# Benjamin Piwowarski MsC PhD

**1st grade researcher at CNRS (Databases and Machine Learning department, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)**

**E-mail: benjamin@bpiwowar.net**
**Web page: http://www.bpiwowar.net/**

## 1 General information

**Research areas:** (Structured) Information Retrieval, Evaluation, Web Log Mining, XML Databases, Machine Learning applied to IR.

**Teaching areas:** Database, XML Databases, Structured Information Retrieval, Networks, Internet, and Operating Systems.

## 2 Education

**10/1999 - 7/2003** Ph.D. Thesis in Computer Science (University Pierre et Marie Curie, Paris, France)

**10/1998 - 30/09/1999** D.E.A. (Master Thesis) (University Pierre et Marie Curie, Paris, France)

**9/1996 - 6/1998** 3rd and 4th year courses in mathematics (University René Descartes, Paris, France)

**9/1994 - 6/1996** 1st and 2nd year courses in mathematics and psychology (University René Descartes, Paris, France)

## 3 Professional Experience

**10/2011 - current** 1st grade researcher in the French Scientific Research National Centre (CNRS) (Databases and Machine Learning department, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)

**9/2008 - 7/2011** Research Associate (Information Retrieval Group, School of Computing science, University of Glasgow, UK)

**5/2008 - 8/2008** Research Assistant (Information Retrieval Group, Department of Computer Science, Queen Mary University of London, UK)

**3/2006 - 3/2008** Researcher position (Yahoo! Research Latin America)

**9/2004 - 2/2006** Post-doctorate position (Departemento de Ciencias de la Computacion (DCC),University of Chile, Santiago)

**9/2002 - 8/2004** Temporary lecturer (ATER, French 1-year full-time position) (Statistical Machine Learning Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)

**4/1999 - 9/1999** DEA Internship (LexiQuest SARL, Montreuil, France)

# Part I

# Teaching and Supervision

## 4 Thesis and Internship Supervision

**09/2013 - current** PhD co-supervision (Machine Learning and Information Access Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)
*Representation Learning for Graphs.*

**01/2012 - current** PhD co-supervision (Machine Learning and Information Access Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)
*Web Archives.*

**01/2012 - current** PhD co-supervision (Machine Learning and Information Access Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)
*Signed Networks.*

**12/2008 - 09/2011** PhD supervision team member – official supervisor: J. Jose (Information Retrieval Group, School of Computing science, University of Glasgow, UK)
*Collaborative Filtering.*

**3/2010 - 7/2010** PhD internship supervision – official supervisor: M. Lalmas, . (Information Retrieval Group, School of Computing science, University of Glasgow, UK)
*Quantum Information Retrieval.*

**2/2010 - 2/2011** Research Associate supervision team member – official supervisor: . (Information Retrieval Group, School of Computing science, University of Glasgow, UK)
*Quantum Information Retrieval and Polyrepresentation.*

**7/2009 - 12/2010** PhD supervision team member – official supervisor: . (Information Retrieval Group, School of Computing science, University of Glasgow, UK)
*Mining web search intents.*

**9/2008 - 12/2008** Master supervision team member – official supervisor: . (Information Retrieval Group, School of Computing science, University of Glasgow, UK)
*Evaluation in XML Retrieval.*

**1/2007 - 4/2008** PhD supervision team member (University of Chile, Santiago, Chile)
*Usage of Learning-Object Metadata during Lesson Authoring.*

**1/2007 - 4/2008** Master internship supervisor (University of Chile, Santiago, Chile)
*XML Database with an XQuery engine based on Proximal Nodes.*

**4/2001 - 9/2001** PhD supervision team member – official supervisor: . (Statistical Machine Learning Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)
*Structure, Machine Learning and Categorisation.*

**4/2003 - 9/2003** D.E.A. (Master) internship team member – official supervisor: . , A. Doucet (Statistical Machine Learning Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)
*Indexation and Storage for Structured Information Retrieval.*

**10/2002 - 2/2003** Engineer internship supervisor (Statistical Machine Learning Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)
*XML editor for letters (in Java).*

