

VITA
November 12, 2008

DAVID W. EMBLEY
Department of Computer Science
Brigham Young University
Provo, UT 84602

Born: October 30, 1946
Place: Salt Lake City
Married, 14 children
1090 E. 650 N.
Orem, UT 84097

Education

Ph.D. in Computer Science, University of Illinois, Urbana, October 1976.

M.S. in Computer Science, University of Utah, June 1972.

B.A. in Mathematics, University of Utah, June 1970.

Experience

Professor, Brigham Young University, since September 1986 (granted continuing status 1987).
Teach graduate and undergraduate classes in database systems and theory, discrete mathematics, information extraction and integration, and the semantic web. Research interests include database systems, model-driven software development, information extraction from semistructured and unstructured documents, and the semantic web.

Sabbatical, Dortmund University, January 1999 - May 1999.

Associate Professor, Brigham Young University, August 1982 - August 1986.

Associate Professor, University of Nebraska, August 1981 - August 1982 (granted tenure 1982).

Assistant Professor, University of Nebraska, August 1976 - August 1981.

Recent Grants

NSF: TANGO: Table Analysis for Semiautomatic Generation of Ontologies (with G. Nagy, D.W. Lonsdale, Y. Tijerino), 15 August 2005 - 31 July 2009, \$260,000.

NSF: Target-Based Document-Independent Information Extraction, (with D.M. Campbell, S.W. Liddle, D. Lonsdale, Y.-K. Ng), 1 April 2001 - 30 June 2005, \$429,549.

Plus 4 prior NSF grants and 6 other external grants, plus many internal grants, all prior to 2000.

Recent Professional Involvement

Editorial Board for *Data & Knowledge Engineering* (2005-present).

Editorial Board for *Journal of Data Semantics* (2002-present).

Editorial Board for *World Wide Web: Internet and Web Information Systems* (2000-2004).

Program Co-chair for the 26th International Conference on Conceptual Modeling (ER'06).

Chair of the Steering Committee for the International Conference on Conceptual Modeling (Oct 2002 - Oct 2004).

Vice Chair of the Steering Committee for the International Conference on Conceptual Modeling (Oct 2000 - Oct 2002).

Steering Committee Member for the International Conference on Conceptual Modeling (1997-present).

General Chair for the 19th International Conference on Conceptual Modeling (ER2000), Salt Lake City, Utah, 9-12 October, 2000.

Program Committee Member for dozens of conferences/workshops since 2000.

Plus much more service prior to 2000.

Recent Conference, Seminar, Colloquia, and Workshop Appearances

Panelist for "Conceptualizing Data Integration: Been there, Done that?" at the 27th International Conference on Conceptual Modeling (ER2008), Barcelona, Spain, 20–23 October 2008.

Seminar talk, "WoK: A Web of Knowledge," at Dagstuhl Seminar: The Evolution of Conceptual Modeling, Dagstuhl Schloss, Germany, 28 April, 2008.

Invited talk, "Foundational Data Modeling and Schema Transformations for XML Data Engineering," at the 2nd International United Information Systems Conference (UNISCON'08), Klagenfurt, Austria, 24 April 2008.

Colloquium talk, "WoK: A Web of Knowledge," at Innsbruck University, STI (Semantic Technology Institute?), Innsbruck, Austria, 21 April, 2008.

Colloquium talk, "WoK: A Web of Knowledge," at the General Earth Environmental Research Laboratories, Research Institute for Humanity and Nature, in Kyoto, Japan, 19 January 2008.

Speaker and panelist for the "Next Stage" Mid-career Faculty Development workshop, BYU, 4 December 2007

Paper Presentation, "Seed-based Generation of Personalized Bio-Ontologies for Information Extraction," at the First International Workshop on Conceptual Modelling for Life Sciences Applications (CMLSA'07), Auckland, New Zealand, 5–9 November 2007.

Paper Presentation, "Automatic Hidden-Web Table Interpretation by Sibling Page Comparison," at the 26th International Conference on Conceptual Modeling (ER2007), Auckland, New Zealand, 5–9 November 2007.

