Bibliography

[Agg12] C. C. Aggarwal and C. Zhai. **A survey of text clustering algorithms**. In *Mining Text Data*, pages 77–128. Springer-Verlag,, 2012. IV.2(b)

[Ahm10] N. K. Ahmed, F. Berchmans, J. Neville and R. Kompella. **Time-based sampling of social network activity graphs**. In *Proceedings of the Eighth Workshop on Mining and Learning with Graphs, MLG '10*, pages 1–9, New York, NY, USA, 2010. ACM. II.1

[Aki08] N. Akiva, E. Greitzer, Y. Krichman and J. Schler. **Mining and visualizing online web content using BAM: Brand association map**. In E. Adar, M. Hurst, T. Finin, N. S. Glance, N. Nicolov and B. L. Tseng, editors, *ICWSM*. The AAAI Press, 2008. I.1

[Ara12] L. L. P. Araújo and R. Rios. **The popularization of the social network sites and the okutinization phenomenon**. In *Proceedings of XIV Congresso de Ciências da Comunicação na Região Nordeste - Recife - PE - 14 a 16/06/2012. Intercom - Sociedade Brasileira de Estudos Interdisciplinares da Comunicação*, pages Available at `http://www.intercom.org.br/papers/regionais/nordeste2012/resumos/R32--0590--1.pdf`, 2012. V.1

[Ben09] F. Benevenuto, T. Rodrigues, M. Cha and V. Almeida. **Characterizing user behavior in online social networks**. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference, IMC '09*, pages 49–62, New York, NY, USA, 2009. ACM. II.1

[Ben11] J. L. Bender, M.-C. Jimenez-Marroquin and A. R. Jadad. **Seeking support on facebook: A content analysis of breast cancer groups**. *Journal of Medical Internet Research*, 13(1):e16, 2011. II.3

[Ber05] M. Berger, T. H. Wagner and L. C. Baker. **Internet use and stigmatized illness**. *Social Science & Medicine*, 61(8):1821–1827, 2005. I.1, II.2, V.1

[Big10] C. A. S. Bigonha, T. N. C. Cardoso, M. M. Moro, V. A. F. Almeida and M. A. Goncalves. **Detecting evangelists and detractors on twitter**. In *Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia)*, pages 107–114, Belo Horizonte, Brazil, 2010. II.1

[Boy07] D. M. Boyd and N. B. Ellison. **Social network sites: Definition, history, and scholarship**. *Journal of Computer-Mediated Communication*, 13(1):article 11, 2007. I.1, II.3, III.1, VI.3

[Bre06] J. G. Breslin, S. Decker, A. Harth and U. Bojars. **SIOC: an approach to connect web-based communities**. *Int. J. Web Based Communities*, 2:133–142, July 2006. IV.1(b)

[Bre12] B. Furneaux. **Task-technology fit theory: A survey and synopsis of the literature**. In Y. K. Dwivedi, M. R. Wade and S. L. Schneberger, editors, *Information Systems Theory*, volume 28, pages 87–106. Springer, New York, 2012. IV.2(d)

[Brin98] S. Brin and L. Page. **The anatomy of a large-scale hypertextual web search engine**. In *Seventh International World-Wide Web Conference (WWW 1998)*, 1998. V.1

[CG10] CGI. **Survey on the use of information and communication technologies in brazil: Ict households and ict enterprises 2010**. Comitê Gestor da Internet no Brasil, 2011. II.2

[Carn09] H. A. Carneiro and E. Mylonakis. **Google trends: A web-based tool for real-time surveillance of disease outbreaks**. *Clinical Infectious Diseases*, 49(10):1557–1564, 2009. I, I.1, II.2

[Carp09] C. Carpineto, S. Osiński, G. Romano and D. Weiss. **A survey of web clustering engines**. *ACM Computing Surveys (CSUR)*, 41(3):17:1–17:38, 2009. II.4

[Car10] D. B. F. Carvalho and C. J. P. Lucena. **Social media savvy: exploiting orkut data**. Technical Report 20/10, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil, dec 2010. I.1, III.2(a), IV.1, IV.1(a)

[Car12] D. Carvalho, W. Madeira, M. Okamura, C. Lucena and S. Zanetta. **A practical approach to exploit public data available on the internet to study healthcare issues**. In *Proceedings of the XXXII Congress of the Brazilian Computer Society, XII Workshop on Medical Informatics*, pages 1–8. SBC, 2012. I.1, III, III.1

