

SB Intuitions

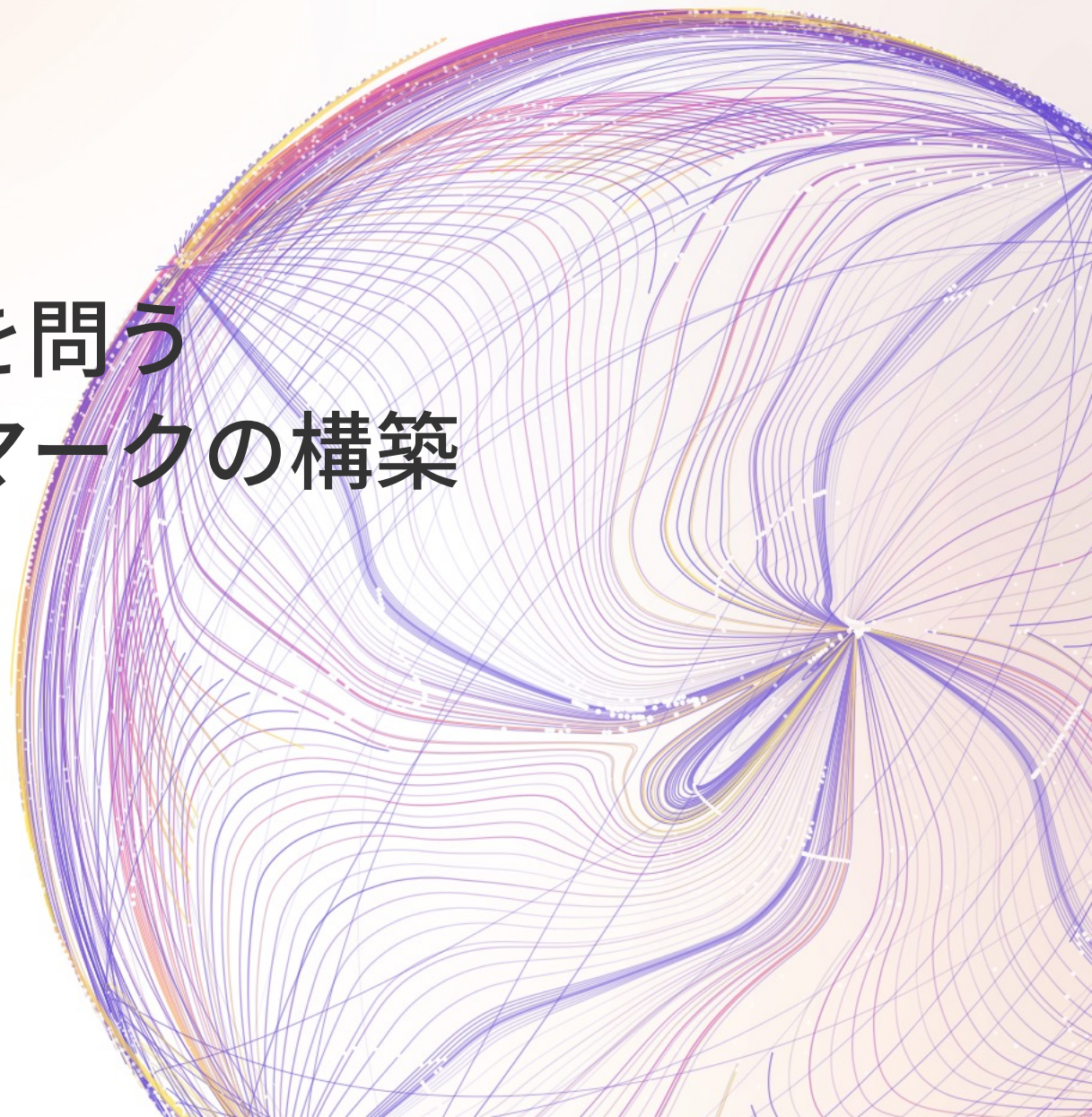
Japanese Multiple Choice QA

JamC-QA: 日本固有の知識を問う 多肢選択式質問応答ベンチマークの構築

2025/03/14 (Fri) @JLR2025

SB Intuitions 株式会社

○岡照晃, 柴田知秀, 吉田奈央



国産LLMを作る意義:

