Nomination Form

Candidates must be nominated by an ACL member (no self-nominations). All materials must be submitted in English.

Candidate Information

Name	Email
Ralph Grishman	**grishman@cs.nyu.edu**
Affiliation	
New York University, Department of Computer Science	

Degree

Highest degree held by candidate, including institution, year conferred and major discipline.

Ph.D., Physics, Columbia University, 1973.

Positions

Relevant positions held by candidate, with dates.

1983-present, Professor (Chairman, 1986-1988), Computer Science Department, Courant Institute of Mathematical Sciences, New York University
1978-1983, Associate Professor, Computer Science Department, Courant Institute of Mathematical Sciences, New York University
(1982-83: on assignment at Navy Center for Applied Research in Artificial Intelligence, Washington, D.C.)
1973-1978, Assistant Professor, Computer Science Department, Courant Institute of Mathematical Sciences, New York University
1973, Associate Research Scientist, Courant Institute of Mathematical Sciences, New York University
1971-1973, Instructor in mathematics, Barnard College, New York
1969-1973, Consultant to Linguistic String Project, New York University.
1968-1971, Associate in Physics, Columbia University, New York.
1968, Consultant to Fides Union Fiduciaire, Zurich
1964-1968, Operations Assistant, Courant Institute of Mathematical Sciences, New York University

Contributions

A summary of the candidate's most significant contributions to the association and the scientific field (within 1500 words if possible), subdivided according to the selection criteria described below:

- Scientific and technical excellence: Important technical contributions to computational linguistics.
- Service excellence: Extended service to the ACL or other scientific organizations whose mission aligns with that of the ACL.
- Enabling excellence: Significant educational and outreach activities with broader impact.
- Other relevant contributions.

Note that candidates are not required to have significant contributions in all areas, so this part of the nomination may include one, two, three or four sections as appropriate.

Scientific and Technical Excellence

Among many contributions made by Prof. Grishman, perhaps, the most impactful is the central role he played in defining Information Extraction (IE) research over the past 30+ years. He is an author of over 200 published research papers with a total of 14038 citations, an h-index of 55 and i10-index of 172.

The biggest impact Prof. Grishman has made was through his deep involvements in the U.S. government-sponsored evaluations, including the MUC (Message Understanding Conferences, 1987-1998), ACE (Automatic Content Extraction, 1999-2008) and Knowledge Base Population (KBP, 2009-present). The success of these multi-site, nationwide evaluations not only shaped the IE we see today, but also helped NLP research continue funded by U.S. government agencies throughout the years. He is the only NLP researcher who has participated in nearly all the evaluations that have been held annually, and most importantly, he also took organizational roles in the task design and data annotation. In 1987, he directed the Proteus project at NYU as one of the six teams to participate in MUC-1 that started the exploration in this new domain (the term “Information Extraction” was only coined later) with 12 narrative paragraphs from naval messages as training and test data and no defined evaluation task or metrics. He continued to participate in all the following MUCs and formed close collaborations with U.S. Defense Advanced Research Projects Agency (DARPA) in creating “Information Extraction” as one of the earliest and most important NLP tasks. In 1995, Prof. Grishman took the leadership role and chaired the MUC-6 planning committee, which innovatively switched away from the increasingly complex template-based IE and introduced several now-standard subtasks and evaluations, including named entity recognition and coreference. The introduction of these subtasks helped IE research remain doable with increasing deeper understanding of complex languages and continue to flourish. His paper on MUC-6 has gathered over 1000 citations. Another organizational role of Prof. Ralph Grishman which had made impact was chairing the Tipster Architecture committee, which developed a standard API for IE. The US contractor to implement this was not successful, but it was picked up by the University of Sheffield and released as the GATE system, widely used in Europe. He and his former student Heng Ji are also among the organization committee for the ongoing KBP evaluations sponsored by U.S. Department of Defense.

Prof. Grishman also made great impact on Information Extraction research through many technical innovations. His group is one of first to apply corpus techniques to IE (Borthwick et al 1998), instead of using hand-crafted rules, in the 1990s. The maximum entropy name tagger (thesis work of his student Andrew Borthwick 1999, cited 480+ times) has several advantage over BBN’s first HMM-based model in terms of the information which could be integrated and was quickly taken up by other researchers. This was followed by early work on semi-supervised methods for event extraction (Yangarber et al. 2000, cited 240+ times). The work on unsupervised relation extraction was another group effort (Hasegawa, Sekine and Grishman 2004, cited 410+ times) which attracted considerable attention and led to work elsewhere involving unsupervised discovery with more refined similarity metrics. His group is also one of the early adopters of word embedding and deep learning methods for IE with a series of publications since 2014. Other notable work on IE by his Proteus group at NYU has included: kernel methods (Zhao et al. 2004), joint inference (Ji and Grishman 2005),

cross-document methods (Ji and Grishman 2008), active learning methods (Fu and Grishman 2013) and distant supervision (Xu et al. 2013; Min et al. 2013; Pershina et al. 2014). He and his group also developed multiple NLP resources and related shared-tasks, including the ACE-2 corpus and COMLEX (a large syntactic dictionary of over 39,000 head word published in 1997) released through LDC, as well as NOMBANK, and NOMLEX (a dictionary of nominalizations). He also has written multiple surveys (e,g. Grishman 1997, cited 690+ times; Grishman 2012) and evaluation overview papers on information extraction at different time periods that guided generations of NLP researchers.

