

# Shenghui Wang

Schipholweg 99  
2316 XA Leiden  
The Netherlands  
+31 71 524 65 00  
shenghui.wang@oclc.org

## Education

- ◊ University of Manchester, Manchester, UK  
Ph.D. in Computer Science, May 2007.  
Supervised by Prof. Dr. David S. Bree  
Thesis: The Semantics of Natural Language Descriptions of Continuous Quantities.
- ◊ University of Science and Technology of China, Hefei, Anhui, P.R.China  
Master [M.Sc.] in Computer Application Technology, July 2003.  
Supervised by Prof. Dr. Xufa Wang  
Thesis: On Several Key Technical Issues in Kansei Image Retrieval.
- ◊ Anhui University, Hefei, Anhui, P.R.China.  
Bachelor [B.Sc.] (Hons) in Computer Science (Minor in English), July 2000.  
Supervised by Prof. Dr. Xiaoping Chen  
Thesis: The Basic Principle of Simulated RoboCup and the Design and Implementation of Anhui Team.

## Research Experience

- ◊ 05/2012 - present: OCLC Research EMEA  
Research Scientist in Data Science
  - Text mining, semantic embedding
  - Data visualization
- ◊ 06/2011 - 04/2012: Wageningen UR Food and Biobased Research  
Researcher in eFoodLab
  - Collective development and maintenance of ontology's in specific application contexts.
  - Extraction, storage, and alignment of vocabularies specialized to particular application domains.
  - Integration of multi-source and multi-format food data.
  - Creating incentives for knowledge sharing
- ◊ 11/2010 - 12/2010: University of Stanford  
Visiting scholar in Stanford Center for Biomedical Informatics Research (BMIR)
  - Analysing change logs of the ongoing development of the new International Disease Classification, ICD-11
  - Applying machine learning techniques for change prediction based on previous changes and social aspects of editors
- ◊ 09/2009 - 05/2011: Vrije Universiteit Amsterdam and Rathenau Institute Postdoc research on Semantically Mapping Science (SMS) Project
  - Applying multilevel analysis on longitudinal Scientometrics network data
- ◊ 06/2009 - 05/2011: Vrije Universiteit Amsterdam  
Postdoc research on social and communication networks
  - Extracting data from online social media
  - Representing social interaction and content communication networks in the communication domain
  - Developing theory, methods and tools for analysing multilevel dynamics of co-evolving networks
  - Developing theory, methods and tools for analysing semantic drift in ontology concepts over time
- ◊ 02/2007 - 05/2009: Vrije Universiteit Amsterdam and Koninklijke Bibliotheek (National Library of the Netherlands)  
Postdoc research in the "Semantic Interoperability To access Cultural Heritage" (STITCH) project and "The European Library Plus" (TELplus) project
  - representing cultural heritage data in RDF/OWL/SKOS
  - developing methods and tools for ontology alignment
  - deploying semantic links in cultural heritage applications
  - organising evaluation with domain experts and for the OAEI campaign (2007, 2008, 2009)

- ◊ 09/2003 - 02/2007: University of Manchester  
PhD research in semantic processing and representation of natural language descriptions of continuous quantities.
  - semantic modelling and similarity measure
  - ontological representation and automated reasoning
  - query answering and information integration
 Funded by the British Overseas Research Students (ORS) Awards and the Computer Science department of the University of Manchester.
- ◊ 09/2001 - 07/2003: University of Science and Technology of China  
Research in the Emotional Information Retrieval System
  - modelling user emotional requirements
  - personalising an image retrieval system based on users' dynamic emotional feedback
 Funded by the National 973 project of China (G1998030500): Intellectual Information Retrieval based on Cognitive Neural Science.
- ◊ 05/2001 - 10/2001: University of Science and Technology of China and the Chinese National Audit Office in Nanjing  
Developing a Heterogeneous Data Retrieval Engine, providing a uniform interface to various data sources, such as net databases (Oracle, SQL Server, DB2, etc), file databases (Access, Excel, XBase, etc.) and text files.
- ◊ 01/2000 - 07/2000: Anhui University  
Leader of the first student research project on RoboCup
  - Designing and implementing a multi-agent system for RoboCup simulation league under Linux in C++ and took part in the National Robocup Competition, held by the 3rd world Congress on Intelligent control and Automation, Hefei, China, 2000.

