

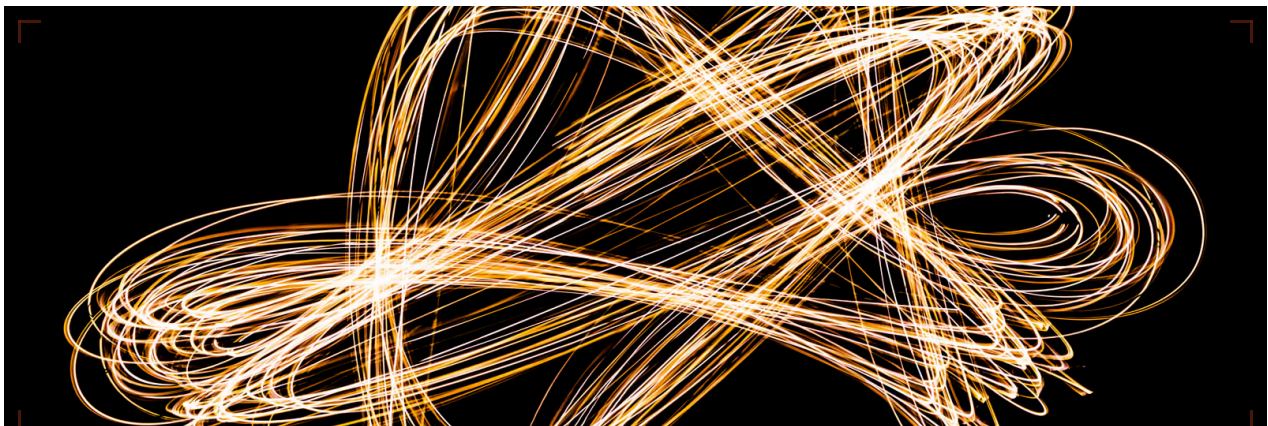
Tipjai Voice 1: model card ฉบับเต็ม

เสียงอ่านข้อความทึปภาษาไทยรุ่นแรกของ Tipjai ใช้ทำอะไร วัดอย่างไร ผล
ข้อจำกัด และการใช้อย่างรับผิดชอบ ฟังผิด 6.0% อันดับ 8 จาก 30 ราคา B0

เผยแพร่
14 ต.ค. 2569

ผู้เขียน
Tipjai Research

PDF
ดาวน์โหลด



I. รูปลิสซาจู เส้นที่การสั่นสองทิศวาดขึ้นเมื่อซ้อนกัน วาดด้วยแสง
Dietmar Rabich / Wikimedia Commons / "Lissajous-Figur -- 2020 -- 7770" / CC BY-SA 4.0 ย่อขนาด สัญญาอนุญาต

1 สรุป

ฟังผิด

6.0%

อันดับ 8 จาก 30 เสียง

เสียงที่ชัดกว่าแบบแยกได้

0

จาก 29 เสียง

ราคาต่อข้อความ

฿0

รันบนเครื่องของสตรีมเมอร์

เวลาต่อข้อความ

1.38 วินาที

ซีพียูแล็ปท็อป

- ฟังผิด 6.0% บนข้อความทิป 60 ข้อที่ไม่มีเสียงไหนเคยเห็น ช่วง 95% คือ 4.6% ถึง 7.5% ไม่มีเสียงไหนใน 29 เสียงที่ชัดกว่าแบบแยกได้ และอีก 22 เสียงสับสนแยกไม่ออก
- จุดที่แพ้ชัดที่สุดคือชื่อผู้ให้ (อันดับ 26 จาก 30) และความเป็นธรรมชาติ (3.17 จาก 5 เทียบกับ 5.00 ของเสียงที่ดีที่สุด ผู้ฟังหนึ่งคน)
- ปัญหาที่รู้แล้ว: จำนวนเงิน 555 บาทถูกอ่านเป็นเสียงหัวเราะ (พบ 14 ตุลาคม 2569 จะแก้ในรุ่นถัดไป)

สำหรับครีเอเตอร์ อ่านข้อความทิปชัดพอๆ กับเสียงที่ดีที่สุดในตลาดที่ต้องจ่ายเงิน โดยไม่มีค่าใช้จ่าย
แต่ยังอ่านชื่อบางชื่อพลาด และฟังออกว่าเป็นเสียงสังเคราะห์