[CAM12] D. Carvalho, H. Fuks and C. Lucena. **Community association map: Processing inter-community relationships**. In *Proceedings of the 8th International Conference on Web Information Systems and Technologies (WEBIST)*, pages 665–670. Springer, 2012. IV.1

[Car12b] D. Carvalho, R. Marcacini, C. Lucena and S. Rezende. **Towards a process to support solving the content selection problem from online community forums**. In *Proceedings of the XXXII Congress of the Brazilian Computer Society, Brazilian Workshop on Social Network Analysis and Mining*, pages 1–4. SBC, 2012. IV.2

[Caa12] M. Carta, J. Angst, M. Moro, G. Mura, M. Hardoy, C. Balestrieri, L. Chessa, G. Serra, M. Lai and P. Farci. **Association of chronic hepatitis c with recurrent brief depression**. *Journal of Affective Disorders*, 141(2–3):361–6, 2012. V.1

[Che08] X. Cheng, C. Dale and J. Liu. **Statistics and Social Network of YouTube Videos**. In *Quality of Service, 2008. IWQoS 2008. 16th International Workshop on*, pages 229–238, 2008. II.1

[Che09] W.-Y. Chen, J.-C. Chu, J. Luan, H. Bai, Y. Wang and E. Y. Chang. **Collaborative filtering for orkut communities: discovery of user latent behavior**. In *Proceedings of the 18th International Conference on World Wide Web*, *WWW '09*, pages 681–690, New York, NY, USA, 2009. ACM. IV.1

[Cor10] G. Cormode, B. Krishnamurthy and W. Willinger. **A manifesto for modeling and measurement in social media**. *First Monday*, 15(9), 2010. III.2(a), IV.1(b)

[Cos07] L. da F. Costa, F. A. Rodrigues, G. Travieso and P. R. V. Boas. **Characterization of complex networks: A survey of measurements**. *Advances In Physics*, 56:167–242, 2007. II.1

[Cscr11] A. Banks, 2011. The Brazilian Online Audience. Feb, 2011. URL: `http://www.comscore.com/por/Press_Events/Presentations_Whitepapers/2011/State_of_the_Internet_in_Brazil`. Accessed: 2012-04-01. I.1, II.2

[Dav12] E. Davoodi, M. Afsharchi and K. Kianmehr. **A social network-based approach to expert recommendation system**. In E. Corchado, V. Snasel, A. Abraham, M. Wozniak, M. Grana and S.-B. Cho, editors,

Hybrid Artificial Intelligent Systems, volume 7208 of *Lecture Notes in Computer Science*, pages 91–102. Springer Berlin, 2012. II.4

[Eys09] G. Eysenbach. **Infodemiology and infoveillance: Framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the internet**. *Journal of Medical Internet Research*, 11(1):e11, 2009. I, I.1, II.2, III.1, III.1(a)

[Fer11] L. Fernandez-Luque, R. Karlsen and J. Bonander. **Review of extracting information from the social web for health personalization**. *Journal of Medical Internet Research*, 13(1):e15, 2011. I.1, V.1

[Gol11] J. Gold, A. E. Pedrana, R. Sacks-Davis, M. E. Hellard, S. Chang, S. Howard, L. Keogh and J. S. H. andMark A Stoove1. **A systematic examination of the use of online social networking sites for sexual health promotion**. *BMC Public Health*, 11(1):583, 2011. II.3

[Gom11] J. Gomide, A. Veloso, W. Meira, V. Almeida, F. Benevenuto, F. Ferraz and M. Teixeira. **Dengue surveillance based on a computational model of spatio-temporal locality of twitter**. In *Proceedings of the ACM WebSci'11*, pages 1–8. ACM, 2011. II.2

[Goo95] D. L. Goodhue. **Understanding user evaluations of information systems**. *Management Science*, 41(12):1827–1844, 1995. IV.2(d), IV.2(d)

[Gre11] J. A. Greene, N. K. Choudhry, E. Kilabuk and W. H. Shrank. **Online social networking by patients with diabetes: a qualitative evaluation of communication with facebook**. *Journal of General Internal Medicine*, 26(3):287–292, 2011. II.3, V.1

