Temporal Artefacts

Can Language Models Handle a Non-Gregorian Calendar?

According to Sasaki, M., Kamoda, G., Takahashi, R., et al.

Calendars

Gregorian Calendar

January, February, March…

Hijri Calendar

Muharram, Safar, Rabi Al-Awwal…

Hebrew Calendar

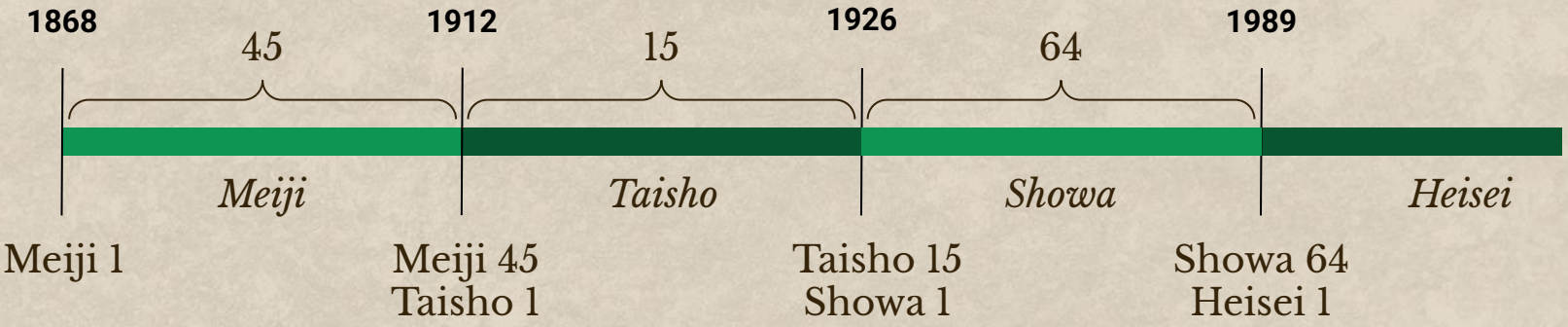
Tishrei, Cheshvan, Kislev…

Japanese Calendar

?