2 ภาพรวม

รุ่น

Tipjai Voice 1 (ภายในเรียกว่า jai-pulse-1) ตีพิมพ์เมื่อ 6 ตุลาคม 2569 เวลา 14:43

ทำอะไร

อ่านข้อความทิปภาษาไทยออกเสียงบนไลฟ์: ชื่อผู้ให้ จำนวนเงิน และข้อความ

ภาษา

ไทย รวมชื่อที่เขียนด้วยอักษรละติน

สถานะ

ทดลองในแล็บ

ทำงานบน

คอมพิวเตอร์ของสตรีมเมอร์เอง ไม่ต้องใช้การ์ดจอ

ราคา

฿0 ต่อข้อความ

ต่อ ยอดจากงานโอเพนซอร์สที่ทำตามสัญญาอนุญาตทุกชิ้น ไม่เคยใช้เสียงของเจ้าอื่นฝึกหรือปรับ และไม่เคยใช้ข้อความที่ปจจริงของผู้ใช้

สิ่งที่เราไม่เผยแพร่: วิธีสร้างเสียงนี้ ทั้งส่วนประกอบ ข้อมูล ขนาด และการตั้งค่า เราเผยแพร่ผลลัพธ์และวิธีวัดให้ทุกคนตรวจได้ ส่วนสูตรเก็บไว้ภายในเพื่อไม่ให้งานถูกลอกไปทั้งชุด

3 ใช้ทำอะไร และไม่ควรใช้ทำอะไร



อ่านข้อความที่ปบนไลฟ์



ชื่อไทยและชื่ออักษรละติน ตัวเลข จำนวนเงิน 555



เลียนเสียงคนจริง



อ่านบทยาว หรือภาษาอื่น



ทำให้คนเข้าใจว่าเป็นเสียงคนจริง

4 วัดอย่างไร

วัดบนข้อความที่ป 60 ข้อที่เขียนแยกและตรึงไว้ก่อนวัด เทียบกับ 29 เสียงจากบริการของแต่ละเจ้าเอง ระบบถอดเสียงภาษาไทย 2 ระบบฟังทุกคลิป คะแนนคือสัดส่วนตัวอักษรที่ได้ยิน ผิด ช่วงความเชื่อมั่นมาจากการสุ่มซ้ำ 2,000 รอบ ความเป็นธรรมชาติมาจากผู้ฟังหนึ่งคนที่ไม่รู้ว่าเสียงไหนเป็นของใคร

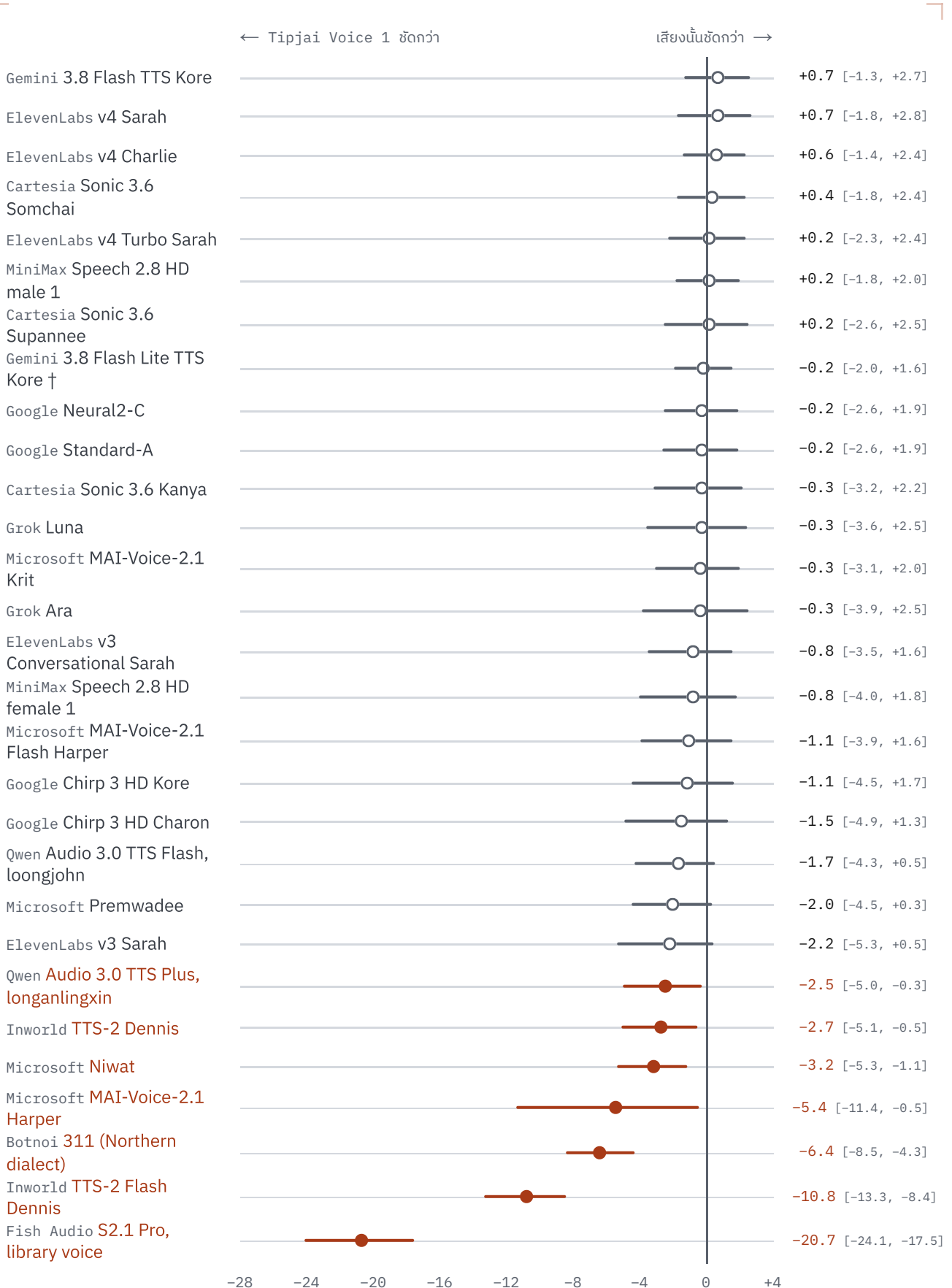
[วิธีวัดฉบับเต็มอยู่ในงานวิจัย พร้อมชุดข้อความให้ดาวน์โหลด \(/research/thai-tip-alert-benchmark/\)](/research/thai-tip-alert-benchmark/)

