

EDUARD CONSTANTIN DRAGUT

506 Ott Rd.
Bala Cynwyd, PA - 19004

(215)-204-0521
edragut@temple.edu
cis.temple.edu/~edragut

RESEARCH INTERESTS

Information retrieval, management of (un)structured data, information extraction, information integration, data quality, opinion mining and retrieval, Web-oriented research, cyber-infrastructure for scientific research

EDUCATION

- Ph.D., University of Illinois at Chicago, Computer Science, July 2010, GPA: 4.00.
 - Dissertation: **A Framework for Transparently Accessing Deep Web Sources**
 - Advisor: Professor Clement Yu; co-Advisor: Professor Prasad Sistla
- M.S., University of Iowa, 2004, GPA: 4.00.
- Bachelor of Science, University of Bucharest, Computer Science, June 1998

FUNDING

- “BIGDATA: Collaborative Research: F: Streaming Architecture for Continuous Entity Linking in Social Media”. NSF. Jan. 1, 2016 – Dec. 31, 2018. \$1,083,339 (Temple University \$783,339 and SUNY at Binghamton \$300,000) (Lead PI, with Co-PI Yuhong Guo (Temple) and PI Weiyi Meng (Binghamton)).

HONORS

Dean's Scholar Award 2007/2008 – “the most distinguished award that UIC offers to graduate students”

EMPLOYMENT

- | | |
|--|-----------------------|
| ▪ Assistant Professor, Temple University | July 2013 - present |
| ▪ Visiting Professor, Purdue University | Jan. 2013 – May 2013 |
| ▪ Post Doctorate, Purdue University | July 2010 – June 2013 |

JOURNAL PUBLICATIONS

1. J. Yuan, L. He, E. Dragut, W. Meng, C. Yu. **Result Merging for Structured Queries on the Deep Web with Active Relevance Weight Estimation**. Information Systems. 64: 93-103 (2017).
2. E. Dragut, H. Wang, C. Yu, S. Prasad, W. Meng. **Polarity Consistency Checking for Domain Independent Sentiment Dictionaries**. In the IEEE Transactions on Knowledge and Data Engineering (TKDE), vol.27, no.3, pp.838-851, March 1 2015 [acceptance rate: unknow].
3. E. Dragut, B. DasGupta, B. Beirne, A. Neyestani, B. Atassi, C. Yu, and W. Meng. **Merging Query Results from Local Search Engines for Geo-referenced Objects**. In the ACM Transactions on the Web (TWEB), Volume 8 Issue 4, October 2014, Article No. 20. 29

pages. 2014 [acceptance rate: unknown].

4. A. Kruger, R. Lawrence, and E. Dragut, **Building a Terabyte NEXRAD Radar Database for Hydrometeorology Research**, Journal of Computers & Geosciences, Vol. 32, Issue 2, 2006, pages 247-258. [acceptance rate: unknown]

PEER-REVIEWED CONFERENCE PROCEEDINGS

5. A. Wang, J. Croft, E. Dragut. **Reflections on Data Integration for SDN**. ACM International Workshop on Security in Software Defined Networks & Network Function Virtualization (SDN-NFV Security'17). Scottsdale, US, March 2017. [acceptance rate: unknown]
6. J. Jupin, J. Shi, E. Dragut. **PSH: A Probabilistic Signature Hash Method with Hash Neighborhood Candidate Generation for Fast Edit-Distance String Comparison on Big Data**. In Proceedings of the IEEE Big Data, Washington. DC, USA, Dec. 2016. [acceptance rate: 18.7%].
7. El Kindi Rezig, E. Dragut, M. Ouzzani, A. Elmagarmid and W. Aref. **ORLF: A Flexible Framework for Online Record Linkage and Fusion**. In Proceedings of the 32st International Conference on Data Engineering (ICDE'16).
8. Q. Liu, E. Dragut., A. Mukherjee, W. Meng. **FLORIN – A System to Support (Near) Real-Time Applications on User Generated Content on Daily News**. In Proceedings of the 41st International Conference on Very Large Data Bases (VLDB'15), Hawaii, Volume 8 Issue 12, August 2015. 1944-1955. [acceptance rate: 33%].
9. A. Schneider, E. Dragut. **Towards Debugging Sentiment Lexicons**. In Proceedings of the 53th Annual Meeting of the Association for Computational Linguistics (ACL'15), Beijing, China, July 2015. 1024-1034. [acceptance rate: 25%]
10. E. Rezig, E. Dragut, M. Ouzzani, and A. Elmagarmid. **Query-Time Record Linkage and Fusion over Web Databases**. In Proceedings of the 31st International Conference on Data Engineering (ICDE'15), Seoul, Korea, April 2015. 42 – 53. [acceptance rate: 20%].
11. D. Salt, M. Ouzzani, E. Dragut, P. Baker, S. Rangarajan. **iHUB - An Information and Collaborative Management Platform for Life Sciences**. In Proceedings of the 23rd International World Wide Web Conference (WWW'14). Seoul, Korea. April 2014. [acceptance rate: 30%]
12. E. Dragut, P. Baker, J. Xu, M. Sarfraz, E. Bertino, A. Madkour, R. Agarwal, A. Mahmood and S. Han. **CRISS - Computational Research Infrastructure for Science**. In Proceedings of 14th IEEE International Conference on Information Reuse and Integration, (IRI'13), San Francisco, USA. August 2013 (invited paper).
13. L. Cen, E. Dragut and L. Si. **Author Disambiguation by Hierarchical Agglomerative Clustering with Adaptive Stopping Criterion**. In Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval (SIGIR'13). Dublin, Ireland. July 2013, 741-744 (short paper) [acceptance rate: 35%].
14. E. Dragut, B. P. Beirne, A. Neyestani, B. Atassi, C. Yu, W. Meng, B. DasGupta, **YumiInt - A Deep Web Integrating System for Local Search Engines for Geo-referenced Objects**, In Proceedings of the 29th IEEE International Conference on Data Engineering (ICDE'13), Brisbane, Australia. April 2013 (demo paper) [acceptance rate: around 30%].
15. E. Dragut, W. Hong, C. Yu, P. Sistla, W. Meng, **Polarity Consistency Checking for**

