

SATOSHI TSUTSUI

Last update: April 19, 2018

E-mail: stsutsui@indiana.edu

Web: <http://homes.soic.indiana.edu/stsutsui/> Github: <https://github.com/apple2373>

EDUCATION

Indiana University Bloomington, USA

September 2015 - present

- Earned Master's degree in Data Science at May 2017.
- Ph.D. student / Major: Informatics / Minor: Computer Science / GPA: 4.0
- Advisors: Professor David Crandall (djeran@indiana.edu), Professor Ying Ding (dingying@indiana.edu)
- AI related courses I took: B551 Elements of Artificial Intelligence, B555 Machine Learning, B657 Computer Vision, E599 Machine Learning for Signal Processing

Keio University, Japan

April 2011 - March 2015

- Bachelor of Engineering in Administration Engineering / Major: Computer Science
- Graduation with highest honors / GPA: 3.96
- Thesis: Extending DBpedia with More Semantics from Wikipedia Articles
- Advisor: Professor Takahira Yamaguchi

RESEARCH EXPERIENCE

Research Intern at Peking University, China

May 2018 - July 2018

- Advisor: Liangcai Gao
- Nail disease recognition from medical images.

Research Intern at Preferred Networks, Japan

May 2017 - July 2017

- Advisor: Tommi Kerola and Shunta Saito
- Semantic Segmentation for Autonomous Driving (Paper 3, 1)

Research with Prof. David Crandall at Indiana University

September 2016 - present

- First-Person Computational Vision
- Recognizing or detecting specific types of figures (Paper 2, 6)
- Automatic generation of image descriptions (Paper 4)
- Key techniques used: deep convolutional neural networks, recurrent neural networks

Research with Prof. Ying Ding at Indiana University

September 2015 - present

- Data-driven knowledge discovery from literature on Alzheimer's Disease (Paper 8, 5)
- Exploit machine reading; Topic modeling for macro view, and open information extraction for micro view
- Extracted medical entities including genes and drugs from more than 200,000 PubMed papers

Research with Prof. Satoshi Sekine at New York University

March 2015 - July 2015

- Created a Japanese question answering dataset
- Analyzed 1,000 questions manually, and converted them to sparql queries

Research with Prof. Takahira Yamaguchi at Keio University

April 2014 - March 2015

- Information extraction from Wikipedia articles as a semi-structured data
- Implemented heuristic methods, and extracted more than 20 million triples with 80% precision

PUBLICATIONS AND PRESENTATIONS

Peer-reviewed publications

1. **Tsutsui, S.**, Kerola, T., Saito, S., Crandall, D. (June 2018) Minimizing Supervision for Free-space Segmentation. *Workshop on Autonomous Driving, In conjunction with the Conference on Computer Vision and Pattern Recognition (CVPR)*. *arXiv:1711.05998*.

2. **Tsutsui, S.**, Crandall, D. (Nov 2017) A Data Driven Approach for Compound Figure Separation Using Convolutional Neural Networks. *The IAPR International Conference on Document Analysis and Recognition (ICDAR)* (oral presentation). *arXiv:1703.05105*. Code: <https://github.com/apple2373/figure-separator>
3. **Tsutsui, S.**, Kerola, T., Saito, S. (Oct 2017) Distantly Supervised Road Segmentation. *Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving, In conjunction with International Conference on Computer Vision (ICCV)*. *arXiv:1708.06118*.
4. **Tsutsui, S.**, Crandall, D. (July 2017) Using Artificial Tokens to Control Languages for Multilingual Image Caption Generation. *Language and Vision Workshop, In conjunction with Conference on Computer Vision and Pattern Recognition (CVPR)*. *arXiv:1706.06275*. Code: <https://github.com/apple2373/chainer-caption>
5. **Tsutsui, S.**, Meng, G., Ding, Y. (Mar 2017). Public Machine Reading System for Alzheimers Disease Literature (poster). *iConference*.
6. **Tsutsui, S.**, Meng, G., Yao, X., Crandall, D., Ding, Y. (Mar 2017) Analyzing Figures of Brain Images from Alzheimer's Disease Papers (poster). *iConference*.
7. Chen, B., **Tsutsui, S.**, Ding, Y., Ma, F. (2017). Understanding the topic evolution in a scientific domain: An exploratory study for the field of information retrieval. *Journal of Informetrics*, 11(4), 1175-1189. (*Impact Factor = 2.920*)
8. **Tsutsui, S.**, Ding, Y., Meng, G. (Oct 2016)., Machine Reading Approach to Understand Alzheimer's Disease Literature, *International Workshop on Data and Text Mining in Biomedical Informatics (DTMBIO)*, In conjunction with *Conference on Information and Knowledge Management (CIKM)*. Demo: http://homes.soic.indiana.edu/stsutsui/machine_reading/

Non peer-reviewed publications and presentations

- **Tsutsui, S.** (Oct 2016). Machine Reading Approach to Understand Alzheimer's Disease Literature. *ILS Doctoral Research Forum* 3rd Best Paper Award
- **Tsutsui, S.** (Nov 2015). Data-Driven Knowledge Discovery in the Alzheimers Disease Literature. *ILS Doctoral Research Forum*
- **Tsutsui, S.**, Morita, T, Yamaguchi, T. (Mar 2015). Extending DBpedia with List Structures in Wikipedia Articles. *The 36th Semantic Web and Ontology Seminar, Japan Society of Artificial Intelligence*.

TEACHING EXPERIENCE

Teaching Assistant in Data Semantics

Fall 2016

- Assisted professor to create course materials for both online and residential class. Created videos to help students with technical troubles. Graded assignments Covered two lectures when professor is traveling.

Teaching Assistant in Social Media Mining

Spring 2016

- Assisted professor to create course materials.

Teaching Assistant in Laboratories in Administration Engineering

July 2014

- Assisted students to create web application with HTML, CSS, and JavaScript

AWARDS

- ICDAR student travel award Nov 2017
- The Scholarship for Study Abroad from Yoshida Scholarship Foundation July 2015 - May 2017
- Keio University Scholarship for Undergraduate Students 2012

SKILLS

- Language: Japanese (native) and English
- Programming: Python, Java, Matlab, SQL, JavaScript, HTML, CSS, R
- Operating systems: UNIX (Mac, CentOS, Ubuntu)
- Semantic Web Technology: Ontology (Protege), RDF (Virtuoso, Jena), SPARQL
- **Deep Learning Framework:** Tensorflow and Chainer, some codes are on my github.
- Other Tools I can use: Open CV, StanfordNLP, Gensim, NLTK, scikit-learn, Network X.