- Paper Presentation, “Ontology-Based Constraint Recognition for Free-Form Service Requests,” at the 23rd International Conference on Data Engineering (ICDE 2007), Istanbul, Turkey, 16–20 April 2007.
- Paper Presentation, “Automatic Creation and Simplified Querying of Semantic Web Content: An Approach Based on Information-Extraction Ontologies,” at the 1st Asian Semantic Web Conference (ASWC 2006) Beijing, China, 3-7 September 2006.
- Colloquium talk, “Toward Semantic Understanding—An Approach Based on Information Extraction Ontologies,” presented at the Universitat Politècnica de Catalunya, Spain, 12 June 2006.
- Paper Presentation, “Resolving Underconstrained and Overconstrained Systems of Conjunctive Constraints for Service Requests,” at the 18th International Conference on Advanced Information Systems Engineering (CAiSE’06) Luxembourg City, Luxembourg, 5-9 June 2006.
- Paper Presentation, “Conceptual Model Based Semantic Web Services,” at the Twenty Fourth International Conference on Conceptual Modeling (ER’05), Klagenfurt, Austria, 26-28 October 2005.
- Colloquium talk, “Toward Semantic Understanding—An Approach Based on Information Extraction Ontologies,” presented at the Technical University of Vienna, Austria, 24 October 2005.
- Panelist for the Family History Technology Workshop panel, Brigham Young University, 24 March 2005.
- Colloquium talk, “Toward Semantic Understanding—An Approach Based on Information Extraction Ontologies,” presented at the Polytechnic University of Valencia, Valencia, Spain, 24 February 2005.
- Colloquium talk, “Toward Semantic Understanding—An Approach Based on Information Extraction Ontologies,” presented at the Swiss Federal Institute of Technology, Lausanne, Switzerland, 21 February 2005.
- Invited talk, “Toward Tomorrow’s Semantic Web—An Approach Based on Information Extraction Ontologies,” presented at the Dagstuhl Seminar, “Machine Learning for the Semantic Web,” Dagstuhl Castle, Germany, 17 February 2005.
- Panel moderator for the panel entitled “Hot Research Topics: How Conceptual Modeling Can Contribute” at the 23rd International Conference on Conceptual Modeling (ER2004), Shanghai, China, 8-12 November 2004.
- Paper Presentation, Query rewriting for extracting data behind HTML forms, at the International Workshop on Conceptual Model-directed Web Information Integration and Mining, Shanghai, China, 8-12 November 2004.
- Keynote Address, “What Do You Want—Semantic Understanding? (You’ve Got to be Kidding)” presented at the Fifteenth Australasian Database Conference, Dunedin, New Zealand, January 18-22, 2004.
- Colloquium talk, “Automating Schema Matching,” Massey University, Palmerston North, New Zealand, 16 January 2004.

Paper Presentation, "On Guaranteeing 'Good' Properties for XML," at the 7th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2003), Orlando, Florida, 27-30 July 2003.

Paper Presentation, "Discovering Direct and Indirect Matches for Schema Elements" at the 8th International Conference on Database Systems for Advanced Applications (DASFAA 2003), Kyoto, Japan, 26-28 March 2003.

Paper Presentation, "Recognizing Records from the Extracted Cells of Microfilm Tables," at the Symposium on Document Engineering (DocEng 2002), McLean, Virginia, 8-9 November 2002.

Paper Presentation, "Automatically Extracting Ontologically Specified Data from HTML Tables with Unknown Structure," at the 21st International Conference on Conceptual Modeling (ER2002), Tampere, Finland, 7-11 October 2002.

Colloquium talk, "Automating Schema Matching for Data Integration," Federal University of Minas Gerais, Belo Horizonte Brazil, 20 June 2002.

Colloquium talk, "Automating Schema Matching for Data Integration," WhizBang!, Provo, Utah and, by teleconferencing, Pittsburgh, Pennsylvania, 7 May 2002.

Paper presentation, "Recognizing Ontology-Applicable Multiple-Record Web Documents" at the Twentieth International Conference on Conceptual Modeling (ER2001), Yokohama, Japan, 27-30 November 2001.

Panelist for the panel: "Semantic B2B Integration: is Modeling the Superior Approach over Programming?" at the Twentieth International Conference on Conceptual Modeling (ER2001), Yokohama, Japan, 27-30 November 2001.

