



INTELLIGENT SYSTEMS PROGRAM

2018-2019 NEWSLETTER

Inside this issue:

PROF. JANYCE WIEBE

ALUMNI HIGHLIGHTS

FACULTY HIGHLIGHTS

STUDENT HIGHLIGHTS

NEW GRADUATES

DIRECTOR'S ADDRESS



DR. DIANE LITMAN

While in 2017-2018 we celebrated ISP's 30 year past, in 2018-2019 we have been taking advantage of the continuing excitement about AI to grow its future.

With respect to faculty, four Pitt faculty members were approved for new ISP secondary appointments during 2018-2019. Although not necessarily new to Pitt, all are new to ISP. In addition, several of their primary appointments represent new Pitt affiliations for ISP (namely, the Graduate School of Public and International Affairs (GSPIA) and the Department of Computational and Systems Biology (School of Medicine).

With respect to graduate students, in Fall 2018 we welcomed one of our largest entering graduate student classes ever (1 MS and 5 PhD students). With another strong applicant pool, in Fall 2019 we will exceed this record and welcome even a larger group of PhD students to ISP!

New graduate student profiles are currently being featured on Facebook, while interviews with newly appointed faculty are later in this newsletter.

ISP also continues to chart its future as part of Pitt's new School of Computing and Information (SCI). Faculty recruiting is a top priority and many exciting academic initiatives are in the works. In addition, the development known as One Bigelow will not only be the future home for SCI, but is a major component of the Pitt Campus Master Plan (<https://www.campusplan.pitt.edu/draft-campus-plan>). A current vision statement has been drafted by SCI's Dean Cohen for the purposes of community engagement (<https://pitt.app.box.com/s/z15xadfsway3ve6jvf6z0gr5vzp159g3>).

Effective 9/1/2019, I will be turning over the Co-Directorship role to Vanathi Gopalakrishnan (Biomedical Informatics) to complete our 2018-2020 elected term. While it has been a privilege to have served as either ISP Director or Co-Director since 2010, I am delighted that Vanathi will be bringing new leadership into the program. Please join me in welcoming Vanathi to her inaugural term!



CONTENT

1	PROF. JANYCE WIEBE	1
2	ALUMNI HIGHLIGHTS	3
3	FACULTY HIGHLIGHTS	
	New book	6
	Honors and Awards	9
	A Selection of New Faculty Grants	11
	ISP People at Conferences	13
	Meet New Co-Director	15
	Meet New Faculty	20
4	STUDENT HIGHLIGHTS	
	Provost Fellowships	35
5	DISSERTATION DEFENSES	
	Ye Ye	38
	Gaurav Trivedi	40
	Zahra Rahimi	42
6	NEW GRADUATES	
	Gaurav Trivedi	44
	Ye Ye	47

The Intelligent Systems Program (ISP) is a multidisciplinary graduate program at the University of Pittsburgh dedicated to applied artificial intelligence (AI).

CONTACT US

Intelligent Systems Program

www.isp.pitt.edu
6129 Sennott Square
210 South Bouquet Street
Pittsburgh, PA 15260

Huihui Xu
Editor



Thanks for Professor Rada Mihalcea's invaluable contribution

<https://naacl2019.org/blog/in-memory-of-jan-wiebe/>

Remembering Jan Wiebe

"Janyce Wiebe - known by her friends and colleagues as "Jan" - was a Professor of Computer Science, the former director of the Intelligent Systems Program at the University of Pittsburgh, and a fellow of the Association for Computational Linguistics. She was an expert in the areas of opinion analysis, discourse processing, pragmatics, and word-sense disambiguation. She was one of the first to carry out research on methodology in text annotation, just in time



for the rapid rise in the need for text annotation for supervised learning methods. She was a pioneer in the research area of "subjectivity analysis" - recognizing and interpreting expressions of opinions and sentiments in text to support NLP applications such as question answering, information extraction, text categorization, and summarization."

Honors and Contributions

"Jan had a long and successful career. She was involved in many professional communities. These roles included ACL Program Co-Chair, NAACL Program Chair, NAACL Executive Board member, Transactions of the ACL Action Editor, Computational Linguistics and Language Resources and Evaluation Editorial Board member, AAAI Workshop Co-Chair, ACM Special Interest Group on Artificial Intelligence (SIGART) Vice-Chair, and ACM-SIGART/AAAI Doctoral Consortium Chair. In 2015, she was named a fellow of the ACL 'for seminal contributions to Subjectivity and Sentiment analysis, Discourse Processing, and Lexical Semantics.'"

Contributions to ISP

"In 2000, Jan moved to become an Associate Professor, and later Professor at the University of Pittsburgh. Jan took on a major Intelligent Systems Program (ISP) administrative and leadership role, serving as co-Director from 2002-2004, Director from 2004-2010, and co-Director from 2010-2016."

She advised several ISP students in the past few years:

Lingia Deng (2016), now a research scientist in Bloomberg;

Cem Akkaya (2013), now a senior research scientist in Yahoo;

Theresa Wilson (2008), now a Visiting Assistant Professor of Computer Science at Hanover College.

Memory from Family Members

"Jan was an influencer for the people who were surrounded by her. For her colleagues, she was a great friend and inspirational role model. For her students, her knowledge and passion empowered the students. For her family, she was warm and supportive."

"She firmly believed in what she did and invested a great deal of passion in it. Not to mention she has set on a path of her own brand of innovation in academia in the computer science field."

Video Link

Memorial Fund

The School of Computing and Information Department of Computer Science and the Intelligent Systems Program have announced that jointly, they are contributing \$10,000 to the Dr. Jan Wiebe Memorial Fund! **(Donation Link)** Dr. Wiebe's legacy will continue to help students of SCI pursue their education. ■

ALUMNI HIGHLIGHTS



Dr. Min Chi (2009) has been promoted to the rank of Associate Professor with Tenure at the Department of Computer Science, North Carolina State University.

Dr. Cristina Conati (1999) has been elected to the Executive Committee of AAAI (Association for the Advancement of AI, Sept. 2018), and was awarded the status of AAAI Senior member (January 2019). She was awarded a grant to run a workshop on Trust in AI systems, as part of the AI & Society initiative sponsored by Canadian Institute for Advanced Research (CIFAR), UK Research and Innovation (UKRI) and France's Centre National de la Recherche Scientifique (CNRS).



Dr. Ilya Goldin (2011) has started a new role as Principal Data Scientist, Phenom People, using AI to help a billion people find their dream job. The work includes NLP, deep learning, recommender systems, and other AI methods. The data science team is growing; graduating PhD students and alumni are welcome to send inquiries to Ilya. Phenom People is headquartered in Ambler, PA, and has additional offices in Hyderabad, India and Rehovot, Israel.

ALUMNI HIGHLIGHTS



Dr. Yun Huang (2018) Postdoc in the Human-Computer Interaction Institute at Carnegie Mellon University. She is also a visiting researcher in Institute of Informatics at the Austral University of Chile.

Dr. Rosta Farzan (2009) Associate Professor at the University of Pittsburgh.

1. Received funding from National Science Foundation to work on a project titled Tangible Privacy: User-Centric Sensor Designs for Assured Privacy , in collaboration with Dr. Adam Lee from Computer Science Department and Dr. Apu Kapadia from Indiana University, Computer Science Department. The project explores tangible designs for sensor devices such as mobile phones or IoT devices (Echo, NestCam) to empower user control of their privacy as they interact with these devices.



2. Received funding from AT&T and Pitt Cyber Institute to research ways in fighting cyberbullying and cyber-hate problems on social media among younger adults. The project involves developing a curriculum along with a support platform to train middle school students how to deal with cyberbullying. The funding also supports development of an educational game to promote empathy among middle school students.

ALUMNI HIGHLIGHTS

Congratulations ISP alumni and ISP faculty for being nominated for the 201 AMIA Doctoral Dissertation Award

Congratulations **ISP alumna Dr. Ye Ye**, and **ISP faculty Dr. Fuchiang (Rich) Tsui** for being nominated for the AMIA Doctoral Dissertation Award. The AMIA Doctoral Dissertation Award nominates and offers high-value and prestigious recognition for the doctoral dissertations that contributes to the science of informatics. Ye's dissertation, "Transfer Learning for Bayesian Case Detection Systems", has attained the fifth place and been highlighted and honored on the AMIA web pages: <https://www.amia.org/amia-doctoral-dissertation-award>.

Congratulations ISP alumni and faculty on receiving the ACL Test-of-Time award (10 years)

Congratulations to **PhD alumna Theresa Wilson**, **MS alumnus Paul Hoffmann**, and family and friends of **Prof. Jan Wiebe** for receiving the recent ACL Test-of-Time award (10 years) for their paper: "**Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis**".

