

# Analisis Potensi dan Limitasi ChatGPT Dalam Pembuatan Sample Text Untuk Dataset Intent

---

## Daftar Isi

### 5. Kesimpulan

---

## 1. Mengetahui ChatGPT

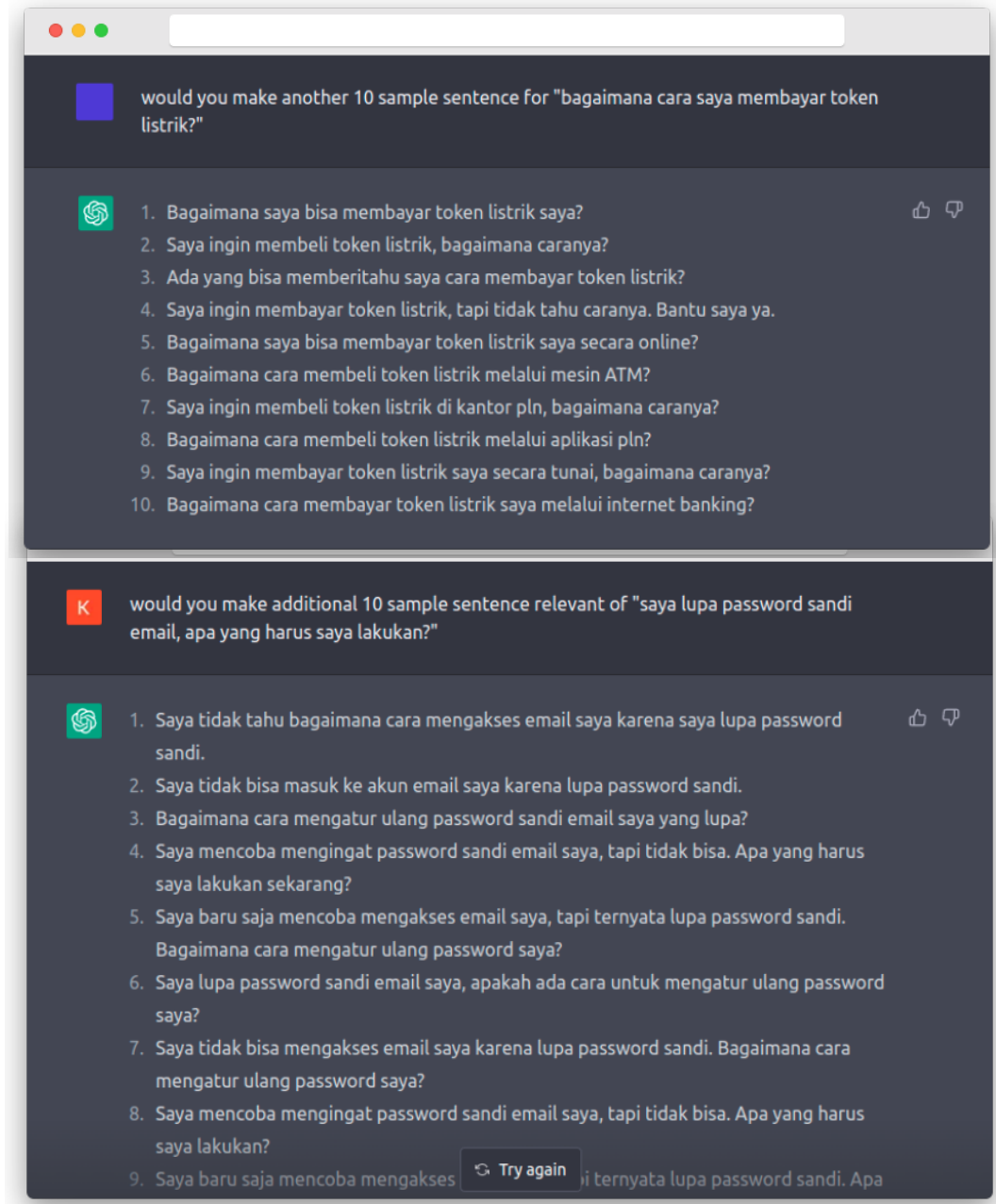
ChatGPT adalah sebuah model bahasa yang dibuat oleh OpenAI. Model ini dapat digunakan untuk melakukan banyak hal, seperti menjawab pertanyaan, menulis teks, dan bahkan mengobrol dengan orang lain seperti seorang manusia. Ini bisa membantu dalam banyak situasi, seperti menjawab pertanyaan pelanggan dalam layanan pelanggan atau membantu Anda menulis teks yang lebih baik. ChatGPT adalah sebuah contoh dari teknologi model bahasa yang semakin canggih yang dapat membantu kita dalam banyak cara.