**4/2001 - 9/2001** D.E.A. (Master) internship team member – official supervisor: . (Statistical Machine Learning Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France)
*Structure, Machine Learning and Categorisation.*

## 5 Teaching Experience

**2014-15** Artificial Intelligence (5th year), École Nationale d'Informatique pour l'Industrie et l'Entreprise, Evry, France. Lecturer (28 hours per year)

**2014-15** Information Retrieval (5th year), University René Descartes, Paris, France. Lecturer (24 hours per year)

**2013-15** Introduction to Artificial Intelligence (4th year), Polytech, University Pierre et Marie Curie, Paris, France. Lecturer (4 hours per year)

**2005-07** XML Databases (5th year), University of Chile, Santiago, Chile. Lecturer (36 hours per year)

**2006-07** Information Retrieval (5th year), University of Chile, Santiago, Chile. Lecturer (part of the course, focusing on XML Information Retrieval) (14h hours per year)

**2002-04** Databases (3rd year), University Pierre et Marie Curie, Paris, France. Teaching assistant (44 hours per year)

**2003-04** Network (3rd year), University Pierre et Marie Curie, Paris, France. Teaching assistant (44 hours per year)

**2000-04** Operating Systems and Internet (1st year), University Pierre et Marie Curie, Paris, France. Teaching assistant (30 hours per year)

**1999-2000, 2003-04** Excel and Pascal (1st year), University Pierre et Marie Curie, Paris, France. Teaching assistant (32 hours per year)

**2001 and 2003** Computer Organization and Assembler Programming (1st year), University Pierre et Marie Curie, Paris, France. Teaching assistant (34 hours per year)

# Part II

# Research

## 6 Themes

**Structured or XML Information Retrieval** Publications: 3 4 5 26 27 33 34 35 42 43 45 46 47

**Evaluation in IR** Publications: 21 22 25 32 44 55 59 60

**Query Log Mining** Publications: 1 12 16 17 18 51 52

**Web Archives** Publications: 36 54

**Representing text** Publications: 53

**Quantum IR** Publications: 2 10 11 13 58 31 38 39 48 50 56 57 58

**Social Networks** Publications: 8 29 30

# 7 Grants and projects

**09/2012 - current** *participating as Principal Investigator for LIP6* – PULSAR, FUI (Regional Grant) (University Pierre et Marie Curie, Paris, France).

**04/2015 - current** *participating as Principal Investigator for LIP6* – Distributed User Models, Yahoo! FREP (Acadamic Collaborations) (University Pierre et Marie Curie, Paris, France).

**2012 - 2015** ARESOS, CNRS challenge (University Pierre et Marie Curie, Paris, France).

**2010 - 2011** *participating as researcher* – Qontext, ERC International Research Staff Exchange Scheme (Information Retrieval Group, School of Computing science, University of Glasgow, UK).

**2003 - 2004** *participating as researcher* – ACIMD, ACI (University Pierre et Marie Curie, Paris, France).

**1999 - 2002** *participating as researcher* – Victor Hugo Letters, (University Pierre et Marie Curie, Paris, France).

# 8 Conference related activies

**Senior Program Committee (International conferences)** SIGIR (2015), AIRS (2015)

**Program committee (International conferences)** ECIR (2012-13), WWW (2012), SIGIR (2007-14), WSDM (2009-15), CIKM (2005, 2007-13), ICTIR (2009,2013,2015), ACL (2011), KDD (2010)

**Program committee (National conferences)** CORIA (2012-15)

**Program committee (Workshops)** Simulation of User Interaction (2010), Information Retrieval Over Query Sessions (2011), Workshop on XML Element Retrieval Methodology (2006), Workshop on Web Search Click Data (2013), INEX (2005)