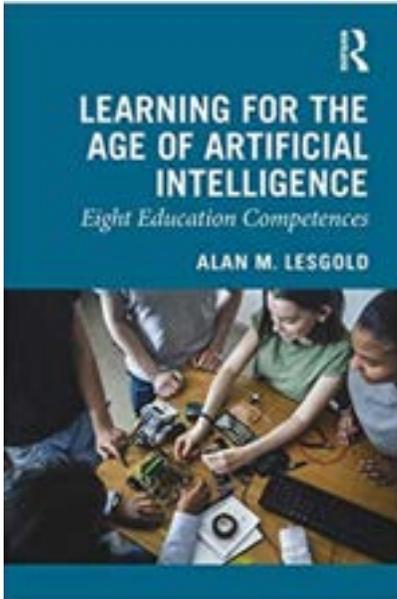
The ACL Test-of-Time Award (10 years) recognizes the authors of an influential paper or series of papers published between ten years ago that has significantly impacted research or applications of NLP field.

ISP publication featured in the Year-In-Review session at AMIA

Congratulations to **ISP alumnus Gaurav Trivedi**, **Dr. Hwa**, **Dr. Wiebe** and **Dr. Hochheiser** for their NLPReViz Paper was featured in the Year-In-Review session at AMIA this year. ■

[Read more](#)

LEARNING FOR THE AGE OF ARTIFICIAL INTELLIGENCE Eight Education Competences by Alan M. Lesgold



Cover photo from Amazon

About the Author

Alan Lesgold is professor and former Ren e and Richard Goldman Dean (2000-2016) of the School of Education at the University of Pittsburgh and also is professor of psychology and intelligent systems. He received his Ph.D. in psychology from Stanford University in 1971 and also holds an honorary



doctorate from the Open University of the Netherlands. He is a fellow of the American Psychological Association (APA), in experimental, applied, and educational psychology, and also of the Association for Psychological Science and the American Educational Research Association. In 2001, he received the APA award for distinguished contributions of applications of psychology to education and training. In 1995, he was awarded the Educom Medal. He was president of the Applied Cognitive Psychology division of the International Association for Applied Psychology 2002-2006. Lesgold is a Lifetime National Associate of the National Research Council (National Academies).

Book Review

The new book addresses the need for change in the education system and helps students to develop skills for the Age of Artificial Intelligence. He draws a picture of how the ubiquity of computers reduces the human routine operations and the urgency of calling the attention of educators and policymakers to this issue.

Computers are getting more competent in jobs that were traditionally held by human beings. White collar jobs, such as high-end jobs at banks are among the many where machines are replacing humans. We are looking towards an unpredictable future even as new jobs are being invented. How do we prepare students for the unpredictable future? Bringing about changes in the education system by teaching students core competencies that allow them to survive in the age of smart machines is perhaps the first course of action. Lesgold lists eight core competences in his book: “the ability to learn efficiently and quickly; socio-emotional skills; skills of civic participation; ability to evaluate information; facility in collaborative activity, including the 4Cs (dealing with complexity, communication, collaboration, and creativity); management of personal finances and some basic economics; confidence; and physical and mental fitness.”

Lesgold uses holes in a cheese slice as a metaphor for education failures where these valuable competencies are not instilled in the students thoroughly. Since the current community education system cannot single-handedly support learning of the eight competences, out-of-school opportunities and home influences, in addition to traditional schooling, would help students achieve learning success. This so called “redundancy” strategy will prevent students from falling through one hole fatally, because other forms of teaching will block the failure at home from previous teaching experience.

He points out that not only does the education system need to change, but also the educators have to learn how to teach the eight competences. He mentioned that teachers “may need to master some of the eight competences themselves before they can help their students acquire them”. Moreover, when parents,

Book Review

political leaders, and business leaders believe deeply that the eight competences are necessary, that is embedding the eight competences into the culture is a crucial decision, “third places”- other than home and the place where one is engaged much of the day- will be a great complement to the traditional schooling. 158pp. Routledge. 47.95. ■

By Huihui u

FACULTY HIGHLIGHTS

Honors and Awards

Dr. Diane Litman and **Dr. Yuru Lin** got a mention in the **list of women amazing researchers** in computational social science. Congratulations to ISP SCI. It's great news that we have two women selected on the list from our school!

[Read more](#)



Dr. Yuru Lin



Dr. Diane Litman



Congratulations to three ISP faculty members for their Pitt Innovator awards from the University of Pittsburgh Innovation Institute - **Professors Gregory F. Cooper (1), Vanathi Gopalakrishnan (2), and Xinghua Lu ()**.

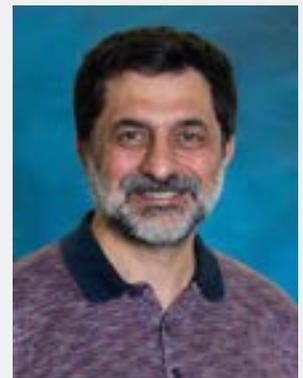
The Pitt Innovator Awards are given to faculty or students who submit invention disclosures to the University via the Innovation Institute, that have resulted in licensing or optioning deals. These deals provide funding to teams for developing prototypes that can subsequently be validated and commercialized through entrepreneurship.

Congratulations Prof. Peter Brusilovsky for receiving **AMiner's Most Influential Scholar Award!** His contributions on the subject of recommender systems have been recognized as some of the most-cited research in the field.

[Read more](#)

Congratulations being elected **chair of ACM Sigweb**, a special interest group on hypertext and the web!

[Read more](#)



Dr. Peter Brusilovsky

Congratulations to Dr. Shandong Wu

Sarah S. Aboutalib, Aly A. Mohamed, Wendie A. Berg, Margarita L. Zuley, Jules H. Sumkin and Shandong Wu, Deep Learning to Distinguish Recalled but Benign Mammography Images in Breast Cancer Screening Clinical Cancer Research, Oct 2018. <http://clincancerres.aacrjournals.org/content/early/2018/09/26/1078-0432.CCR-18-1115>

AACR newsroom:

<https://www.aacr.org/Newsroom/Pages/News-Release-Detail.aspx?ItemID 1231>

ther media coverage:

<https://www.auntminnie.com/index.aspxsec ser&sub def&pag dis&ItemID 122130>

<https://www.eurekaalert.org/pub releases/2018-10/aafc-dlc100918.php>

<https://www.news-medical.net/news/20181011/AI-approach-could-identify-nuanced-imaging-features-specific-for-recalled-benign-mammograms.aspx>

<https://www.healthimaging.com/topics/artificial-intelligence/ai-distinguish-between-false-positive-malignant-negative-mammograms> ■

A Selection of New Faculty Grants

Congratulations to ISP director Dr. Diane Litman for research grant from Institute of Education Sciences

ISP director Dr. Diane Litman has been awarded a research grant from the Institute of Education Sciences to study undergraduate STEM education.

With colleagues Dr. Muhsin Menekse and Dr. Ala Samarapungavan of Purdue University, Dr. Litman will pursue research titled “Enhancing Undergraduate STEM Education by Integrating Mobile Learning Technologies with Natural Language Processing.”

Congratulations ISP faculties and ISP student on their Pitt Cyber Accelerator Grants (PCAG)

The PCAG grants provide initial funding for novel and innovative projects on the critical questions of networks, data, and algorithms, with a focus on the ever-changing gaps among law, policy, and technology.

- Congratulations to Dr. Kevin Ashley and ISP alumnus Dr. Matthias Grabmair for the research project “Technical Support for Course on Applied Legal Data Analytics and AI.”
- Congratulations to Dr. Kevin Ashley and ISP student Jaromir Savelka for their project Annotating Cases for Learning to Summarize.
- Congratulations to Dr. Rosta Farzan and her colleagues' project Fighting Cyberbullying: A Transformative and Educational Game for Promoting Empathic Understanding.

FACULTY HIGHLIGHTS

Congratulations ISP faculty Dr. Rosta Farzan received an NSF grant

Congratulations ! ISP faculty Dr. Rosta Farzan received an NSF grant for the project: SaTC: CORE: Small: Collaborative: Tangible Privacy: User-Centric Sensor Designs for Assured Privacy along with SCI faculty Dr. Adam Lee.

[https://www.nsf.gov/awardsearch/showAward?AWD ID 1814866](https://www.nsf.gov/awardsearch/showAward?AWD_ID_1814866)

Congratulations to ISP Faculty Adriana Kovashka on being selected for an Amazon Research Award

Her proposal Functional objects: How objects foreshadow film plots and explain advertisements has been chosen to receive funding.