日本固有の知識を獲得したLLMを作りたい！



作っても評価できるベンチマークがない！

既存のベンチマークではスコアが頭を打っている！

SB Intuitions

JCommonsenseQA

常識推論能力 [Kurihara+, 22]

水を出すときに捻るものは？

A. 蛇口

B. ハンドル

C. 流し

D. 釘

E. 食器棚

JMMLU (日本問題)

世界知識・問題解決能力

(日本史, 公民, 地理, 熟語)

<https://github.com/nlp-waseda/JMMLU>

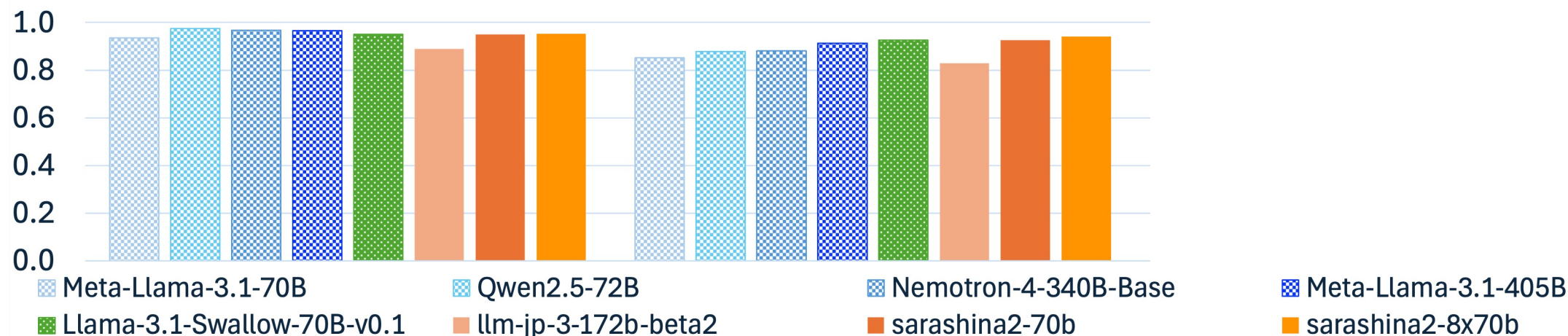
衆議院議員の任期は何年か？

A. 4年

B. 3年

C. 6年

D. 5年



JCommonsenseQA

常識推論能力 [Kurihara+, 22]

水を出すときに捻るものは？

A. 蛇口

B. ハンドル

C. 流し

D. 釘

E. 食器棚

JMMLU (日本問題)

世界知識・問題解決能力
(日本史, 公民, 地理, 熟語)

<https://github.com/nlp-waseda/JMMLU>

衆議院議員の任期は何年か？

A. 4年

B. 3年

C. 6年

D. 5年

JamC-QA

日本固有の知識
(文化, 風習, ...)

New!



