

Bibliographie

[Abu-Salem, 1992]

H., Abu-Salem, A microcomputer based Arabic bibliographic information retrieval system with relational thesauri (Arabic-IRS). Thèse de Doctorat, Illinois Institute of Technology, Chicago, USA, 1992.

[Abu-Salem et al., 1999]

H., Abu-Salem, M., Al-Omari, and M., Evens. Stemming methodologies over individual query words for an Arabic information retrieval system. Journal of the American Society for Information Science, 50:524-529, 1999.

[Adda et al., 1997]

G. Adda, M. Adda-Decker, J. Gauvain, and L. Lamel. Text normalization and speech recognition in french. In Proceedings of the European Conference on Speech Technology, EuroSpeech, Rhodes, 5:2711-2714, 1997.

[Adamson and Boreham, 1974]

G. Adamson and J. Boreham. The use of an association measure based on character structure to identify semantically related pairs of words and document titles. Information storage and retrieval, 10 :253-260, 1974.

[Al-Kabi et al., 2005]

Mohammed Naji Al-Kabi, Ghassan Kanaan, Riad Al-Shalabi, Saja I. Al- Sinjilawi and Ronza S. Al-Mustafa, Al-Hadith Text Classifier, Journal of Applied Sciences 5(3): 584-587, 2005.

[Al-Kabi et Al- Sinjilawi, 2007]

Mohammed Naji Al-Kabi and Saja I. Al- Sinjilawi, A Comparative Study of the Efficiency of Different Measures to Classify Arabic Text, University of Sharjah Journal of Pure & Applied Sciences, 4(2):13-26, 2007.

[Al-Kharashi et Evens, 1994]

I., Al-Kharashi, and M., Evens. Comparing words, stems, and roots as index terms in an Arabic information retrieval system. Journal of the American Society for Information Science, 45(8):548-560, 1994.

[Alkhatib, 2010]

M. Alkhatib, Classification of Al-Hadith Al-Shareef Using Data Mining Algorithm, College of Information Technology, Zayed University, Abu Dhabi, UAE, 2010.

[Amini et al., 1999]

M.R. Amini, H. Zaragoza, and P. Gallinari. Stochastic Models for Surface Information Extraction in Texts, International Conference on Artificial Neural Networks, ICANN'99, pp.892-897, 1999.

[Amini, 2001]

M.R. Amini. Apprentissage Automatique et Recherche de l'Information: application à l'Extraction d'Information de surface et au Résumé de texte. Thèse de Doctorat de l'Université Paris 6, 2001.

- [Anderson et Pérez-Carballo, 2001]
J., Anderson, and J., Pérez-Carballo. The nature of indexing : how humans and machines analyze messages and texts for retrieval, part ii : machine indexing, and the allocation of human versus machine effort. *Information Processing & Management* 37:255-277. Tarrytown, NY, USA, Pergamon Press, Inc. 2001.
- [Arriss, 2001]
Arriss Computer Society, "The Arabian Scientific Encyclopedia: Hal Taâlam", 2001.
- [Attar et Fraenkel, 1977]
R. Attar, and A.S. Fraenkel. Local feedback in full-text retrieval systems. *Journal of the ACM*, 24(3):397- 417, 1977.
- [Attia, 2000]
M., Attia. A large-scale computational processor of the Arabic morphology, Master's thesis, Computer Engineering, Faculty of Engineering, Cairo, Egypt, 2000.
- [Baeza-Yates et Ribeiro-Neto, 1999]
R. Baeza-Yates and B. Ribeiro-Neto, *Modern information retrieval*, ACM Press books, Addison-Wesley, 1999.
- [Beeferman et al. 1999]
D. Beeferman, A. Berger and J. Lafferty. Statistical models for text segmentation, *Machine Learning*, 34:177-210, 1999.
- [Beesley, 1996]
K.R., Beesley. Arabic finite-state morphological analysis and generation. *Proceedings of the 16th International Conference on Computational Linguistics, COLING'96*, 1:89-94, 1996.
- [Bellot, 2000]
P. Bellot. Méthodes de classification et de segmentation locales non supervisées pour la recherche documentaire. Thèse de Doctorat de l'université d'Avignon et des Pays de Vaucluse, France, 2000.
- [Bellot et El-Bèze 2001]
P. Bellot et M. El-Bèze. Classification et Segmentation de Textes par Arbres de Décision, Application à la Recherche Documentaire. *Technique et Science Informatiques*, 20(1): 107-134, 2001.
- [Bishop 1995]
C. M. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.
- [Boughanem et al., 2004]
M. Boughanem, W. Kraaij, et J.-Y. Nie. Modèles de langue pour la recherche d'information. In *Les systèmes de recherche d'informations*, 163-182. Hermes-Lavoisier. 2004.
- [Brants et al., 2002]
T. Brants, F. Chen, And I. Tsochantaridis. Topic-based document segmentation with probabilistic latent semantic analysis, presented at CIKM, McLean, Virginia, USA, 2002.
- [Breiman et al., 1984]
L., Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and regression trees*. Monterey, CA: Wadsworth. 1984.