<https://ara.amazon-ml.com/recipients/> 2018 ■

ISP People at Conferences

The 20th International Conference on Artificial Intelligence in Education (AIED 2019)



From left to right: Peter Brusilovsky (current faculty), Bruce McLaren (PhD alum), Barbara DiEugenio (LRDC alum), Vincent Alevan (PhD alum), Diane Litman (current faculty)

FACULTY HIGHLIGHTS

ACM Federated Computing Research Conference (FCRC 2019)



Leftmost: Sera Linardi (current Faculty)

FACULTY HIGHLIGHTS

Meet New Co-Director



Vanathi Gopalakrishnan

Professional Title:

Associate Professor

Department:

Department of Biomedical Informatics

Website:

Vanathi Gopalakrishnan's Web site

Research Interests:

Bioinformatics

Q: Will ISP be facing some changes in the following years? What are they?

Professor Vanathi Gopalakrishnan (VG): ISP is a premier program and has been around for 30 years plus. My main duty is to run everything smoothly as a co-director. We do have plans to bring in more faculty members from applied AI disciplines across the campus, so the program will continue to grow. We have already had a few new faculty members join the ISP this past year, and two more who have applied for consideration this upcoming Fall. Relevant topics include Business Intelligence and I am hoping to establish a connection with the business school which can give our students more options, while also enhancing innovation for that field of research. Regarding changes to the curriculum, our students can expect to see some Biomedical Informatics track changes with respect to required courses. We encourage students to communicate directly with us to express their needs - this can definitely

help maintain and improve the quality of the program and student experiences herein.

Q: Are there any opportunities and obstacles that you could foresee for ISP?

VG: Yes. A main challenge graduate students face is that of acquiring funding through research assistantships. It has always been an obstacle for continued research. A possible solution, which has worked in the past, is bringing in more faculty members, who are interested in mentoring our students and have ongoing research funding available. We hope that our selective and bright ISP students will have enough interesting projects to learn from when doing lab rotations, while making progress towards their coursework and dissertation.

We also have a lot of opportunities. Some faculty members who recently joined the ISP, have startup packages and funded research projects that are in need of students. Being a part of the newly established School of Computing and Information (SCI) is a great opportunity, as ISP can contribute in many ways to building the culture, while benefitting from the fellowships available for new students joining our program. There is a nice

FACULTY HIGHLIGHTS

opportunity to attend the East Coast Industry Forum (ECIF) in October, which is focused on AI. This event will be held on CMU campus, and will feature ISP faculty on panels. There is also an interest in promoting industry relationships across Germany and America. Since many CMU AI folks will be present, we would have opportunities to network with them. Students can also present demos of their research during this event. The more you do show and tell, the more likely it is that you will get useful feedback to refine your work, by learning from different perspectives. Besides, ISP will have opportunities to form national and international collaborations.

Q: ISP is an interdisciplinary program, and your research requires interdisciplinary collaboration. What do you think about bridging your research experience to your new role?

VG: Nowadays, the AI systems are facing several challenges: explainability, adaptability and context-awareness. Explainable AI is an emerging field that tries to address how the decision of AI systems are made; adaptable AI is the ability to adapt to the ever-changing environment; context-sensitive AI is a way to describe the ability of AI systems to

solve problems based on their knowledge of the situation.

It is insufficient for us to just generate and use models with simple classifications and predictions that cannot explain the why or how it arrived at these conclusions or decisions. As a result, I think that knowledge-intensive and data-driven AI together will be important, and the way in which humans interact with AI and computing will change rapidly due to these needs, with evolving technological modalities and solutions.

I've had the vision of a Center for Informed Biomedical Analytics Research (CIBAR) for a very long time. It may not exist right now, but this is a plan. As you know, data is the new gold'. I'm lucky enough that I've been looking at biomedical data all these years and bring a unique perspective on how to do biomedical data analytics. This doesn't mean our research is limited to the biomedical field. In the ISP, you can obtain data from other domains as well. The key aspect of CIBAR is the word "informed". What does "informed" mean? "Informed" means "knowledge-based." My PRoBE laboratory for **P**attern **R**ecognition from **B**iomedical **E**vidence has been looking at different modes of data: text, ontology, and images. If we can put all of the information

FACULTY HIGHLIGHTS

together as prior knowledge, then this is the “informed” part. How do you use prior knowledge and combine it with data when learning models? This is the focus. CIBAR is a realistic goal because ISP faculty have expertise in developing and using text analysis, knowledge-based systems, computer vision, imaging and data visualization.

Another area of focus for the ISP under my co-directorship with Prof. Litman, could be for us to try to obtain training grants to increase student slots within selected sub-areas of AI, such as NLP, wherein we need adequately trained members from our next generation of students.

Q: What do you expect from current and future ISP students?

VG: First and foremost, to become an outstanding researcher scientist in AI and get international recognition through respectable publications in their subfields of inquiry. Then, to give back to their communities through service that utilizes their unique training to make a positive difference in the lives of others.

Q: What do you think the role of ISP playing in the AI era?

VG: ISP has been a role model in the applied AI field, as an established program that has evolved over three decades, and the future is bright. ISP is internationally famous for AI in Education, and I hope the ISP can make a major mark in AI and Medicine in the future as well. ISP is unique and is probably the first program of its kind to include faculty from so many diverse disciplines united in the need for a common AI curriculum, and that is why we received the German-American Chamber of Commerce invitation to participate in the ECIF this year. It is our responsibility to carry on the legacy that we have received from our program founders, Profs. Alan Lesgold and Rich Thomason. We want to achieve visibility and national level recognition in terms of being able to demonstrate many diverse contributions. These can have an even greater impact if we can focus on creating value through building stronger collaborations that enable AI systems with a deeper understanding of specific domains of knowledge. This is more easily achievable by strong interdisciplinary programs such as ours, that have the necessary infrastructure already built-in.

FACULTY HIGHLIGHTS

Q: For those ISP students who are willing to work in the industry, what makes ISP graduates unique in the job market from your point of view?

VG: The uniqueness of ISP is that you can do foundational research from an AI point of view, and also contribute heavily to the domain. ISP is interdisciplinary by nature, and ISP graduates are excellently trained in fundamental AI methods as well as their application to one or more real-world problems, which makes them rather attractive for intensive research jobs in both academia and industry.

Q: Do you have any advice for those ISP students who are stepping into the job market?

VG: One thing I want our students to learn about is that we have an Innovation Institute at the University of Pittsburgh (UPitt). This institute can help students with an entrepreneurial spirit along the path toward creating their own startups, if they file their intellectual property (IP) along with their faculty mentor(s), using online forms. Many ISP faculty, including myself, have experience with this process and can help guide students. There are a lot of interesting things happening in the AI industry. Keeping industry in mind is

always a good idea. As we know, some companies have explored certain domains in-depth, and their experience or products could help advance our research goals without reinventing the wheel.

Some students choose to go to industry, and others decide to go to academia, based on their background and interests. I feel that pursuing AI in academia is good at this time, since there is government backing and also, several positions opening up in Universities across the nation. We would like to see our students well situated in tenure-track positions. Working towards a tenured position in academia needs at least seven years of hard work, but it is worth it. When you get tenure, you can be sure that your work has been recognized. Freedom is another advantage of working in academia since you are allowed to pick up a project and take full control of it. This kind of freedom is hard to find in the industry. Of course, staying in academia is never an easy job learning to figure out what you know and what you do not know is key here. Then, you need to decide which portions of what you do not know is important in your line of research, and who can help fill this gap in knowledge and skills. After which, significant collaborations and research contributions can arise. For example, I was

FACULTY HIGHLIGHTS

trained by Prof. Bruce Buchanan in symbolic/rule-based AI, and as a faculty, I was mentored by Prof. Gregory Cooper in probabilistic AI, and we were able to synthesize a hybrid method combining the strengths of Bayesian networks with rule-based interpretability.

Q: How do you spend your spare time?

VG: Professors generally read a lot, and therefore, I found complementary activities, viz. classical Indian dancing and biking as hobbies that have stayed with me since my youth. I think dancing is my natural talent. Biking keeps you so in touch with nature, and I especially love hearing the birds chirping. It also gives me a clear appreciation for a quote that is attributed to Albert Einstein, “Life is like riding a bicycle. To keep your balance, you must keep moving.”

I would also recommend that our students have hobbies outside of research. It is so important to spend a little bit of time growing good social and professional networks, with people who can be your support backbone when research gets tough or you need more creativity!