### **Editorial**

- ◊ Member of the editorial board of the International Journal on Advances in Intelligent Systems (2012 -)

### **Organiser**

- ◊ Joint International Semantic Technology Conference (JIST2011), Tutorial chair
- ◊ Joint organiser of the SIKS advanced course on Semantic web, 2009, 2011

### **Community Member**

- ◊ Europeana Task force on Multilingual and Semantic Enrichment Strategy, 2013
- ◊ COAR Interest Group Controlled Vocabularies for Repository Assets, member of the Controlled Vocabularies Authority. 2014 - 2017

### **PC Member**

- ◊ The International Conference on Semantic Systems (SEMANTiCS), 2017, 2018, 2019
- ◊ The International Semantic Web Conference (ISWC), 2010, 2011, 2012, 2013, 2014, 2015, 2017
- ◊ The Extended Semantic Web Conference (ESWC), 2011, 2012, 2014, 2015, 2016, 2017
- ◊ The International Workshop on Ontology Matching (OM), 2010, 2011, 2012, 2013, 2014, 2015
- ◊ Joint International Semantic Technology Conference (JIST), 2011, 2012, 2013, 2014, 2015, 2017, 2018
- ◊ Chinese Semantic Web & Web Science Conference, 2013, 2014
- ◊ The workshop on Making Sense of Microposts (MSM), 2011, 2012, 2013, 2014, 2015
- ◊ The 36th Annual IEEE Computer Software and Applications Conference (COMPSAC 2012), Web Semantic Track
- ◊ The International Conference on Advances in Semantic Processing (SEMAPRO) 2009, 2010, 2011, 2012
- ◊ The Workshop on Interdependent Networks: Quality, Influence and Evolution in Social and Information Networks (WIN2011), at the 22nd ACM Hypertext 2011 track on Social Media
- ◊ The 4th International Workshop on Ontology Dynamics (IWOD-10)
- ◊ The 17th biennial European Conference on Artificial Intelligence (ECAI 2008)
- ◊ The AI and the Web Special Track in the 23rd AAAI Conference on Artificial Intelligence (AAAI-08 AIW)
- ◊ The 9th Annual CLUK Research Colloquium (CLUK 2006)

**Reviewer**

- ◊ International Journal of Digital Libraries 2015
- ◊ PLOS One, 2014
- ◊ Swiss National Science Foundation, 2013
- ◊ Journal of Web Semantics, 2010, 2012
- ◊ Semantic Web Journal, 2010
- ◊ Journal of Computer Science and Technology, 2010, 2011
- ◊ Journal of Data & Knowledge Engineering, 2009
- ◊ IEEE Intelligent Systems, Special Issue on AI and Cultural Heritage, ISSI-2008-08-0119.R1
- ◊ Journal of the American Society for Information Science and Technology, 2008
- ◊ Posters of the 17th International World Wide Web Conference (WWW 2008)
- ◊ External reviewer for ASWC2008, EKAW2008, AIMSAS2008, ESWC2009, OM2009, AAAI2010, EKAW2010, ESWC2010

## Book

- ◊ Carol Jean Godby, Shenghui Wang, Jeffery K. Mixture. Library Linked Data in the Cloud: OCLCs Experiments with New Models of Resource Description. Synthesis Lectures on the Semantic Web: Theory and Practice. Morgan & Claypool Publishers. 2015
- ◊ Shenghui Wang. Semantics of natural language descriptions of continuous quantities. LAP LAMBERT Academic Publishing. 2012.