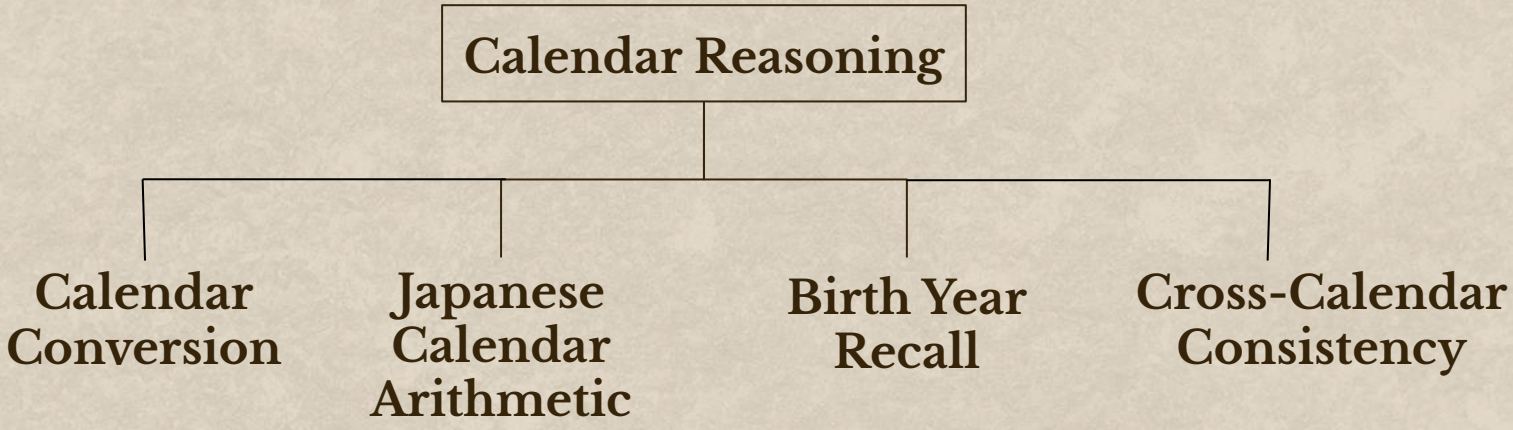
Japanese *Wareki* Calendar

Gregorian Calendar Years



Japanese Wareki Eras and Years

Gregorian-Wareki Reasoning



Language Models

- LLM-JP-3-13B
- Sarashina2-13B
- Swallow-13B
- Swallow-MS-7B
- LLaMA3-Swallow-8B

- LLaMA-2-7B
- LLaMA-2-13B
- Mistral-7B
- LLaMA3-8B

1. Calendar Conversion

Accuracy-Announcer Functions

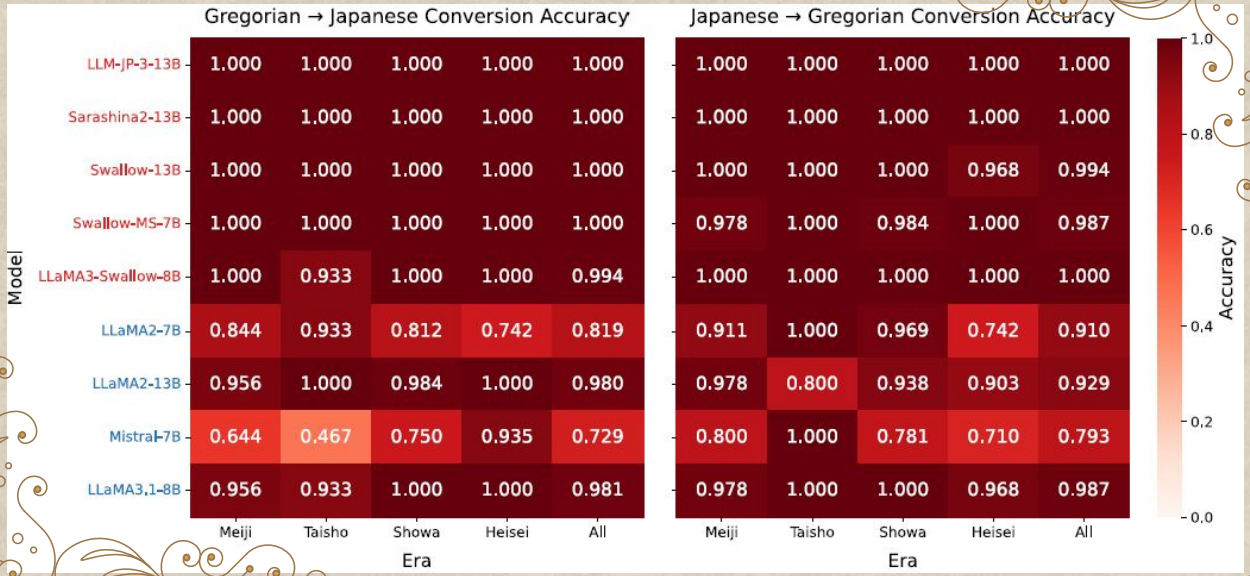
Japanese to Gregorian

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(\hat{y}_i = y_i)$$

Gregorian to Japanese

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(Ei = \hat{\mathrm{E}}_i \wedge x_i = \hat{x}_i)$$

