



Exploring the Effect of Multiple Natural Languages on Code Suggestion Using GitHub Copilot

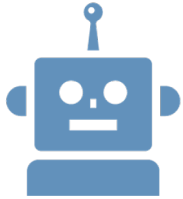
©Kei Koyanagi, Dong Wang, Kotaro Noguchi, Masanari Kondo
Alexander Serebrenik, Yasutaka Kamei, Naoyasu Ubayashi



九州大学
KYUSHU UNIVERSITY

TU/e EINDHOVEN
UNIVERSITY OF
TECHNOLOGY

Background: Development using AI



GitHub Copilot (Coding Support)

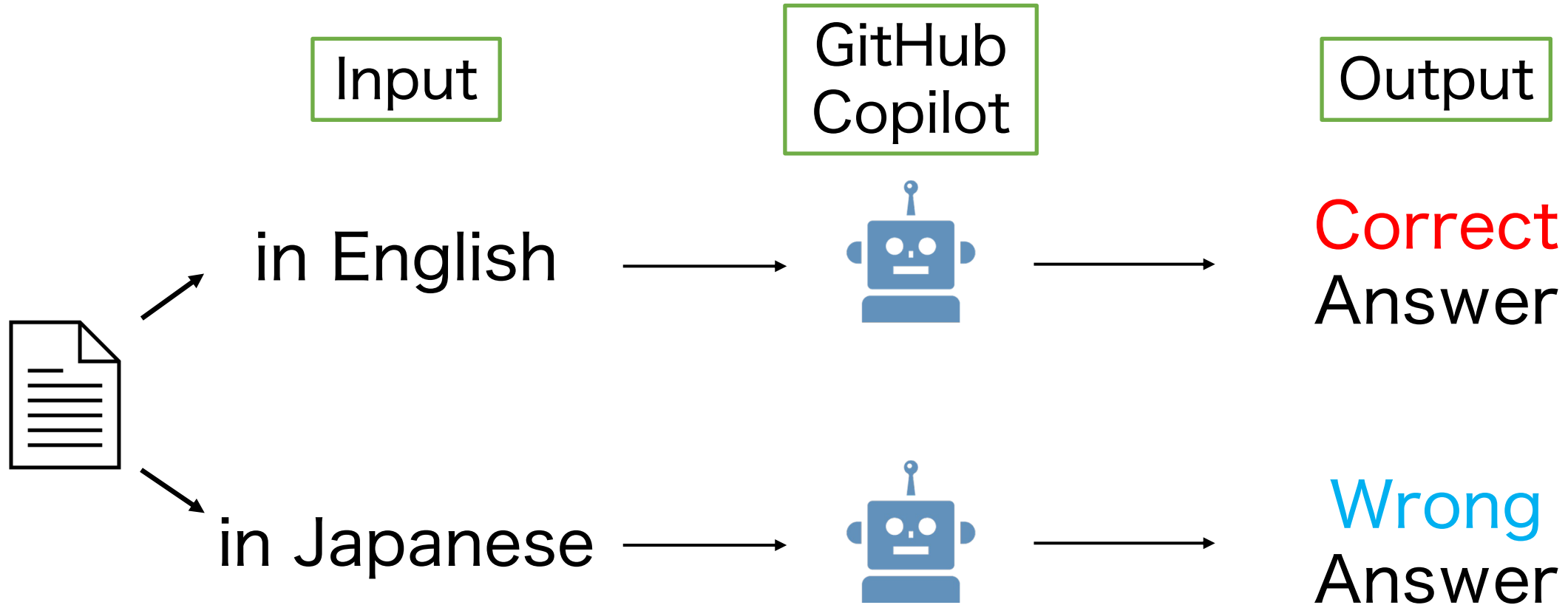
```
1 import datetime
2
3 def parse_expenses(expenses_string):
4     """Parse the list of expenses and return the list of triples (date, value, currency).
5     Ignore lines starting with #.
6     Parse the date using datetime.
7     Example expenses_string:
8         2016-01-02 -34.01 USD
9         2016-01-03 2.59 DKK
10        2016-01-03 -2.72 EUR
11    """
```

Input

```
12    expenses = []
13    for line in expenses_string.splitlines():
14        if line.startswith("#"):
15            continue
16        date, value, currency = line.split(" ")
17        expenses.append((datetime.datetime.strptime(date, "%Y-%m-%d"),
18                        float(value),
19                        currency))
20    return expenses
```

