## Citation

For pioneering contributions to research in computational linguistics and natural language processing including machine translation and text mining

# Dr. Junichi Tsujii

**Positions and Organizations :**
Director, Artificial Intelligence Research Center,
National Institute of Advanced Industrial Science and Technology
Professor Emeritus, The University of Tokyo

**Doctorate :** Ph. D. (Kyoto Univ., 1978)

**Date of Birth :** February 7, 1949

**Brief Biography :**
1971  Graduated from Kyoto Univ., Faculty of Engineering, Undergraduate School of Electrical and Electronic Engineering
1973  Completed postgraduate master's program, Kyoto Univ.
1973  Research Associate, Faculty of Engineering, Kyoto Univ.
1978  Obtained Ph.D., Kyoto Univ.
1979  Associate Professor, Graduate School of Engineering, Kyoto Univ.
1988  Professor, Centre for Computational Linguistics, Univ. of Manchester
1995  Professor, Univ. of Tokyo
2005  Professor, Univ. of Manchester (concurrent post)
2011  Principal Researcher, Microsoft Research Asia
2015  Director, Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology, Japan

**Main Awards and Honors:**
1988  Japan IBM Science Award
2000  Hong Kong SEYMF Visiting Professorship Award
2004  Daiwa Adrian Prize
2005  IBM Faculty Award (Eclipse Innovation Award)
2008  Japanese Society for Artificial Intelligence (JSAI) Achievement Award
2010  Medal with Purple Ribbon
2013  Information Processing Society of Japan (IPSJ) Contribution Award
2014  Funai Achievement Award

Chair of the International Committee for Computational Linguistics (ICCL), Fellow of the IPSJ, Fellow of the ACL (Association of Computational Linguistics)

**Main Achievements :**

Dr. Junichi Tsujii is achieving major results in research on computational linguistics and natural language processing, particularly in the area of sentence structure parsing, which serves as the underlying technology of such research, and its applications in the fields of machine translation and text mining. In recent years he has also endeavored to utilize his research outcomes to create new scientific research methodologies through the development of artificial intelligence systems for the promotion of Big Science, which extracts and utilizes information from scientific literature.
Three of his representative achievements are cited below:

### 1) Research and development of machine translation
Dr. Tsujii played core roles in two machine translation projects – the Government of Japan's Mu Project at Kyoto University in the 1980s, and the European Community's Eurotra project at the University of Manchester in the 1990s. In the former, he undertook the construction of a practical Japanese/English, English/Japanese translation system for addressing article abstracts, and in the latter he achieved pioneering results that included defining the framework (Eurotra-7) for declarative grammatical descriptions incorporating the outcomes of computational linguistics. In addition, he has also been a driving force in the field as a worldwide leader, presenting keynote addresses on the results of these projects at two international conferences – Coling '86 (Bonn) and ACL '91 (Auckland).

### 2) Sentence structure parsing using unification grammar
Deep sentence parsing using grammar based on HPSG, CCG and other feature structures, and their unification, had long been considered impractical in terms of processing speed and tolerance. In response to this, a research group at the University of Tokyo led by Dr. Tsujii developed an abstract machine for unification, a probability model of feature structures, a beam-search technique, CFG filtering, supertagging and other methods, and demonstrated that deep sentence parsing displays superior performance to shallow sentence parsing not only in terms of accuracy but also in terms of processing speed. The group also proposed an empirical method for corpus-based grammar development and established methodologies for handling a wide range of linguistic phenomena.

### 3) Text mining for biomedicine
In contrast to conventional text mining, which perceives text as gatherings of words, Dr. Tsujii proposed a form of text mining that perceives the structure of sentences explicitly, and he verified the effectiveness of this new form by applying it to information extraction and text mining from a large collection of biomedical dissertations. In the process, he revealed a close relationship between the extraction of information from text and the definition of the ontology of the field concerned. He then defined the ontology of the field of biomedicine and carried out text annotation based on that definition. The outcome of this, the GENIA Corpus, has been widely adopted internationally as the gold standard of biomedical text mining. As an extension of this research, Dr. Tsujii has since created a system that connects the information in the text to a more global knowledge structure (biomedical pathway), and is proposing new methodologies for science based on text parsing.

As a result of these accomplishments, Dr. Tsujii has received a large number of awards, among them the Japan IBM Science Award, Daiwa Adrian Prize, Japanese Society for Artificial Intelligence (JSAI) Achievement Award, Information Processing Society of Japan (IPSJ) Contribution Award, and the Medal with Purple Ribbon. Furthermore, as a world leader in the fields of machine translation and language processing, in 1992 he was made a permanent member of the International Committee for Computational Linguistics (ICCL), and was elected as its Chair in 2014. Additionally, he has served as President for the International Association for Machine Translation (IAMT) and the Association for Computational Linguistics (ACL), and established the Asian Federation of Natural Language Processing (AFNLP), where he also served as President, as an organization that brings together researchers in Asia.

On the personnel training front as well, Dr. Tsujii has produced a large number of outstanding personnel through his teaching posts at Kyoto University, the University of Tokyo and the University of Manchester. He has also nurtured young researchers around the world as Head of the University of Manchester Centre for Computational Linguistics (CCL) (1990-95), Head of the University of Manchester National Centre for Text Mining (2005-08), Principal Researcher at Microsoft Research Asia (Beijing) (2011-15), and Director of the National Institute of Advanced Industrial Science and Technology's Artificial Intelligence Research Center.

As can be seen, along with conducting pioneering research in the fields of machine translation, language processing and text mining, Dr. Junichi Tsujii has contributed to training a succession of personnel, as well as the implementation of national projects in those fields. His contributions to science, industry and society are considerable, and he is presented with this Okawa Prize in celebration of those achievements.