- Sentiment Dictionaries**, In Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (ACL'12), Jeju, South Korea, July 2012 [acceptance rate: 19%].
16. E. Dragut, M. Ouzzani, A. Madkour, N. Mohamed, P. Baker, D. Salt. **Ionomics Atlas - A Tool To Explore Interconnected Ionomic, Genomic and Environmental Data**. In Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM '12), Maui, USA, November 2012 (demo paper) [acceptance rate: 40%]
 17. E. Dragut, C. Yu, P. Sistla, W. Meng, **Construction of a Sentiment Word Dictionary**, In Proceedings of the 19th International Conference on Information and Knowledge Management (CIKM'10), Toronto, Canada, October 2010 (short paper) [acceptance rate: 31.3%]
 18. T. Kabisch, E. Dragut, C. Yu, U. Leser, **Deep Web Integration with VisQI**, In Proceedings of the 36th International Conference on Very Large Data Bases (VLDB'10), Singapore, Sept 2010 (demo paper) [acceptance rate: 36%].
 19. E. Dragut, F. Fang, P. Sistla, C. Yu, W. Meng, **Stop Word and Related Problems in Web Interface Integration**. In Proceedings the of 35th International Conference on Very Large Data Bases (VLDB'09), Pages 349-360, Lyon, France, August 2009 [acceptance rate: 17.9%]
 20. E. Dragut, T. Kabisch, C. Yu, U. Leser, **A Hierarchical Approach to Model Web Query Interfaces for Web Source Integration**. In Proceedings of the 35th International Conference on Very Large Data Bases (VLDB'09), Pages 325-336, Lyon, France, August 2009 [acceptance rate: 17.9%]
 21. E. Dragut, F. Fang, C. Yu, W. Meng, **Deriving Customized Integrated Web Query Interfaces**. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI'09). Milan, Italy, September 2009 [acceptance rate: 34%]
 22. E. Dragut, C. Yu, W. Meng, **Meaningful Labeling of Integrated Query Interfaces**. In Proceedings of the 32nd International Conference on Very Large Data Bases (VLDB'06), Pages 679-690, Seoul, Korea, September 2006 [acceptance rate: 13.2%]
 23. E. Dragut, W. Wu, C. Yu, P. Sistla, W. Meng, **Merging Source Query Interfaces on Web Databases**. In Proceedings of the 22nd International Conference on Data Engineering (ICDE'06), Pages 679-690, Atlanta, USA, April 2006 [acceptance rate: 19%]
 24. E. Dragut and R. Lawrence, **Reducing the Cost of Validating Mapping Compositions by Exploiting Semantic Relationships**. In Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics (ODBASE'06), Pages 882-890, Montpellier, France, October 2006 [34.1%]
 25. E. Dragut and R. Lawrence, **Composing Mappings between Schemas using a Reference Ontology**. In Proceedings of the International Conference on Ontologies, Databases and Applications of Semantics (ODBASE'04), Pages 783-800, Larnaca, Cyprus, 2004 [acceptance rate: 25%]

BOOKS

26. E. Dragut, W. Meng and C. Yu, **Deep Web Query Interface Understanding and Integration**, Synthesis Lectures on Data Management, Morgan & Claypool Publishers, 2012.