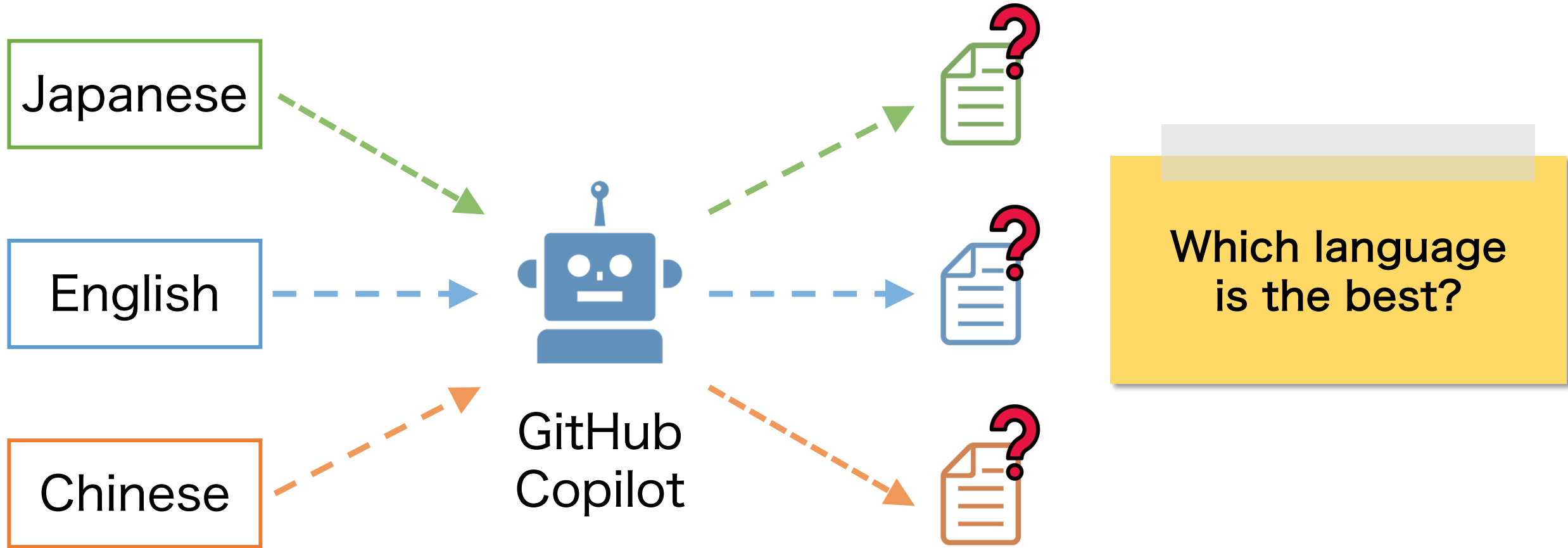
Recommendation

Background: The effect of input language

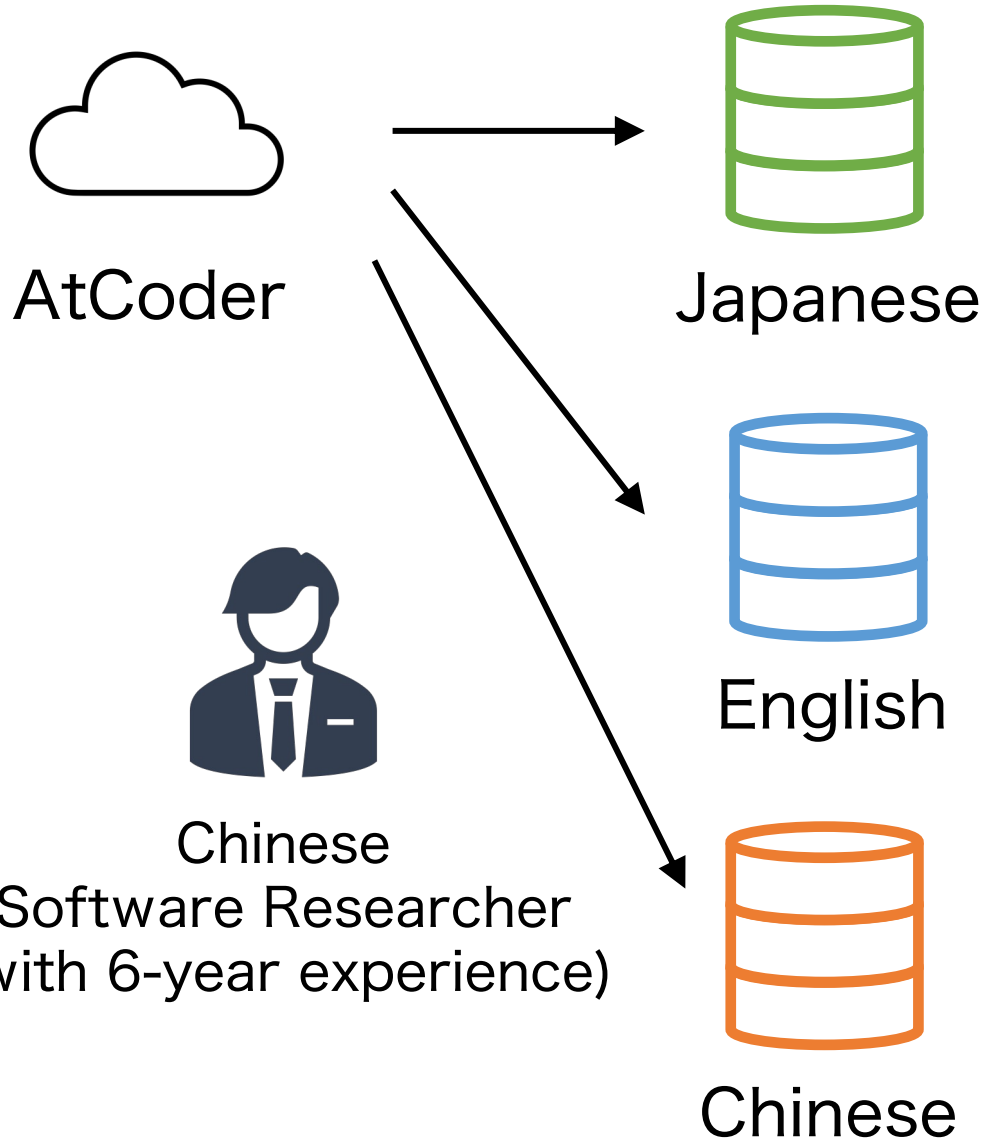


We should give appropriate input to use AI efficiently.