**Organizing Committee (workshops)** Information retrieval and applications of graphical models (2008) , INEX (2005-07)

**Organizing Committee (international conferences)** ECDL (2010)

**Organizing Committee (national conference)** CORIA (2015)

**Chairing (conferences)** ICTIR (2009) , AIRS (2010)

# 9 Invited talks

1. 2015-11-09 – *Centre International de Mathématiques et d'Informatique de Toulouse, Toulouse, France.* Representation learning for text processing
2. May 2014 – *AFIA-ATALA, Paris, France.* Representation learning for text processing
3. January 2014 – *Xerox, Grenoble, France.* The Quantum Information Access framework
4. January 2014 – *Université Joseph Fourrier, Grenoble, France.* The Quantum Information Access framework
5. October 2010 – *Institute for Language, Cognition and Computation, University of Edinburgh, Scotland, UK.* The Quantum Information Access framework
6. October 2010 – *CNRC ITI– NRCC IIT, Gatineau, Canada.* A Quantum (inspired) Information Retrieval Framework
7. October 2010 – *University of Montréal, Canada.* A Quantum (inspired) Information Retrieval Framework

8. June 2010 – *Microsoft Research Cambridge, UK.* A Quantum (inspired) Information Retrieval Framework
9. May 2010 – *Yahoo! Research Barcelona, Spain.* A Quantum (inspired) Information Retrieval Framework
10. February 2009 – *University of Granada, Spain.* Quantum Information Retrieval
11. June 2006 – *University Pompeus Fabro, Barcelona, Spain.* XML Databases

# 10 Publications

In the Information Retrieval field, the lead author is the (most of the times) the first author of the publication. In the following,

♦ denotes publications for which I was lead author, ◊ those for which I was a co-lead author, and △ those for which I was supervising the work. For selected publications, the number of citations (as estimated by google) is given. In the IR field, major conferences have an impact which is equivalent to major journals. The major journals in IR are Transactions on Information Systems (TOIS), Information Retrieval and the major conferences are SIGIR, CIKM, and WSDM. In the following, major conferences and journals are denoted by † .

## 10.1 Journal Papers: Refereed

1. G. Dupret and B. Piwowarski. Model Based Comparison of Discounted Cumulative Gain and Average Precision. *Journal of Discrete Algorithms*, 18:49–62, Jan. 2013.

2. ♦† B. Piwowarski, M. Amini, and M. Lalmas. On using a quantum physics formalism for multidocument summarization. *Journal of the American Society for Information Science and Technology*, 63(5):865–888, May 2012.

3. ♦† B. Piwowarski, A. Trotman, and M. Lalmas. Sound and complete relevance assessments for XML retrieval. *ACM Transactions On Information Systems*, 27(1), jan 2009.

4. ♦† B. Piwowarski, P. Gallinari, and G. Dupret. An extension of precision-recall with user modelling (PRUM): Application to XML retrieval. *ACM Transactions On Information Systems*, 25(1), 2007.

5. ♦† B. Piwowarski and P. Gallinari. A bayesian network for XML information retrieval: Searching and learning with the INEX collection. *Information Retrieval*, 8(4):655–681, December 2005.

## 10.2 Journal Papers: Not Refereed

6. J. M. Fernández-Luna, J. F. Huete, and B. Piwowarski. Introduction to the special issue on Graphical Models and Information Retrieval. *International Journal of Approximate Reasoning*, 50(7):929–931, July 2009.

## 10.3 Book chapter

7. ♦† B. Piwowarski and R. Blanco. Introducción a la recuperación de información. In F. C. Sijo, J. M. Fernández Luna, and J. F. Huete Guadix, editors, *Recuperación de Información. Un enfoque práctico y multidisciplinar*, chapter 1. RA-MA, 2011.

## 10.4 International Conferences: Refereed

8. △ L.-A. Gauthier, P. Gallinari, and B. Piwowarski. Leveraging rating behavior to predict negative social ties. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, 2015.

9. △† Y. Moshfeghi, B. Piwowarski, and J. M. Jose. Handling data sparsity in collaborative filtering using

emotion and semantic based features. In *Proceedings of the 34th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, July 2011.

