Apr;60:334-41.
81. Li M, Carrell D, Aberdeen J, Hirschman L, Kirby J, Li B, et al. Optimizing annotation resources for natural language de-identification via a game theoretic framework. J Biomed Inform 2016 Jun;61:97-109.
82. * Scaiano M, Middleton G, Arbuckle L, Kolhatkar V, Peyton L, Dowling M, et al K. A unified framework for evaluating the risk of re-identification of text de-identification tools. J Biomed Inform 2016 Oct;63:174-83.
83. Carrell DS, Cronkite DJ, Malin BA, Aberdeen JS, Hirschman L. Is the Juice Worth the Squeeze? Costs and Benefits of Multiple Human Annotators for Clinical Text De-identification. Methods Inf Med 2016 Aug 5;55(4):356-64.
84. Kilicoglu H, Rosemblat G, Fiszman M, Rindflesch TC. Sortal anaphora resolution to enhance relation extraction from biomedical literature. BMC Bioinformatics 2016 Apr 14;17:163.
85. Liu S, Liu H, Chaudhary V, Li D. An Infinite Mixture Model for Coreference Resolution in Clinical Notes. AMIA Jt Summits Transl Sci Proc 2016 Jul 22;2016:428-37.
86. Tseytlin E, Mitchell K, Legowski E, Corrigan J, Chavan G, Jacobson RS. NOBLE - Flexible concept recognition for large-scale biomedical natural language processing. BMC Bioinformatics 2016 Jan 14;17:32.
87. Lin C, Dligach D, Miller TA, Bethard S, Savova GK. Multilayered temporal modeling for the clinical domain. J Am Med Inform Assoc 2016 Mar;23(2):387-95.
88. Kaggal VC, Elayavilli RK, Mehrabi S, Pankratz JJ, Sohn S, Wang Y, et al. Toward a Learning Healthcare System - Knowledge Delivery at the Point of Care Empowered by Big Data and NLP. Biomed Inform Insights 2016 Jun 23;8(Suppl 1):13-22.
89. Wallace BC, Kuiper J, Sharma A, Zhu MB, Marshall IJ. Extracting PICO Sentences from Clinical Trial Reports using Supervised Distant Supervision. J Mach Learn Res 2016;17.
90. Clark DG, McLaughlin PM, Woo E, Hwang K, Hurtz S, Ramirez L, et al. Novel verbal fluency scores and structural brain imaging for prediction of cognitive outcome in mild cognitive impairment. Alzheimers Dement (Amst) 2016 Feb 15;2:113-22.
91. Takano K, Ueno M, Moriya J, Mori M, Nishiguchi Y, Raes F. Unraveling the linguistic nature of specific autobiographical memories using a computerized classification algorithm. Behav Res Methods 2017 Jun;49(3):835-52.
92. Luo SX, Shinall JA, Peterson BS, Gerber AJ. Semantic mapping reveals distinct patterns in descriptions of social relations in adults with autism spectrum disorder. Autism Res 2016 Aug;9(8):846-53.

Correspondence to:
A. Névéol, P. Zweigenbaum
LIMSI
Campus universitaire, bât 508
Rue John von Neumann
F-91405 Orsay cedex
France
E-mail: {neveol,pz}@limsi.fr

Appendix: Content Summaries of Selected Best Papers for the 2017 IMIA Yearbook, Section Clinical Natural Language Processing

Althoff, T, Clark K, Leskovec, J

Large-scale Analysis of Counseling Conversations: An Application of Natural Language Processing to Mental Health

Trans Assoc Comput Linguist 2016;4:463-47

This paper presents an analysis of counselor/patient interactions in text message conversations collected from a free hotline. The authors characterize counseling conversation relying on a wide range of methods including discourse analysis, statistical language modeling, and sentiment analysis. They evidence actionable conversation strategies that are associated with better conversation outcomes: successful counselors exhibit an ability to adapt creatively to each new counseling situation, their reaction to ambiguity with check questions and appreciation language is well received, they conduct conversations efficiently by spending less time to understand the patient's issue and more time in problem-solving, and they facilitate a change in perspective. These results may be used towards improving practice recommendations and counselor training.

Kilicoglu, H, Demner-Fushman, D

Bio-SCoRes: A Smorgasbord Architecture for Coreference Resolution in Biomedical Text

PLoS One 2016 Mar 2;11(3):e0148538

This paper addresses coreference resolution, which is a fundamental and challenging task in natural language processing. The authors offer a very instructive and informed definition of coreference as well as a comprehensive literature review of the state of the art in coreference resolution. They describe and evaluate a general method for biomedical coreference resolution called BioSCoRes, implemented in a publicly available toolkit. The evaluation broadly covers the biomedical domain by relying on two existing corpora (one clinical, one biological), as well as a newly developed and shared corpus of drug label inserts. It also offers a detailed performance report on each type of coreference. BioSCoRes obtains overall results that come close (on the clinical corpus) or exceed (on the other two corpora) the state-of-the art.

Morid, MA, Fiszman, M, Raja, K, Jonnalagadda, SR, Del Fiol, G

Classification of clinically useful sentences in clinical evidence resources

J Biomed Inform 2016 Apr;60:14-22

This paper presents a method for classifying sentences from a variety of evidence-based clinical decision support knowledge sources according to clinical usefulness. This work offers a specific definition of actionable, clinically useful sentences. Then, it proceeds to explore advanced NLP methods to extract rich features for sentence classification. A feature ablation study supports the proposed feature-rich approach and shows that the system performs with an F-measure of at least 73% on different text genres. This work is exemplary in exploring fundamental approaches towards the practical goal of providing real-time clinical information at the point of care, while setting up a technical framework that will facilitate the integration of the research results in a clinical setting.

Shivade C, de Marneffe MC, Fosler-Lussier E, Lai AM

Identification, characterization, and grounding of gradable terms in clinical text

Proceedings of the 15th Workshop on Biomedical Natural Language Processing 2016:17-26

This paper presents an analysis of gradable adjectives found in clinical text to qualify medical findings. The authors use existing methods to automatically identify gradable adjectives in clinical corpora and estimate prevalence at about 30% of adjectives.

Focusing on four clinical phenomena, the authors show that the gradable adjectives used to qualify these phenomena in a clinical corpus can be reliably associated to numerical value intervals using a probabilistic model. This very original work relies on a regular expression-based analysis of a clinical description of selected phenomena comprising a gradable adjective along with a numerical value. It offers a first step towards the interpretation of statements using gradable adjectives by grounding their meaning to value intervals, which would facilitate clinical decision-making.

Wu, Y, Denny, JC, Rosenbloom, ST, Miller, RA, Giuse, DA, Wang, L, Blanquicett, C, Soysal, E, Xu, J, Xu, H

A long journey to short abbreviations: developing an open-source framework for clinical abbreviation recognition and disambiguation (CARD)

J Am Med Inform Assoc 2017 Apr 1; 24(e1):e79-e86

This paper presents an open source framework for clinical abbreviation recognition and disambiguation via entity linking to the Unified Medical Language System (UMLS). This work builds on a large body of research on different aspects of abbreviation resolution including recognition of abbreviated entities in clinical text, the extraction of possible long forms or senses from knowledge bases, and the disambiguation of a given entity that leverages context to identify the unique sense of a given abbreviated entity. The overall framework offers performance that exceeds the state-of-the-art on two shared datasets. In addition, the tool may be tailored to specific needs by allowing the use of customized resources.