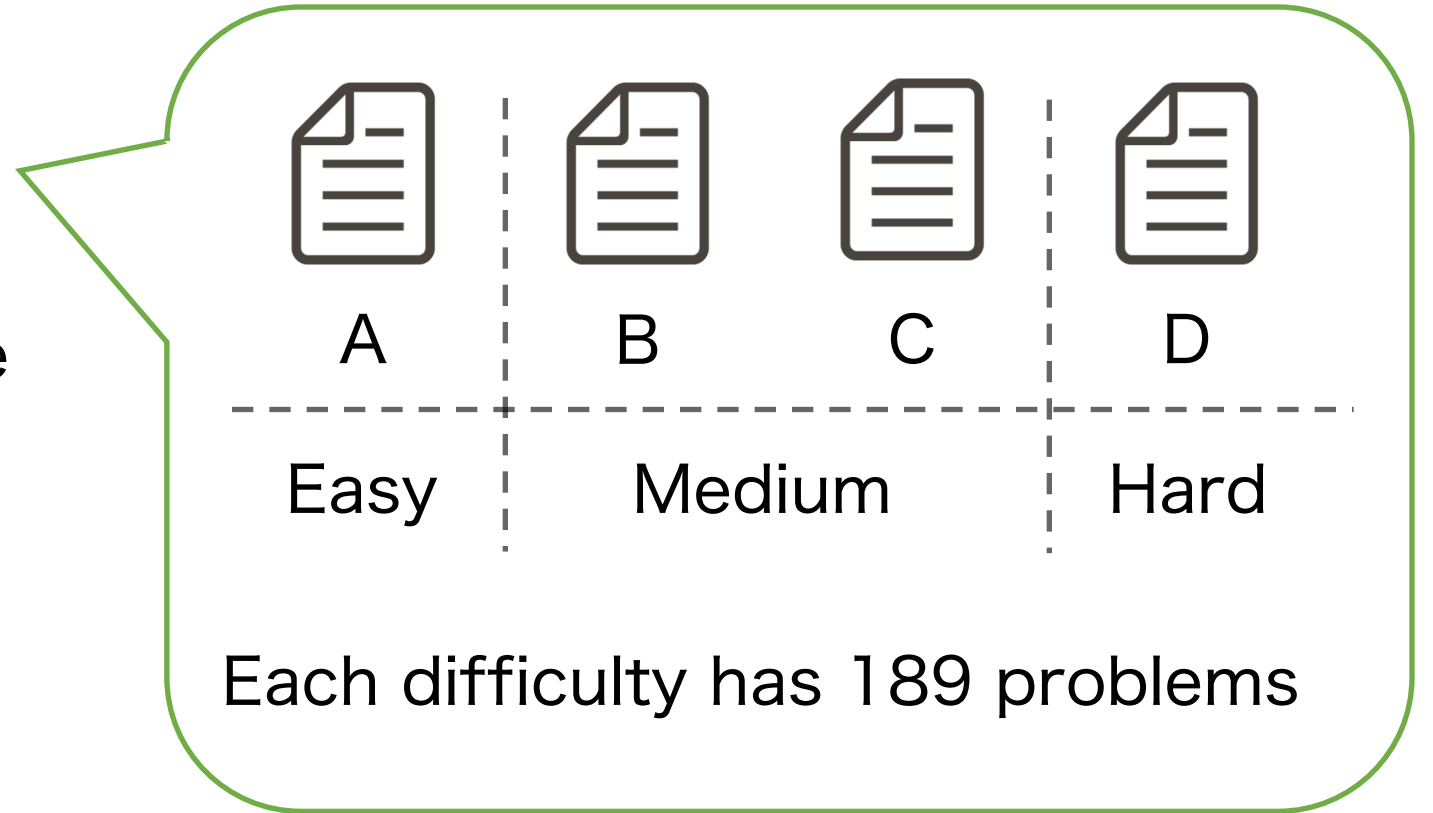
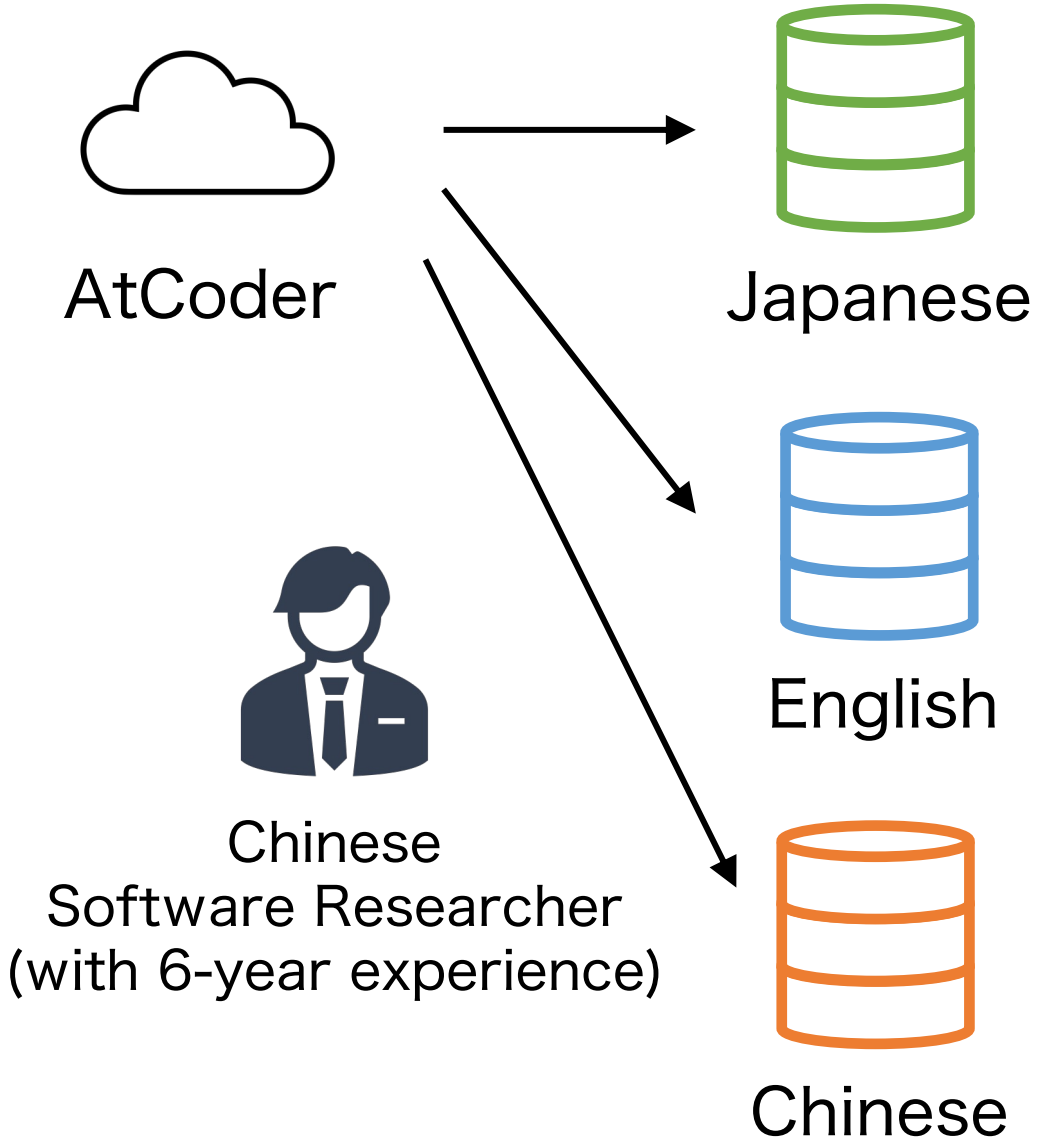
Goal: Understand how language differences affect GitHub Copilot performance