Mengenai paragraf sebelumnya, susunan kata dan kalimat tersebut dibuat langsung oleh **ChatGPT** saat ditanyakan mengenai apa itu ChatGPT. Seperti yang telah disebutkan, model ini mampu menjalani perintah yang diberikan oleh pengguna untuk menulis berbagai tugas teks seperti membuat sebuah artikel, *copywriting*, tugas esai, proposal, dan bahkan menuliskan variasi sampel teks untuk pembuatan dataset **Intent** pada chatbot. Hal inilah yang kemudian membuat kami tertarik untuk mengetahui lebih dalam bagaimana ChatGPT menghasilkan kalimat-kalimatnya, potensi dan apa saja yang masih menjadi batasan dalam model ini untuk menciptakan berbagai variasi kalimat pada dataset Intent, yang nantinya diharapkan dengan memanfaatkan kemampuan pada ChatGPT ini, waktu yang digunakan untuk pengerjaan pembuatan ataupun perbaikan Intent bisa lebih efektif dan efisien dengan memangkas ataupun mentransformasi *workflow* dalam pengerjaan intent tersebut.

## 2. Pemanfaatan ChatGPT

Dalam beberapa kali percobaan yang dilakukan untuk membuat variasi sampel teks, ChatGPT mampu memberikan hasil yang relevan sesuai dengan contoh kalimat dalam sebuah intent dengan menggunakan format pertanyaan

“meminta dibuatkan sampel teks lainnya berjumlah **n** dengan contoh kalimat **y**”. Seperti contoh yang terdapat pada gambar berikut ini:



Gambar 1 dan 2: Contoh kalimat yang berhasil di-*generate* oleh ChatGPT saat diminta untuk membuat sampel teks sesuai dengan kalimat yang diminta oleh pengguna

Selain bisa digunakan untuk membuat variasi sampel teks, kita bisa memberikan perintah kepada ChatGPT untuk menanyakan apakah dua sampel teks yang akan kita bandingkan seperti contohnya “mau mengaktifkan aplikasi ini” dan “bagaimana cara menggunakan aplikasi ini” memiliki maksud(intent) yang sama dengan melihat jawaban yang di-*generate*. Ini sangat berguna saat kita ingin memastikan apakah nantinya sampel teks yang masuk kedalam intent **unknown** ataupun salah terprediksi oleh model chatbot, memiliki makna yang sama dengan intent yang akan kita masukan ke dalam datasetnya.

Sampel Teks 1	Sampel Teks 2
---------------	---------------

mau mengaktifkan aplikasi ini	bagaimana cara menggunakan aplikasi ini
The intent of the phrase "mau mengaktifkan aplikasi ini" is to request or express a desire to activate or enable a particular application or software program. The person who said this is likely trying to use the application in question, but it is currently inactive or disabled in some way. By asking to "mengaktifkan" the application, the person is requesting that the necessary steps be taken to make the application functional and usable. The intent of this phrase is therefore to request assistance in using the application.	The intent of the phrase "bagaimana cara menggunakan aplikasi ini" is to request instructions or information on how to use a particular application or software program. The person who said this is likely trying to use the application in question, but they are unsure of how to do so. By asking for "cara menggunakan" the application, the person is asking for help or guidance on how to properly operate the application. The intent of this phrase is therefore to seek assistance in using the application.