Anecdote:

Here is an anecdote of how I came into the bioinformatics field. My husband, Dr. Ganesh Mani, and I first met at AAI 1990 on MIT campus when I was still a first-year graduate student in UPitt CS, with Prof. Bruce Buchanan as my advisor. Ganesh was studying at the University of Wisconsin(UW)-Madison back then, with another famous CS pioneer, Prof. Leonard Uhr. I visited UW-Madison for a summer because of Ganesh, and Prof. Jude Shavlik gave me an office with Mark Craven as my office mate. Prof. Craven currently serves as the Director of the Center for Predictive Computational Phenotyping, one of the NIH's Centers of Excellence for Big Data Computing. All the people that I met there were generous about their experience and knowledge. What I learned from them led me to launch

my first course at the University of Pittsburgh Introduction to Bioinformatics, which was co-taught by me and Dr. Paul Hodor, who was then a postdoctoral fellow at the Center for Biomedical Informatics. This course was the first formal course in Bioinformatics at UPitt, and it launched my career here in the Bioinformatics field over the past two decades as faculty. The serendipitous events at AAI-90 led to the creation of my current family and career who would have guessed?

FACULTY HIGHLIGHTS

Meet New Faculty

(Alphabetize by Last Name)



Panayiotis (Takis) Benos

Professional Title:

Professor and Vice Chair

Department:

Department of Computational and Systems biology

Website:

Panayiotis (Takis) Benos's Web site

Research Interests:

Causal modeling algorithms

Systems Biology

Probabilistic Graphical Models

Machine Learning applications in Healthcare

David Ryan Koes

Professional Title:

Assistant Professor

Department:

Department of Computational and Systems Biology

Website:

David Ryan Koes's Web site

Research Interests:

Computational Drug Discovery

Deep learning

Discrete Algorithms



FACULTY HIGHLIGHTS

Meet New Faculty



Sera Linardi

Professional Title:

Associate Professor

Department:

GSPIA

Website:

Sera Linardi's Web site

Research Interests:

Prosocial Behavior

Econ Social Work

Econ CS

Yalini Senathirajah

Professional Title:

Associate Professor

Department:

Department of Biomedical Informatics

Website:

Yalini Senathirajah's Web site

Research Interests:

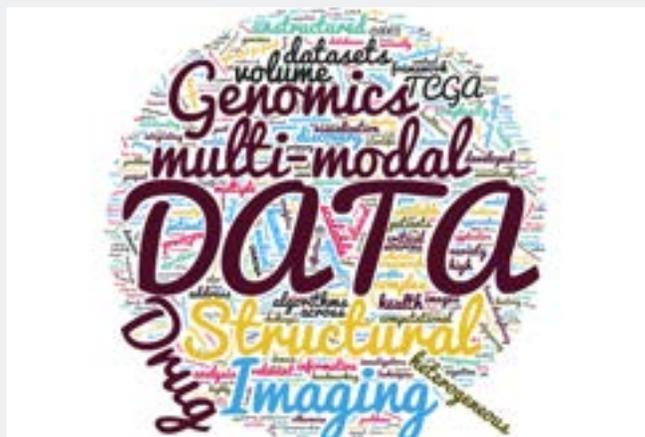
Bioinformatics

EHR



Getting What You Came for

An Interview with Professor David Ryan Koes



Huihui Xu (HX): What's the reason for joining ISP?

Professor David Koes (DK): The reason for joining ISP is to get a larger pool of potential graduate students with computational ability.

HX: How would ML methods impact the academic field?

DK: Machine learning makes a lot of sense when we have a lot of data, and there is a growing amount of biological data. We are particularly interested in applying deep learning to protein structures in the form of 3D images so that we can leverage the success of image recognition approaches in

structure based drug design. Deep learning is moving quickly with ground-breaking papers constantly being published. For example, there are so many different types of GANs, generative adversarial models. It is difficult to keep on top of everything, but it is also very exciting, because there is potential to build something even better on top all this ground-breaking work.

One exciting direction of machine learning research right now is generative modeling, which is using neural networks to generate something that doesn't exist. If we could do this for drugs, this would be revolutionary for drug discovery. If it is feasible, we might just need to click a button and let the computer to create drug candidates. This research is still quite preliminary, but I'm very excited about it.

Another avenue of investigation is how to combine deep scientific knowledge and deep learning. For example, we know how atoms work in the physical

FACULTY HIGHLIGHTS

world though quantum mechanics. However, it is not feasible to apply quantum chemical calculations to biological systems. Machine learning methods can learn molecular properties from a training set, but they have difficulty generalizing to new domains.

Can we combine these two types of knowledge to get the best of both worlds?

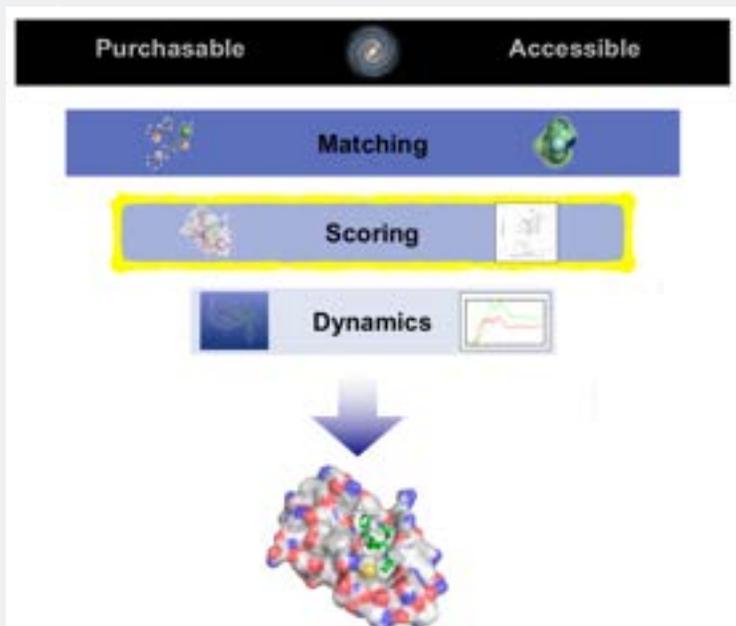
One challenge with generative modeling drug discovery research is validating predictions. In order to validate a drug candidate, we need a chemist to synthesize and test it which is expensive and time-consuming. Most of publications attempt to compare the similarity of the candidate to existing drugs.

HX: How could ML methods impact the industry in the future?

DK: The most expensive part of drug discovery is the failure: things don't work. The pharmaceutical companies spend a lot of time and effort on clinical trials. But only 10 percent of clinical candidates actually become FDA approved drugs. Generating the right drug at the start of the process would dramatically reduce the cost of drug development.

HX: I read that you are an advocate of open science and software. How do you define open science? Why do you think that it is important?

DK: Releasing your data is an example of open science. Publishing data,



software, and protocols let other people to reproduce it. Our lab releases all our software under open source licenses. It is important to remember that publishing your work is contributing to science as whole, not just adding a line to your CV.

HX: Scalable machine learning for big data biology sounds very appealing. What is the motivation for designing this course? What makes this course uni ue? How is the feedback from students?

FACULTY HIGHLIGHTS

DK: We teach this course every year and the students have provided positive feedback. This is an interesting course, and maybe a little ambitious. We teach machine learning on big data sets, and also do a little bit distributed computing at the same time. We use Google Cloud and Pyspark for this course.

One thing I've been thinking about is instead of going in depth with one approach, each assignment has a different approach to scalable computation. For example, you would use Pyspark in one assignment and NVIDIA RAPIDS for another.

HX: Do you have any new courses and publications you wish ISP students to know?

DK: No new courses other than the Scalable machine learning for big data biology course. Our CNN scoring for protein-ligand interactions paper is still our most exciting paper.

HX: How many graduate students are you advising now?

DK: Three PhD students.

HX: Are you accepting new students at the moment? If yes, do you have any messages for the ISP students who wish to join your research group?

DK: Yes. I am planning on recruiting a student next semester, but it also depends on the funding.

Willing to learn new things is important for graduate school in general. I am a computer scientist, and my thesis is about compiler optimization. I feel like you could learn biology, chemistry and physics. If you come with a biology, chemistry or physics background, you can learn computer science too. But you need to have some subset of those four things.

HX: Do you have any advice for the students who are in the job market?

DK: Do internships. And academic conferences are good places to meet people. In my field, there are a lot of industry people at conferences. Attending conferences is a good way to talk to people and meet potential employers.

FACULTY HIGHLIGHTS

HX: Is there anything you wish you knew or did when you were a PhD student?

DK: I really enjoyed being a PhD student. For some students who have doubts, there is a book called *Getting What You Came For*. I feel like the title is the most important part: knowing what you want and working for it is your PhD journey.

HX: How do you spend your spare time? hobbies etc.

DK: I like to bake and bring cookies and brownies to group meetings. I have two kids and enjoy spending time with my family. ■

Helping Vulnerable Population

An Interview with Professor Sera Linardi



Source: <https://thepositivecontrarian.com/wp-content/uploads/pophealth.png>

HuiHui Xu (HX): What is the motivation for joining ISP?

