T., Oria, V. and Ng, R. "An Extendible Hash for Multi-precision Similarity Querying of Image Databases", *Proceedings of 27th International Conference on Very Large Databases*, pp 221-230, September, 2001. (Acceptance rate: 59/339).
53. (+) Ng, R., Wagner, A. and Y. Yu., "Iceberg-cube computation with PC clusters", *Proceedings of ACM SIGMOD*, pp. 25-36, May 2001 (Acceptance rate: 44/293).
54. Chen, Z., Jagadish, H.V., Korn, F., Koudas, N., Muthukrishnan, M., Ng, R., and Srivastava, D., "Counting Twig Matches in a Tree", *Proceedings of International Conference on data engineering*, pp. 595-604, April, 2001 (Acceptance rate: approximately 1 out of 5).

55. Tung, A., Ng, R., Han, J., and Lakshmanan, L., "Constraint-based clustering in large databases," *Proceedings 2001 International Conference on Database Theory*, pp. 405–419, January, 2001 (Acceptance rate: approximately 1 out of 4).
56. Johnson, T., Lakshmanan, L., and Ng, R., "The 3W Model and Algebra for Unified Data Mining," *Proceedings of 26th International Conference on Very Large Databases*, pp. 21–32, September, 2000 (Acceptance rate: approximately 1 out of 6).
57. (+) Breunig, M., Kriegel, H., Ng, R. and Sander J., "LOF: Identifying Density-based Local Outliers," *Proceedings of ACM SIGMOD*, pp. 93–104, May 2000 (Acceptance rate: 42/248).
58. Jagadish, H.V., and Ng, R., "Incompleteness in Data Mining," *Proceedings SIGMOD Workshop on Research Issues on Data Mining*, pp. 1-10, May 2000.
59. Kapitskaia, O., Ng, R. and Srivastava, D., "Evolution and Revolutions in LDAP Directory Caches," *Proceedings of 2000 Conference on Extending Database Technologies*, pp. 202–216, March 2000.
60. (+) Breunig, M., Kriegel, H., Ng, R., and Sander, J., "OPTICS-OF: Identifying Local Outliers", *Proceedings of the 3rd European Conference on Principles and Practice of Knowledge Discovery in Databases*, pp. 262–270, October 1999.
61. (+) Jagadish, H., Madar, J. and Ng, R., "Semantic Compression and Pattern Extraction with Fascicles", *Proceedings of 25th International Conference on Very Large Databases*, pp. 186-197, September 1999. (Acceptance rate: 51/387).
62. (+) Knorr, E. and Ng, R., "Finding Intentional Knowledge of Distance-based Outliers," *Proceedings of 25th International Conference on Very Large Databases*, pp. 211-222, September 1999. (Acceptance rate: 51/387).
63. Jagadish, H., Kapitskaia, O., Ng, R., and Srivastava, D., "Multi-dimensional Substring Selectivity Estimation," *Proceedings of 25th International Conference on Very Large Databases*, pp. 387-398, September 1999. (Acceptance rate: 51/387).
64. Wijzen, J. and Ng, R., "Temporal Dependencies Generalized for Spatial and Other Dimensions", *Proceedings of 1999 VLDB Workshop on Spatio-Temporal Databases Management (STDBM 99)*, September 1999.
65. Wijzen, J., Ng, R., and T. Calders, "Discovering Roll-Up Dependencies," *Proceedings of the 5th International Conference on Knowledge Discovery & Data Mining*, pp. 213–222, August 1999.
66. (+) Lakshmanan, L., Ng, R., Han, J. and Pang, A., "Optimization of Constrained Frequent Set Queries with 2-variable Constraints", *Proceedings of ACM SIGMOD*, pp. 157–168, June 1999. (Acceptance rate: 42/205).
67. Jagadish, H., Ng, R. and Srivastava, D., "Substring Selectivity Estimation", *Proceedings of 18th ACM PODS*, pp. 249–260, June 1999. (Acceptance rate: 35/116).
68. Johnson, T., Lakshmanan, L. and Ng, R., "Toward a Toolkit for Data Analysis and Mining", *Proceedings SIGMOD Workshop on Research Issues on Data Mining*, June 1999.
69. (+) Knorr, E. and Ng, R., "Algorithms for Mining Distance-based Outliers from Large Datasets," *Proceedings of 24th International Conference on Very Large Data Bases*, pp. 392–403, August 1998. Full version published under the title "Distance-based Outliers: Algorithms and Applications". (Acceptance rate: 51/328).