- [Broomhead and Loewe, 1988]
D., Broomhead, and D., Loewe. Multivariate functional interpolation and adaptive networks. *Complex Systems*, 2:321-355, 1988.
- [Caillet et al., 2004]
M. Caillet, J.F. Pessiot, M. Amini, P. Gallinari. Unsupervised Learning with Term Clustering for Thematic Segmentation of Texts. 7th Proceedings of Recherche d'Information Assistée par Ordinateur RIAO'04, 2004.
- [Carletta, 1996]
J. Carletta. Assessing agreement on classification tasks: The kappa statistic. *Computational Linguistics*, 22(2):249-254, 1996.
- [Celeux, 1992]
G. Celeux, and G. Govaert. A classification em algorithm for clustering and two stochastic versions. *Computational Statistics and Data Analysis*, 14(3):315-332, 1992.
- [Chathura et al., 2008]
R. D. Chathura, R. Surendra and C.D Liyanage, Cloud Basis Function Neural Network: a Modified RBF Network Architecture for Holistic Facial Expression Recognition, *Pattern Recognition*, 41:1241-1253, 2008.
- [Cherfi, 2004]
Hacène Cherfi. Etude et réalisation d'un système d'extraction de connaissance à partir de textes. Thèse de Doctorat, INRIA, France, 2004.
- [Choi, 2000]
F. Choi. Advances in domain independent linear text segmentation," presented at the first conference on North American chapter of the Association for Computational Linguistics (NAACL), Seattle, Washington, 2000.
- [Chuang & Chien, 2005]
S.L., Chuang, L.F., Chien. Taxonomy Generation for Text Segments: A Practical Web-Based Approach, *ACM Transactions on Information Systems*, 23(4):363-396, 2005.
- [Church et Hanks, 1990]
K.W. Church, and P. Hanks. Word association norms, mutual information, and lexicography, *Computational Linguistics*, 16(1):22-29, 1990.
- [Cover et Hart, 1967]
T.M., Cover, and P.E. Hart, Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13:21-27, 1967.
- [Croft, 1983]
W.B., Croft. Experiments with representation in a document retrieval system. *Information Technology: Research and Development*, 2(1):1-21, 1983.
- [Darwish, 2002]
K. Darwish. Building a Shallow Arabic Morphological Analyzer in One Day, Proceedings of the workshop on Computational Approaches to Semitic Language, in the 40th Annual Meeting of the Association for the Computational Linguistics, (ACL-02), pp. 01-08, 2002.
- [Dempster, 1977]
Dempster, N. Laird and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistical Society*, 39(1):1-38, 1977.