Professor Sera Linardi (SL): I did my PhD in Social Science at Caltech where I was part of the former Yahoo! Social and Information Science Lab (SISL). We all come from various engineering, computational and social science disciplines but were able to use quantitative and mathematical methods to make both theoretical and practical advances, and so I have seen how fruitful this interdisciplinary collaboration can be and am starting the Center for Analytics for Social Innovation (CASI) at GSPIA. I am

interested in bringing to light the struggle of populations that are living in the margins, and in using technologies and algorithms to make existing institutions work better for them. While economics is great at big picture institutional design, computer science is much better at measuring and accounting for real life frictions. We are having an Econ-CS reading group in the Fall and a class in the spring on this exact topic. I'll say more about it later. My hope is that by putting the two together, we can quantify the difficulties of living in the margins such as getting an ID if you do not have a permanent address, or find a job when you are under SSDI income limits and think about its implication for institutional design (e.g. employment services). I have not started any of these projects yet and am looking for the right collaborators!

HX: What role is data playing in your research? What kinds of data are your labs collecting?

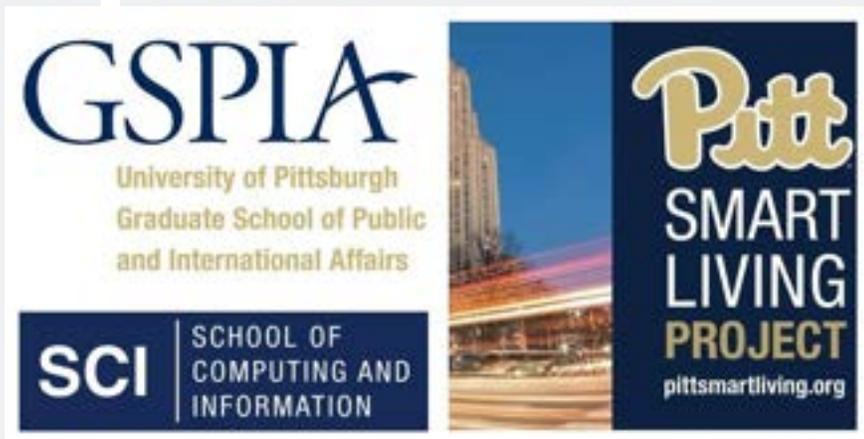
FACULTY HIGHLIGHTS

SL: I am a behavioral / experimental economist, and so a lot of our research generate primary data through laboratory and field experiments. In the past we have worked a lot on people's willingness to help others, but now we are doing more work on the design of public and social services. In one of our projects, we study the demand of ex-inmates in Pittsburgh for social services. Recidivism rates are notoriously high in the US 2/3 of released prisoners are rearrested within three years. Something we often hear from ex-inmates is that they have to reoffend in order to survive life outside of prison. We study whether encouraging more frequent use of available social services can alter their trajectory and reduce recidivism. We randomly offer frequent user

cards that reward them after certain number of services. We now have detailed data on social service utilization of about 300 or so people who used to be incarcerated we will combine this in the future with their criminal record data, and hopefully gain exciting insights into the role of social services in preventing recidivism.

HX: I learned that you are leading the PittSmartLiving project, which is organized within three research labs: Data System, Human Behavior and Business Integration. What are the advantages of interdisciplinary collaboration? Are there obstacles you have already experienced or could foresee in the future?

SL: Professor Alex Labrinidis is the lead PI for PSL NSF grant he is the one who came up with the idea of creating a market ecosystem of businesses and



transportation. I am a co-PI and in charge of the social science side through the Human Behavior lab.

There are many advantages of interdisciplinary collaboration. We wouldn't have the initial ideas without the computer scientists, and even if we came up with it, we wouldn't be able to create the multitude of operational

FACULTY HIGHLIGHTS

components needed such as prediction for bus capacity, or the mobile apps that they are programming for the field experiment. Without the social science team, we won't have a good way of testing our intuition about people, and may be building a site-specific application instead of a general and portable social science infrastructure. Right now, different members of the PSL Human Behavior lab are working on various components of this infrastructure: the substitutability of time and money in travelers, promotional and queue management strategies in businesses, and the market that connects them. I gave a talk on this at CITRIS (UC Berkeley) on this lab here is the link: <https://www.youtube.com/watch?v=s90dp-8oA6o>

One issue in disciplinary collaboration is that what one discipline need from another are often considered applications instead of answers to fundamental questions. For example, we might need a web app from the CS team, and the CS team might need us to interview local businesses. As a result, both parties have to allocate some time to deliver less research-intensive applications so that the other

team can make advances.

Another issue is publication. In computer science, one paper can have many authors. But in economics, even big projects usually have very few authors. Our paper takes 3-4 years to publish and can be 40 pages with a 60-page Appendix full of robustness checks. CS publications are much faster and shorter. So there is a lot to negotiate and discuss.

HX: Do you have any new courses or are involved in any conferences you wish ISP students to know?

SL: My postdoc (Jinyong Jeong) and I will have a reading group this fall and a class in the spring. The class / reading group is called "Econ Meets CS: Mechanism Design and Applications" and will be in GSPIA but will be open to all. It is about the design of markets and institutions and how computation help bridge the gap between theoretical results and practical implementation. Whenever possible, we will try to highlight applications that are targeted towards social good. We hope you will join us!

(Email jinyong.jeong@pitt.edu to be added on the reading group list.)

FACULTY HIGHLIGHTS

I am involved with the ACM EC (Economics and Computation) conference. I work with the Mechanism Design for Social Good (MD4SG) group, which consist of researchers in AI, Operations Research, and Economics, among other fields. I was program chair and moderated a panel bridging theory and practice at EC'19 in Phoenix.

HX: How many graduate students are you advising now?

SL: I have about 6 graduate students and a postdoc that are working with me, but I'm not an advisor to all of them.

HX: Are you accepting new students at the moment? If yes, do you have any messages for the ISP student who wish to join your research group?

SL: Yes. Yes. I'm always open to work with students who are interested in mathematical models and about bringing the perspective of marginalized population into research.

HX: Do you have experience collaborating with industry?

SL: I collaborate with organizations a lot: shelters, unions, non-profits, ACDHS (Department of Human Services), WPRDC (Western Pennsylvania

Regional Data Center), etc. I work less frequently with for-profit companies, but the talk I presented at ISP last year (on accounting for stochastic noise in information aggregation) is actually motivated by internal prediction markets in SV tech firms. I also work with large consulting firms on my R Data Visualization class in GSPIA.

HX: Do you have any advice for the students who are on the job market?

SL: I would say that in a world where everyone is technically savvy, your ability to understand what motivates people will be key. This requires a student to be able to put herself in someone else's shoes, be it their employers, collaborators, or clients. I saw some brilliant people get stuck in their career because they can't understand and listen to other people's perspectives.

I would really encourage students who want to have a job to interact with the people who they are working for. For example, if you are trying to work for an education company, then try to work with kids. This experience will bring them a really valuable perspective of the end users.

FACULTY HIGHLIGHTS

HX: Is there anything you wish you knew or did when you were a PhD student?

SL: I really enjoyed my PhD experience, since Caltech fosters an open collaboration environment across disciplines.

One thing I wish I'd done was starting the research early. I didn't fully understand the class material until I collected my own data and started to do research. During the exploration, things I learned from class suddenly make sense to me.

HX: How do you spend your spare time?

Professor Linardi: I do Muay Thai and meditation. I enjoy spending time with my kids. ■

Learning to Connect Dots

An Interview with Professor Yalini Senathirajah

Huihui Xu(HX): What's the motivation for joining ISP?

Professor Yalini Senathirajah (YS): The main reason for joining ISP is there are a lot of students who are doing work related to intelligent systems. I believe there is possibility to leverage what they learn to solve the problems we have in healthcare. I gave a talk last semester and had positive feedback from students. So, I'm interested in working with ISP students.

HX: I learned that you have a background from different fields. The most impressive thing is you were studying in medical school to become a veterinarian. What led you from becoming a veterinarian to a scientist who has a specialized interest in designing better electronic health records and health IT systems?

YS: I always liked mathematics and physics since I was a little kid. I set my

path to study animals at a very early age, because I like animals. I was fascinated by physiological structures of animals, since some animals could adapt to extreme environments. As a result, I studied animals when I was an undergraduate. I got into veterinary school to study animal medicine to become a vet. However, I eventually leave this field because of some family issues.

After leaving veterinary school, I went back to where I did my undergraduate and started to work as a technical translator. I also worked at MIT as a secretary, but I got a chance to learn programming to organize the library. I was surrounded by people who were into computing, and I became more and more interested in computing too. Then I moved to New York City and got an entry level job as a programmer at the Columbia health sciences campus. I got promoted as a webmaster for the whole health sciences campus in the next year and continued doing this job for five

FACULTY HIGHLIGHTS

years. It was a very heavy job and I learned a lot of programming.