Colloquium talk, "Computer Science Research for Family History and Genealogy," introduced a coordinated series of student presentations (Heath Nielson, Mike Rimer, Luke Hutchison, Ken Tubbs, Doug Kennard, Tom Finnigan), Brigham Young University, 8 November 2001.

Presentation, "Milestones in the Research Proposal Process", New Faculty Seminar, session for new faculty in the College of Physical and Mathematical Sciences, 8 May 2001.

Paper presentation, "Multifaceted exploitation of metadata for attribute match discovery in information integration," at the International Workshop on Information Integration on the Web (WIIW'01), Rio de Janeiro, Brazil, 9-11 April 2001.

Colloquium talk, "Extracting and Structuring Web Data," University of Nebraska, Lincoln, Nebraska, 3 November 2000.

Panelist for the session, "Metadata: Adding Value to Content," at the 19th International Conference on Conceptual Modeling, Salt Lake City, Utah, 9-12 October 2000.

Paper presentation, "Record Location and Reconfiguration in Unstructured Multiple-Record Web Documents," at the Third International Workshop on the Web and Databases (WebDB2000), Dallas, Texas, 19 May 2000.

Plus many more appearances prior to 2000.

Publications—Books

- Object Database Development: Concepts and Principles*, Addison-Wesley, Reading Massachusetts, 1998.
- D.W. Embley, B.D. Kurtz and S.N. Woodfield, *Object-oriented Systems Analysis: A Model-driven Approach*, Japanese translation, Tokyo Denki University Press, Tokyo, 1998.
- D.W. Embley, B.D. Kurtz and S.N. Woodfield, *Object-oriented Systems Analysis: A Model-driven Approach*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1992.

Publications—Recent Articles

- L. Xu and D.W. Embley, Categorization of Web Documents using Extraction Ontologies, *International Journal of Metadata, Semantics and Ontologies*, Vol. 3, No. 1, 2008, 3–20.
- D.W. Embley, S.W. Liddle, D. Lonsdale, G. Nagy, Y. Tijerino, R. Clawson, J. Crabtree, Y. Ding, P. Jha, Z. Lian, S. Lynn, R.K. Padmanabhan, J. Peters, C. Tao, R. Watts, C. Woodbury, and A. Zitzelberger, A Conceptual-Model-Based Computational Alembic for a Web of Knowledge, *Proceedings of the 27th International Conference on Conceptual Modeling (ER2008)* Barcelona, Catalonia, Spain, 20–23 October 2008, 532–533.
- D.W. Lonsdale, C. Tustison, C. Parker, and D.W. Embley, Assessing clinical trial eligibility with logic expression queries, *Data & Knowledge Engineering*, Vol. 66, Nr. 1, July 2008, 3–17.
- R. Al-Kamha, D.W. Embley, and S.W. Liddle, Foundational Data Modeling and Schema Transformations for XML Data Engineering, *2nd International United Information Systems Conference*, (UNISCON'08), Klagenfurt, Austria, April 2008, Pages 25–36. (invited paper)
- J. Biskup, D.W. Embley, and J.-H. Lochner, Reducing Inference Control to Access Control for Normalized Database Schemas, *Information Processing Letters*, Vol. 106, Nr. 1, March 2008, 8–12.
- S.W. Liddle and D.W. Embley, A Common Core for Active Conceptual Modeling for Learning from Surprises, in *Active Conceptual Modeling of Learning: Next Generation Learning-Base System Development*, Lecture Notes in Computer Science, Vol. 4512, Pages 47–56, Springer Verlag, Berlin, 2008.
- C. Tao and D.W. Embley, Seed-based Generation of Personalized Bio-Ontologies for Information Extraction, *Proceedings of the First International Workshop on Conceptual Modelling for Life Sciences Applications (CMLSA '07)*, Auckland, New Zealand, 5–9 November 2007, 74–84.
- C. Tao and D.W. Embley, Automatic Hidden-Web Table Interpretation by Sibling Page Comparison, *Proceedings of the 26th International Conference on Conceptual Modeling (ER2007)*, Auckland, New Zealand, 5–9 November 2007, 566–581.
- R. Al-Kamha, D.W. Embley, and S.W. Liddle, Augmenting Traditional Conceptual Models to Accommodate XML Structural Constructs, *Proceedings of the 26th International Conference on Conceptual Modeling (ER2007)*, Auckland, New Zealand, 5–9 November 2007, 518–533.

