

CURRICULUM VITAE

February 11, 2014

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

Born: October 30, 1946
Place: Salt Lake City
Married, 14 children, 38 grandchildren
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

Service Missionary, since September 2013, working with the Engineering Research Team in the Family History Department of The Church of Jesus Christ of Latter-day Saints. Current research includes automated information extraction from obituaries and from the scanned-and-OCR'd book collection. Research activities also include automated table understanding and conceptualizing relational databases for the semantic web.

Professor, Brigham Young University, September 1986 – August 2013 (granted continuing status 1987). Taught graduate and undergraduate classes in database systems and theory, discrete mathematics, conceptual modeling, information extraction and integration, and the semantic web. Research interests include database systems, query languages and systems, model-driven software development, information integration, and information extraction from semistructured and unstructured documents.

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.

Grants

NSF: TANGO: Table Analysis for Semiautomatic Generation of Ontologies (with G. Nagy, D.W. Lonsdale, Y. Tijerino), 15 August 2005 - 31 July 2008, \$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 7 other external grants, plus many internal grants.

Patents and Copyrights

Patent: “Form-based Ontology Creation and Information Harvesting” (TISP and TISP++), granted 24 January 2012, Patent Number 8,103,962.

Patent Application: FOCIH (Forms-based Ontology Creation and Information Harvesting), December 2008.

Copyright Claim: WoK (Web of Knowledge), November 2008.

Patent Application: TISP (Table Interpretation by Sibling Pages), November 2008.

Recent Service—Professional Involvement

Area Editor for the *Encyclopedia of Database Systems*

Editorial Board for *Data & Knowledge Engineering* (2005-2013).

Editorial Board for *Journal on Data Semantics* (2002-2013).

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 and Chair-elect 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 (1997-2013)

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

Plus service in various capacities for many conferences/workshops.

Honors and Awards

College of Physical and Mathematical Sciences Distinguished Citizenship Award, 2012.

Peter P. Chen Award, 2009.

Elected as an ER Fellow, 2009.

Karl G. Maeser Research and Creative Arts Award at BYU, 2006.

Wesley P. Lloyd Award for Distinction in Graduate Education at BYU, 2001.

Recent Conference, Seminar, Colloquia, and Workshop Appearances

Swan-song colloquium talk, “A Conceptual Modeling Approach to Harvesting, Organizing, and Analyzing Scanned & OCRed Family History Documents,” Department of Computer Science, Brigham Young University, 21 November 2013.

Keynote speech, “Big Data—Conceptual Modeling to the Rescue,” at the 32nd International Conference on Conceptual Modeling (ER 2013), Hong Kong, 11–13 November 2013.

Seminar talk, “Automated Reasoning on Conceptual Schemas in FamilySearch—a Large-Scale Reasoning Application,” at Dagstuhl Seminar: Automated Reasoning on Conceptual Schemas, Dagstuhl Schloss, Germany, 20 May, 2013.

Participated in an invitation-only Dagstuhl Seminar, “Automated Reasoning on Conceptual Schemas,” Dagstuhl Castle, Germany, 19-24 May 2013.

Invited presentation, “Cross-Language Hybrid Keyword and Semantic Search” at the February LDS Church NLP Luncheon, Brigham Young University, Provo, Utah, 21 February 2013.

Paper presentation, “Cross-Language Hybrid Keyword and Semantic Search” at the 31st International Conference on Conceptual Modeling (ER 2012), Florence, Italy, 15–18 October 2012.

Invited presentation, “A Web of Knowledge for Historical Documents” at the June LDS Church NLP Luncheon, Salt Lake City, Utah, 21 June 2012.

Invited presentation, “A Web of Knowledge for Historical Documents” at RootsTech 2012, 2 February 2012.

Paper presentation, “Principled Pragmatism: A Guide to the Adaption of Ideas from Philosophical Disciplines to Conceptual Modeling” at the ER2011 Workshop on Ontologies and Conceptual Modeling (Onto-CoM’11), Brussels, Belgium, 2 November 2011.

Paper presentation, “Enabling Search for Facts and Implied Facts in Historical Documents” at the International Workshop on Historical Document Imaging and Processing (HIP’11), Beijing, China, 16–17 September 2011.

Keynote presentation, “The Wizard Words and Worlds of ACM-L” at the Third International Workshop on Active Conceptual Modeling of Learning (ACM-L 2010), Vancouver, British Columbia, Canada, 3 November 2010.

Paper presentation, “Theoretical Foundations for Enabling a Web of Knowledge,” at the Sixth International Symposium on Foundations of Information and Knowledge Systems (FoIKS’10), Sophia, Bulgaria, 15–19 February 2010.

Paper presentation, “KBB: A Knowledge-Bundle Builder for Research Studies” at the 2nd International Workshop on Active Conceptual Modeling of Learning (ACM-L 2009), Gramado, Brazil, 9–12 November 2009.

Paper presentation, “Semantically Conceptualizing and Annotating Tables,” at the 3rd Asian Semantic Web Conference (ASWC08), Bangkok, Thailand, 2–5 February 2009.

Plus many more appearances prior to 2009.