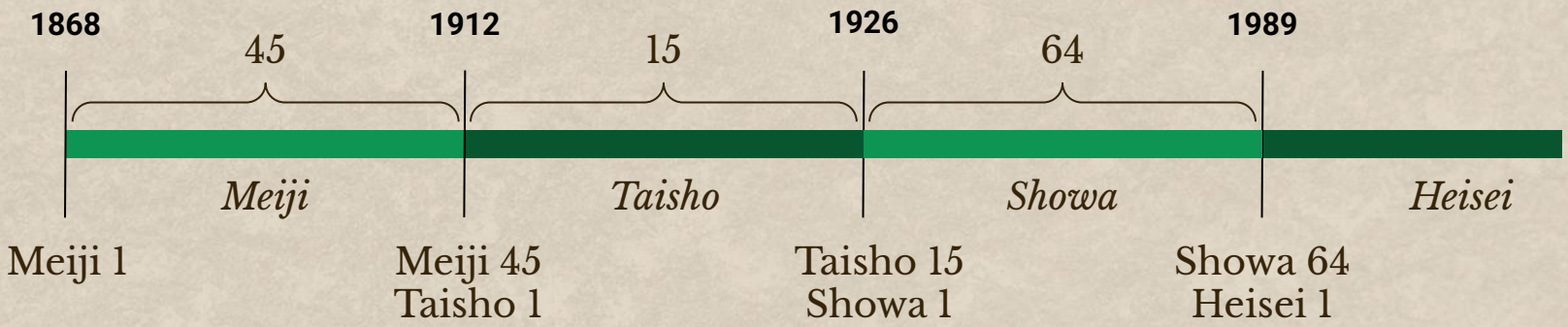
1. Calendar Conversion



2. Japanese Calendar Arithmetic

Question: What year in the Japanese calendar is 10 years after Meiji 44?

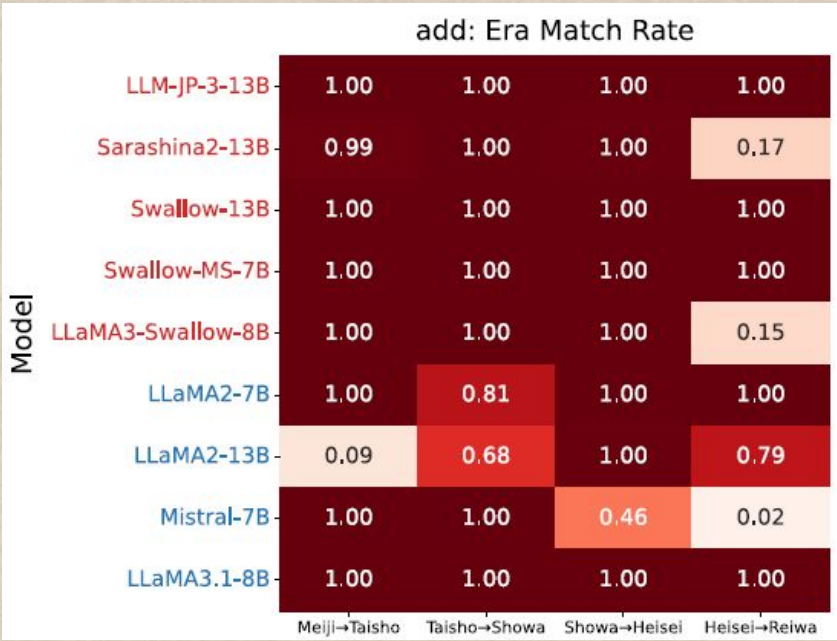
Link to Mentimeter: https://www.menti.com/al2zeieuj88p
Quiz Code: 7531 8685



2. Japanese Calendar Arithmetic

Ratio of outputs with the correct era

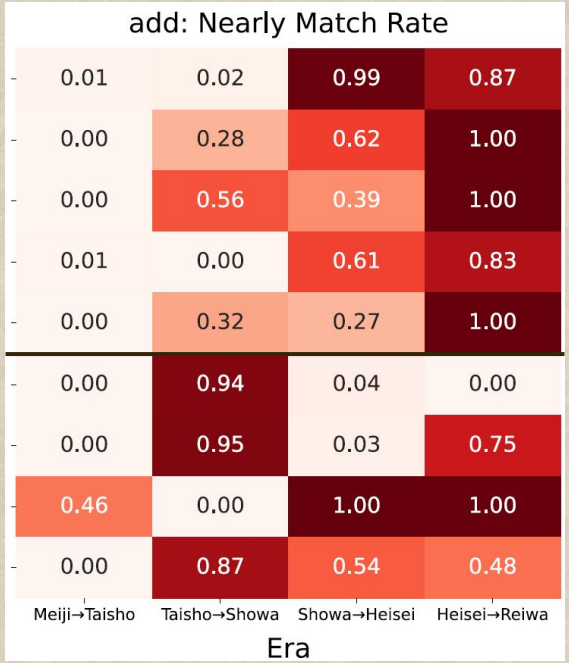
$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(\hat{E}_i = E_i)$$