-
- [El-Kourdi et al., 2004]
M. El-Kourdi, A. Bensaid, and T. Rachidi, Automatic Arabic Document Categorization Based on the Naïve Bayes Algorithm, 20th International Conference on Computational Linguistics, Geneva, August 2004.
- [El Shayeb et al. 2007]
M. A. El-Shayeb, S. R. El-Beltagy and A. Rafea. Comparative Analysis of Different Text Segmentation Algorithms on Arabic News Stories, In Proc. IEEE International Conference on Information Reuse and Integration, pp. 441 - 446, Aug, 2007.
- [Fellbaum, 1998]
C. Fellbaum, WordNet: An Electronic Lexical Database (Language, Speech, and Communication), The MIT Press, 1998.
- [Ferret et Grau, 1998]
O. Ferret and B. Grau. A Thematic Segmentation Procedure for Extracting Semantic Domains from Texts, ECAI'98, 1998.
- [Ferret et al., 1998]
O. Ferret, B. Grau, and N. Masson. Thematic Segmentation of Texts: Two Methods for Two Kinds of Text. COLING-ACL'98, 392-396, 1998.
- [Finan et al., 1996]
R. Finan, A. Sapeluk, and R. Damper, Comparison of Multilayer and Radial Basis Function Neural Networks for Text-dependent Speaker Recognition, IEEE International Conference on Neural Networks (IJCNN'96), Washington DC, USA, 1996.
- [Frakes, 1992]
W. B. Frakes. Stemming Algorithms. Frakes W B, Baeza-Yates R (eds), 131-160, Prentice Hall, New jersey, 1992.
- [Galley et al, 2003]
M. Galley, K. Mckeown, E. Fosler-Lussier, And H. Jing. Discourse segmentation of multi-party conversation. In: Proceedings of the 41st Annual Meeting of ACL, Sapporo, Japan, 2003.
- [Ghazizadah et al., 2008]
M. Ghazizadeh, M. H. Zahedi, M. Kahani and B. Minaei Bidgoli, Fuzzy Expert System in Determining Hadith Validity, T. Sobh (ed.), Advances in Computer and Information Sciences and Engineering, 354-359. Springer Science+Business Media B.V. 2008.
- [Gilbert, 2006]
S., Gilbert. Data Mining : une nouvelle façon de faire de la statistique, chaire de statistique appliquée, conservatoire national des arts et métiers, Paris, 2006.
- [Gilleron et Tommasi, 2000]
R., Gilleron and M., Tommasi. Découverte de connaissances à partir de données. Cours donné à l'IUP MIAE troisième année à Lille 1, France, 2000.
- [Grosz et Sidner, 1986]
B. J. Grosz And C. L. Sidner. Attention, Intentions and the Structure of Discourse, Computational Linguistics, 12:175-204, 1986.
- [Halliday et Hasan, 1976]
M. A. K. Halliday and R. Hasan, Cohesion in English, Longman, London, 1976.

- [Hammed et Tolba, 2000]
H., Hammed, H., Tolba. Automated Narrator Chain Validation for Hadith Studies: Mathematical Modeling and Artificial Intelligence Implementation, National Telecommunication Institute, Le Caire, Egypte, 2000.
- [Harf, 1997]
Harf Company, The Encyclopedia of the Nine Books for the Honorable Prophetic Traditions, 1997, <http://www.Harf.com>.
- [Harman, 1986]
D.K., Harman. An exeperimental study of factors important in document ranking. Actes de ACM/SIGIR Conference on Research and Development in Information Retrieval, 186-193, Pise, Italie, 1986.
- [Harman, 1992]
D.K. Harman, Ranking algorithms, in [Frakes & Baeza-Yates, 1992], 363-392, 1992.
- [Harrag, 2005]
F. Harrag, Modélisation de l'expertise dans les bases de données : application à l'authentification des traditions prophétiques "Hadith". Mémoire de Magistère de l'université Farhat Abbas de Sétif, Algérie, 2005.
- [Hasnah, 1996]
A., Hasnah. Full text processing and retrieval: Weight ranking, text structuring, and passage retrieval for Arabic documents. PhD dissertation, Illinois Institute of Technology, Chicago, USA, 1996.
- [Hastie et al., 2001]
T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer, 2001.
- [Hayes et Weinstein, 1990]
P.J., Hayes and S.P. Weinstein, Construe: A system for content-based indexing of a database of new stories. In: Second Annual Conference on Innovative Applications of Artificial Intelligence, 1990.
- [Haykin, 1999]
S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice-Hall, Englewood Cliffs, NJ, USA, 2nd Edition, Chapter 4, MLP, 156-252, Chapter 5 RBF 256-312, 1999.
- [Hearst, 1994]
M.A. Hearst, Context and structure in automated full-text information access, PhD dissertation, Université de Californie, Berkeley, USA, rapport n° UCB/CSD-94/386, 1994.
- [Hearst, 1997]
M. A. Hearst. TextTiling: Segmenting text into multi-paragraph subtopic passages, Computational Linguistics, 23:33-64, 1997.
- [Helfman, 1996]
J.I. Helfman. Dotplot pattern: A literal look at pattern languages, Theory and Practice of object Systems, 2(1):31-41, 1996.