Publications—Books

D.W. Embley, *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—Book Chapters and Supplements

D.W. Embley, S.W. Liddle, D.W. Lonsdale, and Y. Tijerino, Multilingual Extraction Ontologies, *Towards the Multilingual Semantic Web*, P. Buitellar and Philipp Cimiano (Eds.), Springer, 2014 (to appear).

S.W. Clyde, D.W. Embley, S.W. Liddle, and S.N. Woodfield, OSM-Logic: A Fact-Oriented, Time-Dependent Formalization of Object-oriented Systems Modeling, in *Conceptual Modelling and Its Theoretical Foundations* A. Dösterhöft, M. Klettke, and K.-D. Schewe (Eds.), LNCS 7260, Springer, Heidelberg, 151–172, 2012.

D.W. Embley, S.W. Liddle, and O. Pastor, Conceptual-Model Programming: A Manifesto, Chapter 1 in *The Handbook of Conceptual Modeling: Theory, Practice, and Research Challenges*, Springer, 2011, 3-16.

D.W. Embley and W.Y. Mok, Mapping Conceptual Models to Database Schemas, Chapter 5 in *The Handbook of Conceptual Modeling: Theory, Practice, and Research Challenges*, Springer, 2011, 123-163.

D.W. Embley, S.W. Liddle, and D.W. Lonsdale, Conceptual Modeling Foundations for a Web of Knowledge, Chapter 15 in *The Handbook of Conceptual Modeling: Theory, Practice, and Research Challenges*, Springer, 2011, 477-516.

D.W. Embley, S.W. Liddle, and C. Tao, A Web of Knowledge: A Conceptual-Modeling Perspective, in *The Evolution of Conceptual Modeling: From a Historical Perspective towards the Future of Conceptual Modeling*, Lecture Notes in Computer Science, R. Kaschek and L. Delcambre (eds.), Volume 6520, Springer Verlag, 2011, 137-160.

Plus several more prior to 2011.

Publications—Books Edited

D.W. Embley and Bernhard Thalheim (Eds.), *The Handbook of Conceptual Modeling: Theory, Practice, and Research Challenges*, Springer Verlag, Heidelberg, Germany, 2011.

D.W. Embley, A. Olive, and S. Ram (Eds.), *Conceptual Modeling – ER 2006: Proceedings of the 25th International Conference on Conceptual Modeling (LNCS 4215)*, 6-9 November, Tucson, Arizona, Springer Verlag, Berlin, Germany, 2006.

P.P. Chen, D.W. Embley, J. Kouloumdjian, S.W. Liddle, and J.F. Roddick, (Eds.), *Advances in Conceptual Modeling: Proceedings of the ER'99 Workshops (LNCS 1727)*, 15-18 November, Paris, France, Springer Verlag, Heidelberg, Germany, 1999.

D.W. Embley and Robert C. Goldstein, (Eds.), *Conceptual Modeling – ER'97: Proceedings of the 16th International Conference on Conceptual Modeling (LNCS 1331)*, 3-6 November, Los Angeles, California, Springer Verlag, Heidelberg, Germany, 1997.

Recent Publications—Articles

- G. Nagy, S. Seth, and D.W. Embley, End-to-End Conversion of HTML Tables for Populating a Relational Database, *Proceedings of the 11th IAPR Workshop on Document Analysis Systems*, Tours, France, 7–10 April 2014, (to appear).
- W.Y. Mok, J. Fong, and D.W. Embley, Generating the Fewest Redundancy-Free XML Scheme Trees from Acyclic Conceptual-Model Hypergraphs in Polynomial Time, *Information Systems*, Volume 41, May 2014, 20–44.
- D.W. Embley and S.W. Liddle, Big Data—Conceptual Modeling to the Rescue, *Proceedings of the 32nd International Conference on Conceptual Modeling (ER 2013)*, Hong Kong, China, 11–13 November, 2013, 1–8.
- T.L. Packer and D.W. Embley, Cost Effective Ontology Population with Data from Lists in OCRed Historical Documents, *Proceedings of the 2nd International Workshop on Historical Document Imaging and Processing (HIP 2013)*, Washington D.C., 24 August, 2013, 1–8.
- G. Nagy, D.W. Embley, and D. Lopresti, Boxy, Semi-structured Document Elements, *Proceedings of the Tenth IAPR International Workshop on Graphics Recognition (GREC 2013)*, Lehigh University, Bethlehem, Pennsylvania, 20–21 August, 2013.
- D.W. Embley, S.W. Liddle, D.W. Lonsdale, J.S. Park, B-J. Shin, A.J. Zitzelberger, Cross-Language Hybrid Keyword and Semantic Search, *Proceedings of the 31st International Conference on Conceptual Modeling (ER 2012)*, Florence, Italy, 15 – 18 October, 2012, 190-203.
- D.W. Embley, S.W. Liddle, and D.W. Lonsdale, Principled Pragmatism: A Guide to the Adaptation of Ideas from Philosophical Disciplines to Conceptual Modeling, *Proceedings of the ER 2011 Workshop on Ontologies and Conceptual Modeling (Onto-CoM 2011)*, Brussels, Belgium, 2 November, 2011, 183-192.
- D.W. Embley, S.W. Liddle, D.W. Lonsdale, and Y. Tijerino, Multilingual Ontologies for Cross-Language Information Extraction and Semantic Search, *Proceedings of the 30th International Conference on Conceptual Modeling (ER 2011)*, Brussels, Belgium, 31 October – 3 November, 2011, 147-160.
- D.W. Embley, S.W. Liddle, D.W. Lonsdale, S. Machado, T. Packer, J. Park, N. Tate, and A. Zitzelberger, Enabling Search for Facts and Implied Facts in Historical Documents, *Proceedings of the International Workshop on Historical Document Imaging and Processing (HIP 2011)*, Beijing, China, 16–17 September, 2011, 59–66.
- G. Nagy, S. Seth, D.W. Embley, M. Krishnamoorthy, D. Jin, and S. Machado, Data Extraction from Web Tables: the Devil is in the Details, *Proceedings of the 11th International Conference on Document Analysis and Recognition (ICDAR 2011)*, Beijing, China, 18–21 September 2011, 242-246.
- D.W. Embley, M. Krishnamoorthy, G. Nagy, and S. Seth, Factoring Web Tables, *Proceedings of The Twenty Fourth International Conference on Industrial Engineering & Other Applications of Applied Intelligent Systems*, Syracuse, New York, 28 June – 1 July 2011, 253–263.
- T.L. Packer, J.F. Lutes, A.P. Stewart, D.W. Embley, E.K. Ringger, K.D. Seppi, and L.S. Jensen, Extracting Person Names from Diverse and Noisy OCR Text, *Proceedings of the Fourth*