## Journal articles

- ◊ Shenghui Wang and Rob Koopman. Embed first, then predict. *Knowledge Organizatio*, 2019 (Accepted for publication)
- ◊ T. Castermans, K. Verbeek, B. Speckmann, M.A. Westenberg, R. Koopman, S. Wang, H. van den Berg, and A. Betti. SolarView: Low Distortion Radial Embeddings with a Focus. In *IEEE Transactions on Visualization and Computer Graphics*, 2018 (Accepted for publication).
- ◊ Koopman, R., Wang, S., Scharnhorst, A.: Contextualization of topics – browsing through the universe of bibliographic information. Gläser, A. Scharnhorst, W. Glänzel (eds.) Same data – different results? Towards a comparative approach to the identification of thematic structures in science, Special Issue of Scientometrics (2017) DOI 10.1007/s11192-017-2303-4 <http://rdcu.be/pFC2>
- ◊ Koopman, R., Wang, S.: Mutual information based labelling and comparing clusters. Gläser, A. Scharnhorst, W. Glänzel (eds.) Same data – different results? Towards a comparative approach to the identification of thematic structures in science, Special Issue of Scientometrics (2017) DOI 10.1007/s11192-017-2305-2 <http://rdcu.be/pEmc>
- ◊ Wang, S., Koopman, R.: Clustering articles based on semantic similarity. Gläser, A. Scharnhorst, W. Glänzel (eds.) Same data – different results? Towards a comparative approach to the identification of thematic structures in science, Special Issue of Scientometrics (2017) DOI 10.1007/s11192-017-2298-x <http://rdcu.be/pDZH>
- ◊ Velden, T., Boyack, K., van Eck, N., Glänzel, W., Gläser, J., Havemann, F., Heinz, M., Koopman, R., Scharnhorst, A., Thijs, B., Wang, S.: Comparison of topic extraction approaches and their results. Gläser, A. Scharnhorst, W. Glänzel (eds.) Same data – different results? Towards a comparative approach to the identification of thematic structures in science, Special Issue of Scientometrics (2017) DOI 10.1007/s11192-017-2306-1.
- ◊ Balthasar A. C. Schopman, Shenghui Wang, Antoine Isaac, Stefan Schlobach: Instance- Based Ontology Matching by Instance Enrichment. *Journal of Data Semantics* 1(4): 219- 236. 2012
- ◊ Shenghui Wang, Antoine Isaac, Stefan Schlobach, Lourens van der Meij and Balthasar Schopman. Instance-based Semantic Interoperability in the Cultural Heritage. *Semantic Web* 3(1): 45-64. 2011
- ◊ Christophe Gueret, Shenghui Wang, Paul Groth and Stefan Schlobach. Multi-scale analysis of the Web of Data: a challenge to the Complex System’s community. *Advances in Complex Systems* 14(4). 2011
- ◊ Shenghui Wang, Stefan Schlobach and Michel Klein. Concept drift and how to identify it. *Web Semantics: Science, Services and Agents on the World Wide Web* 9 (2011), pp. 247-265.
- ◊ Antoine Isaac, Shenghui Wang, Claus Zinn, Henk Mattheizing, Lourens van der Meij and Stefan Schlobach. Evaluating Thesaurus Alignments for Semantic Interoperability in the Library Domain. *IEEE Intelligent Systems Special Issue on AI and Cultural Heritage* 24(2) - Mar/Apr 2009
- ◊ Shenghui Wang, David Rydeheard and Jeff Z. Pan. The Semantics Processing of Continuous Quantities for Discrete Terms in Ontologies. *Journal of Logic and Computation* 18(3): 341- 359. 2008.
- ◊ Shenghui Wang and Jeff Z. Pan. Semantically Processing Parallel Colour Descriptions. *Journal of Data Semantics* X:212-236, 2008.
- ◊ Shenghui Wang, Shangfei Wang and Xufa Wang. Visualised Interactive Genetic Algorithm and its Application in Kansei-based Image Retrieval. *Mini-Micro Systems*, Vol.25, No.3, P.399-403, 2004.
- ◊ Shangfei Wang, Shenghui Wang and Xufa Wang. Improved Interactive Genetic Algorithm Incorporating with SVM and Its Application. *Journal of Data Acquisition and Processing*, Vol.18, No.4, P.429-433, 2003.
- ◊ Shangfei Wang, Enhong Chen, Shenghui Wang and Xufa Wang. A Kansei Image Retrieval Based on Emotion Model. *Journal of Circuits and Systems*, Vol.8, No. 6, P.48-52, 2003.

## Conference Papers

- ◊ Shenghui Wang, Rob Koopman and Gwenn Englebienne. Non-parametric Subject Prediction. In *Proceedings of the International Conference on Theory and Practice of Digital Libraries (TPDL2019)*.