- [Hmeidi, 1995]
I., Hmeidi. Design and implementation of automatic word and phrase indexing for information retrieval with Arabic documents. Thèse de Doctorat, Illinois Institute of Technology, Chicago, USA, 1995.
- [Hmeidi et al., 1997]
I., Hmeidi, G., Kanaan, and M., Evens. Design and implementation of automatic indexing for information retrieval with Arabic documents. *Journal of the American Society for Information Science*, 48:867-881, 1997.
- [Huang et al., 2004]
T. M. Huang and V. Kecman. Bias term b in SVMs again. In *Proceedings of the Twelfth European Symposium on Artificial Neural Networks (ESANN 2004)*, 441-844, Bruges, Belgium, 2004.
- [Husbands et al., 2000]
P. Husbands, H. Simon, and C. Ding, On the Use of the Singular Value Decomposition for Text Retrieval, In *Proceedings of 1st SIAM Computational Information Retrieval Workshop*, 2000.
- [Isard et Carletta, 1995]
A. Isard and J. Carletta. Replicability of transaction and action coding in the map task corpus. In Johanna Moore and Marilyn Walker, editors, *Empirical Methods in Discourse: Interpretation & Generation*, AAAI Technical Report SS-95~06. AAAI Press, Menlo Park, CA, 1995.
- [Jalam, 2003]
Radwan JALAM. Apprentissage automatique et catégorisation de textes multilingues, Thèse de doctorat, de l'université lumière lyon2, France, 2003.
- [Jianping et al., 2002]
D. Jianping, N. Sundararajan, and P. Saratchandran, "Communication Channel Equalization Using Complex-Valued Minimal Radial Basis Function Neural Networks", *IEEE Transactions on Neural Networks*, 13(3):687-696, 2002.
- [Joachims 1998]
T. Joachims. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. *Proceedings of the Tenth European Conference on Machine Learning (ECML'98)*, Springer Verlag, 137-142, 1998.
- [Jones et Furnas, 1987]
W.P. Jones, G.W. Furnas, Pictures of relevance: a geometric analysis of similarity measures, *Journal of the American Society for Information Science*, 38(6):420-442, 1987.
- [Kan et al., 1998]
M.Y. Kan, J.L. Klavans, and K.R. Mckeown. Linear segmentation and segment significance. In *Sixth Workshop on Very Large Corpora*. Association for Computational Linguistics, 1998.
- [Khalis, 2006]
Z. Khalis. La segmentation thématique Application à la campagne DEFT'06, 2006.
- [Kleinberg, 1999]
J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604-632. September 1999.