Workshop on Analytics for Noisy Unstructured Text Data Documents (AND 2010), Toronto, Ontario, Canada, 26 October 2010, 19–26.

- D.W. Embley, S.W. Liddle, D.W. Lonsdale, A.P. Stewart, and C. Tao, KBB: A Knowledge-Bundle Builder for Research Studies, *Advances in Conceptual Modeling—Applications and Challenges: Proceedings of the ACM-L 2010 Workshop*, Vancouver, British Columbia, Canada, 1–4 November 2010, 148–157.
- W.Y. Mok, J. Fong, and D.W. Embley, Extracting a Largest Redundancy-Free XML Storage Structure from an Acyclic Hypergraph in Polynomial Time, *Information Systems*, Volume 35, Number 7, November 2010, 804–824.
- D.W. Lonsdale, D.W. Embley, and S.W. Liddle, Ontologies for Multilingual Extraction, *Proceedings of the World Wide Web Conference’s 1st International Workshop on MultiLingual Ontologies (MSW 2010)*, P. Buitelaar, P. Cimiano, and E. Montiel-Ponsoda (Eds.), CEUR Workshop Proceedings, Volume 571, pages 1-4, on line at: <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-571/>, Raleigh, North Carolina, 26–27 April 2010.
- D.W. Lonsdale, D.W. Embley, Y. Ding, L. Xu, and M. Hepp, Reusing Ontologies and Language Components for Ontology Generation, *Data & Knowledge Engineering*, Volume 69, Number 4, April 2010, 318–330.
- D.W. Embley and A. Zitzelberger, Theoretical Foundations for Enabling a Web of Knowledge, *Proceedings of the Sixth International Symposium on Foundations of Information and Knowledge Systems (FoIKS10)*, Sophia, Bulgaria, 15–19 February 2010, 211–229.
- D. Embley, Semantic Data Model, definitional entry in the *Encyclopedia of Database Systems*, L. Liu and M. Tamer Özsu (eds.), Volume 4, Springer Verlag, 2009, Pages 2559–2561.
- D. Embley, Relational Data Model, regular entry in the *Encyclopedia of Database Systems*, L. Liu and M. Tamer Özsu (eds.), Volume 4, Springer Verlag, 2009, Pages 2372–2376.
- D. Embley, Key, definitional entry in the *Encyclopedia of Database Systems*, L. Liu and M. Tamer Özsu (eds.), Volume 3, Springer Verlag, 2009, Pages 1587–1588.
- C. Tao, D.W. Embley, and S.W. Liddle, FOCIH: Form-based Ontology Creation and Information Harvesting, *Proceedings of the 28th International Conference on Conceptual Modeling (ER’09)*, Gramado, Brazil, 9–12 November 2009, 346–359.
- C. Tao and D.W. Embley, Automatic Hidden-Web Table Interpretation, Conceptualization, and Semantic Annotation, *Data & Knowledge Engineering*, Vol. 68, Nr. 7, July 2009, 683–703.
- Y. Ding, L. Xu, and D.W. Embley, A Model of World Wide Web Evolution, *Proceedings of the Web Science Conference 2009 (WebSci’09)*, Athens, Greece, 18–20 March 2009, <http://www.websci09.org/proceedings/>, (6 pages).
- S. Lynn and D.W. Embley, Semantically Conceptualizing and Annotating Tables, *Proceedings of the Third Asian Semantic Web Conference (ASWC 2008)*, Bangkok, Thailand, 2–5 February 2009, 345–359.

Plus more than 150 prior to 2009.