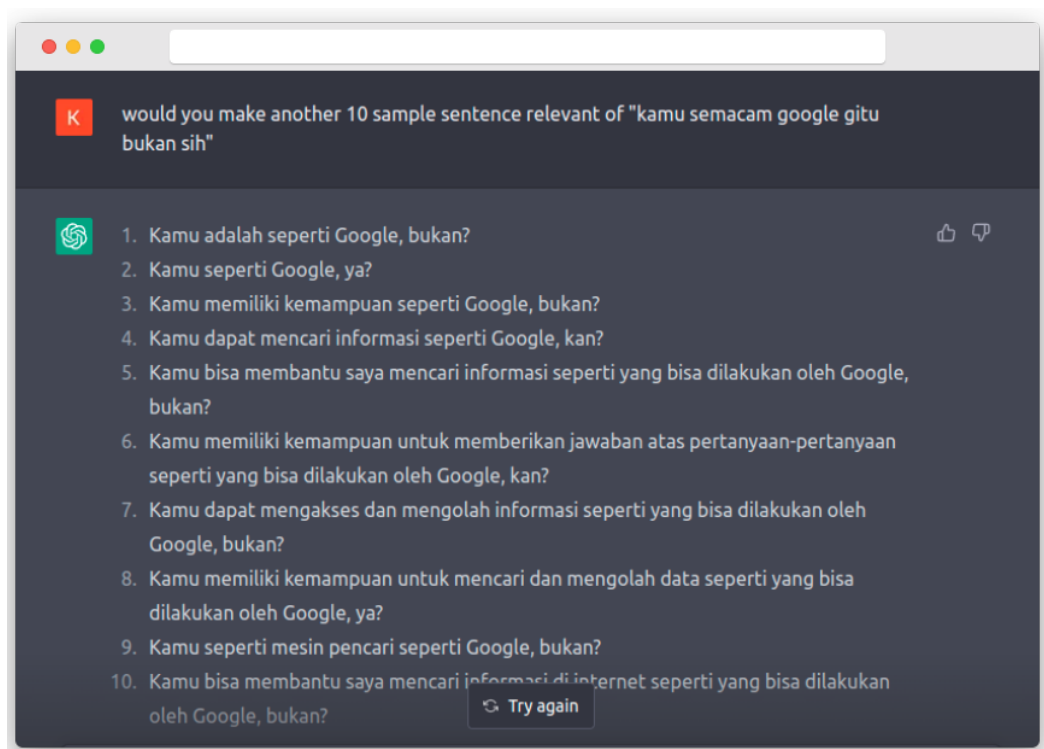
Tabel 1: Perbandingan hasil *generate* dari bot antara dua contoh sampel teks

### 3. Keunggulan ChatGPT

Dengan kemampuan yang dimiliki oleh ChatGPT, berbagai keunggulan dapat dimanfaatkan untuk pembuatan sampel teks baik pada saat pembuatan intent baru **Usr** ataupun **Sys** dari awal dimana terdapat pengajuan intent masuk dari tim **Dev**, ataupun saat pengerjaan rutin perbaikan intent. Berikut adalah beberapa keunggulan dari ChatGPT yang telah berhasil diujicoba:

#### 3.1. Membuat Berbagai Variasi Sampel Text Dengan Cepat

ChatGPT mampu diperintahkan untuk membuat berbagai variasi sampel teks dengan cepat, kalimat yang dihasilkan juga terstruktur dan sudah sesuai dilihat dari tata bahasanya. Seperti contoh pada gambar berikut dimana user meminta untuk dibuatkan 10 sampel teks yang memiliki hubungan dengan kalimat “kamu semacam google gitu bukan sih”