- Y. Ding, D.W. Embley, and S.W. Liddle, Enriching OWL with Instance Recognition Semantics for Automated Semantic Annotation, *Proceedings of the 1st International Workshop on Ontologies and Information Systems for the Semantic Web (ONISW 2007)*, Auckland, New Zealand, 5–9 November 2007, 160–169.
- M.J. Al-Muhammed, D.W. Embley, S.W. Liddle, and Y. Tijerino, Bringing Web Principles to Services: Ontology-Based Web Services, *Proceedings of the Fourth International Workshop on Semantic Web for Services and Processes (SWSP'07)*, Salt Lake City, Utah, 9 July 2007, 73–80.
- Y. Ding, D. Lonsdale, D.W. Embley, M. Hepp, and L. Xu, Generating Ontologies via Language Components and Ontology Reuse, *Proceedings of the 12th International Conference on Applications of Natural Language to Information Systems (NLDB'07)*, Paris, France, 27–29 June 2007, 131–142.
- M. Al-Muhammed and D.W. Embley, Ontology-Based Constraint Recognition for Free-Form Service Requests, *Proceedings of the 23rd International Conference on Data Engineering (ICDE 2007)*, Istanbul, Turkey, 16–20 April 2007, 366–375.
- L. Xu, D.W. Embley, and Y. Ding, Combining declarative and procedural knowledge to automate and represent ontology mapping, *1st International Workshop on Semantic Web Applications: Theory and Practice (SWAT 2006)* in *Advances in Conceptual Modeling – Theory and Practice (LNCS 4231)*, Tucson, Arizona, 6–9 November 2006, 395–404.
- L. Xu and D.W. Embley, A composite approach to automating direct and indirect schema mappings, *Information Systems*, Vol. 31, Nr. 8, December 2006, 697–732.
- Y. Ding, D.W. Embley, and S.W. Liddle, Automatic creation and simplified querying of semantic web content: An approach based on information-extraction ontologies, *Proceedings of the 1st Asian Semantic Web Conference (ASWC 2006)*, Beijing, China, 3–7 September 2006, 400–414.
- W.Y. Mok and D.W. Embley, Generating compact redundancy-free XML documents from conceptual-model hypergraphs, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 18, Nr. 8, August 2006, 1082–1096.
- D.W. Embley, M. Hurst, D. Lopresti, and G. Nagy, Table processing paradigms: A research survey, *International Journal on Document Analysis and Recognition*, Vol. 8, Nrs. 2–3, June 2006, 66–86.
- M.J. Al-Muhammed and D.W. Embley, Resolving underconstrained and overconstrained systems of conjunctive constraints for service requests, *Proceedings of the 18th International Conference on Advanced Information Systems Engineering (CAiSE'06, LNCS 4001)*, Luxembourg City, Luxembourg, 5–9 June 2006, 223–238, (in press).
- D.W. Lonsdale, C. Tustison, C.G. Parker, and D.W. Embley, Formulating Queries for Assessing Clinical Trial Eligibility, *Proceedings of the 11th International Conference on Applications of Natural Languages to Information Systems (NLDB 2006)*, Klagenfurt, Austria, 31 May – 2 June 2006, (in press).
- L. Xu and D.W. Embley, A composite approach to automating direct and indirect schema mappings, *Information Systems*, (in press).