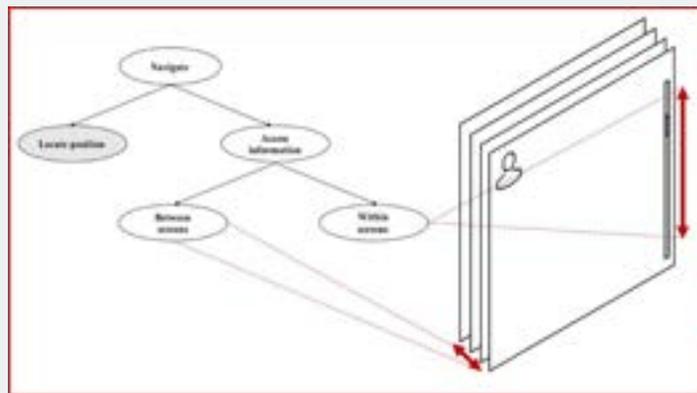
Ted Shortliffe was one of the pioneers who did AI in medicine, and he was also the new informatics department head at Columbia University. I went to a public lecture and got interested in AI and medicine. I asked Ted if I could just audit this class, and he agreed if and only if I did all the homework and exams. I did and got an A for that class. Then he recommended me to take a master's program in informatics. Since I did well in that program, I applied for a PhD in informatics at Columbia University.

My past experience exposed me to the problem of information overload and programming, and I saw the big gap between doctors and programmers due to their different backgrounds. As a result, I came up with an idea about leveraging the computer power to let doctors build their system on their own. I spent two years to build an experimental system and it worked. That's why I got into this field.

HX: Your short term goal is identifying design patterns that reduce usability problems in electronic health records. Could you

explain a little about some usability problems?

YS: Electronic health records have old-fashioned UI. Doctors can easily make mistakes, for example since the medicine list is always very long, and some drugs' names are similar to each other. The target audiences are both doctors and nurses, but they have different jobs. We have a modular type of approach, and you could sample from the data.



truncated from Navigation in the electronic health record: A review of the safety and usability literature

HX: Could you explain medical cognition for us?

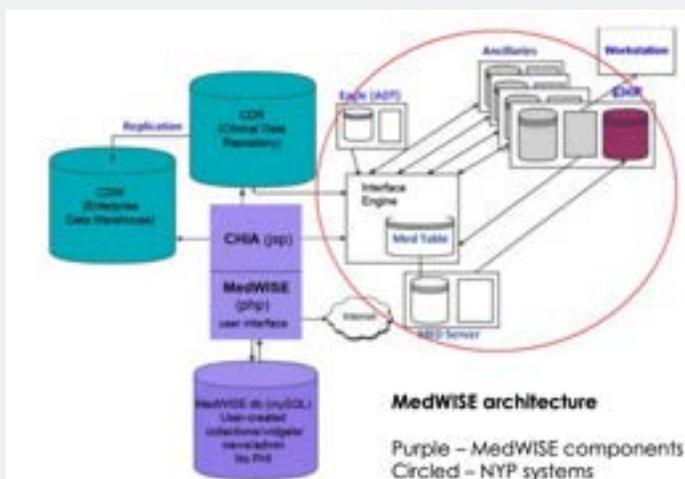
YS: Medical cognition: How does an expert solve a problem in their field? They are presented certain information about a patient and to decide how to treat them. It requires a long-time training. We noticed that there exists a

FACULTY HIGHLIGHTS

performance curve among medical students: when students don't have enough information, they only have few choices to make a decision. After a certain amount of time, students learn more and their brains start to reorganize all the information, which we called [FACETS]. When the brain is reorganizing the information, the performance is actually going down. The performance is going up again when the student learns more. This is an interesting experimental fact that intermediate students perform worst.

HX: How does the interaction design impact medical cognition?

YS: We are doing it by testing. We develop new UI features and let the user test them. We learn how those features affect the performance of diagnoses.



HX: MedWise is the research project you are currently doing, which allows clinician users to take the charge of the interfaces' composition. Are there any patterns you already observed about clinician users' preferences? Is there any correlation between the preferences and medical cognition you mentioned before?

YS: Yes. I noticed that a lot of doctors like to put orienting material on the left, the numerical analysis in the middle and study report on the right. There is a common pattern we observed: they will move things in different ways to fit their preferences. Doctors also have different diagnosis styles.

HX: Do you have any new courses and publications you wish the ISP students to know about?

YS: Not at the moment.

HX: How many graduate students are you advising now?

YS: Summer graduate students. They are Informal advisees.

FACULTY HIGHLIGHTS

HX: Are you accepting new students at the moment? If yes, do you have any messages for the ISP student who wish to join your research group?

YS: Yes, I'm accepting one or two graduate students for next semester. I would like to have students who are interested in health care, computational techniques. Students who can also think out of the box.

HX: Do you have any advice for the students who are on the job market?

YS: It is important to be able to communicate with people who are from different disciplines. Students who are self-motivated and have the ability to connect dots can overcome obstacles.

HX: Is there anything you wish you knew or did when you were a PhD student?

YS: It is important to know a PhD advisor's advising style.

HX: How do you spend your spare time?

YS: In New York City, I liked to walk

through all the different neighborhoods. Now, I'm doing this and exploring Pittsburgh a bit more. I also enjoy cooking with students. I have had students who are from Finland, India and China, we travel to conferences, sometimes cooked for each other or explore restaurants and have fun together. ■

Provost Fellowship Recipients

The Provost Fellowships in Intelligent Systems are awarded to two students of exceptional ability and promise. With the move to SCI, these Fellowships were created to replace the Dietrich School's Mellon Fellowship program.

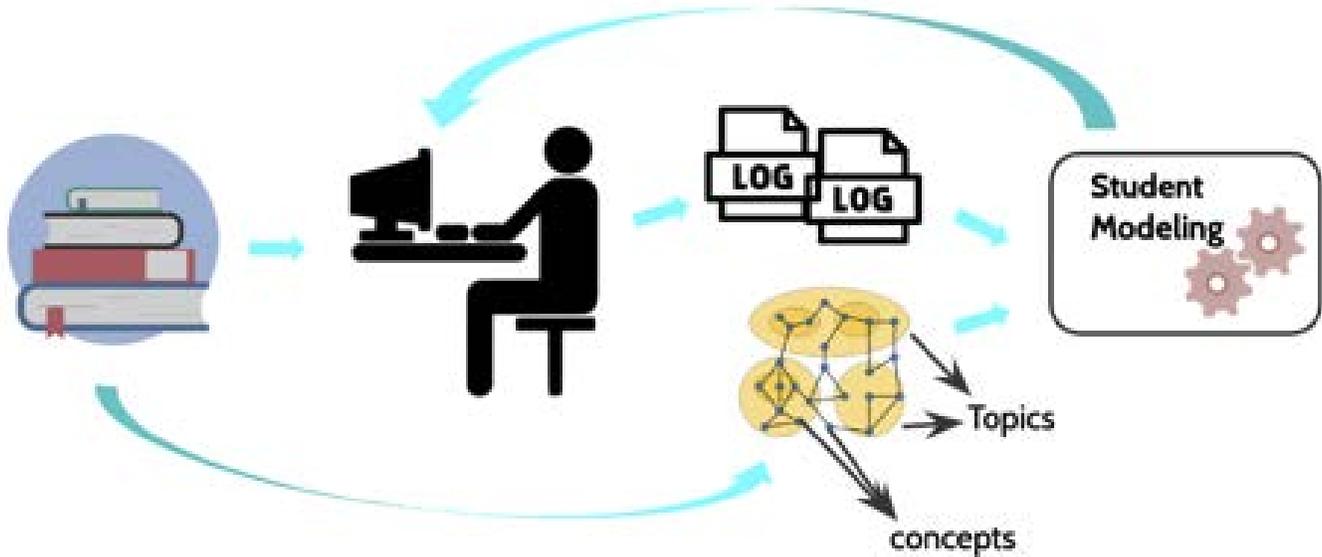


PhD Student Since 2016
Khushboo Thaker

Adaptive online Textbooks (AoT) recommend the most relevant pages and practice activities based on students' current knowledge state. The primary goal of AoTs is to provide effective learning and reduce the total time spent on skill acquisition through adaptive feedback. AoTs use student interaction data to infer the current state of student knowledge through student modeling (SM). The knowledge is inferred on knowledge components (KCs) associated with textbook material (sections/pages, practice activities, and quizzes). Recent research has explored the use of automatic KC extraction to annotate textbook sections

with KCs. However, none of the automatic KC extraction techniques are perfect and introduce noisy as well as correlated KCs. These correlated KCs, breaks the underlying assumption of student models, which expect a set of independent KCs annotated with each textbook material and makes SM impractical for AoT. To bridge this gap between practical and theoretical SM models, in my research my goal is to relax the assumption of independent KCs in SMs by introducing the notion of topic-based hierarchical SM. This will ease SM based adaptation in practical systems and make SM feasible for AoT.