- [Kowalski, 1997]
G., Kowalski. Information Retrieval Systems-Theory and implementation, Kluwer Academic Publishers, 1997.
- [Kozima, 1993]
H. Kozima. Text Segmentation Based on Similarity between Words, In Proceedings of ACL'93, Ohio, Japan, pp. 286-288, 1993.
- [Lam & Lee, 1999]
S.L.Y. Lam and D.L. Lee, Feature Reduction for Neural Network Based Text Categorization, in: Sixth International Conference on Database Systems for Advanced Applications (DASFAA'99), 195-203, 1999.
- [Lamprier, 2008]
Sylvain Lamprier, Vers la conception de documents composites : extraction et organisation de l'information pertinente. Thèse de Doctorat de l'université d'Angers, France, 2008.
- [Larkey et al. 2002]
L., Larkey, L., Ballesteros, and M., Connell. Improving stemming for Arabic information retrieval: light stemming and occurrence analysis. In Proc. of the 25th ACM International Conference on Research and Development in Information Retrieval (SIGIR'02), 275-282, Tampere, Finland, 2002.
- [Lee et al., 1997]
D.L. Lee, H. Chuang, K. Seamons. Document ranking and the Vector Space model. IEEE Computing, 67-75, 1997.
- [Lewis et al., 1998]
D., Lewis, A., McCallum & K., Nigam. The independence assumption in information retrieval. In proceedings of ECML-98, 10th European conference on Machine learning, 1998.
- [Litman et Passonneau, 1995]
D.J. Litman and J.R. Passonneau. Combining multiple knowledge sources for discourse segmentation. In Proceedings of the 33rd Meeting of Association for Computational Linguistics, pp.108-115, 1995.
- [Liu et al., 2003]
T. Liu, S. Liu, Z. Chen, and M.A. Wei-Ying, An Evaluation on Feature Selection for Text Clustering, Proceedings of the 12th International Conference ICML'03, 488-495, Washington, DC, USA, 2003.
- [Maamouri & Cieri, 2002]
M., Maamouri, C., Cieri. Resources for Arabic natural language processing at the Linguistic Data Consortium. In A. Braham (Ed.), Proceedings of the International Symposium on the Processing of Arabic, 125-146, Manouba, Tunisia, 2002.
- [Maron and Kuhns, 1960]
Maron, M. and Kuhns, J. On relevance, probabilistic indexing and information retrieval. Journal of the Association for Computing Machinery, 7:216-244. 1960.
- [Massih, 2007]
R., Massih, Apprentissage de Fonctions de Classification et d'Ordonnement avec des Données Partiellement Étiquetées, Habilitation à diriger des recherches de l'université Pierre et Marie Curie, Paris, France, 2006.

- [Masson, 1995]
N. Masson. An Automatic Method for Document Structuring," In Proceedings of the 18th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, Seattle, Washington, USA, 1995.
- [Mataoui, 2007]
M. Mataoui, Reformulation de requêtes dans les systèmes de recherche d'information dans des documents XML. Mémoire de Magistère de l'université M'hamed Bougara de Boumerdes, Algérie, 2007.
- [Mekhaldi et al., 2004]
D. Mekhaldi, D. Lalanne et R. Ingold. Using bi-modal alignment and clustering techniques for documents and speech thematic segmentations, 2004.
- [Moody and Darken, 1989]
J., Moody, and C., Darken. Fast learning in networks of locally-tuned processing units. *Neural Computation*, 1(2):281-294, 1989.
- [Morris et Hirst, 1991]
J. Morris And G. Hirst. Lexical cohesion computed by thesaurus relations as an indicator of the structure of text, *Computational Linguistics*, 17(1):21-48, 1991.
- [Park et al., 2002]
J. Park, R. Harley, and G. Venayagamoorthy, Comparison of MLP and RBF Neural Networks using Deviation Signals for On-Line Identification of a Synchronous Generator, IEEE Power Engineering Society Winter Meeting, New York, USA, 2002.
- [Papaioannou et al., 2006]
I. V. Papaioannou, I. G. Roussaki, M. E. Anagnostou, Comparing the Performance of MLP and RBF Neural Networks Employed by Negotiating Intelligent Agents, Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'06), 2006.
- [Passonneau et Litman, 1993]
J.R. Passonneau and D.J. Litman. Intention-based segmentation: Human reliability and correlation with linguistic cues. In Proceedings of the 31st Annual Meeting, 148-155, 1993.
- [Pevzner et Hearst 2002]
L. Pevzner and M. Hearst. A critique and improvement of an evaluation metric for text segmentation, *Computational Linguistics*, 28(1):19-36, 2002.
- [Pessiot et al, 2004]
J.F. Pessiot, M. Caillet, M.R. Amini et P. Gallinari: Apprentissage non-supervisé pour la segmentation automatique de textes. CORIA'04, 213-228, 2004.
- [Platt, 1999]
J.C. Platt. Fast Training of Support Vector Machines Using Sequential Minimal Optimization. *Advances in Kernel Methods - Support Vector Learning*, MIT Press, 185-208, 1999.
- [Poggio and Girosi, 1989]
T., Poggio, and F., Girosi. A theory of networks for approximation and learning. Technical Report AIM-1140, 1989. Available from World Wide Web : <http://citeseer.nj.nec.com/poggio89theory.html>.