Prof. Grishman is also one of earliest researchers to work on NLP in the 1960s and pioneered the development of natural language processing for computers under Naomi Sager (https://en.wikipedia.org/wiki/Naomi_Sager). Between 1969 and 1973, Grishman worked on the Linguistic String Project at New York University (NYU), which drew on the noted linguist Zellig Harris (https://en.wikipedia.org/wiki/Zellig_Harris)'s theory: linguistic string theory, transformation analysis, and sublanguage grammar. The influence of the structuralist tradition from the Linguistic String Project prompted some of the first quantitative studies of sublanguages (languages as used in individual scientific and technical domains) based on distributional analysis (Hirschman et al. 1975, Grishman et al. 1986). The volume on sublanguages Grishman co-edited (Grishman and Kittredge 1986) continues to be cited and was recently republished in 2014. Since 1985, Grishman has directed the Proteus Project at NYU, which is active till today and has conducted a wide variety of research in NLP with a particular focus on information extraction (IE). He also wrote one of the first NLP textbooks "Computational Linguistics: An Introduction (Studies in Natural Language Processing)", published by Cambridge University Press (1986), which was translated into Spanish, Italian, and Japanese and educated a generation of NLP researchers in Europe and Japan.

Service Excellence

Prof. Grishman served as president of the Association for Computational Linguistics in 1991 and was program chair of the ACL's Applied NLP conference in 1997. He served as the Chairman of Computer Science Department at NYU between 1986 and 1988.

His Service for Association for Computational Linguistics (ACL):

- Member of the Executive Committee, 1989
- Vice President, 1990
- President, 1991
- Member of the Nominating Committee, 1992-1994
- Program Chair, Conference on Applied Natural Language Processing, 1997
- Member of the Executive Committee, North American Chapter of the ACL, 2000-2001

His Service on US Government Committees:

- Advanced Research Projects Agency (Department of Defense):
  Member, ARPA Speech & Natural Language Standing Committee, 1992-1994
  Chair, Tipster Phase II Architecture Working Group, 1994-1998
- National Institutes of Standards and Technology
  Member, Organizing Committee, Text Analysis Conference, 2010-2015

Enabling Excellence

Prof. Grishman has graduated 24 PhD students, 8 of which are women and 4 female students are holding faculty positions:

- Carol Friedman: Professor of Biomedical Informatics at Columbia University;
- Patricia Teller: Professor of Computer Science at The University of Texas at El Paso;
- Heng Ji: Associate Professor of Computer Science at Rensselaer Polytechnic Institute;
- Wei Xu: Assistant Professor of Computer Science and Engineering at The Ohio State University.

He also supported and encouraged many international students to pursue a high-profile career in US, including quite a few students from Asian countries. According to a new report on diversity (http://goo.gl/HeTLzs) based on 2013 data filed with the U.S. Equal Employment Opportunity Commission, Asian Americans are often well-represented in lower-level positions — but, in comparison, severely underrepresented at the high-level management of US Tech companies (as well as NLP faculty positions in top US universities). The nominator is aware of 5 Asian-born students who came to US to pursue phd with Grishman and now hold senior management or faculty positions in US:

- Satoshi Sekine
- Kiyoshi Sudo
- Heng Ji
- Wei Xu
- Thien Huu Nguyen

Citation

Suggested citation if the candidate is selected. This should be a concise statement (maximum of 20 words) describing the key contributions for which the candidate merits this distinction. For example, "For significant contributions to unification-based parsing, finite-state morphology, and example-based machine translation."

For significant contributions to information extraction.

Basic information of nominator

Basic information about the nominator, including how he/she knows the candidate. Note that no single person may nominate more than 2 candidates for Fellow in each year.

Dr. Wei Xu received her PhD in Computer Science with a MacCracken Fellowship from New York University. Her research interests include semantics, data-driven paraphrasing and natural language generation. She was a postdoctoral researcher at the University of Pennsylvania. She is organizing the Workshop on Noisy User-generated Text, serving as a workshop chair for ACL 2017, an area chair for EMNLP 2016 and the publicity chair for NAACL 2016 and 2018. Prof. Ralph Grishman was her PhD advisor.