5 ผล

	Tipjai Voice 1 ของเรา	Gemini 3.8 Flash TTS Kore ชุดที่สุดบนกระดาน	Google Standard-A เสียงเสียเงินที่ถูกต้องที่สุด
ฟังผิด	6.0%	5.3%	6.3%
ช่วง 95%	4.6% ถึง 7.5%	3.4% ถึง 7.7%	3.9% ถึง 8.8%
อันดับที่เป็นไปได้	1 ถึง 22	1 ถึง 14	3 ถึง 18
ได้ยื่นซื้อถูก (ระบบ 1 / 2)	29 / 30	41 / 43	39 / 40
ถูกทั้งข้อความ	16 / 60	25 / 60	24 / 60
ความเป็นธรรมชาติ (1 ถึง 5)	3.17	-	-
ต่อ 1,000 ข้อความ	฿0	฿35	฿6.7

ตัวหนาคือค่าที่ดีที่สุดในแต่ละแถว ไม่ว่าของใคร

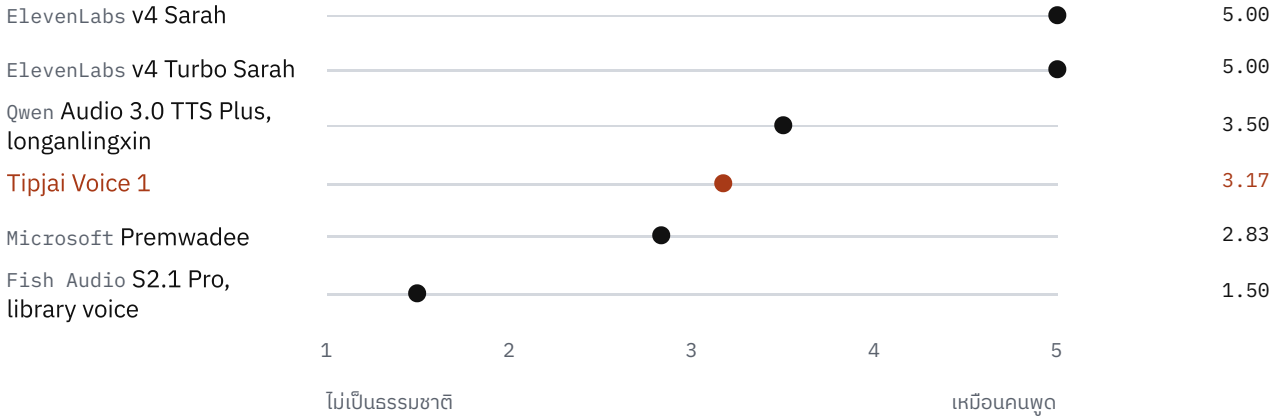
Tipjai Voice 1 ชัดกว่า 7 เสียงแบบแยกได้ ไม่มีเสียงไหนชัดกว่ามันแบบแยกได้ และสถิติ 22 เสียง