September 2019, Oslo Norway.

- ◊ Shenghui Wang, Rob Koopman. A Two-step approach towards subject prediction. In Proceedings of the 17<sup>th</sup> International Conference on Scientometrics & Informetrics. Research-in-progress papers. September 2019, Rome Italy
- ◊ Rob Koopman and Shenghui Wang. An innovative approach to scalable semantic embedding. Accepted to present at the IFLA WLIC 2019 pre-conference satellite meeting: “Data intelligence in libraries: The actual and artificial perspectives”.
- ◊ Rob Koopman and Shenghui Wang, Gwenn Englebienne.: Fast and discriminative semantic embedding. In: Proceedings of the 13th International Conference on Computational Semantics - Long Papers. pp. 235–246. ACL, Gothenburg, Sweden (23–27 May 2019)
- ◊ Shenghui Wang, Rob Koopman, Titia van der Werf and Jean Godby. An innovative approach to scalable semantic search. Abstract accepted and oral presentation at the NSF-supported conference Artificial Intelligence for Data Discovery and Reuse. May 2019, Carnegie Mellon University.
- ◊ Wang, Shenghui and Rob Koopman. Exploring a world of networked information built from free-text metadata. In proceedings of ELAG2015.
- ◊ Koopman, Rob, Shenghui Wang, Andrea Scharnhorst, and Gwenn Englebienne. 2015. Ariadnes Thread Interactive Navigation in a World of Networked Information. In CHI’15 Extended Abstracts on Human Factors in Computing Systems. ACM, Seoul, South Korea.
- ◊ Koopman, Rob, Shenghui Wang and Andrea Scharnhorst. 2015. Contextualization of topics - browsing through terms, authors, journals and cluster allocations. In Proceedings of 15th International Conference on Scientometrics & Informetrics. Istanbul, Turkey.
- ◊ Marc Bron, Titia van der Werf, Shenghui Wang, Maarten de Rijke: A social bookmarking system to support cluster driven archival arrangement. In Proceedings of The Information Interaction in Context conference (IiX2014): 295-298
- ◊ Koopman, Rob, and Shenghui Wang. 2014. Where Should I Publish? Detecting Journal Similarity Based on What Has Been Published There. In Proceedings of Digital Libraries 2014, 483-484. London, United Kingdom
- ◊ Shenghui Wang, Antoine Isaac, Valentine Charles, Rob Koopman, Anthi Agoropoulou, Titia van der Werf: Hierarchical Structuring of Cultural Heritage Objects within Large Aggregations. In Proceedings of the International Conference on Theory and Practice of Digital Libraries (TPDL2013): 247-259
- ◊ Christine Moser, Julie M. Birkholz, Dirk Deichmann, Iina Hellsten, Shenghui Wang: Exploring Ideation: Knowledge Development in Science through the Lens of Semantic and Social Networks. HICSS 2013: 235-243
- ◊ Christian Meilicke, Raul Garcia-Castro, Fred Freitas, Willem Robert van Hage, Elena Montiel-Ponsoda, Ryan Ribeiro de Azevedo, Heiner Stuckenschmidt, Ondrej Svoboda, Vojtech Svttek, Andrei Tamilin, Cássia Trojahn dos Santos, Shenghui Wang: MultiFarm: A benchmark for multilingual ontology matching. Journal of Web Sem. 15: 62-68 (2012)
- ◊ Julie M. Birkholz, Shenghui Wang, Paul T. Groth, Sara Magliacane: Who Are We Talking About? Identifying Scientific Populations Online. CSWS 2012: 237-250
- ◊ Shenghui Wang and Paul Groth. A Framework for Longitudinal Influence Measurement between Communication Content and Social Networks. Best Paper Track at the IJCAI 2011 conference.
- ◊ Shenghui Wang, Jan Kleinijenhuis and Dirk Oegema. Network dynamics in online discussions and the proliferation of negative emotions about immigrants: a reciprocal longitudinal random coefficients model. In Proceedings of 2011 Conference of the International Communication Association (ICA2011), May 2011, Boston
- ◊ Shenghui Wang, Paul Groth, Jan Kleinijenhuis and Dirk Oegema. Multilevel longitudinal analysis for studying influence between co-evolving social and content networks. In SUNBELT 2011: 31st annual conference of the International Network for Social Network Analysis (INSNA). St. Pete Beach, Florida, USA
- ◊ Shenghui Wang and Paul Groth. Measuring the dynamic bi-directional influence between content and social networks. In Proceedings of the 9th International Semantic Web Conference (ISWC 2010). Shanghai, China. **(Best research paper award)**
- ◊ Shenghui Wang, Gwenn Englebienne, Christophe Gueret, Stefan Schlobach, Antoine Isaac and Martijn Schut. Similarity Features, and their Role in Concept Alignment Learning. In Proceedings of the Fourth International Conference on Advances in Semantic Processing (SEMAPRO2010). October 2010, Florence, Italy **(Best paper award)**

