

Natural Language Processing State of Science Conference
September 9th and 10th, 2015
Veterans Health Administration National Conference Center
2011 Crystal Drive, Crystal City, Virginia 22202

Background

Research and development using Natural Language Processing (NLP) has accelerated over the past 3 decades with the explosion of electronic text data in medical, commercial, and public domains. The use of NLP expands the ability to apply “big data” and machine learning techniques to a combination of structured and unstructured data. The growth of electronic health records has accelerated the need for and value of NLP in healthcare as a way to extract information from text notes and other unstructured elements (e.g. pathology and microbiology reports) of the EHR. Both operations and research elements of VA and other health systems have invested in NLP research and development. As one of the leading federal research funders, the VA Health Services Research and Development has been actively supporting NLP research over the last ten years. While commercial approaches in healthcare have concentrated on extracting information on clinical conditions and procedures for billing purposes, tools developed in VA and elsewhere have often been aimed at solving discrete problems rather than developing broader solutions. More innovative and more broadly applicable approaches are needed and may have been pursued outside of VA and outside of healthcare.

This conference seeks to bring leading NLP researchers from non-health care areas (including industry, other scientific fields and other government agencies) along with leading health care NLP researchers from in VA and outside VA, to explore new research areas and approaches to accelerate VA research in NLP. The goal will be to identify a research agenda that emphasizes developing broadly applicable solutions rather than narrowly targeted tools (this has been characterized as “trunk” and “branch” vs. “leaf” approaches.) A proposed conference will invite a wide array of NLP researchers and industry to discuss the state of NLP science and effective use of NLP research products.

Purpose

- Facilitate collaborations and cross-fertilization among NLP experts from within and outside of healthcare.
- Identify NLP research agenda for VA HSR&D community; and
- Articulate a pathway for developing enterprise system tools from NLP research products.

Conference schedule

Day 1, September 9th, 2015	
8:30 am – 9:30 am	Welcome remarks Tim O’Leary, CRADO, VA ORD David Atkins, Director, VA ORD/HSRD Miho Tanaka, Scientific Program Manager, VA ORD/HSRD Self-introduction
9:30 am – 10:00 am	Key note presentation “State of Art of Health NLP” Guergana Savova, Harvard Medical School
10:00 am – 10:20 am	“NLP research trend in the last five years and VA NLP research accomplishments” Wendy Chapman, Salt Lake City VAMC
10:20 am – 10:35 am	Break
10:35 am – 11: 10 am	"Using Natural Language Processing to Uncover Signals of Mental State" Philip Resnik, Professor, Department of Linguistics and Institute for Advanced Computer Studies, University of Maryland
11:10 am – 11: 45 am	“Linguistic Structured Sparsity in Text Categorization” Noah Smith, Associate Professor, University of Washington, Department of Computer Science and Engineering
11: 45 am – 12:30 pm	Microsoft Research presentations: “Processing Text for HealthCare” “Machine Reading for Cancer Panomics”, presented by Hoifung Poon “The Structure of Free Text in EMRs”, presented by Lucy Vanderwende
12:30 pm – 1:30 pm	Lunch
1:30 pm – 2:30 pm	Panel Discussions: NLP innovations – what does the healthcare industry see as knowledge gap in the NLP research? <u>Moderator:</u> Hongfang Liu, Mayo Clinic, Open Health NLP Consortium <u>Panelists:</u> Eric Brown, IBM Watson, Richard Wolniewicz, 3M Health Information Systems, Brian Hazlehurst, Kaiser Permanente Northwest Center for Health Research Scott DuVall, VA Salt Lake City VAMC

	Ruth Reeves, VA Tennessee Valley VAMC
2:30 pm – 3:30 pm	Breakout sessions – Collaboration and knowledge exchanges between biomedical NLP and general domain NLP
3:30 pm – 3:45 pm	Break
3:45 pm – 4:30 pm	Breakout session report back
4:30 pm – 5:00 pm	Wrap up for the first day
Day 2, September 10th, 2015	
8:30 am – 8:45 am	A summary from the first day, the plan for the second day
8:45 am – 10:15 am	<p>Panel Discussion: State of VA NLP research – accomplishments, opportunities, and future challenges</p> <p>Moderator: Matt Samore, Salt Lake City VAMC</p> <p><u>Panelists:</u> Michael Matheny, Tennessee Valley VAMC Wendy Chapman, Salt Lake City VAMC Qing Zeng, Salt Lake City VAMC Hong Yu, Bedford VAMC</p>
10:15 am – 10:30 am	Break
10:30 am – 12:00 pm	<p>Panel Discussion: Recommendations for VA NLP research agenda</p> <p><u>Moderator:</u> Miho Tanaka, VA ORD/HSRD</p> <p><u>Panelists:</u> Ani Nenkova, University of Pennsylvania Eric Brown, IBM Watson Noah Smith, University of Washington Brian Hazelhurst, Kaiser Permanente Northwest Center for Health Research Hoifung Poon or Lucy Vanderwende, Microsoft Research John Conroy, IDA Center for Computing Sciences Michael Matheny, Tennessee Valley VAMC Jennifer Garvin, Salt Lake City VAMC</p>
12:00 pm – 12:45 pm	<p>A summary of the conference, action items</p> <p>David Atkins, Director, VA ORD/HSRD Miho Tanaka, Scientific Portfolio Manager, VA ORD/HSRD</p>
12:45 pm	Meeting Adjourn