[HSL12] H. S.-L. P. Release, 2012. Pesquisa com 2.427 comunidades virtuais identifica motivações para o uso de drogas e o abandono do vício. Hospital Sírio-Libanês. URL:`http://www.hospitalsiriolibanes.org.br/Lists/Noticias/Visualizador.aspx?ID=213`, Archived by WebCite® at `http://www.webcitation.org/6EZfhWupI`. I.1

[Hey09] T. Hey, S. Tansley and K. Tolle, editors. **The fourth paradigm: Data-intensive scientific discovery**. Microsoft Research, Redmond, Washington, 2009. I

[Hub08] B. Huberman, D. Romero and F. Wu. **Social networks that matter: Twitter under the microscope**. *First Monday*, 14(1), 2008. II.1

[IRSSL2012] I. de Responsabilidade Social Sírio-Libanês Press Release, 2012. Pesquisa aponta perfil de usuário de drogas nas comunidades virtuais. Instituto de Responsabilidade Social Sírio-Libanês. URL:`http://www.irssl.org.br/noticias/noticias_detalhe.php?cod=15`, Archived by WebCite® at `http://www.webcitation.org/6EZfxyFga`. I.1

[IBGE11] IBGE. **Sinopse do censo demográfico 2010**. Instituto Brasileiro de Geografia e Estatística, 2011. II.2

[Jan91] N. W. Jankowski. *A Handbook of Qualitative Methodologies for mass communication research*, chapter Qualitative research and community media, pages 163—174. Routledge, 1991. III.1, V.1

[Jup06] V. Jupp. **The sage dictionary of social research methods**. SAGE, London, 2006. III.1, V.1

[Kad12] C. Kadushin. **Understanding social networks: Theories, concepts, and findings**. Oxford Univ Pr, 2012. II.4

[Kan07] I.-S. Kang, S.-H. Na, J. Kim and J.-H. Lee. **Cluster-based patent retrieval**. *Information Processing and Management*, 43(5):1173–1182, Sept. 2007. II.4

[Koz09] R. V. Kozinets. **Netnography: Doing Ethnographic Research Online**. Sage Publications Ltd, 2009. I, I.1, II.1, II.3, III.1, III.1(a), IV.1(b), V.1, VI.2

[Krz09] M. I. Krzywinsk, J. E. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S. J. Jones and M. A. Marra. **Circos: an information aesthetic for comparative genomics**. *Genome Res.*, 19(9):1639–1645, 2009. IV.1(a)

[Las05] J. N. Lasker, E. D. Sogolow and R. R. Sharim. **The role of an online community for people with a rare disease: Content analysis of messages posted on a primary biliary cirrhosis mailinglist**. *Journal of Medical Internet Research*, 7(1):e10, 2005. II.3

[Laz09] D. Lazer, A. Pentland, L. Adamic, S. Aral, A.-L. Barabási, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy and M. Van Alstyne. **Computational Social Science**. *Science*, 323(5915):721–723, Feb. 2009. I, I.1, I.3, II.1

[Lef05] F. Lefevre and A. M. C. Lefevre. **Depoimentos e discursos – uma proposta de análise em pesquisa social – vol.12**. Liber Livros, 2005. I.1

[Lef06] F. Lefevre and A. M. C. Lefevre. **The collective subject that speaks**. *Interface — Comunicação, Saúde, Educação*, 10(20):517–524, 2006. I.1, II.3, III.1, IV.2

[Lit12] H. Litwin. **Physical activity, social network type, and depressive symptoms in late life: An analysis of data from the national social life, health and aging project**. *Aging & Mental Health*, 16(5):608–16, 2012. V.1

[Liu11] B. Liu. *Web Data Mining - Exploring Hyperlinks, Contents, and Usage Data*, chapter Unsupervised Learning, pages 133–166. Springer-Heidelberg,, 2nd edition, 2011. II.4

[Mad11] W. Madeira. **Transformar é preciso: transformações na relação de poder estabelecida entre médico e paciente (um estudo em comunidades virtuais)**. PhD thesis, *Faculdade de Saúde Pública, Universidade de São Paulo, BR*, 2011. I.1, I.2, II.3, III.1, IV.2(c), V, V.1, VI.3

[Mar10] R. M. Marcacini and S. O. Rezende. **Torch: a tool for building topic hierarchies from growing text collection**. In *WTA'2010: IX Workshop on Tools and Applications. In 8th Brazilian Symposium on Multimedia and the Web (Webmedia)*, pages 133–135, 2010. I.1, II.4, IV.2, IV.2(b), IV.2(b), IV.2(c)