STUDENT HIGHLIGHTS



Provost fellowship will support me for the research and will help me in my goal to develop better student models which can help Adaptive Education. I would like to make this research part of my dissertation research, thus it also helps me towards my PhD. I am thankful to Prof. Peter Brusilovsky and Prof. Daqing He for their encouragement and support in pursuing for this fellowship. And our director Dr. Litman who has always been an encouragement. ■

Provost Fellowships Recipients

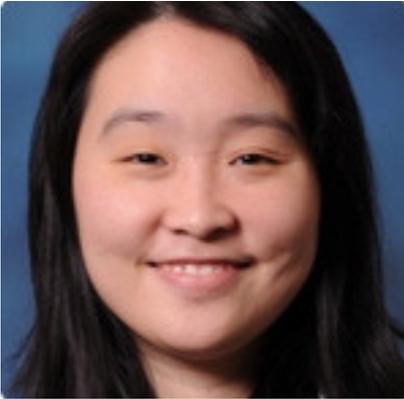


**PhD Student in
Computer Science and
Intelligent Systems
Program Since 2014
Mahbaneh Eshaghzadeh
Torbati**

Cardiovascular disease (CVD) is the leading cause of morbidity and mortality and causes 1 in 3 deaths in the US (a total of 800,000 annually). CVD accounts for about one-third of the disparity in potential life-years lost between African-Americans and Caucasians. Current CVD screening methods rely on two main types of risk assessments: the CVD risk scores and noninvasive peripheral tests. While the risk scores mainly consist of the traditional biomarkers (e.g., age, cholesterol, blood pressure), the peripheral tests are mainly noninvasive tests referred to as non-traditional risk factors, such as indicators of endothelial dysfunction or artery blockage.

In general, people with a higher risk of CVD benefit more from screening and treatment. However, both risk scores, and peripheral tests have their own shortcomings. Framingham risk score (FRS), Reynolds Risk Score, and pooled cohort risk equations are the state-of-the-art CVD risk scores, all of which have high Recall but suffer from low Precision. This implies that these risk scores tend to incorrectly predict a significant number of patients at high risk of CVD. Such a behavior question is the practicality of the current risk scores for the CVD screening purposes.

The Provost fellowship gave me the opportunity to study how both of these assessments can be merged into an improved risk score. In my research, I answer whether peripheral tests can be used to improve FRS over the group of patients with intermediate CVD health. ■



Ye Ye

Employer: DBMI, University of Pittsburgh

Defense date: November 12, 2018

Research Field: Public health informatics; Outbreak detection; Decision support systems

Transfer Learning for Bayesian Case Detection Systems

In this age of big biomedical data, massive amounts of data have been produced worldwide. If we could effectively share all the information accumulated from all existing resources, we may develop a deeper understanding of biomedical problems and find better solutions.

Compared to traditional machine learning techniques, transfer learning techniques fully consider differences between shared parties in order to provide a smooth transfer of knowledge from source party to target party. Most well-established techniques focus on sharing data, while recent techniques have begun to explore the possibility of sharing models. Model-sharing techniques are especially appealing for biomedical area because of much less privacy risks. Unfortunately, most model-transferring techniques are unable to handle heterogeneous scenarios where feature spaces, marginal and conditional distributions differ among shared parties, which commonly exist in biomedical data.

My dissertation developed an innovative transfer learning framework to share data or model under heterogeneous scenarios. Heuristic scores have been designed to integrate source information with target data, while allowing injections of target-specific features for a better localization. Both synthetic and real-world datasets were used to test two hypotheses: 1) Transfer learning is

DISSERTATION DEFENSES

better than using the model constructed with target data only; 2) Transfer learning is better than direct adoption of the source model. A comprehensive analysis was conducted to investigate conditions where these two hypotheses hold, and more generally the factors that affect the effectiveness of transfer learning, providing empirical opinions about when and what to share.

My research contributes to the fields of machine learning, medical informatics and disease surveillance. It enables knowledge sharing under heterogeneous scenarios and provides methodologies for diagnosing transfer learning performance under tasks varying degrees of feature space overlapping, similarities of distributions, and sample sizes. The model-transferring algorithm can be viewed as a new Bayesian network learning algorithm with a flexible representation of prior knowledge allowing partial feature coverage. To the best of my knowledge, this is the first exploration on model-transferring for biomedical data in heterogeneous scenarios. My work shows the potential of quick development of a case detection system for an emergent unknown disease and demonstrates its transferability and adaptability. ■

Committee: Dr. Fuchiang (Rich) Tsui, Dr. Michael M. Wagner, Dr. Gregory F. Cooper, Dr. Jeremy Weiss

DISSERTATION DEFENSES



Gaurav Trivedi

Employer: Health Technology Company

Defense date: April 29, 2019

Research Field: Intelligent Interfaces; Human-Computer Interaction; Biomedical Informatics

Interactive Natural Language Processing for Clinical Text

Clinicians use free-text to conveniently capture rich information about patients. Care providers are likely to continue using narratives and first-person stories in Electronic Medical Records (EMRs) due to their convenience and utility, which complicates information extraction for computation and analysis. Despite advances in Natural Language Processing (NLP) techniques, building models is often expensive and time-consuming. Current approaches require a long collaboration between clinicians and data-scientists. Clinicians provide annotations and training data, while data-scientists build the models. With the current approaches, the domain experts - clinicians and clinical researchers - do not have provisions to inspect these models and give feedback. This forms a barrier to NLP adoption in the clinical domain by limiting power and utility of real-world applications.

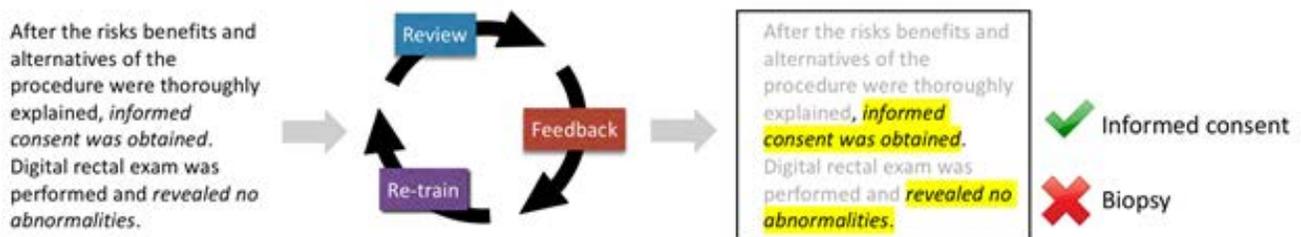


Figure 1 Interactive Natural Language Processing allows domain experts, without machine learning experience to build models on their own, and also reduce or eliminate the need for collecting prior annotations and training data.

NEW GRADUATES

Building models interactively can help narrow the gap between clinicians and data-scientists (Figure 1). Interactive learning systems may allow clinicians, without machine learning experience, to build NLP models on their own and also reduce the need for prior annotations upfront. These systems make it feasible to extract understanding from unstructured text in patient records; classifying documents against clinical concepts, summarizing records and other sophisticated NLP tasks. Interactive systems enable end-users to review model outputs and make corrections to build model revisions within an interactive feedback loop.

Interactive methods are particularly attractive for clinical text due to the diversity of tasks that need customized training data. In my dissertation, I demonstrate this approach by building and evaluating prototype systems for both clinical care and research applications. I built NLPReViz as an interactive tool for clinicians to train and build binary NLP models on their own for retrospective review of colonoscopy procedure note. Next, I extended this effort to design an intelligent tool to identify incidental findings from radiology notes as clinicians review patient notes during their regular workflow. I follow a two-step evaluation with clinicians as study participants: a usability evaluation to demonstrate feasibility and overall usefulness of the tool, followed by an empirical evaluation to evaluate model correctness and utility. Lessons learned from the development and evaluation of these prototypes will provide insight into the generalized design of interactive NLP systems for wider clinical applications. ■

Committee: Dr. Harry Hochheiser, Dr. Shyam Visweswaran, Dr. Rebecca Hwa, Dr. Wendy Chapman



Zahra Rahimi

Employer: Pandora

Defense date: May 28, 2019

*Research Field: Automatic Essay Assessment, Natural Language Processing, Spoken Dialogue Systems
Text Mining*

Linguistic Entrainment in Multi-party Spoken Dialogues

Entrainment is the propensity of speakers to begin behaving like one another in conversations. Evidence of entrainment has been found in multiple aspects of speech, including acoustic-prosodic and lexical. More interestingly, the strength of entrainment has been shown to be associated with numerous conversational qualities, such as social variables. These two characteristics make entrainment an interesting research area for multiple disciplines, such as natural language processing and psychology. To date, mainly simple methods such as unweighted averaging have been used to move from pairs to groups, and the focus of prior multi-party work has been on text rather than speech (e.g., Wikipedia, Twitter, online forums, and corporate emails). The focus of this research, unlike previous studies, is multi-party spoken dialogues. The goal of this work is to develop, validate, and evaluate multi-party entrainment measures that incorporate characteristics of multi-party interactions, and are associated with measures of team outcomes.

