

# Автоматическая рекомендация имен методов в IntelliJ IDEA

Автор: З.И. Курбатова, 471 группа

Научный руководитель: к. т. н., доц. Т.А. Брыксин

Рецензент: аналитик ООО "Интеллиджей Лабс" Н.И.Поваров

Санкт-Петербургский Государственный Университет

Кафедра системного программирования

# Мотивация

Плохо подобранные имена идентификаторов

- Усложняют понимание кода
- Допускают двусмысленность
- Приводят к ошибкам
  - Статья “Relating identifier naming flaws and code quality: An empirical study”, Butler et al

# Постановка задачи

Цель: разработать плагин для IntelliJ IDEA, который автоматически рекомендует имена для методов

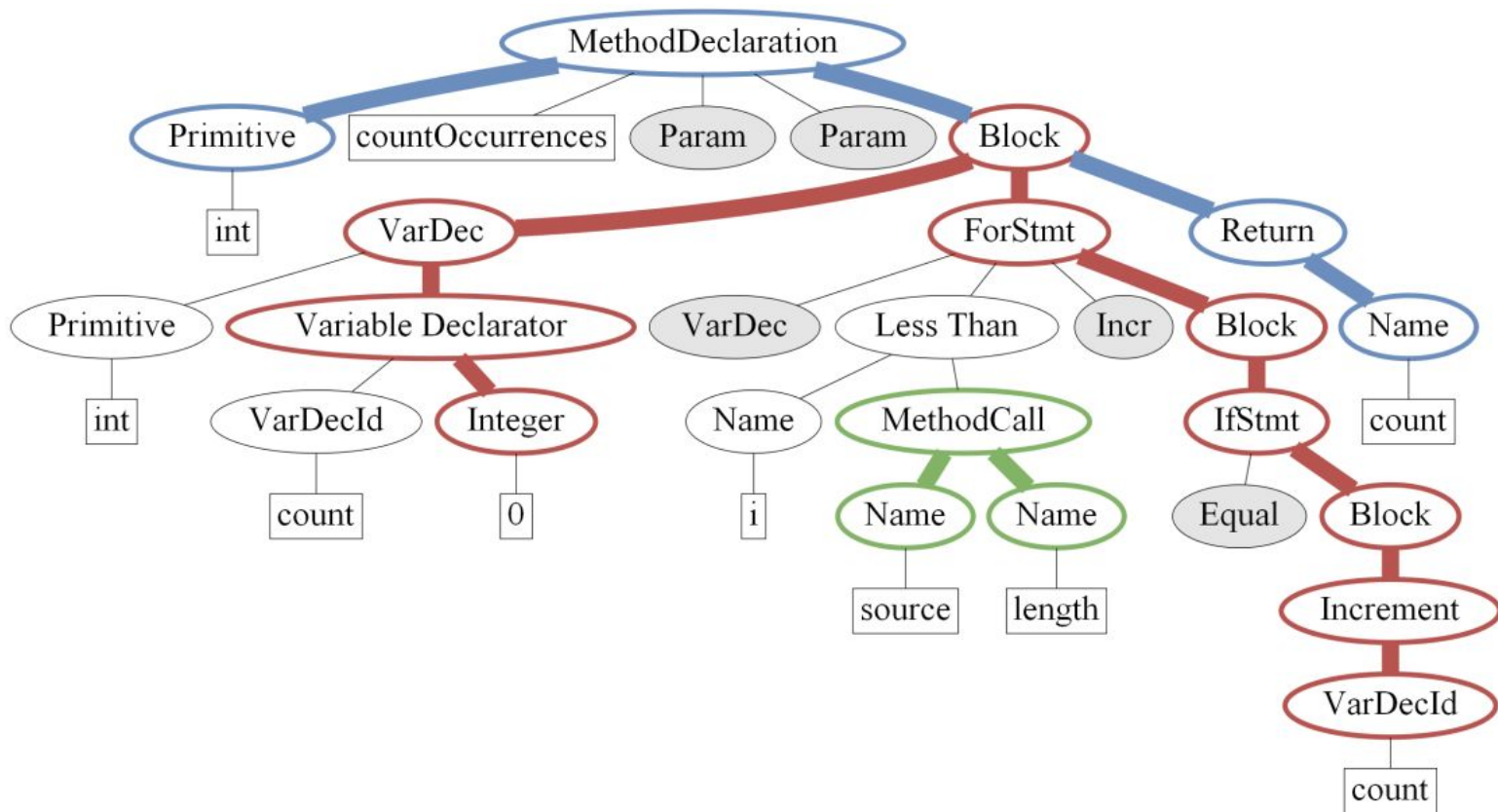
Задачи:

- Провести обзор предметной области
- Выбрать алгоритм рекомендации имен методов
- Разработать архитектуру плагина
- Выполнить апробацию

