4th Workshop on NLP for Software Engineering

Co-located with ESEC/FSE - https://2018.fseconference.org/
November 4th, 2018
Lake Buena Vista, FL, USA
https://nl4se2018.github.io/

Important Dates

August 31st, 2018:	**Paper submission**
October 1st, 2018:	Author notification
October 15th, 2018:	Camera-ready

Organizing Committee

Yijun Yu (The Open University, UK)
Erik Fredericks (Oakland University, USA)
Premkumar Devanbu (UC Davis, USA)

Research Topics

This interdisciplinary workshop will explore issues related to the statistical modeling of software corpora, including topics such as:

- Modeling repetitiveness ("naturalness") in source code
- Applications to code suggestion in IDEs
- Mining programming idioms
- Statistical inference of types and other annotations
- Applications of Statistical Machine Translation for porting and reverse engineering
- Statistical methods for bug localization
- Statistical methods for automatic code patching, code summarization, code retrieval, code annotation, or test generation
- Formal and informal methods for enhancing assurance via NLP techniques

We are pleased to announce the 4th workshop on natural language processing for software engineering (NL4SE) to be held at **Lake Buena Vista, FL**.

This interdisciplinary workshop will explore issues related to the statistical modeling of software corpora, including topics such as: modeling repetitiveness in source code; use of language models for the code suggestion in IDEs; using probabilistic grammars to mine programming idioms; statistical methods for type inference in a dynamically typed languages; statistical machine translation for porting applications between programming languages, or “mini-fying”JavaScript; using statistical language models to find bugs; or statistical methods for automatic code patching, code summarization, code retrieval, code annotation, or test generation.

The workshop follows several earlier workshops on this topic at Microsoft Research, Dagstuhl event, SIGSOFT FSE, and AAAI.

We invite short position papers or early-stage research papers of at most 4 pages in length. Several submissions will be invited for presentation.

Direct link to EasyChair submission website:
https://easychair.org/conferences/?conf=nl4se18

Software Corpora

If you are looking for software corpora to study, we recommend the following:

- Conala Corpus (https://conala-corpus.github.io/), curated by Carnegie Mellon University
- NL2Bash (https://github.com/TellinaTool/nl2bash), curated by the University of Washington

You are encouraged to use a dataset to demonstrate your technology.

More details about the workshop may be found on our website: https://nl4se2018.github.io/