In this thesis, first, I explore the relation between entrainment on acoustic-prosodic and lexical features and show that they correlate. In addition, I show that a multi-modal model using entrainment features from both of these modalities outperforms the uni-modal model at predicting team outcomes.

DISSERTATION DEFENSES

Moreover, I present enhanced multi-party entrainment measures which utilize dynamics of entrainment in groups for both global and local settings. As for the global entrainment, I present a weighted convergence based on group dynamics. As the first step toward the development of local multi-party measures, I investigate whether local entrainment occurs within a time-lag in groups using a temporal window approach. Next, I propose a novel approach to learn a vector representation of multi-party local entrainment by encoding the structure of the presented multi-party entrainment graphs. The positive results of both the global and local settings indicate the importance of incorporating entrainment dynamics in groups. Finally, I propose a novel approach to incorporate a team-level factor of gender-composition to enhance multi-party entrainment measures. All of the proposed works are in the direction of enhancing multi-party entrainment measures with the focus on spoken dialogues although they can also be employed on text-based communications. ■

Committee: Dr. Diane Litman, Dr. Rebecca Hwa, Dr. Kevin Ashely, Dr. Louis-Philippe Morency

CONGRATULATIONS TO OUR GRADUATING STUDENTS!

Gaurav Trivedi and Ye Ye shared their experience and memory with ISP

An Interview with Gaurav Trivedi

Q: What is your dissertation topic? Why do you think it is important and unique?

Gaurav Trivedi (GT): While preparing my statement of purpose, I had plans to work on AI systems that work in collaboration with human experts. However, I was not sure about the application areas such systems when applying for the program. I was interested in the human-computer interaction and intelligent interfaces in general at that time. During my first year, I had the opportunity to join a project with Drs. Hochheiser, Wiebe, Hwa, and Chapman, who also served as my preliminary exam committee. That project was about incorporating clinician feedback to build NLP models. It helped me to form the core idea of my dissertation: Interactive natural language processing.

Current approaches require a long collaboration between clinicians and data-scientists. Clinicians provide annotations and training data, while data-scientists build the models. With

the current approaches, the domain experts - clinicians and clinical researchers - do not have provisions to inspect these models or give direct feedback. This forms a barrier to NLP adoption and limits its power and utility for real-world clinical applications. I explored interactive methods to allow clinicians without machine learning experience to build NLP models on their own. This approach may make it feasible to extract understanding from unstructured text in patient records; classifying documents against clinical concepts, summarizing records and other sophisticated NLP tasks while reducing the need for prior annotations and training data upfront. In my dissertation, I demonstrate this approach by building and evaluating prototype systems for both clinical care and research applications.

Q: What were the obstacles you encountered when you were preparing the dissertation? How did you overcome those?

NEW GRADUATES

GT: My research is about using Natural language processing (NLP) methods on data from Electronic Medical Records (EMR). The use of real-world data would be necessary to verify the feasibility of applying NLP. One challenge during my dissertation was to find problems where I could replicate the ideas defined in my first project on interactive NLP. For my dissertation, I extended the interactive tool we developed for reducing clinicians' workload in clinical care applications. Pursuing my Ph.D. program in the Intelligent systems program allowed me to form good collaborations at the Department of Biomedical Informatics, with Dr. Visweswaran's group, as well as clinicians from UPMC. I took an applied clinical informatics course at DBMI where the TA was Dr. Rob Handzel, who is a surgeon. I talked to him about my ideas, and he got on board to develop the ideas further and also helped with getting access to the required data. I even had a chance to shadow trauma surgeons for initial validation of my ideas. These collaborations also made it possible to run my evaluation studies with physicians as participants for my dissertation.

Q: What was your experience when

studied in ISP? What did you learn from this experience?

GT: ISP is an excellent program for applied AI. The founders were definitely visionaries in starting a program dedicated to AI applications over thirty years ago. Now everybody is talking about using machine learning (and more recently deep learning) for applications in medicine & health, education and law. ISP provides an environment for interdisciplinary collaboration. I definitely benefited a lot from these collaborations my work during my dissertation.

Q: What were some things you wish you had done when you were a PhD student?

GT: My whole experience was very enjoyable. In hindsight, I could have completed my projects a bit faster. But it totally depends on the situation you are in and it is certainly easier to know the right thing to do after they are done

Q: What was your most proud moment?

GT: Our work getting featured in American Medical Informatics Association (AMIA)'s Year-in-review in 2018.

NEW GRADUATES

Q: What is your future plan?

GT: I will work for a health technology company in Pittsburgh. ■

An Interview with Ye Ye

Q: What is your dissertation topic? Why do you think it is important and unique?

Ye Ye (YY): My dissertation is “transfer learning for Bayesian case detection systems.” Transfer learning is a type of machine learning techniques. Different from traditional machine learning, transfer learning analyzed the difference among different origins of data to provide a smooth knowledge share. I developed a transfer learning framework to facilitate data sharing and model sharing between different settings (e.g. hospitals). My research enables knowledge sharing under heterogeneous scenarios and provides an approach for understanding transfer learning performance in terms of differences of features, distributions, and sample sizes between two settings. The model transfer algorithm can be viewed as a new Bayesian network learning algorithm with a flexible representation of prior knowledge. In concrete terms, this work shows the potential for transfer learning to assist in the rapid development of a case detection system for an emergent unknown disease. More generally, to my knowledge, this research is the first investigation of

model-based transfer learning in biomedicine under heterogeneous scenarios.

Q: What were the obstacles you encountered when you were preparing the dissertation? How did you overcome those?

YY: Finishing a PhD dissertation is challenging. Rather than finding a tailored solution for a specific problem, the goal is to develop a fundamental methodology that is innovative and generalizable. I sincerely appreciate all the guidance provided by my dissertation committee. Dr. Tsui suggested the dissertation topic. Dr. Cooper worked very closely with me to discuss the theoretical part; he met with me as often as I needed. Dr. Wagner guaranteed my usage of real-world datasets and provided me critical suggestions on dissertation writing. Dr. Weiss suggested me to conduct simulation experiments to analyze different transfer learning scenarios. I finished the most important algorithm implementations in summer 2018, when I took a walk in Shadyside every

NEW GRADUATES

afternoon with my best friend, Yun Huang (also an ISP alumnus), who was finishing her dissertation. We shared positive energy every day. Lastly, my family provided unconditional love and support. They allowed me to finish the trip I have chosen on my own pace.

Q: What was your experience when studied in ISP? What did you learn from this experience?

YY: The ISP program opens a door for students with diverse backgrounds toward a new field, Artificial Intelligence (AI). I enrolled in ISP with medical and population health background, and statistical training. The seven-year study in ISP provided me fundamental training in artificial intelligence, which enables me to conduct research on biomedicine using advanced AI techniques. Working in the Real-time Outbreak and Disease Surveillance Laboratory (RODS) in the Department of Biomedical Informatics was a great experience. RODS provided me an opportunity to engage and initiate in multiple research topics and collaborations. Moreover, through PhD study, I developed logic thinking, research independence, meticulous attention to detail and accuracy, and patience and confidence

in meeting uncertainties.

Q: What were some things you wish you had done when you were a PhD student?

YY: Take an internship to gain a new perspective.

Q: What was your most proud moment?

YY: In 2013, our team was selected as one of finalists in Michael G. Wells Student Health Care Entrepreneurship Competition.

In 2014, my first journal article finally got accepted after two rounds of revisions.

In 2017, I received the Andrew Mellon Fellowship.

In July 2017, I gave birth to a very healthy baby girl, Jane.

In November 2018, I successfully defend my dissertation.

In June 2019, my dissertation was selected as finalists (top 5) in American Medical Informatics Association Doctoral dissertation competition.

NEW GRADUATES

Q: What is your future plan?

YY: I decide to stay in academia. Ever since I was 6 years old, I have dreamed about becoming an elementary school teacher. After enrolling in college, I've always been impressed by the creativity and ingenuity of students and dedications of faculty. In annual conferences, I enjoyed meeting with colleagues, and I was always inspired by other researchers' achievements. ■



Pitt

School of Computing
and Information