# Рекомендация имен методов

- ConvAttention, 2016
  - Предлагаем неологизмы
  - Плохо масштабируется
- Code2vec, 2018
  - На вход принимает множество путей в AST
  - Предлагает значения из словаря -> размер модели 1.5Gb
- **Code2seq**, 2019
  - На вход принимает множество путей в AST
  - Предлагает неологизмы

# code2seq: Построение AST

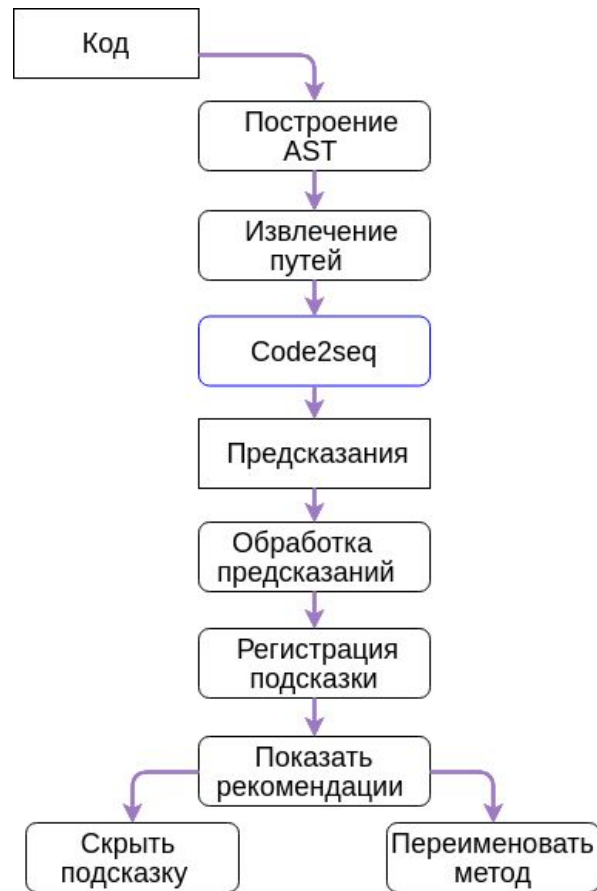


# Режимы работы

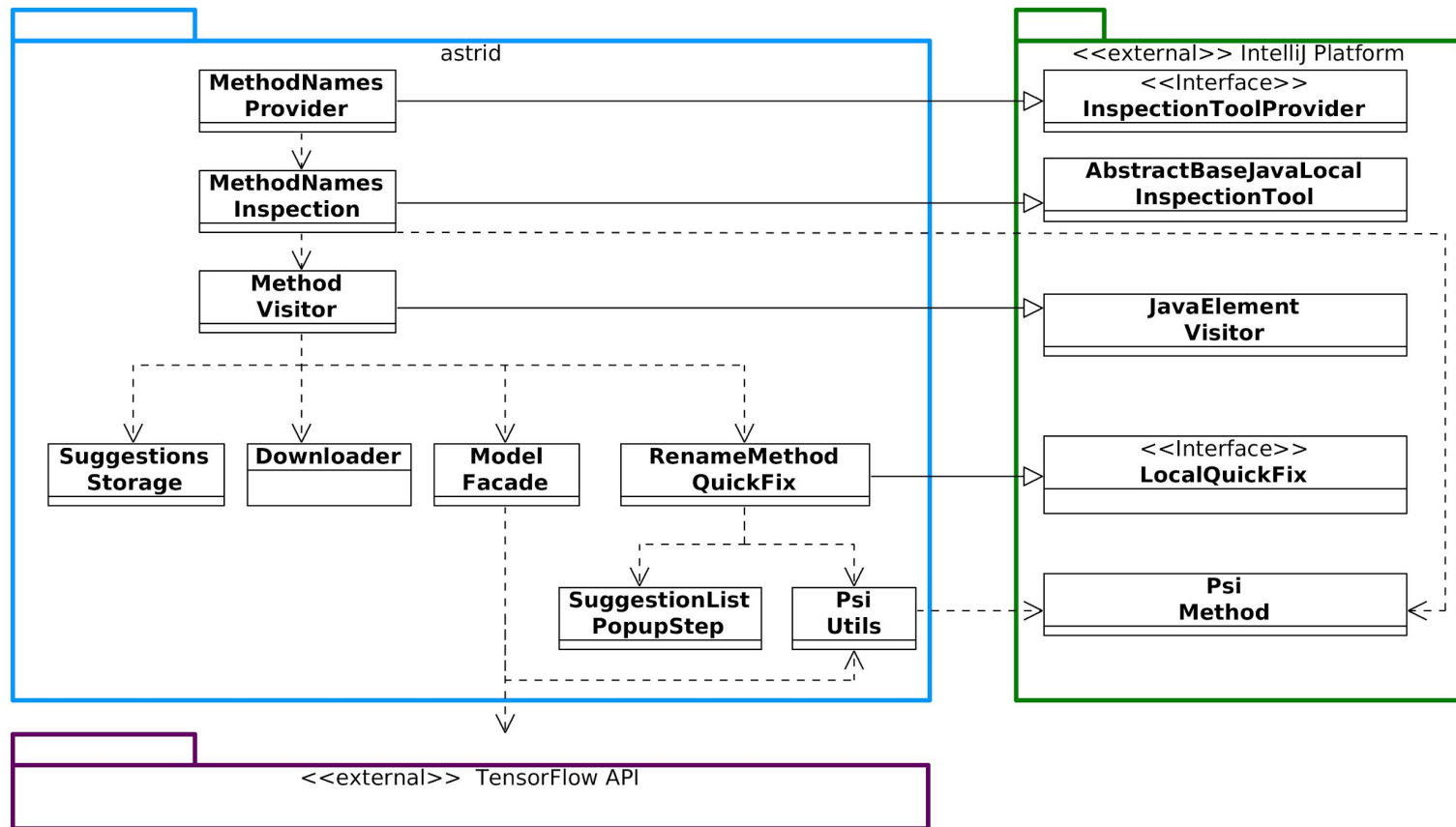
- Генерация рекомендаций по запросу пользователя
- Автоматическая инспекция для рекомендации имен
  - Генерация рекомендаций в фоновом режиме
  - Подсветка имени метода и список рекомендаций
  - При выборе рекомендации совершается rename refactoring
- Автоматическая инспекция для поиска больших условных выражений