- ◊ Shenghui Wang, Stefan Schlobach and Michel Klein. What is concept drift and how to measure it? In Proceedings of the 17th International Conference on Knowledge Engineering and Knowledge Management (EKAW2010). October 2010, Lisbon
- ◊ Yiwen Wang, Shenghui Wang, Natalia Stash, Lora Aroyo, and Guus Schreiber. Enhancing Content-based Recommendation with the Task Model of Classification. In Proceedings of the 17th International Conference on Knowledge Engineering and Knowledge Management (EKAW2010) short paper. October 2010, Lisbon
- ◊ Christophe Gueret, Shenghui Wang and Stefan Schlobach. The Web of Data as a Complex System. In Proceedings of European Conference on Complex Systems (ECCS2010) . September 2010, Lisbon
- ◊ Shenghui Wang, Janet Heleen Takens, Wouter van Atteveldt and Stefan Schlobach. Extensional Mapping-Chains for Studying Concept Drift in Political Ontologies. In Proceedings of 2010 Conference of the International Communication Association (ICA2010), June 2010, Singapore (**Best paper award in the Information System division**)
- ◊ Dirk Oegema, Shenghui Wang and Jan Kleinnijenhuis. Online discussions about politics as a function of news attention. In Proceedings of the 2nd Web Science Conference (poster), April 2010, Raleigh, NC, USA.

### Workshop Papers

- ◊ Hein van den Berg, Arianna Betti, Thom Castermans, Rob Koopman, Bettina Speckmann, Kevin Verbeek, Titia van der Werf, Shenghui Wang, and Michel A. Westenberg. A Philosophical Perspective on Visualization for Digital Humanities. In proceedings of 2018 Workshop on Visualization for the Digital Humanities (VIS4DH).
- ◊ Pauline van Wierst, Steven Hofstede, Yvette Oortwijn, Thom Castermans, Rob Koopman, Shenghui Wang, Michel A. Westenberg, and Arianna Betti. BolVis: Visualization for Text-based Research in Philosophy. In proceedings of 2018 Workshop on Visualization for the Digital Humanities (VIS4DH).
- ◊ Wang, S., Koopman, R.: Semantic Embedding for Information Retrieval. In: Proc. of the 5th Workshop on Bibliometric-enhanced Information Retrieval (BIR 2017), CEUR-WS.org (2017) 122–132
- ◊ Shenghui Wang, Stefan Schlobach, Janet Takens and Wouter van Atteveldt. Large Scale Content Analysis, Ontology Mapping, and Concept Drift. In Het Etmaal van de Communicatiewetenschap 2010, Gent, Belgium.
- ◊ Antoine Isaac, Dirk Kramer, Lourens van der Meij, Shenghui Wang, Stefan Schlobach and Johan Stapel. Vocabulary Matching for Book Indexing Suggestion in Linked Libraries A Prototype Implementation & Evaluation. In Proceedings of the 8th International Semantic Web Conference (ISWC 2009). Washington, US.
- ◊ Shenghui Wang, Antoine Isaac, Balthasar Schopman, Stefan Schlobach and Lourens van der Meij. Matching multi-lingual subject vocabularies. In Proceedings of the 13th European Conference on Digital Libraries (ECDL2009). Corfu, Greece.
- ◊ Balthasar Schopman, Shenghui Wang and Stefan Schlobach. Deriving concept mappings through instance mappings. In Proceedings of the 3rd Asian Semantic Web Conference (ASWC2008): 122-136. Bangkok, Thailand, December, 2008.
- ◊ Shenghui Wang, Gwenn Englebienne and Stefan Schlobach. Learning Concept Mappings from Instance Similarity. In Proceedings of the 7th International Semantic Web Conference (ISWC2008):339-355. Karlsruhe, Germany, October 2008.
- ◊ Antoine Isaac, Lourens van der Meij, Stefan Schlobach, and Shenghui Wang. An empirical study of instance-based ontology matching. In Proceedings of the 20th Belgium-Netherlands Conference on Artificial Intelligence (BNAIC 2008). Enschede, The Netherlands, October, 2008.
- ◊ Antoine Isaac, Henk Matthezing, Lourens van der Meij, Stefan Schlobach, Shenghui Wang and Claus Zinn. Putting ontology alignment in context: usage scenarios, deployment and evaluation in a library case. In Proceedings of the 5th European Semantic Web Conference (ESWC 2008): 402-417. Tenerife, Spain, June, 2008.
- ◊ Laura Hollink, Mark van Assem, Shenghui Wang, Antoine Isaac and Guus Schreiber. Two Variations on Ontology Alignment Evaluation: Methodological Issues. In Proceedings of the 5th European Semantic Web Conference (ESWC 2008): 388-401. Tenerife, Spain, June, 2008.
- ◊ Shenghui Wang and Jeff Z. Pan. Ontology-based Integration and Retrieval over Multiple Quantities — What if “Ovate leaves and often blue to purple flowers?” In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence: 388-394. Silicon Valley, CA, USA, 2-5 Nov, 2007.