[Porter, 1980]

M. F. Porter. An algorithm for suffix stripping. Program 14. 1980.

[Powell, 1987]

M., Powell, Algorithms for Approximation, chapter Radial basis functions for multivariable interpolation: A review, 143-167. Clarendon Press, New York, 1987.

[Preece, 1973]

S.E. Preece. Clustering as an output option. In Proceedings of the American Society for Information Science, 10:189–190, 1973.

[Qiu & Frei, 1993]

Y., Qiu, H.P., Frei. Concept based query expansion. Actes de ACM/SIGIR'93 Conference on Research and Development in Information Retrieval, 160-169, Pittsburgh PA, USA, 1993.

[Quinlan, 1993]

J.R., Quinlan, C4.5: Programs for machine learning. Morgan Kaufmann, San Francisco, USA, 1993.

[Raghavan et Wong, 1986]

V.V., Raghavan, S.K.M., Wong. A critical analysis of vector space model for information retrieval, Journal of the American Society for Information Science, 37(5):279-287, 1986.

[Rajan et al., 2009]

K. Rajan, V. Ramalingam, M. Ganesan, S. Palanivel, and B. Palaniappan, Automatic Classification of Tamil Documents Using Vector Space Model and Artificial Neural Network. Expert Systems with Applications, 2009.

[Rakotomalala & Mhamdi, 2006]

R. Rakotomalala, and F. Mhamdi, Combining Feature Selection and Feature Reduction for Protein Classification, Proceedings of the 6th WSEAS International Conference on Simulation, Modeling and Optimization, Lisbon, Portugal, 444-451, September 2006.

[Réhel, 2005]

S. Réhel, Catégorisation automatique de textes et cooccurrence de mots provenant de documents non étiquetés, Mémoire de maîtrise de l'université Laval, Quebec, Canada, 2005.

[Reynar, 2000]

J.C. Reynar. Topic Segmentation: Algorithms and Application, Ph.D. thesis, Computer and Information Science. University of Pennsylvania, Pennsylvania, USA, 2000.

[Richmond et al, 1997]

K. Richmond, A. Smith, and E. Amitay. Detecting subject boundaries withintext: A language independent statistical approach. In Proceedings of Second Conference on Empirical Methods in Natural Language Processing, ACL, pp. 47–54, 1997.

[Rijsbergen, 1979]

C. van Rijsbergen. Information retrieval. Butterworths. 1979.

[Rizzo, 2004]

Hassan Mudhafar Rizzo. Modèle informatique pour la Tradition Prophétique, Dix-septième Conférence Nationale en Informatique, Université du Roi Abdelaziz, Médine, Arabie Saoudite, 597-610, 2004.