70. (+) Johnson, T., Kwok, I. and Ng, R., "Fast Computation of 2-Dimensional Depth Contours," *Proceedings of 4nd International Conference on Knowledge Discovery & Data Mining*, pp. 224–228, August 1998.
71. (*) (+) Ng, R., Lakshmanan, L., Han, J. and Pang A., "Exploratory Mining and Pruning Optimizations of Constrained Association Rules," *Proceedings of ACM SIGMOD*, pp. 13–24, June 1998. (Acceptance rate: 42/246).
72. (+) Tucakov, V. and Ng, R., "Identifying Unusual People Behaviour: a Case Study of Mining Outliers in Spatio-Temporal Trajectory Databases," *SIGMOD 98 Workshop on Research Issues on Data Mining and Knowledge Discovery*, June, 1998.
73. (+) Leung, K., Ng R., "Multiscale Similarity Matching for Subimage Queries of Arbitrary Size," *International Conference on Visual Database Systems*, May 1998.
74. (+) Leung, K., Ng R., "Multi-resolution Sub-image Similarity Matching for Large Image Databases", *Proceedings of SPIE Conference on Storage and Retrieval for Image & Video Databases*, February 1998.
75. (+) Knorr, E., Ng R., "A Unified Approach for Mining Outliers", *Proceedings of CASCON '97*, Toronto, November 1997.
76. (+) Knorr, E., Ng R., "A Unified Notion of Outliers: Properties & Computation", *Proceedings of 1997 International Conference on Knowledge Discovery & Data Mining*, Los Angeles, August 1997.
77. (+) Knorr, E., Ng R., "Finding Bounding Shape Matching Relationships in Spatial Data", *Proceedings of 1997 International Conference on Spatial Data Handling*, Germany, July 1997.
78. Ng, R., "Discovering Strong, Common & Discriminating Characteristics of Clusters from Thematic Maps", *Proceedings of Eleventh Annual Symposium on Geographic Information Systems*, pp. 392-394, February 1997.
79. (+) Ng, R. and Tam, D., "An Analysis of Multi-level Color Histograms", *Proceedings of SPIE Conference on Storage and Retrieval for Image & Video Databases*, pp. 22-33, February 1997. Full journal version published under the title "An Analysis of Multi-level Filtering for Image Queries".
80. (+) Knorr, E., Ng, R. and Shilvock, D., "Performing Boundary Shape Matching in Spatial Data", *Proceedings of CASCON '96*, pp. 115-122, November 1996.
81. (+) Ng, R. and Shum, P., "Optimal Chip Ordering for News On-Demand Queries", *Proceedings of 2nd International Workshop on Multimedia Information Systems*, pp. 1-5, September 1996. Full journal version published under the title "Optimal Clip Ordering for News On-Demand Systems".
82. (+) Knorr, E. and Ng, R., "Extraction of Spatial Proximity Patterns by Concept Generalization", *Proceedings of 2nd International Conference on Knowledge Discovery & Data Mining*, pp. 347-351, August 1996.
83. Ng, R., "Spatial Data Mining: Discovery Knowledge of Clusters from Maps", *Proceedings of SIGMOD Workshop on Research Issues on Data Mining & Knowledge Discovery*, pp. 79-84, Montreal, June 1996.
84. (+) Ng, R. and Sedighian, A., "Multi-dimensional Indexing Structures for Images Transformed by Principal Component Analysis", *Proceedings of SPIE Conference on Storage & Retrieval for Image & Video Databases V*, pp. 50-61, February 1996.

85. (+) Ng, R. and Faulus, D., "EXQUISI: an Expressive Query Interface for Similar Images", *Proceedings of SPIE Conference on Storage & Retrieval for Image & Video Databases V*, pp. 215-227, February 1996. Full journal version published under the title "An Expressive Language & Interface for Image Querying".
86. (+) Ng, R. and Tian, X., "Incremental Methods for Optimizing Partial Instantiation", *Proceedings of 3rd International Workshop on Non-monotonic Reasoning and Logic Programming*, pp. 287-301, 1995. Full journal version published under the title of "Incremental Algorithms for Optimizing Model Computation Based on Partial Instantiation".
87. (+) Ng, R. and Yang, J., "Maximizing Buffer and Disk Utilizations for News On-Demand", *Proceedings of 20th International Conference on Very Large Data Bases*, pp. 451-462, September 1994. Full journal version published under the title of "An Analysis of Buffer Sharing and Prefetching Techniques for Multimedia Systems". (Acceptance rate: 54/293).
88. Ng, R. and Han, J., "Efficient and Effective Clustering Methods for Spatial Data Mining", *Proceedings of 20th International Conference on Very Large Data Bases*, pp. 144-155, September 1994. (Acceptance rate: 54/293).
89. Fu, Y., Han, J. and Ng, R., "Cooperative Query Answering Using Multiple Layered Databases", *Proceedings of 2nd International Conference on Cooperative Information Systems*, pp. 47-58, May 1994.
90. Ng, R., "Semantics and Consistency of Empirical Databases", *Proceedings of 10th International Conference of Logic Programming*, pp. 812-826, June 1993. Full journal version published under the title of "Semantics, Consistency & Query Processing of Empirical Deductive Databases".
91. Bell, C. Ng, R., Nerode, A. and Subrahmanian, V.S., "Implementing Stable Semantics by Linear Programming", *Proceedings of 2nd International Workshop on Non-monotonic Reasoning and Logic Programming*, pp. 23-42, May 1993. Full journal version published under the title of "Mixed Integer Programming Methods for Computing Nonmonotonic Deductive Databases".
92. Ioannidis, Y., Ng, R., Sellis, T. and Shim, K., "Parametric Query Optimization", *Proceedings of the 18th International Conference on Very Large Data Bases*, pp. 103-114, August 1992. Full journal version published under the title of "Parametric Query Optimization". (Acceptance rate: 49/373).
93. Ng, R. and Subrahmanian, V.S., "Empirical Probabilities in Monadic Deductive Databases", *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, pp. 215-222, July 1992.
94. Bell, C., Ng, R., Nerode, A. and Subrahmanian, V.S., "Implementing Deductive Databases by Linear Programming", *Proceedings of the 11th Symposium on Principles of Database Systems*, pp. 283-291, June 1992. Full journal version published under the title of "Implementing Deductive Databases by Mixed Integer Programming". (Acceptance rate: 32/105).
95. Faloutsos, C., Ng, R. and Sellis, T., "Predictive Load Control for Flexible Buffer Allocation", *Proceedings of the 17th International Conference on Very Large Data Bases*, pp. 265-274, September 1991. Full journal version published under the title of "Adaptive and Flexible Buffer Allocation". (Acceptance rate: 50/318)
96. Ng, R. and Subrahmanian, V.S., "Relating Dempster-Shafer Theory to Stable Semantics", *Proceedings of the 1991 International Logic Programming Symposium*, pp. 551-565, September 1991.