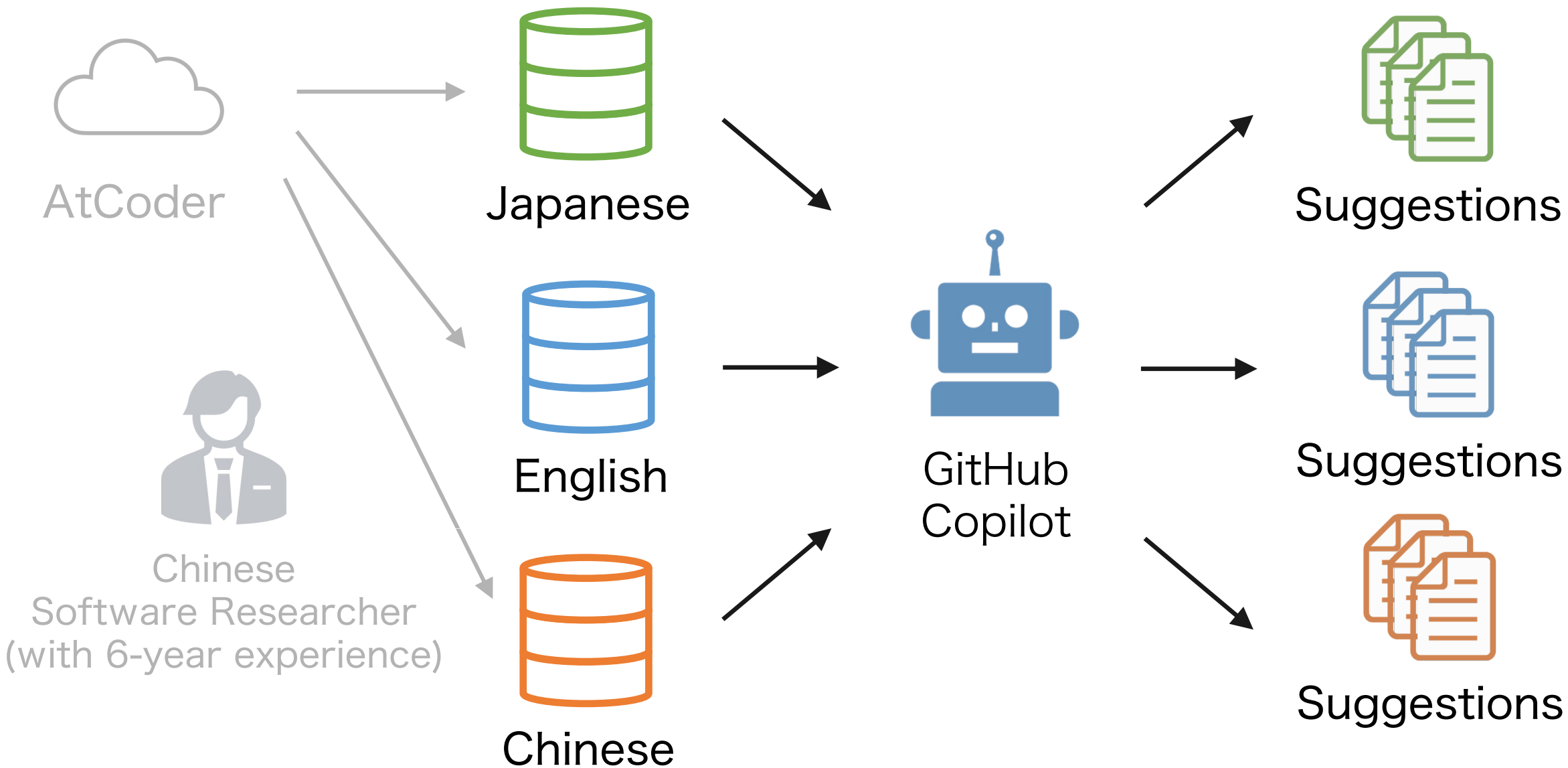
Experiment



Experiment

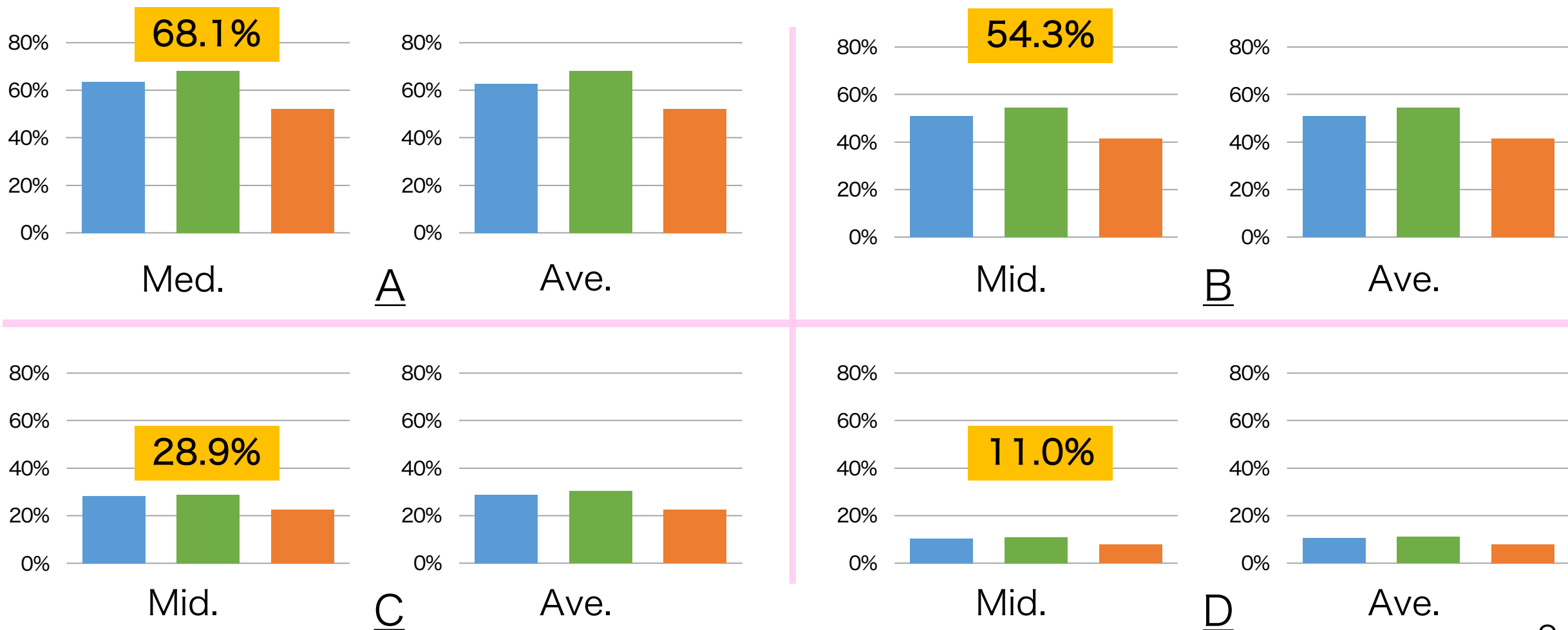