10. ♦† B. Piwowarski, I. Frommholz, M. Lalmas, and K. van Rijsbergen. What can quantum theory bring to IR? In J. Huang, N. Koudas, G. Jones, X. Wu, K. Collins-Thompson, and A. An, editors, *CIKM'10: Proceedings of the nineteenth ACM conference on Conference on information and knowledge management*. ACM, 2010 *[Acceptance rate: 13.4]*.

11. △ I. Frommholz, B. Larsen, B. Piwowarski, M. Lalmas, P. Ingwersen, and K. van Rijsbergen. Supporting polyrepresentation in a quantum-inspired geometrical retrieval framework. In *Proceedings of the 3rd iiiX symposium*, aug 2010 *[Award: Short-listed for best paper]*.

12. † G. Dupret and B. Piwowarski. A user behavior model for average precision and its generalization to graded judgments. In F. Crestania, M.-M. Stéphane, H.-H. Chen, E. N. Efthimiadis, and J. Savoy, editors, *Proceedings of the 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2010.

13. ♦ B. Piwowarski, I. Frommholz, M. Lalmas, and K. van Rijsbergen. Exploring a multidimensional representation of documents and queries. In *Proceedings of RIAO*, 2010.

14. △ Y. Moshfeghi, D. Agarwal, B. Piwowarski, and J. M. Jose. Movie recommender: Semantically enriched unified relevance model for rating prediction in collaborative filtering. In M. Boughanem, C. Berrut, J. Mothe, and C. Soulé-Dupuy, editors, *ECIR*, volume 5478 of *Lecture Notes in Computer Science*, pages 54–65, Toulouse, France, mar 2009. Springer.

15. ♦ B. Piwowarski and M. Lalmas. A quantum-based model for interactive information retrieval. In L. Azzopardi, G. Kazai, S. E. Robertson, S. M. Rüger, M. Shokouhi, D. Song, and E. Yilmaz, editors, *Proceeedings of the 2nd International Conference on the Theory of Information Retrieval*, volume 5766. Springer, Sep 2009.

16. ♦† B. Piwowarski, G. Dupret, and R. Jones. Mining user web search activity with layered bayesian networks or how to capture a click in its context. In R. A. Baeza-Yates, P. Boldi, B. A. Ribeiro-Neto, and B. B. Cambazoglu, editors, *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, Barcelona, Spain, February 2009. ACM.

17. † G. Dupret and B. Piwowarski. A user browsing model to predict search engine click data from past observations. In S.-H. Myaeng, D. W. Oard, F. Sebastiani, T.-S. Chua, and M.-K. Leong, editors, *SIGIR 2008*, Singapore, July 2008. ACM.

18. ♦† B. Piwowarski and H. Zaragoza. Predictive user click models based on click-through history. In *Proceedings of the Sixteenth Conference on Information and Knowledge Management (CIKM 2007)*, pages 175–182, Lisbon, Portugal, November 2007. ACM.

19. △ O. Motelet, B. Piwowarski, G. Dupret, J. A. Pino, and N. Baloian. Enhancing educational-material retrieval using authored-lesson metadata. In *Fourteenth String Processing and Information Retrieval Symposium (SPIRE 2007)*, Santiago, Chile, October 2007.

20. △ O. Motelet, N. Baloian, B. Piwowarski, and J. A. Pino. Taking advantage of the semantics of a lesson graph based on learning objects. In *The 13th International Conference on Artificial Intelligence in Education (AIED 2007)*. IOS Press, July 2007.

21. G. Kazai, B. Piwowarski, and S. E. Robertson. Effort precision and gain-recall based on a probabilistic navigation model. In *1st International Conference on the Theory of Information Retrieval*, October 2007.

22. ♦† B. Piwowarski and G. Dupret. Evaluation in (XML) information retrieval: Expected precision-recall with user modelling (EPRUM). In E. N. Efthimiadis, S. T. Dumais, D. Hawking, and K. Järvelin, editors, *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 260–267, Seattle, Washington, USA, aug 2006. ACM.

23. G. Dupret, B. Piwowarski, C. Hurtado, and M. Mendoza. A statistical model of query log generation. In *Proceedings of the 13th International Symposium on String Processing and Information Retrieval (SPIRE 2006)*, LNCS 4209, pages 217–228. Springer, 2006.