2. Japanese Calendar Arithmetic

Ratio of outputs of a margin of error of a year

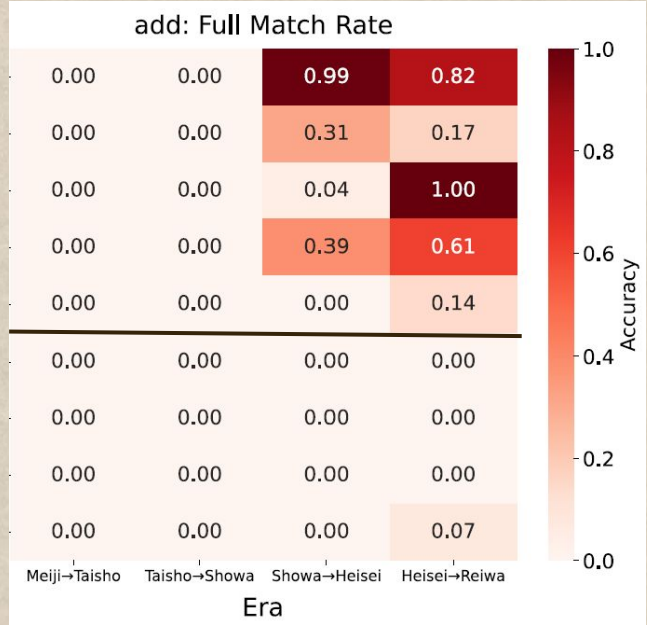
$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(|G(E_i,\ x_i) - G(\hat{E}_i, \hat{x}_i)|\ \leq\ 1)$$



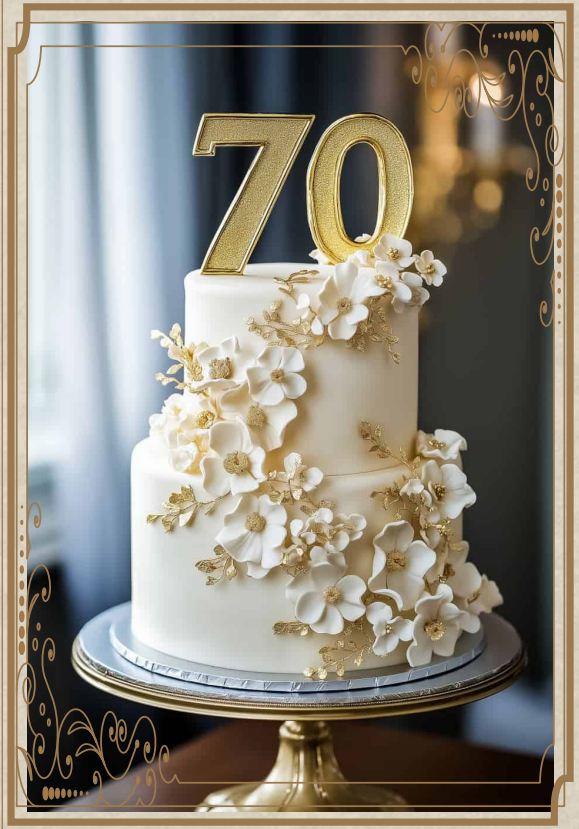
2. Japanese Calendar Arithmetic

Ratio of outputs with the correct ***exact year***

$$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(Ei = \hat{\mathrm{E}}_i \wedge x_i = \hat{\mathrm{x}}_i)$$