Experiment



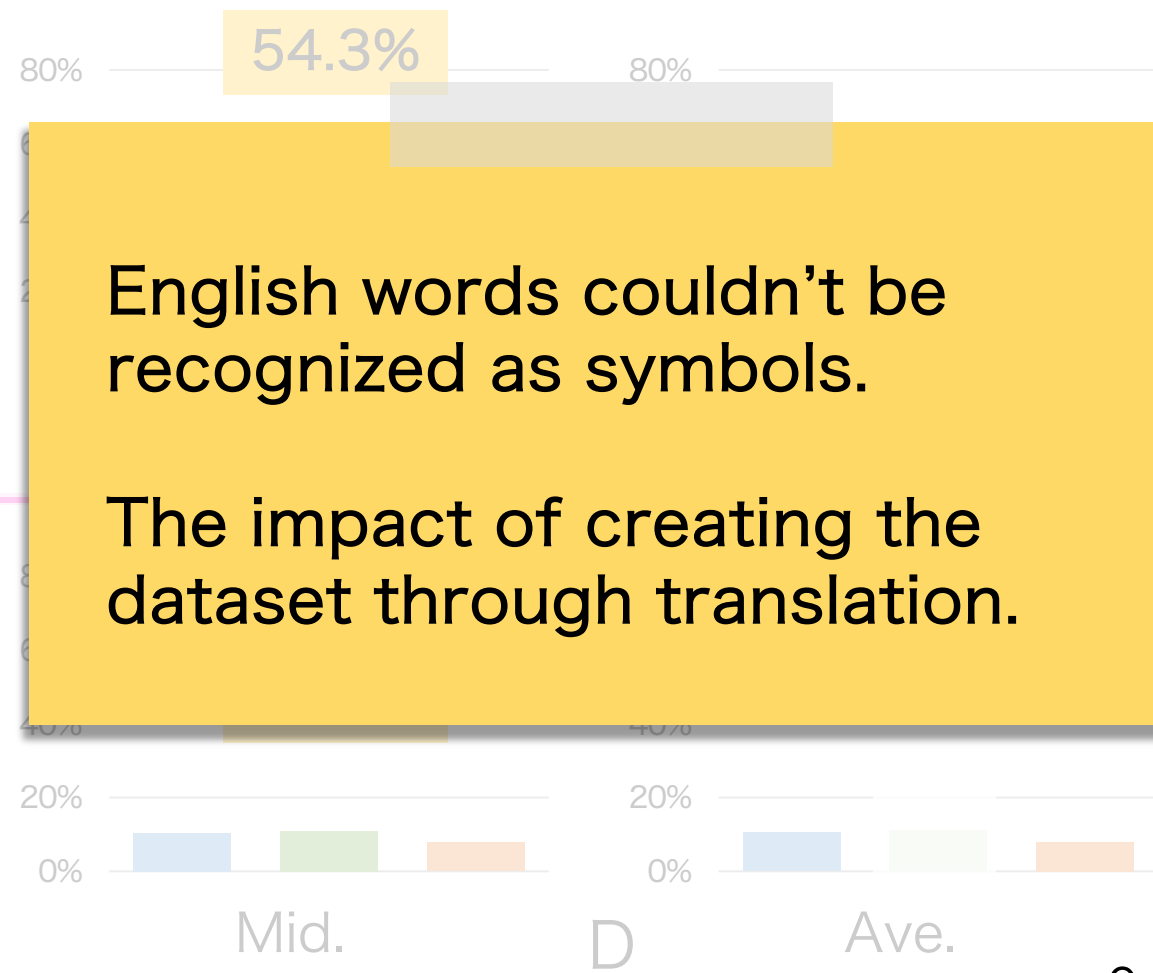
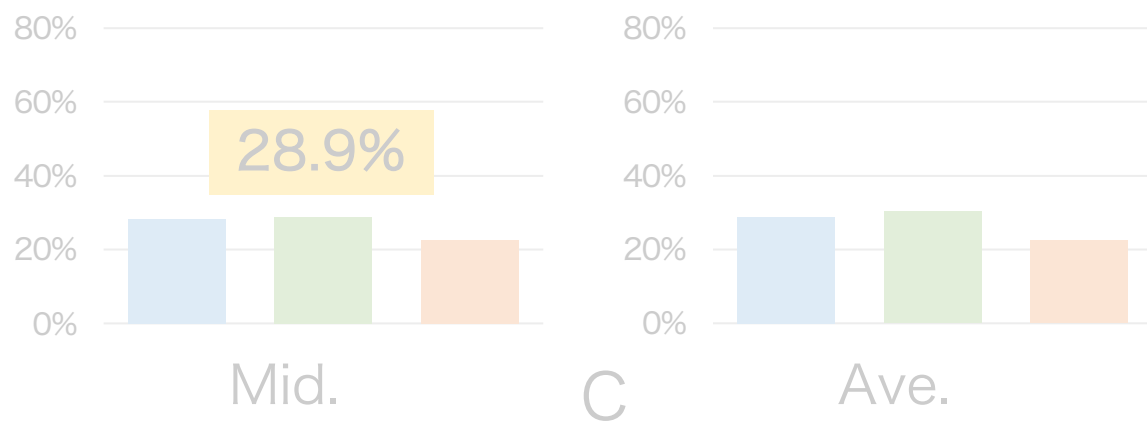