# Предлагаемый подход

- Перед посещением метода проверка
  - Не переопределяет метод наследуемого класса
  - Не реализует метод интерфейса
- Обработка предсказаний
  - Удаление коротких слов, лишних символов
- Регистрация подсказки
  - Если предсказания не содержат текущее имя
  - Подсветка имени метода



# Архитектура плагина





# Сбор статистики использования плагина

- Сохранение в формате XML
- Длина метода (количество строк)
- Логарифм вероятности предсказания модели
  - Условная вероятность -> маленькие значения, проще работать с логарифмом

$$p(y_1, \dots, y_n | X) = \prod_{i=1}^n p(y_i | y_{<i}, X)$$

# Пример: подсказка

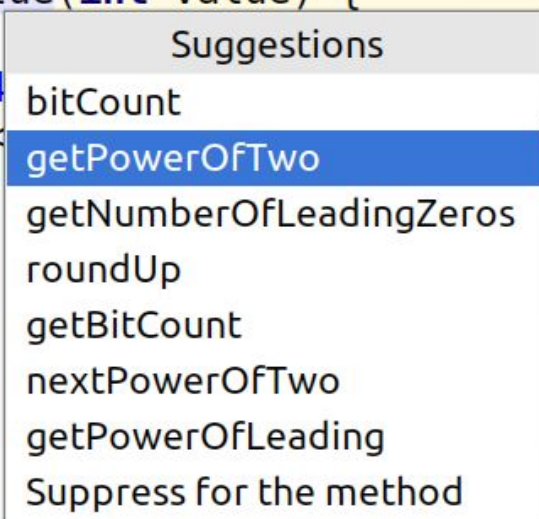
```
public static int getValue(int value) {  
    return value <= 0 ? 1 :  
           value >= 0x40000000 ? 0x40000000 :  
           1 << (32 - Integer.numberOfLeadingZeros(i: value - 1));  
}
```

```
public static int getValue(int value) {  
    return value <= 0 ?  
           value >= 0x4  
           1 <<  
}
```

- 💡 Get method name suggestions ▶
- 🔗 Bind method parameters to fields ▶
- 🔗 Generate missed test methods ▶
- 🔗 Insert '@Contract(pure = true)' ▶
- 🔗 Add Javadoc ▶
- 🔗 Generate overloaded method with default parameter values ▶
- 🔗 Hide gutter icon for annotations inferred from source code ▶
- 🔗 Make 'private' ▶
- 🔗 Make 'protected' ▶
- 🔗 Make package-private ▶