[Mor11] L. Morgado, 2011. Conectados pela hepatite C. Ciência Hoje Online. URL:`http://cienciahoje.uol.com.br/blogues/bussola/2011/09/conectados-pela-hepatite-c`, Archived by WebCite® at `http://www.webcitation.org/6EZfK4GdQ`. I.1

[NCHS10] R. A. Cohen and B. Stussman. **Health information technology use among men and women aged 18-64: Early release of estimates from the national health interview survey, january-june 2009**. Health E-Stats. National Center for Health Statistics, February 2010. II.2

[Naz08] A. Nazir, S. Raza and C.-N. Chuah. **Unveiling facebook: a measurement study of social network based applications**. In *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement, IMC '08*, pages 43–56, New York, NY, USA, 2008. ACM. II.1

[Nog08] B. M. Nogueira, M. F. Moura, M. S. Conrado, R. G. Rossi, R. M. Marcacini and S. O. Rezende. **Winning some of the document preprocessing challenges in a text mining process**. In *IV Workshop on*

Algorithms and Data Mining Applications, XXIV Brazilian Symposium on Database, pages 10–18, 2008. IV.2(c)

[Por09] M. Porter. **The Porter Stemming Algorithm**, 2009. Accessible at `http://www.tartarus.org/martin/PorterStemmer`, 2012-04-20. IV.2(c)

[Pre05] J. Preece and D. Maloney-Krichmar. **Online communities: Design, theory, and practice**. *Journal of Computer-Mediated Communication*, 10(4):article 1, 2005. II.1, II.3, III.1

[Rez11] S. O. Rezende, R. M. Marcacini and M. F. Moura. **O uso da mineração de textos para extração e organização não supervisionada de conhecimento**. *Revista de Sistemas de informação da Faculdade Salesiana Maria Auxiliadora*, 1(7):7–21, 2011. II.4

[Shi12] S. Shiovitz-Ezra and H. Litwin. **Social network type and health-related behaviors: Evidence from an american national survey**. *Social Science & Medicine*, 75(5):901–4, 2012. V.1

[Shm10] M. Shmueli-Scheuer, H. Roitman, D. Carmel, Y. Mass and D. Konopnicki. **Extracting user profiles from large scale data**. In *Proceedings of the 2010 Workshop on Massive Data Analytics on the Cloud*, page 4. ACM, 2010. II.4

[Shn11] B. Shneiderman, J. Preece and P. Pirolli. **Realizing the value of social media requires innovative computing research**. *Commun. ACM*, 54:34–37, September 2011. I, I.1

[Sou08] V. V. Sousa and K. P. S. Cruvinel. **Being a hepatitis c carrier: feelings and expectatives**. *Texto & Contexto - Enfermagem*, 17(4):689–695, 2008. V.1, V.1

[Tel10] T. Telegraph, 2010. Archived by WebCite® at `http://www.webcitation.org/66b3jO4Fc`. II.2

[WON12] Words and Networks: Language Use in Socio-Technical Networks (WON), 2012. at ACM Web Science Conference. II.3

[Whi07] L. C. Whitehead. **Methodological and ethical issues in internet-mediated research in the field of health: an integrated review of the literature**. *Social Science & Medicine*, 57(4):782–791, 2007. I, I.1, II.3, III.1(a)

[Wil09] C. Wilson, B. Boe, A. Sala, K. P. Puttaswamy and B. Y. Zhao. **User interactions in social networks and their implications**. In *Proceedings of the 4th ACM European conference on Computer systems, EuroSys '09*, pages 205–218, New York, NY, USA, 2009. ACM. II.1

[Woo01] M. Woolridge and M. J. Wooldridge. **Introduction to multiagent systems**. John Wiley & Sons, Inc., New York, NY, USA, 2001. III.2

[Zha05] Y. Zhao, G. Karypis and U. Fayyad. **Hierarchical clustering algorithms for document datasets**. *Data Mining and Knowledge Discovery*, 10(2):141–168, 2005. II.4

[Zha07] Q. Zhao, P. Mitra and B. Chen. **Temporal and information flow based event detection from social text streams**. In *Proceedings of the 22nd national conference on Artificial intelligence, AAAI'07*, pages 1501–1506. AAAI Press, 2007. II.4