24. G. Dupret and B. Piwowarski. Principal components for automatic term hierarchy building. In *Proceedings of the 13th International Symposium on String Processing and Information Retrieval (SPIRE 2006)*, LNCS

4209, pages 37–48. Springer, 2006.

25. ♦† B. Piwowarski and M. Lalmas. Providing consistent and exhaustive relevance assessments for XML retrieval evaluation. In *Proceedings of the Thirteenth Conference on Information and Knowledge Management (CIKM 2004)*, Washington D.C., U.S.A., November 2004.

26. ♦ B. Piwowarski and P. Gallinari. A machine learning model for information retrieval with structured documents. In P. Perner, editor, *Machine Learning and Data Mining in Pattern Recognition*, pages 425–438, Leipzig, Germany, July 2003. Springer Verlag.

27. ♦ B. Piwowarski. Learning in information retrieval: a probabilistic differential approach. In *Proceedings of the BCS-IRSG, 22nd Annual Colloquium on Information Retrieval Research*, Sidney Sussex College, Cambridge, England, April 2000.

## 10.5 International Conferences: Not Refereed

28. ♦ B. Piwowarski and M. Massimo. Quantum mechanics and information retrieval: From theory to application. In *Proceedings of the 2013 Conference on the Theory of Information Retrieval*, pages 1–1, 2013.

## 10.6 National Conference: Refereed

29. △ L. Gauthier, B. Piwowarski, and P. Gallinari. Polarité des jugements et des interactions pour le filtrage collaboratif et la prédiction de liens sociaux. In *CORIA 2015 - Conférence en Recherche d'Infomations et Applications - 12th French Information Retrieval Conference, Paris, France, March 18-20, 2015.*, pages 139–154, 2015.

30. △ L.-A. Gauthier, B. Piwowarski, and P. Gallinari. Filtrage collaboratif et intégration de la polarité des notes. In *Conférence en Recherche d'Infomations et Applications - 11th French Information Retrieval Conference*, 2014.

31. ♦ B. Piwowarski. Méthodologie pour une représentation multi-dimensionnelle des documents. In *Conférence en Recherche d'Infomations et Applications - 10th French Information Retrieval Conference*, 2013.

32. ♦ B. Piwowarski and M. Lalmas. Interface pour l'évaluation de systèmes de recherche sur des documents xml. In *Premiere COnference en Recherche d'Information et Applications (CORIA'04)*, Toulouse, France, March 2004.

33. ♦ B. Piwowarski and P. Gallinari. Structure, recherche d'information et apprentissage. Lyon, France, January 2003.

34. ♦ B. Piwowarski, L. Denoyer, and P. Gallinari. Un modèle pour la recherche d'information sur des documents structurés. In *JADT*, Saint-Malo, France, March 2002.

35. ♦ B. Piwowarski. Apprentissage et recherche documentaire : une approche probabiliste différentielle. In *Colloque Francophone sur l'Apprentissage Automatique (CAP'2000)*, Saint-Etienne, France, June 2000.

## 10.7 Workshop Papers: Refereed

36. △ Z. Pehlivan, B. Piwowarski, and S. Gançarski. A comparison of static index pruning methods with temporal queries. In *SIGIR 2013 Workshop on Time-aware Information Access*. LIP6, 2013.

37. S. Attfield, G. Kazai, M. Lalmas, and B. Piwowarski. Towards a science of user engagement (position paper). In *WSDM Workshop on User Modelling for Web Applications*, February 2011.

38. △ A. Caputo, B. Piwowarski, and M. Lalmas. A query algebra for quantum information retrieval. In *Proceedings of the 2nd Italian Information Retrieval Workshop*, January 2011.

39. ♦ B. Piwowarski and M. Lalmas. Structured information retrieval and quantum theory. In P. Bruza,

D. Sofge, W. Lawless, C. J. van Rijsbergen, and M. Klusch, editors, *Proceedings of the Third Quantum Interaction Symposium*, volume 5494 of *Lecture Notes in Artificial Intelligence*. Springer, March 2009.