จุดคือผลต่างของการฟังผิด (Tipjai Voice 1 ลบเสียงนั้น หน่วยเป็นจุดเปอร์เซ็นต์) บนข้อความเดียวกัน
เส้นคือช่วง 95%

รูปที่ 2

ความเป็นธรรมชาติ: งานหลักของรุ่นถัดไป



ผู้ฟัง 1 คน ฟังแบบไม่รู้ชื่อ 6 ข้อความ ตัวอย่างเล็ก ใช้ดูทิศทาง

6 ความเร็วและความเสถียร

บนซีพียูแล็ปท็อป ใช้ 1.38 วินาทีต่อข้อความ (ค่ากลางของ 20 ข้อความ) อ่านได้สูงสุด 76
ข้อความต่อนาทีเมื่ออ่านพร้อมกัน ข้อความเดิมได้เสียงเดิมทุกครั้ง

รูปที่ 3

400 ข้อความทดสอบความทน ไม่ล่ม ไม่เจียบ ไม่แตก



ชื่อและข้อความยาวที่สุดที่ Tipjai รับ อีโมจิ ลิงก์ ระบบเขียนอื่น ตัวเลขทุกรูปแบบ ตัวอักษรที่มองไม่เห็น โค้ด
HTML และตัวอักษรลากยาว

7 ความปลอดภัยและการใช้ผิดทาง

การเลียนเสียง	Tipjai Voice 1 เป็นเสียงสังเคราะห์เสียงเดียว ตั้งให้เป็นเสียงของคนใดคนหนึ่งไม่ได้
การเปิดเผย	เป็นเสียงสังเคราะห์ ไม่ใช้การอัดเสียงคน ห้ามใช้ทำให้คนเข้าใจว่าเป็นเสียงคนจริง
ข้อความหายาบ	เสียงอ่านสิ่งที่ได้รับ คำหายาบถูกปิดกั้นถึงเสียงในระบบแจ้งเตือนของ Tipjai
แจ้งปัญหา	พบเสียงอ่านผิดหรือใช้ในทางที่ไม่ควร แจ้งได้ที่ tipjai.com/feedback

8 ความเป็นส่วนตัว

ไม่มีข้อความที่ปจจริงของผู้ใช้อยู่ในข้อมูลที่ใช้สร้างหรือทดสอบเสียงนี้ ข้อความทดสอบทั้งหมดเขียนขึ้นเพื่อการทดสอบโดยเฉพาะ

9 ข้อจำกัดที่รู้แล้ว

ชื่อ	ได้ยินชื่อถูกน้อยกว่าเสียงส่วนใหญ่ (อันดับ 26 จาก 30)
ความเป็นธรรมชาติ	ยังฟังออกว่าเป็นเสียงสังเคราะห์ และวัดจากผู้ฟังคนเดียว
555 บาท	จำนวนเงิน 555 บาทถูกอ่านเป็นเสียงห้าเราะ พบเมื่อ 14 ตุลาคม 2569 การแก้ไขเผยแพร่พร้อมผลวัดซ้ำ
กลุ่มหัวแถว	60 ข้อความแยกกลุ่มหัวแถวออกจากกันไม่ได้

ระบบถอดเสียง

สองระบบให้อันดับต่างกันได้ถึง 10 อันดับ คะแนนจึงเป็นค่าประมาณของสิ่งที่
คนได้ยิน

ภาษา

ภาษาไทยเท่านั้น

10 รุ่นและบันทึกการแก้ไข

14 ต.ค. 2569

พบปัญหา: จำนวนเงิน 555 บาทถูกอ่านเป็นเสียงหัวเราะ เพิ่มในข้อจำกัด

14 ต.ค. 2569

เพิ่มความเป็นธรรมชาติรอบที่ 2 (ผู้ฟัง 1 คน 6 ข้อความ)

6 ต.ค. 2569

ลองวิธีอ่านแบบใหม่บนชุดทดสอบปิดชุดที่ 3: แย่ลง จึงไม่ใช่

6 ต.ค. 2569

ข้อความเดิมได้เสียงเดิมทุกครั้ง และอ่านข้อความยาวได้ครบ

6 ต.ค. 2569

ตรึง jai-pulse-1 เวลา 14:43 ตัวเลขทั้งหมดในหน้านี้วัดจากรุ่นนี้

```
@misc{tipjai2026voice1card,  
  title = {Tipjai Voice 1: model card},  
  author = {{Tipjai Research}},  
  year = {2026},  
  month = oct,  
  url = {https://research.tipjai.com/en/models/tipjai-voice-1/card/}  
}
```