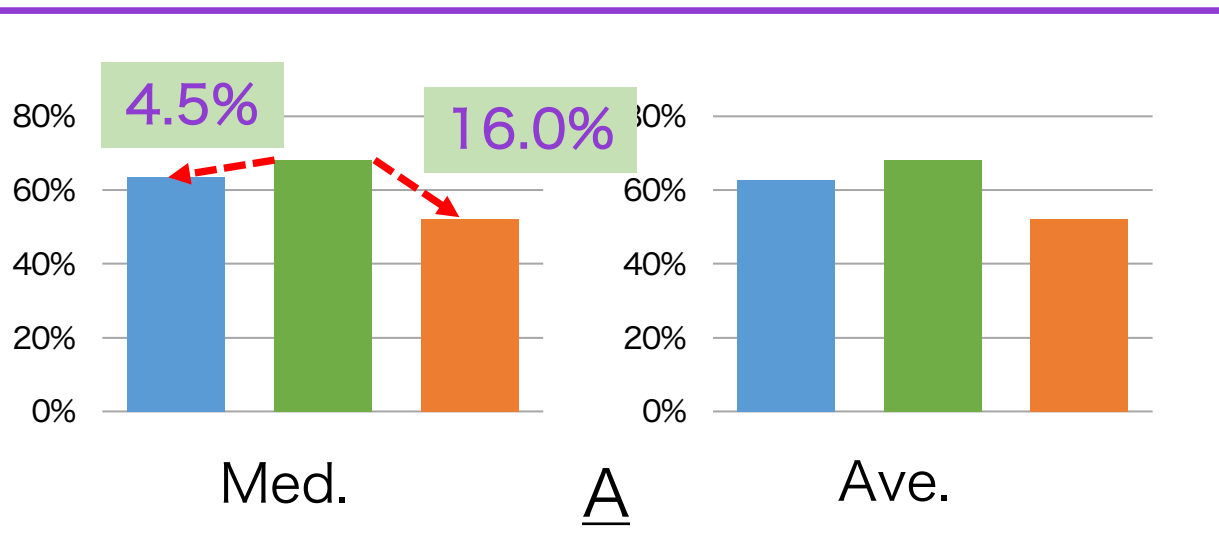
Results: Japanese recorded the highest *Accuracy*