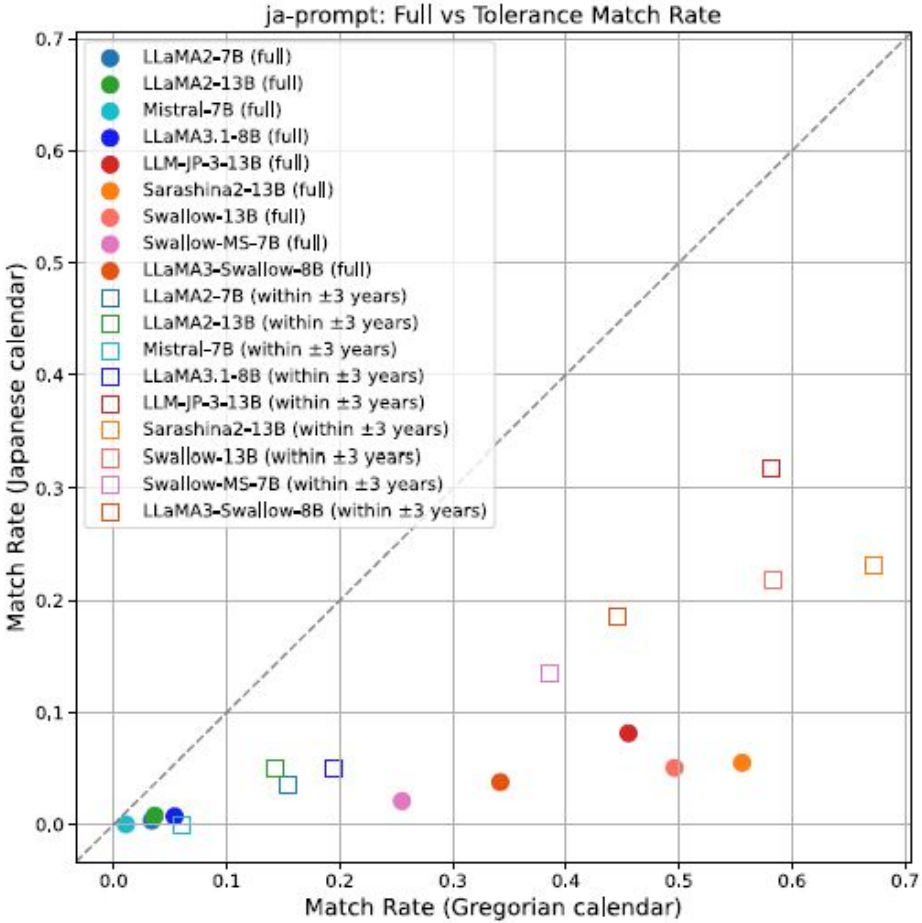


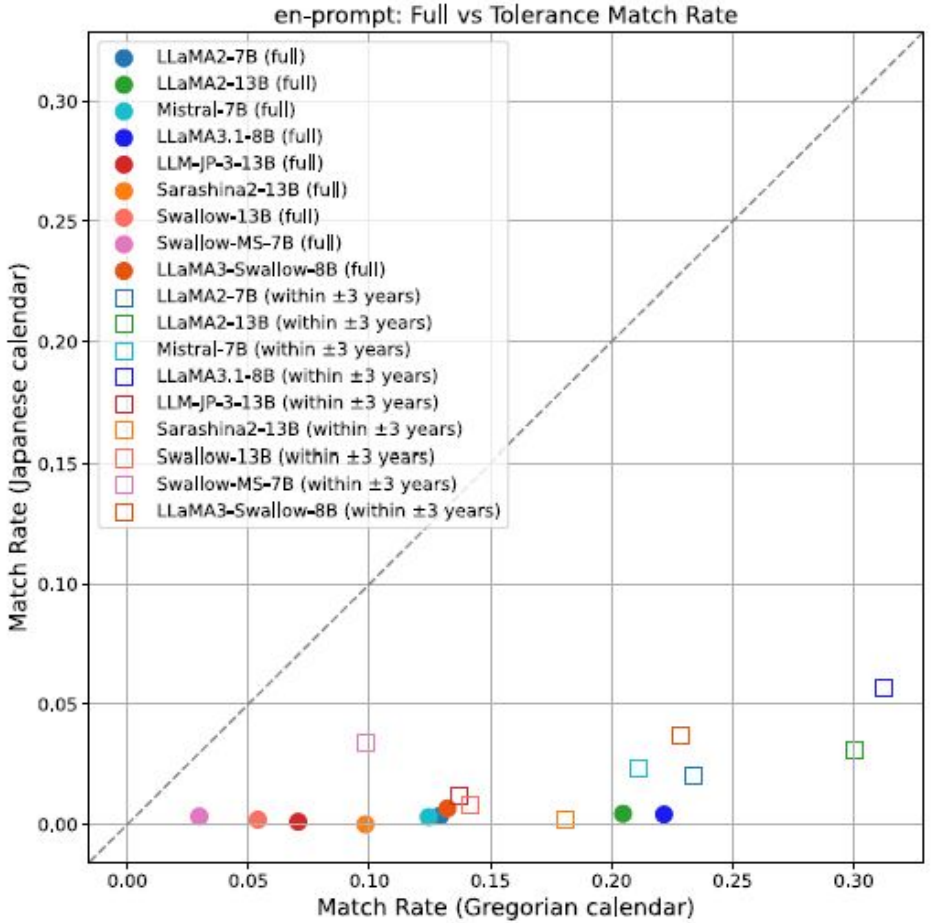
3. Birth Year Recall



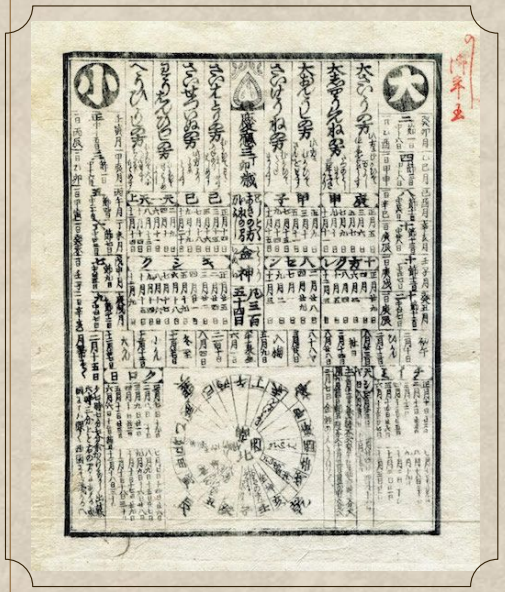
3. Birth Year Recall

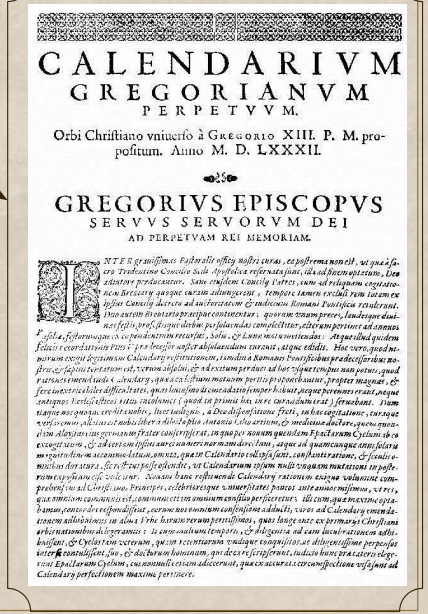
	Japanese to Gregorian	*Gregorian to Japanese*
Full matches	$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(\hat{y}_i = y_i)$	$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(Ei = \hat{E}_i \wedge x_i = \hat{x}_i)$
Tolerance matches	$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(\lvert\hat{y}_i = y_i\rvert \leq 3)$	$\frac{1}{N}\sum_{i=1}^{N}\mathbb{1}(Ei = \hat{E}_i \wedge \lvert x_i = \hat{x}_i\rvert \leq 3)$