WORKSHOPS

27. S. Wu, J. Male, E. Dragut. **Spatial-Temporal Campus Crime Pattern Mining from Historical Alert Messages**. The 3rd National Workshop for REU Research in Networking and Systems (REUNS'17). Silicon Valley USA. January, 2017.
28. E. Dragut and C. Fellbaum. **The Role of Adverbs in Sentiment Analysis**. Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929-2014). Association for Computational Linguistics. Baltimore, MD, USA. June 2014 [acceptance rate: unknown]
29. A. R. Mahmood, W. G. Aref, E. C. Dragut, S. Basalamah. **The Palm-tree Index: Indexing with the crowd**. 1st VLDB Workshop on Databases and Crowdsourcing (DBCrowd'13). Trento, Italy. August 2013.

OTHER PUBLICATIONS

30. X. L. Dong and E. Dragut. **The 10th International Workshop on Quality in Databases – QDB 2012** (workshop report). ACM SIGMOD Record. Volume 41 Issue 4, December 2012 Pages 55-59.

PRESENTATIONS

- “Entity Matching via Word Embedding,” Invited Talk at NSF BDSpoke Initiative Workshop at Purdue, West Lafayette, March 29, 2017.
- “Towards a Formal Understanding of Sentiment Annotation in Sentiment Lexicons,” Invited Talk: CLUNCH (Computational Linguistics and Lunch) - University of Pennsylvania. Philadelphia, USA. November 11, 2015.
- “FLORIN – A System to Support (Near) Real-Time Applications on User Generated Content on Daily News.” demo presentation, at the 41st International Conference on Very Large Data Bases (VLDB'15), September, 2015.
- “Merging Query Results from Local Search Engines for Geo-referenced Objects”. Invited Talk. Computer & Information Science - University of Pennsylvania. Philadelphia, USA. August 6, 2014.
- “Merging Query Results from Local Search Engines for Geo-referenced Objects”. Invited Talk. Orbitz. Chicago, USA. April 2, 2014.
- “Ionomics Atlas - A Tool to Explore Interconnected Ionomic, Genomic and Environmental Data”, demo presentation, at the 21st ACM International Conference on Information and Knowledge Management (CIKM '12), Maui, USA, November 2012.
- “Polarity Consistency Checking for Sentiment Dictionaries”, poster presentation, at the 50th Annual Meeting of the Association for Computational Linguistics (ACL'12), Jeju, South Korea, July 2012
- “Construction of a Sentimental Word Dictionary”, poster presentation, at the 19th International Conference on Information and Knowledge Management (CIKM'10), Toronto, Canada, October 2010
- “A Framework for Transparently Accessing Deep Web Sources”, Purdue University, Cyber Center Seminars, March 11th, 2010
- “Stop Word and Related Problems in Web Interface Integration”, paper presentation at the 35th International Conference on Very Large Data Bases (VLDB'09). Lyon, France, August 2009

- “Meaningful Labeling of Integrated Query Interfaces”, paper presentation at the 32nd International Conference on Very Large Data Bases (VLDB’06). Seoul, Korea, September 2006
- “Merging Source Query Interfaces on Web Databases”, paper presentation at the 22nd International Conference on Data Engineering (ICDE’06). Atlanta, USA, April 2006
- “Metasearch Engines for e-commerce”, poster presentation at the Industrial Advisory Board Meeting, Chicago, October 2005
- “Merging Source Query Interfaces on the Deep Web”, poster presentation, Second Midwest Database Research Symposium, Chicago, April 16, 2005.

STUDENTS SUPERVISED

Ph.D. Students

Student	Status	Role	Year	Outcome
Andrew Schneider	In progress	Thesis Supervisor	3 rd	1 ACL full publication.
Lihong He	In progress	Thesis Supervisor	2 nd	1 IS Journal
Sergei Iakhnitskii	Start Aug. 2016	Thesis Supervisor		

M.S. Students

Student	Status	Role	Semester	Outcome
Radheshyam Karnani	Terminal semester	Project/ Independent Study Supervisor	Fall 2016, Spr. 2017	Author Disambiguation
Yagni Patel	Terminal semester	Project/ Independent Study Supervisor	Fall 2016, Spr. 2017	Develop a system that automatically collects data from 25 universities about campus crimes.
Vasisht Pabathi	Graduated	Project Supervisor	Spr. 2016	Develop crawlers to collect large datasets for experiments.
Qingyuan Liu	Graduated	Independent Study Supervisor	Spr. 2015	1 VLDB demo publication.
Qingyuan Liu	Graduated	Project Supervisor	Fall 2014	
Malavkumar Patel	Graduated	Independent Study Supervisor	Fall 2014	Dataset Search Engine.