English Japanese Chinese

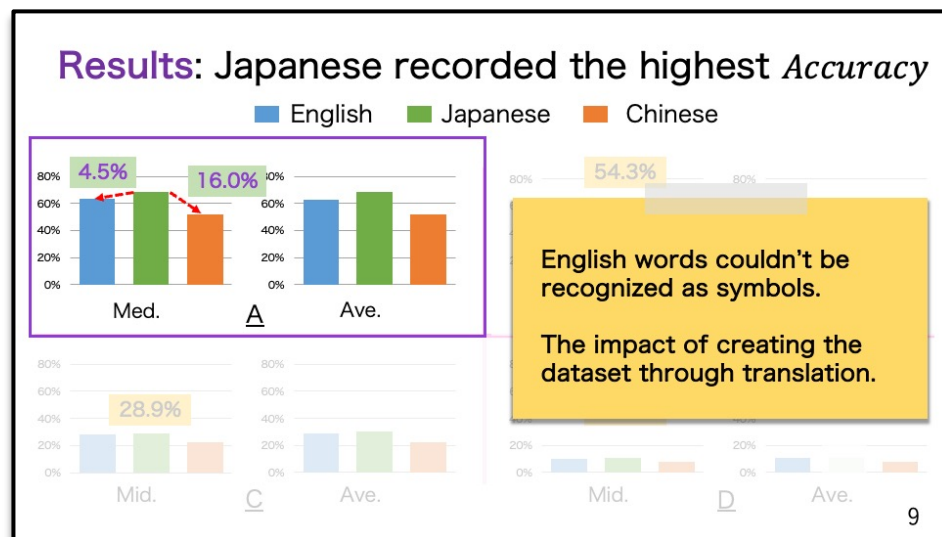
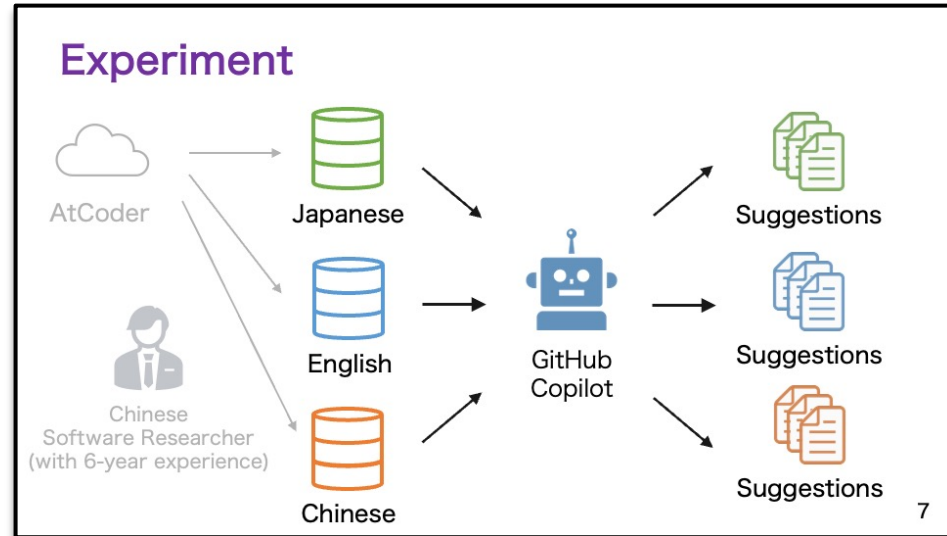
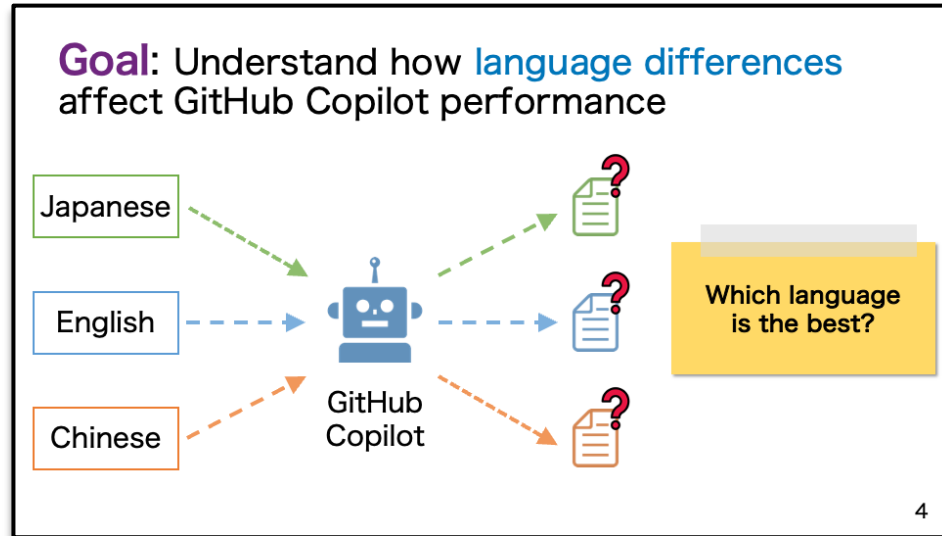


Results: Japanese recorded the highest *Accuracy*

English Japanese Chinese



Summary: The capability varies across natural languages.



Contact:

Email: kondo@ait.kyushu-u.ac.jp

Twitter: @masanari_kondo