- D.W. Embley, D. Lopresti, and G. Nagy, Notes on contemporary table recognition, *Proceedings of the Seventh International Association for Pattern Recognition Workshop on Document Analysis Systems, (DAS 2006, LNCS 3872)*, Nelson, New Zealand, February 2006, 164-175.
- Y.A. Tijerino, D.W. Embley, D.W. Lonsdale, Y. Ding and G. Nagy, Toward ontology generation from tables, *World Wide Web: Internet and Web Information Systems*, Volume 8, Number 3, September 2005, 261-285.
- M. Al-Muhammed, D.W. Embley, and S.W. Liddle, Conceptual model based semantic web services, *Proceedings of the Twenty Fourth International Conference on Conceptual Modeling (ER'05)*, Klagenfurt, Austria, 26-28 October 2005, 288-303.
- R. Al-Kamha, D.W. Embley, and S.W. Liddle, Representing generalization/specialization in XML schema, *Proceedings of the Workshop on Enterprise Modeling and Information Systems Architectures (EMISA'05)*, Klagenfurt, Austria, 24-25 October 2005, 250-263.
- D.W. Embley, C. Tao and S.W. Liddle, Automating the extraction of data from HTML tables with unknown structure, *Data & Knowledge Engineering*, Volume 54, Number 1, July 2005, 3-28.
- A. Wessman, S.W. Liddle and D.W. Embley, A Generalized Framework for an Ontology-Based Data-Extraction System, *Proceedings of the 4th International Conference on Information Systems Technology and its Applications (ISTA'05)*, Palmerston North, New Zealand, 23-25 May 2005, 239-253.
- D.W. Embley, L. Xu and Y. Ding, Automatic direct and indirect schema mapping: Experiences and lessons learned, *SIGMOD Record*, Volume 33, Number 4, December 2004, 14-19.
- Y.A. Tijerino, M. Al-Muhammed and D.W. Embley, Toward a flexible human-agent collaboration framework with mediating domain ontologies for the semantic web, *Proceedings of the ISWC-04 Workshop on Meaning Coordination and Negotiation* Hiroshima, Japan, 8 November 2004, 131-142.
- R. Al-Kamha and D.W. Embley, Grouping search-engine returned citations for person-name queries, *Proceedings of the ACM Sixth International Workshop on Web Information and Data Management (WIDM 2004)*, Washington, DC, 12-13 November 2004, 96-103.
- D.W. Embley, S.W. Liddle and R. Al-Kamha, Enterprise modeling with conceptual XML, *Proceedings of the 23rd International Conference on Conceptual Modeling (ER'04)*, Shanghai, China, 8-12 November 2004, 150-165.
- X. Chen, D.W. Embley and S.W. Liddle, Query rewriting for extracting data behind HTML forms, *Proceedings of the International Workshop on Conceptual Model-directed Web Information Integration and Mining*, Shanghai, China, 8-12 November 2004, 335-345.
- T. Walker and D.W. Embley, Automatic location and separation of records: A case study in the genealogical domain, *Proceedings of the International Workshop on Conceptual Model-directed Web Information Integration and Mining*, Shanghai, China, 8-12 November 2004, 302-313.
- D.W. Embley and L. Xu, Combining the best of global-as-view and local-as-view for data integration, *Proceedings of the 3rd International Conference on Information Systems Technology and its Applications*, Salt Lake City, Utah, 15-17 July 2004, 123-136.

- D.W. Embley, Toward semantic understanding—an approach based on information extraction ontologies, *Proceedings of the Fifteenth Australasian Database Conference* Dunedin, New Zealand, 18-22 January 2004, 3-12.
- Y.A. Tijerino, D.W. Embley, D.W. Lonsdale and G. Nagy, Ontology generation from tables, *Proceedings of the 4th International Conference on Web Information Systems Engineering*, Rome, Italy, 10-12 December 2003, 242-249.
- L. Xu and D.W. Embley, Using domain ontologies to discover direct and indirect matches for schema elements, *Workshop on Semantic Integration*, 20 October 2003, Sanibel Island, Florida, 105-110.
- D.W. Embley, D. Jackman and L. Xu, Attribute match discovery in information integration: exploiting multiple facets of metadata, *Journal of the Brazilian Computing Society*, Volume 8, Number 2, November 2002, 32-43.
- D.W. Embley and W.Y. Mok, On guaranteeing “good” properties for XML, invited session on Dependency Theory on the Web at *Proceedings of the 7th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2003)*, Volume IX, Orlando, Florida, 27-30 July 2003, 195-198.
- S.W. Liddle, K.A. Hewett and D.W. Embley, An integrated ontology development environment for data extraction, *Proceedings of the 2nd International Conference on Information Systems Technology and its Applications (ISTA2003)*, *Lecture Notes in Informatics*, Vol. P-30, Kharkhiv, The Ukraine, 19-21 June 2003, 21-33.
- L. Xu and D.W. Embley, Discovering direct and indirect matches for schema elements, *Proceedings of the 8th International Conference on Database Systems for Advanced Applications (DASFAA 2003)*, Kyoto, Japan, 26-28 March 2003, 39-46.
- J. Biskup and D.W. Embley, Extracting information from heterogeneous information sources using ontologically specified target views, *Information Systems*, Volume 28, Number 3, 2003, 169-212.
- K.M. Tubbs and D.W. Embley, Recognizing records from the extracted cells of microfilm tables, *Proceedings of the Symposium on Document Engineering (DocEng 2002)*, McLean, Virginia, 8-9 November 2002, 149-156.
- D.W. Embley, S.W. Liddle, D.T. Scott, and S.H. Yau, Extracting data behind web forms, *Proceedings of the Workshop on Conceptual Modeling Approaches for e-Business: A Web Service Perspective (eCOMO'2002)*, Tampere, Finland, 11 October 2002, 38-49.
- D.W. Embley, C. Tao and S.W. Liddle, Automatically extracting ontologically specified data from HTML tables with unknown structure, *Proceedings of the 21st International Conference on Conceptual Modeling (ER'02)*, Tampere, Finland, 7-11 October 2002, 322-327.
- I.M.E. Filha, A.S. da Silva, A.H.F. Laender and D.W. Embley, Representing and querying semistructured web data using nested tables with structural variants, *Proceedings of the 21st International Conference on Conceptual Modeling (ER'02)*, Tampere, Finland, 7-11 October 2002, 135-151.