97. Ng, R. and Subrahmanian, V.S., “Non-monotonic Negation in Probabilistic Deductive Databases”, *Proceedings of the 7th International Conference on Uncertainty in Artificial Intelligence*, pp. 249-256, July 1991. Full journal version published under the title of “Stable Semantics for Probabilistic Deductive Databases”.
98. Ng, R., Faloutsos, C. and Sellis, T., “Flexible Buffer Allocation Based on Marginal Gains”, *Proceedings of ACM SIGMOD International Conference on Management of Data*, pp. 387-396, June 1991. Full journal version published under the title of “Adaptive and Flexible Buffer Allocation”. (Acceptance rate: 45/201).
99. Ng, R. and Subrahmanian, V.S., “A Semantical Framework for Supporting Subjective and Conditional Probabilities in Deductive Databases”, *Proceedings of the 8th International Conference on Logic Programming*, pp. 565-580, June 1991. Full journal version published under the title of “A Semantical Framework for Supporting Subjective and Conditional Probabilities in Deductive Databases”.
100. Ng, R. and Subrahmanian, V.S., “Stable Model Semantics for Probabilistic Deductive Databases”, *Proceedings of the 6th International Symposium on Methodologies for Intelligent Systems*, pp. 162-171, May 1991. Full journal version published under the title of “Stable Semantics for Probabilistic Deductive Databases”.
101. Ng, R. and Subrahmanian, V.S., “Probabilistic Reasoning in Logic Programming”, *Proceedings of the 5th International Symposium on Methodologies for Intelligent Systems*, pp. 9-16, May 1990. Full journal version published under the title of “Probabilistic Logic Programming”.

(c) *Other*

2. NON-REFEREED PUBLICATIONS

(a) *Journals*

1. Christodoulakis, S., Ng, R., et al.: “Optical Mass Storage Systems and Their Performance”, *IEEE Database Engineering*, March 1988.

(b) *Conference Proceedings*

(c) *Other*

3. BOOKS

(a) *Authored*

1. Carenini, G., Murray, G. and Ng, R. “Methods for mining and summarizing text conversations,” Morgan Claypool, June 2011.

(b) *Edited*

(c) *Chapters*

1. Han, J., Ng, R., Fu, Y. and Dao, S., "Dealing with Semantic Heterogeneity by Data Mining Techniques", *Cooperative Information Systems: Current Trends & Directions*, Academic Press, pp. 207-231, 1998.
2. Ng, R. and Dilek, R., "Buffer Allocation with User-Selectable Non-Overflow Guarantee Levels", *Multimedia Information Storage & Management KAP*, pp. 147-162, 1997.
3. Ng, R. and Subrahmanian, V.S., "Dempster-Shafer Logic Programs and Stable Semantics", *Logical Methods: In Honor of Anil Nerode's Sixtieth Birthday*, Birkhauser, pp. 654-704, 1993.

4. PATENTS**5. SPECIAL COPYRIGHTS****6. ARTISTIC WORKS, PERFORMANCES, DESIGNS****7. OTHER WORKS****8. WORK SUBMITTED (including publisher and date of submission)****9. WORK IN PROGRESS (including degree of completion)**