ja-prompt: Full vs Tolerance Match Rate
LLaMA2-7B (full)
LLaMA2-13B (full)
Mistral-7B (full)
LLaMA3.1-8B (full)
LLM-JP-3-13B (full)
Sarashina2-13B (full)
Swallow-13B (full)
Swallow-MS-7B (full)
LLaMA3-Swallow-8B (full)
LLaMA2-7B (within ±3 years)
LLaMA2-13B (within ±3 years)
Mistral-7B (within ±3 years)
LLaMA3.1-8B (within ±3 years)
LLM-JP-3-13B (within ±3 years)
Sarashina2-13B (within ±3 years)
Swallow-13B (within ±3 years)
Swallow-MS-7B (within ±3 years)
LLaMA3-Swallow-8B (within ±3 years)
Match Rate (Japanese calendar)
Match Rate (Gregorian calendar)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7


en-prompt: Full vs Tolerance Match Rate
LLaMA2-7B (full)
LLaMA2-13B (full)
Mistral-7B (full)
LLaMA3.1-8B (full)
LLM-JP-3-13B (full)
Sarashina2-13B (full)
Swallow-13B (full)
Swallow-MS-7B (full)
LLaMA3-Swallow-8B (full)
LLaMA2-7B (within ±3 years)
LLaMA2-13B (within ±3 years)
Mistral-7B (within ±3 years)
LLaMA3.1-8B (within ±3 years)
LLM-JP-3-13B (within ±3 years)
Sarashina2-13B (within ±3 years)
Swallow-13B (within ±3 years)
Swallow-MS-7B (within ±3 years)
LLaMA3-Swallow-8B (within ±3 years)
Match Rate (Japanese calendar)
Match Rate (Gregorian calendar)
0.00
0.05
0.10
0.15
0.20
0.25
0.30