Undergraduate Students

Student	Institution	Role	Semester	Outcome
Aidan Shea	Temple University	Supervisor	Fall 2016 – Spr. 2017	SIGMOD SRC paper.
John Male	Temple University	Supervisor	Fall 2015 – Sum. 2016	- Workshop paper. - Special mention at Temple URP Poster presentation.

				- Prepares a demo for ACL.
Sharon Greenblatt	Temple University	Supervisor	Sum. 2016	Poster presentation.
Adam Barbato	UIUC	NSF REU Supervisor	Sum. 2015	Poster presentation.
Kendall Lane	Hampton College	NSF REU Supervisor	Sum. 2015	Poster presentation.

TEACHING METHODS RESEARCH

- **Using Algorithm Animation Technology to Demystify Algorithms.** To be delivered at the Teaching with Technology Symposium, Temple University. March 31st.

COOURSES TAUGHT

Temple University

- CIS 5590 - Topics in Computer Science: Introduction to Information Retrieval - Fall 2017.
- CIS 5516 - Principles of Data Management - Spring 2017.
- CIS 5590 - Topics in Computer Science: Introduction to Information Retrieval - Fall 2016.
- CIS 5516 Principles of Data Management - Spring 2016.
- CIS 4331/5301 Principles of Database Systems – Fall 2015.
- CIS 4362/5301 Application System Development Using Relational Technology – Spring 2015.
- CIS 5516 - Principles of Data Management – Fall 2014
- CIS2168 - Data Structures and Algorithms - Fall 2013

Purdue University

- CS448 - Introduction to Relational Database Systems - Spring 2013.

COURSE DEVELOPMENT

- **CIS 5590 - Topics in Computer Science: Introduction to Information Retrieval.** The course covers models for information retrieval, techniques for indexing and searching, query expansion, latent semantic indexing, and link analysis and ranking.
- **CIS 4517/5517 - Data-Intensive and Cloud Computing.** The course covers recently emerged and fast moving technology of big data and cloud computing. It covers a spectrum of topics from core techniques in data management and analysis to highly-scalable data processing using parallel database systems. It introduces big data ecosystems such as Hadoop, Spark, Storm and MapReduce; cloud technologies such as Amazon EC2, Microsoft Azure and Google Cloud; data management tailored to cloud and big data such as No SQL, Google Big Table/Apache HBase, and introductory applications to Big Data and cloud environment.

CURRICULUM DEVELOPMENT

- **Master of Science in Data Science.** (Part of the committee.) It provides students with a strong foundation in algorithmic, computational, and statistical thinking, as well as a firm

understanding of computer systems.

PROFESSIONAL ACTIVITIES

Program Committee (by year)

- 2017
 - SIGMOD Student Research Competition, SIGMOD, ACL, SocialNLP@ EACL.
- 2016
 - SIGMOD, WWW, ACL.
 - Chair SIGMOD/PODS PhD Symposium.
- 2015
 - SIGMOD, WWW, DASFAA, NAACL-HLT, EMNLP
- 2014
 - Chair ICDE PhD Symposium, Chicago, USA.
 - DASFAA, AICCSA, COLING, ADBIS
- 2013
 - ACL, CIKM (demo track), WWW (Data Mining track); DASFAA; AICCSA.
- 2012
 - ICDE; CIKM (session chair);
 - Co-chair (with Dr. Xin Luna Dong, AT&T) International Workshop on Quality in Databases (QDB) in conjunction with VLDB 2012 (report in SIGMOD Record);

Journal Reviewer:

Transactions on Knowledge and Data Engineering (TKDE); Transactions on the Web (TWEB); Distributed and Parallel Databases; Transactions on Services Computing (TSC); Information Systems Journal (ISJ); Transactions on Knowledge Discovery from Data (TKDD); International Journal of Information Security (IJIS); ACM Transactions on Interactive Intelligent Systems (TiiS), Computers & Geosciences.

Panelist:

- NSF SBIR/STTR 2017 (2 panels)
- NSF IIS 2016, NSF SBIR/STTR
- NSF IIS 2015 – 3 panels
- DoE 2015 – 1 panel

NON-ACADEMIC EXPERIENCE

TRONN SOFTWARE, July 2000 – November 2001

Project manager

- Managed the development of an internet e-commerce web site; supervised 4 people.
- Designed and built an online car-commerce application prototype (imports, supply, distribution and spare parts management) for PORSCHE ROMANIA

IPACRI ROMANIA SRL, September 1998 – May 2000

- **Analyst-programmer consultant** at ELSAG BANKLAB SPA in Rome, Italy. Participated

in two large banking projects for Cassa di Risparmio di Perugia; involved in analysis, development, deployment and maintenance.

SINORG ROUMANIE SA, May 1998 – September 1998

- **Programmer** - Developed applications for local public administrative areas in French provinces.