40. G. Dupret, V. Murdock, and B. Piwowarski. Web search engine evaluation using clickthrough data and a user model. In *Query Log Analysis: Social and Technological Challenges*, 2007.

41. H. Bast, G. Dupret, D. Majumdar, and B. Piwowarski. Discovering a term taxonomy from term similarities using principal component analysis. In M. Ackermann, B. Berendt, M. Grobelnik, A. Hotho, D. Mladenic, G. Semeraro, M. Spiliopoulou, G. Stumme, V. Svátek, and M. van Someren, editors, *Semantics, Web and Mining, Joint International Workshops, EWMF 2005 and KDO 2005*, volume 4289 of *Lecture Notes in Computer Science*, pages 103–120, Porto, Portugal, 2006. Springer.

42. △ H.-T. Vu, B. Piwowarski, and P. Gallinari. Filtering in XML retrieval: a prospective analysis. In *XML and Information Retrieval workshop of SIGIR 2004*, University of Sheffield, UK, July 2004.

43. ♦ B. Piwowarski and P. Gallinari. An algebra for probabilistic XML retrieval. In *The First Twente Data Management Workshop*, Enschede, The Netherlands, June 2004. SIKS.

44. ♦ B. Piwowarski and P. Gallinari. Expected ratio of relevant units: A measure for structured information retrieval. In N. Fuhr, M. Lalmas, and S. Malik, editors, *INitiative for the Evaluation of XML Retrieval (INEX). Proceedings of the Second INEX Workshop*, Dagstuhl, France, December 2003.

45. ♦ B. Piwowarski, H.-T. Vu, and P. Gallinari. Bayesian networks and INEX'03. In N. Fuhr, M. Lalmas, and S. Malik, editors, *INitiative for the Evaluation of XML Retrieval (INEX). Proceedings of the Second INEX Workshop*, Dagstuhl, Germany, December 2003.

46. ♦ B. Piwowarski, G.-E. Faure, and P. Gallinari. Bayesian networks and INEX. In *Proceedings of the First Annual Workshop of the Initiative for the Evaluation of XML retrieval (INEX)*, DELOS workshop, Dagstuhl, Germany, December 2002. ERCIM.

47. ♦ B. Piwowarski and P. Gallinari. A bayesian network model for page retrieval in a hierarchically structured collection. In *XML Workshop of the 25th ACM SIGIR Conference*, Tampere, Finland, August 2002.

## 10.8 Posters: Refereed

48. △ I. Frommholz, B. Piwowarski, M. Lalmas, and K. van Rijsbergen. Processing queries in session in a quantum-inspired ir framework. In *Proceedings of ECIR 2011*, March 2011. Poster.

49. △ S. Sushmita, B. Piwowarski, and M. Lalmas. Dynamics of genre and domain intents. In *Proceedings of The Sixth Asia Information Retrieval Society Conference*, 2010.

50. ♦ B. Piwowarski, I. Frommholz, Y. Moshfeghi, M. Lalmas, and K. van Rijsbergen. Filtering documents with subspaces. In C. Gurrin, Y. He, G. Kazai, U. Kruschwitz, S. Little, T. Roelleke, S. Rüger, and K. van Rijsbergen, editors, *Advances in Information Retrieval*, volume 5993 of *Lecture Notes in Computer Science*. Springer, 2010 *[Award: Best poster award]*.

## 10.9 Patents

51. B. Piwowarski and G. Dupret. System and method for deducing user interaction patterns based on limited activities. United States Patent Application 20100082605, August 2010.

52. B. Piwowarski and H. Zaragoza. System and method for creating and applying predictive user click models to predict a target page associated with a search query. United States Patent Application 20090094196, July 2009.

### 10.10 Other

53. △ B. Piwowarski, S. Lamprier, and N. Despres. Parameterized Neural Network Language Models for Information Retrieval. arxiv 1510.01562, Oct. 2015.

54. △ Z. Pehlivan, B. Piwowarski, and S. Gançarski. Diversification Based Static Index Pruning - Application to Temporal Collections. *arXiv.org*, Aug. 2013.