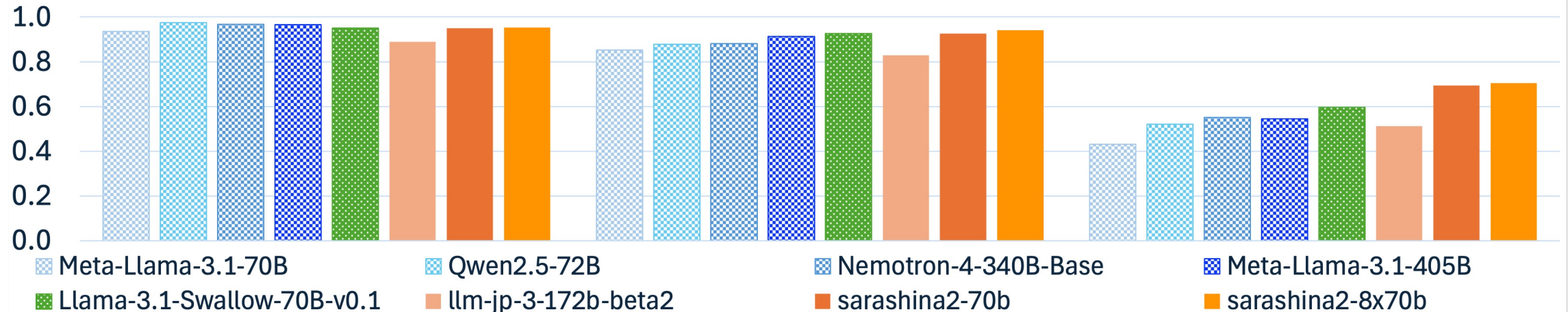
お墓参りに適切でない花は？

A. 菊

B. キンセンカ

C. カーネーション

D. バラ

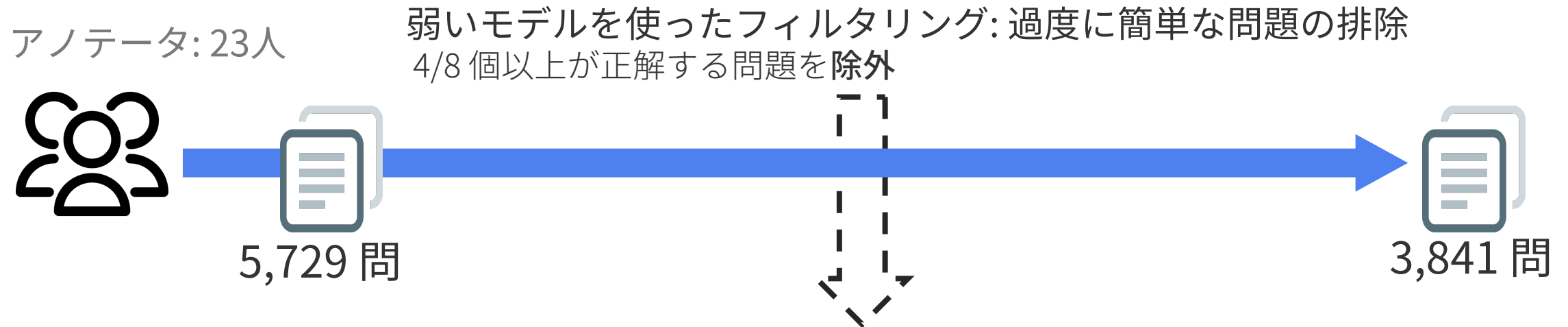


🇯🇵 日本の文化や風習に特化した四択問題

- 文化, 風習, 風土, 日本史, 行政, 医療, 法律の 7カテゴリ

✎ 社内専属の日本人アノテータがスクラッチから問題を作成

- Web 上の文章 (e.g., Wikipedia) をそのまま問題・選択肢にしない
- 弱いモデルの正解率をアノテータにフィードバックし、難易度を統制



「東北四大祭り」のひとつに数えられる「仙台七夕まつり」は、どこの県で行われるか

A. 山形県

B. 青森県

C. 秋田県

D. 宮城県

文化（日本語）：

小春日和が表す季節を選べ			
A. 夏から秋	B. 冬から春	C. 春から夏	D. 秋から冬

風習（マナー）：

お歳暮として送っていけないものは次のうちどれ？			
A. ハンカチ	B. ジャム	C. 鮭	D. タオル

法律：

ドラッグストアなどで、指定第2類の医薬品は薬剤師等がいるカウンターから何メートルの範囲内にある棚に置かなくてはならない？			
A. 5メートル	B. 7メートル	C. 3メートル	D. 10メートル

- 七草粥に入っていない草は
- 七草粥はせり・なずな・ごぎょう・はこべら・すずな・すずしろともう一つの植物は何？
- 春の七草は七草粥で有名だが秋の七草でなでしこ・はぎ・おばな・くず・おみなえし・ふじばかまとその他一つは何か？
- 春の七草粥でセリ・ナズナ・ゴギョウ・ハコベラ・ホトケノザ・スズシロとあとひとつは何
- 「仏体」を表す縁起物の意味の春の七草は次のうちどれ？



問題のネタになる情報を Slack 上で出し合い共有

👍 JamC-QA はデータ整備の後、公開を予定しています



直感を、知性へ