- D. Lonsdale, Y. Ding, D.W. Embley and A. Melby, Peppering knowledge sources with SALT: Boosting conceptual content for ontology generation, *Proceedings of the AAAI Workshop: Semantic Web Meets Language Resources* Edmonton, Alberta, Canada, 28 July 2002, 30-36.
- I.M.E. Filha, A.S. da Silva, A.H.F. Laender and D.W. Embley, Using nested tables for representing and querying semistructured web data, *Proceedings of the Fourteenth International Conference on Advanced Information Systems Engineering (CAiSE 2002)*, and also *Lecture Notes in Computer Science #2348*, A.B. Pidduck, J. Mylopoulos, C.C. Woo, and M.T. Ozsu (editors), Toronto, Ontario, Canada, 27-31 May 2002, 719-723.
- D.W. Embley and W.Y. Mok, Developing XML documents with guaranteed “good” properties, *Proceedings of the 20th International Conference on Conceptual Modeling (ER’01)*, Yokohama, Japan, 27-30 November 2001, 426-441.
- D.W. Embley and Y.-K. Ng and L. Xu, Recognizing ontology-applicable multiple-record web documents, *Proceedings of the 20th International Conference on Conceptual Modeling (ER’01)*, Yokohama, Japan, 27-30 November 2001, 555-570.
- D.W. Embley, S.W. Liddle and S.H. Yau, On the automatic extraction of data from the hidden web, *Proceedings of the International Workshop on Data Semantics in Web Information Systems (DASWIS-2001)*, Yokohama, Japan, 27-30 November 2001, 106-119, also printed as *Conceptual Modeling for New Information Systems Technologies (LNCS 2465)*, Springer Verlag, 2002, 212-226.
- D.W. Embley, D. Jackman and L. Xu, Multifaceted exploitation of metadata for attribute match discovery in information integration, *Proceedings of the International Workshop on Information Integration on the Web (WIIW’01)*, Rio de Janeiro, Brazil, 9-11 April 2001, 110-117.
- D.W. Embley and L. Xu, Locating and reconfiguring records in unstructured multiple-record web documents, *The World Wide Web and Databases, Lecture Notes in Computer Science*, Volume 1997, D. Suciu and G. Vossen (eds.), Springer Verlag, 2001, 256-274.
- J. Biskup and D.W. Embley, Mediated Information Gain, *International Database Engineering and Applications Symposium (IDEAS2000)*, Yokohama, Japan, 18-20 September 2000, 360-370.
- D.W. Embley and L. Xu, Record location and reconfiguration in unstructured multiple-record web documents, *Proceedings of the Third International Workshop on the Web and Databases (WebDB2000)*, Dallas, Texas, 18-19 May 2000, 123-128.

Plus over 100 publications prior to 2000.