55. ♦ B. Piwowarski, G. Dupret, and M. Lalmas. Beyond cumulated gain and average precision: Beyond cumulated gain and average precision: Including willingness and expectation in the user model. Technical Report 1209.4479, arXiv, 2012.

56. ♦ B. Piwowarski. The Kernel Quantum Probabilities (KQP) Library. Technical report, arXiv, March 2012.

57. ♦ B. Piwowarski, I. Frommholz, M. Lalmas, and K. van Rijsbergen. Exploring a multidimensional representation of documents and queries (extended version). Technical Report 1002.3238, arXiv, 2010.

58. ♦ B. Piwowarski and M. Lalmas. A quantum-based model for interactive information retrieval. In L. Azzopardi, G. Kazai, S. E. Robertson, S. M. Rüger, M. Shokouhi, D. Song, and E. Yilmaz, editors, *Proceeedings of the 2nd International Conference on the Theory of Information Retrieval*, volume 5766. Springer, Sep 2009.

59. ◊ J. Pehcevski and B. Piwowarski. Evaluation metrics for structured text retrieval. In M. T. Özsu and L. Liu, editors, *Encyclopedia of Database Systems*, Encyclopedia of Database Systems, pages 1015–1024. Springer, May 2009.

60. ♦ B. Piwowarski. Working group report: the assessment tool. In N. Fuhr, M. Lalmas, and S. Malik, editors, *INitiative for the Evaluation of XML Retrieval (INEX). Proceedings of the Second INEX Workshop*, Dagstuhl, Germany, December 2003.

61. G. Kazai, M. Lalmas, and B. Piwowarski. Inex guidelines for topic development. In N. Fuhr, M. Lalmas, and S. Malik, editors, *Proceedings of INEX 2003*, 2003.

62. ♦ B. Piwowarski. *Techniques d'apprentissage pour le traitement d'informations structurées : application à la recherche d'information.* PhD thesis, University Paris 6, Paris, France, July 2003.

# Part III
# Contributions and skills

## 11 Realisations

**Experimastro** Experimaestro - experiment manager for computer science experiments and simulations
https://github.com/bpiwowar/experimaestro/
main developer – Python, C++

**mg4j-ir-extensions** Extension for the MG4J framework: new IR models and tools
https://github.com/bpiwowar/mg4j-ir-extensions/
main developer – Java

**datasets** Utilities and tools for managing standard datasets (in IR)
https://github.com/bpiwowar/datasets/
main developer – Python

**KQP** Kernel Quantum Probabilities
http://kqp.sf.net/
main developer – Java, C++

**QIR** Quantum Information Retrieval framework
http://experimaestro.sf.net/
main developer – Java, Bash, C++

**XPN** Implementation of XQuery based on the Proximal Node data model
main developer – C++

**EvalJ** Evaluation software for Structured Information Retrieval
main developer – Java

**EPSIR** Experimental Platform for Structured Information Retrieval
main developer – C++

**C-shellexec** small cross-platform web server
main developer – C

## 12 Management and administration

**09/2012 - current** Seminars organiser for the DAPA department
Databases and Machine Learning department, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France

**09/2008 - 08/2010** Seminars organiser (with M. Lalmas)
Information Retrieval Group, School of Computing science, University of Glasgow, UK

**09/1999 - 12/2002** Network and System Administration for the group, maintenance of the team website (internet and intranet)
Statistical Machine Learning Group, Laboratoire d'Informatique de Paris 6 - CNRS, University Pierre et Marie Curie, France

## 13 Main Computing Skills

**Programmation** C/C++, Java, Python, Bash, Lua, Julia

**Scientific** Cran R, Octave

**Parallel computing** Hadoop, Spark, Graphlab

**Databases** SQL, MySQL, Postgresql

**Web** HTML, Javascript, PHP

**XML technologies** XML, XML Schema, XSLT, XQuery, XML Databases

## 14 Languages

**French** Native

**English** Very good

**Spanish** Very good