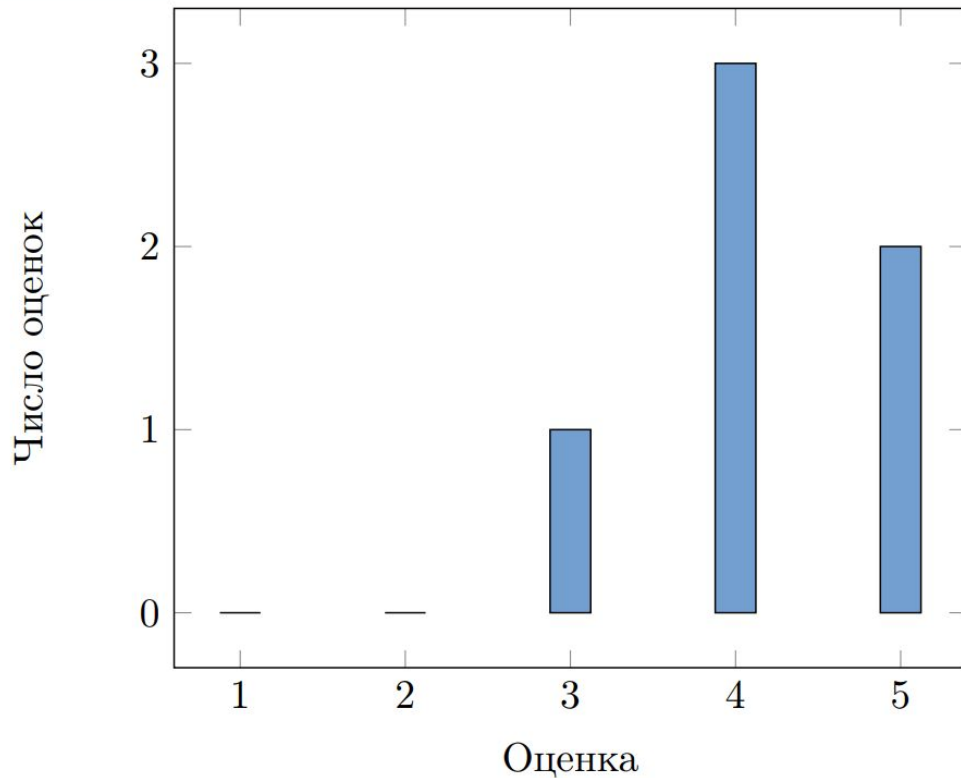
# Пример: список рекомендаций

```
public static int getValue(int value) {  
    return value <= 0 ?  
        value >= 0x40000000 ?  
            1 << (Integer.numberOfLeadingZeros(i: value - 1));  
}
```



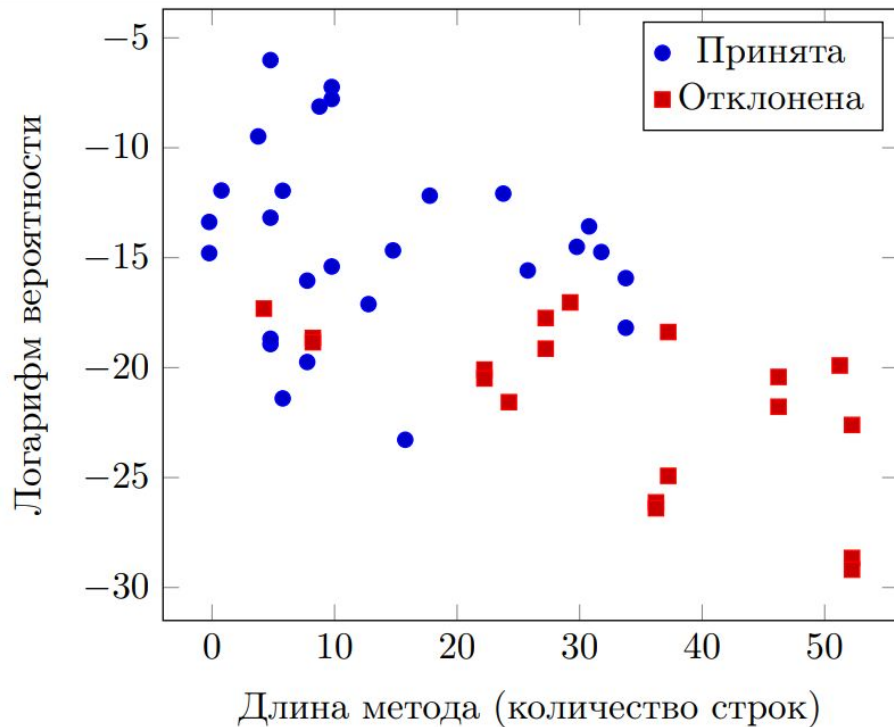
# Апробация

- Язык — Java
- 6 разработчиков
- Оценка
  - Полезность



# Апробация

Релевантность рекомендаций в зависимости от длины метода



# Результаты

- Выбран алгоритм рекомендации имен методов
- Разработана архитектура плагина к IntelliJ IDEA
- Реализовано два инструмента рекомендации имени метода
  - Выдача рекомендаций по запросу пользователя
  - Анализ кода в фоновом режиме и регистрация подсказок
- Реализован модуль сбора статистики использования плагина
- Проведена апробация плагина