- [Robertson et Sparck-Jones, 1976]
S.E., Robertson, K., Sparck-Jones, Relevance weighting of search terms, *Journal of the American Society for Information Science*, 27(3):129-146, 1976.
- [Robertson, 1977]
S. Robertson. The probability ranking principle in IR. *Journal of Documentation*, 33(4):294-304. 1977.
- [Robertson et al., 1994a]
S. E. Robertson and S. Walker. Some simple effective approximations to the 2-Poisson model for probabilistic weighted retrieval. In *Proceedings of SIGIR 1994*, 232-241. 1994.
- [Robertson et al., 1994b]
S. Robertson, S. Walker, S. Jones, and M. H.-B. and M. Gatford. Okapi at TREC 3. In *Proceedings of the 3rd Text REtrieval Conference (TREC-3)*, 109-126. 1994.
- [Rocchio 1971]
J.J. Rocchio. The SMART Retrieval System: Experiments in Automatic Document Processing, chapter 14, *Relevance Feedback in Information Retrieval*, 313-323. Gerard Salton (editor), Prentice-Hall Inc. : New Jersey, 1971.
- [Roget, 1977]
Roget's International Thesaurus. 4th Edition, 1977.
- [Sahami 1998]
M. Sahami. Using Machine Learning to Improve Information Access. Ph.D. Dissertation, Stanford, 1998.
- [Salton, 1970]
G. Salton. The SMART retrieval system : Experiments in automatic document processing. Prentice Hall. 1970.
- [Salton, 1983]
G. Salton. *Introduction to Modern Information Retrieval*. Mc Graw-Hill, New York, 1983.
- [Salton et al., 1983]
Salton, G., Fox, E., and Wu, H. Extended boolean information retrieval. *Communications of the ACM*, 26(11) :1022-1036. 1983.
- [Salton & Buckley, 1988]
G., Salton, C., Buckley. Term weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5):513-523, 1988.
- [Salton & Allan, 1994]
G. Salton & J. Allan. Automatic text decomposition and structuring. *Actes de RIAO'94*, 6-29, New York, NY, USA, 1994.
- [Salton et al., 1996]
G. Salton, A. Singhal, C. Buckley, M. Mitra, « Automatic text decomposition using text segments and text themes », *actes de ACM/Hypertext'96*, Washington D.C., USA, pp. 53-65.
- [Sanjuan.2004]
Ibekwe Sanjuan. *Ingénierie linguistique et fouille de textes*, 2004.

- [Sauvagnat, 2005]
Karen Sauvagnat, *Modèle flexible pour la Recherche d'Information dans des corpus de documents semi-structurés*, thèse de Doctorat de l'Université Paul Sabatier de Toulouse, France, 2005.
- [Sawaf et al., 2001]
H. Sawaf, J. Zaplo, and H. Ney, *Statistical Classification Methods for Arabic News Articles*, Workshop on Arabic Natural Language Processing, ACL'01, Toulouse, France, July 2001.
- [Sebastiani, 2002]
Fabresio Sebastiani, *Machine Learning in Automated Text Categorization*. ACM Computing Surveys, 34(1):1-47, 2002.
- [Selamat & Omatu, 2003]
A. Selamat, and S. Omatu, *Neural Networks for Web Page Classification Based on Augmented PCA*, Proceedings of the International Joint Conference on Neural Networks, 1792-1797, 2003.
- [Sentürk, 2007]
Sentürk, R. *Hadith: Origins and Developments*. Edited by Harald Motzki. Book review in *Journal of Islamic Studies* 18:97-99; 2007.
- [Stéphane, 2005]
T., Stéphane. *Data Mining et statistique décisionnelle, L'intelligence dans les bases de données*, Edition TECHNIP, Paris, France, 2005.
- [Singhal et al., 1996]
A. Singhal, G. Salton, M. Mitra, and C. Buckley. *Document length normalization*. *Information Processing & Management*, 32(5) :619-633. 1996.
- [Sitbon et Bellot, 2005]
L. Sitbon, P. Bellot. *Segmentation thématique par chaînes lexicales pondérées*, Laboratoire d'Informatique d'Avignon, Université d'Avignon France, 2005.
- [Sparck-Jones, 1972]
K., Sparck-Jones. *A statistical interpretation of term specificity and its application in retrieval*. *Journal of Documentation*, 28(1):11-20, 1972.
- [Stein et Schmid, 1995]
A. Stein et H. Schmid. *Étiquetage Morphologique de Textes Français avec un Arbre de Décisions*, *Traitement Automatique des Langues*, 36:1-2, 1995.
- [Syed Ifran et Syed Ghazanfar, 2008]
Hyder, Syed Irfan & Ahmad, Syed Ghazanfar. *Towards a Database Oriented Hadith Research Using Relational, Algorithmic and Data-Warehousing Techniques*, Shaikh Zayed Islamic Research Center, University of Karachi, Pakistan, 2008.
- [Syiam et al. 2006]
M.M., Syiam, Z.T., Fayed, and M.B., Habib. *An Intelligent System for Arabic Text Categorization*. *IJICIS*, 6(1):1-19, 2006.
- [Tombros et al., 2002]