- ◊ Antoine Isaac, Lourens van der Meij, Stefan Schlobach and Shenghui Wang. An empirical study of instance-based ontology matching. In Proceedings of the 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference, (ISWC/ASWC2007): 253-266. Busan, Korea, Nov 2007.
- ◊ Shenghui Wang and Jeff Z. Pan. Integrating and Querying Parallel Leaf Shape Descriptions. In Proceedings of the 5th International Semantic Web Conference (ISWC2006): 668-681. Athens, Georgia, USA, 2006.
- ◊ Shenghui Wang and Jeff Z. Pan. Ontology-based Representation and Query of Colour Descriptions from Botanical Documents. OTM Conferences (2) 2005: 1279-1295.
- ◊ Julie M. Birkholz and Shenghui Wang. Who are we talking about?: the validity of online metrics for commenting on science. In Proceedings of the altmetrics11: Tracking scholarly impact on the social Web. Koblenz, Germany, 2011.
- ◊ Shenghui Wang, Stefan Schlobach, Janet Takens and Wouter van Atteveldt. Studying Concept Shift in Political Ontologies. In Proceedings of the Workshop on Matching and Meaning of the 36th Annual Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB'10), 29th March 1st April 2010, Leicester, UK
- ◊ Shenghui Wang, Stefan Schlobach, Janet Takens and Wouter van Atteveldt. Mapping- Chains for studying Concept Shift in Political Ontologies. In Proceedings of the Third International Workshop on Ontology Matching (OM2009). Washington, US.
- ◊ Antoine Isaac, Cassia Trojahn, Shenghui Wang and Paulo Quaresma. Using quantitative aspects of alignment generation for argumentation on mappings. In Proceedings of the Third International Workshop on Ontology Matching, collocated with the 7th International Semantic Web Conference (ISWC2008). Karlsruhe, Germany, October, 2008.
- ◊ Shenghui Wang, Antoine Issa, Lourens van der Meij and Stefan Schlobach. Multi-Concept Alignment and Evaluation. In Proceedings of the Second International Workshop on Ontology Matching, collocated with the 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference (ISWC/ASWC2007). Busan, Korea, Nov 2007.
- ◊ Shenghui Wang. Logic-based representation and integration of colour descriptions from botanical documents. In Proceedings of Tenth ESSLLI Student Session: 411-422, Edinburgh, Scotland, 2005.
- ◊ Mary McGee Wood and Shenghui Wang. Motivation for “ontology” in parallel-text information extraction. In Proceedings of ECAI-2004 Workshop on Ontology Learning and Population (ECAI-OLP), Poster, Valencia, Spain, 2004.