4. Cross-Calendar Consistency



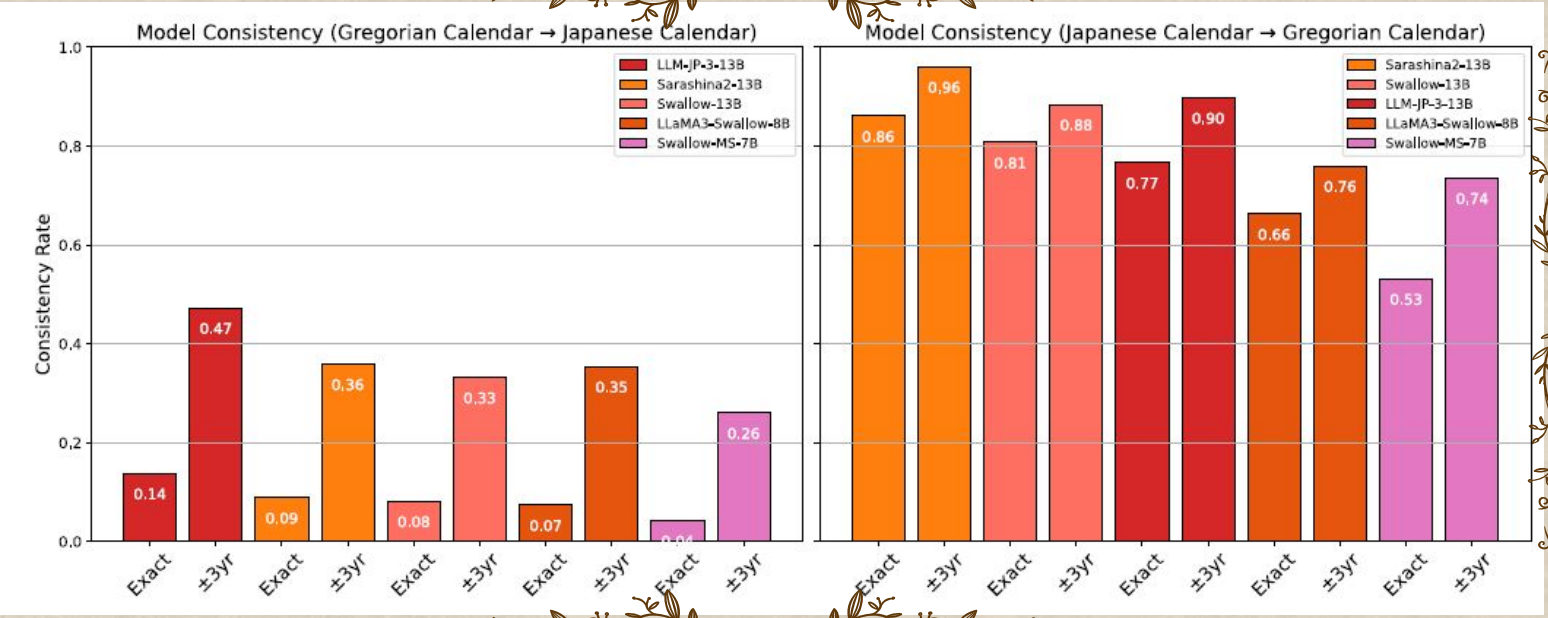

CALENDARIVM
GREGORIANVM
PERPETVVM.

Orbi Christiano vniuerſo à GREGORIO XIII. P. M. propoſitum. Anno M. D. LXXXII.

GREGORIVS EPISCOPVS
SERVVS SERVORVM DEI
AD PERPETVAM REI MEMORIAM.

Heisei 13 ≟ 1932

4. Cross-Calendar Consistency



Ultimate Words

- Can Language Models Handle a Non-Gregorian Calendar?
- What purpose can this paper serve?

References

Academic References

- Aizawa, A., Aramaki, E., Chen, B., Cheng, F., Deguchi, H., Enomoto, R., ... & Yoshino, K. (2024). *Llm-jp: A cross-organizational project for the research and development of fully open japanese llms.* arXiv preprint arXiv:2407.03963.
- Chaplot, D. S. (2023). Albert q. jiang, alexandre sablayrolles, arthur mensch, chris bamford, devendra singh chaplot, diego de las casas, florian bressand, gianna lengyel, guillaume lample, lucile saulnier, lélio renard lavaud, marie-anne lachaux, pierre stock, teven le scao, thibaut lavril, thomas wang, timothée lacroix, william el sayed. arXiv preprint arXiv:2310.06825, 3.
- Das, D., De Langis, K., Martin-Boyle, A., Kim, J., Lee, M., Kim, Z. M., ... & Kang, D. (2024). *Under the surface: Tracking the artifactuality of llm-generated data*. arXiv preprint arXiv:2401.14698.

References

Academic References

- Fujii, K., Nakamura, T., Loem, M., Iida, H., Ohi, M., Hattori, K., ... & Okazaki, N. (2024). *Continual pre-training for cross-lingual llm adaptation: Enhancing japanese language capabilities.* arXiv preprint arXiv:2404.17790.
- Grattafiori, A., Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., ... & Vasic, P. (2024). The llama 3 herd of models. arXiv preprint arXiv:2407.21783.
- Okazaki, N., Hattori, K., Shota, H., Iida, H., Ohi, M., Fujii, K., ... & Mizuki, S. (2024). *Building a large japanese web corpus for large language models*. arXiv preprint arXiv:2404.17733.
- Sasaki, M., Kamoda, G., Takahashi, R., Sato, K., Inui, K., Sakaguchi, K., & Heinzerling, B. (2025). *Can Language Models Handle a Non-Gregorian Calendar?*. arXiv preprint arXiv:2509.04432.

References

Academic References

- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M. A., Lacroix, T., ... & Lample, G. (2023). Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971.

Images

- Birth Certificate, Birthday Cake (70th), Birthday Cake (90th), Disintegrating Clock, Flag of Japan, Gregorian Calendar, Japanese Calendar, Union Jack

Thank you!

Questions Are Welcome

rosing@cl.uni-heidelberg.de