- Anastasios Tombros, Robert Villa, and C. J. Van Rijsbergen. The effectiveness of query-specific hierarchic clustering in information retrieval. *Information Processing & Management*, 38(4): 559-582, 2002.
- [Touir et al., 2008]
A.A. Touir, H. Makhtour and W. Al-Sanea. Semantic-based segmentation of Arabic texts, *Information Technology Journal*, 7(7):1009-1015, 2008.
- [Vapnik, 1995]
V., Vapnik. *The Nature of Statistical Learning*. Springer, 1995.
- [Vinot, 2004]
R., Vinot. Classification automatique de textes dans des catégories non thématiques, Thèse de Doctorat de l'école national supérieur des télécommunications, France, 2004.
- [Voorhees & Harman, 2005]
E.M., Voorhees, D., Harman, *TREC: experiment and evaluation in information retrieval*. Cambridge (Mass.) MIT Press, 2005.
- [Wall et al., 2003]
M., Wall, A. Rechtsteiner, and L. Rocha, *A Practical Approach to Microarray Data Analysis*, Chapter: Singular Value Decomposition and Principal Component Analysis. Kluwer, 91-109, 2003.
- [Wasserman, 1995]
P. D. Wasserman, "Advanced Methods in Neural Computing", Van Nostrand Reinhold, New York, NY, USA, 1995.
- [Wensinck, 1936-1988]
Arent Jan Wensinck, Witkam et Ravan. *Concordance et Indices de la Tradition Musulmane*. Editions E.J. Brill, Leiden, Vol.1-Vol.8, 1936-1988. {Concours de 66 chercheurs}.
- [Wong et al., 1985]
S. Wong, W. Ziarko, and P. Wong. Generalized vector space model in information retrieval. In *Proceedings of the 8th ACM SIGIR Conference on Research and Development in information retrieval*, New-York, USA, 18-25. 1985.
- [Wermeter, 2000]
S. Wermeter. Neural Network Agents for Learning Semantic Text Classification, *Information Retrieval*, 3(2):87-103, 2000.
- [Xu et al., 2001]
J., Xu, A., Fraser, J., Makhoul, M., Noamany, & G., Osman. UN Arabic English parallel text version 1.0 beta [CD-ROM]. University of Pennsylvania, Linguistic Data Consortium, Philadelphia, USA, 2001.
- [Yang 1999]
Y. Yang, J. Carbonell, R. Brown, T. Pierce, B. T. Archibald, and X. Liu. Learning approaches for detecting and tracking news events. *IEEE Intelligent Systems*, 14(4):32-43, 1999.
- [Zipf 1949]
G. K. Zipf. *Human Behavior and the Principle of Least Effort*. Addison-Wesley, 1949. et <http://users.info.unicaen.fr/~giguette/java/zipf.html>.