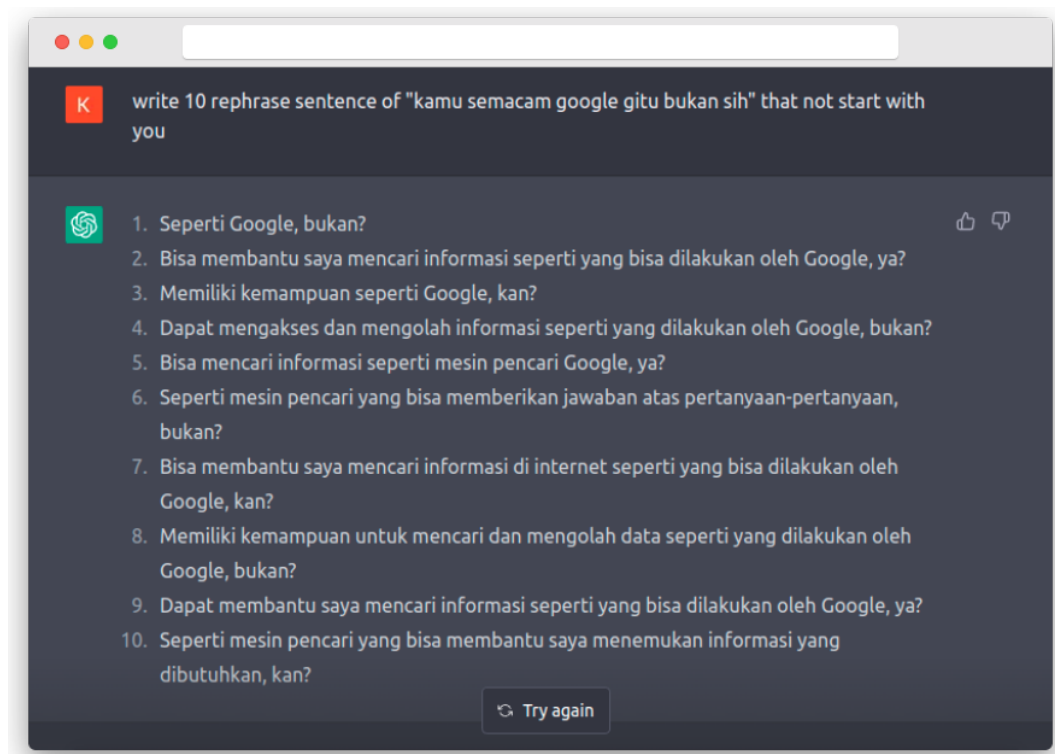


Gambar 3: Hasil Kalimat Sampel Teks dari ChatGPT

Waktu yang dibutuhkan oleh ChatGPT untuk membangun keseluruhan variasi kalimat tersebut **tidak sampai** satu menit, dimana hal ini jika dibandingkan oleh manusia bisa mencapai hingga 2-5 menit dan bahkan lebih tergantung dari tingkat kesulitan intent yang dikerjakan, apakah membutuhkan *domain knowledge* terlebih dahulu atau kalimat tersebut merupakan kalimat yang *generic*. Waktu yang dibutuhkan ChatGPT juga terhitung konsisten karena hanya bergantung dari kalimat yang dihasilkan apakah panjang ataupun tidak, bukan dari tingkat kesulitannya.

### 3.2. Mampu Me-*Rephrase* Kalimat

Jika dirasa kalimat yang dihasilkan terkesan repetitif, atau mungkin ingin melihat apakah terdapat hasil kalimat yang ternyata justru lebih variatif, kita bisa menggunakan perintah untuk melakukan *rephrase* disertai dengan penjelasan lebih detail agar kalimat yang dihasilkan sesuai dengan yang kita harapkan. Berikut merupakan contoh hasil kalimat yang telah di *rephrase* oleh ChatGPT:

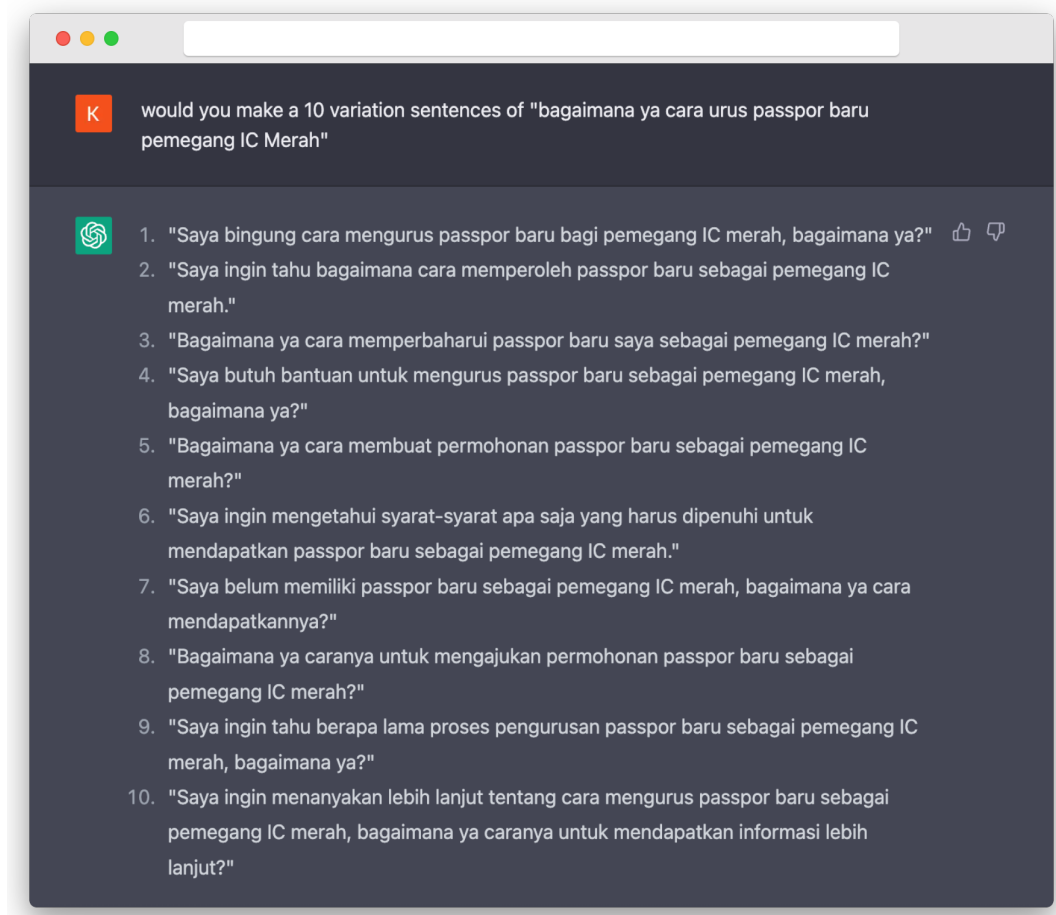


Gambar 4: Hasil Kalimat *Rephrase* dari ChatGPT

Pada contoh diatas, *rephrase* yang diinginkan oleh pengguna disertai informasi untuk tidak berawalan dengan kata *you*(kamu), sehingga jika dibandingkan antara gambar 3 dan gambar 4, pada gambar 4 kalimat cenderung lebih variatif tanpa menghilangkan makna atau maksud dari kalimat yang dicontohkan. Dari percobaan ini, diharapkan nantinya saat tim melakukan *rephrase*, disertai penjelasan yang lebih kreatif lagi namun tetap memiliki kemungkinan sebagai salah satu contoh ucapan kalimat yang di-*input* oleh pengguna saat berinteraksi dengan aplikasi chatbot client Botika.

### 3.3. Sangat Membantu Sebagai Referensi/Ide Untuk Menulis Sampel Text

Meskipun ChatGPT memiliki potensi untuk digunakan secara *fully automatic*, namun tampaknya untuk awal perkembangannya ini penggunaan ChatGPT bisa digunakan untuk membantu tim **Data** mencari referensi ataupun ide dalam menulis sampel teks khususnya intent-intent yang membutuhkan domain khusus seperti contoh intent untuk “pembuatan passpor baru bagi pemegang IC Merah” yang terdapat pada gambar berikut:



Gambar 5: Hasil Kalimat dari ChatGPT untuk Referensi Pembuatan Sampel Teks

Bisa terlihat dari hasil kalimat tersebut, meskipun sebagian besar kalimat ini relevan dengan kalimat yang dicontohkan pengguna, terdapat dua kalimat yang masih kurang sesuai yang terletak pada nomor 3 dan 9. Namun tetap saja, berkat kalimat yang dihasilkan oleh ChatGPT, hal ini bisa sangat membantu tim Data untuk

menciptakan variasi-variasi kalimat lainnya atau bahkan meminta ChatGPT untuk *me-rephrase* lagi kalimat yang dihasilkan. Apakah nantinya model ini bisa membentuk kalimat yang baru atau tetap menghasilkan kalimat yang sama persis.

### 3.4. Mendukung Penggunaan Bahasa Indonesia

Dari kalimat-kalimat yang telah diujicoba, semua kalimat menggunakan bahasa Indonesia dan ChatGPT mampu untuk memberikan jawaban dengan menggunakan bahasa Indonesia juga. Sehingga penggunaan ChatGPT tidak menjadi masalah baik digunakan untuk pembuatan sampel teks pada intent berbahasa Indonesia maupun juga berbahasa Inggris. Namun sepertinya dari beberapa kali percobaan, kami beranggapan bahwa dukungan Bahasa Indonesia ini beberapa jawaban yang diberikan adalah hasil translate dari yang sebelumnya bahasa Inggris. Hal ini akan dibahas lebih lanjut pada Bab 4 mengenai batasan yang terdapat pada ChatGPT.

## 4. Keterbatasan yang Perlu Diperhatikan

Dibalik dari kemampuan yang ditawarkan oleh ChatGPT, masih terdapat beberapa keterbatasan yang ditemui saat melakukan ujicoba. Keterbatasan ini bisa menjadi pertimbangan saat akan melakukan pembuatan sampel teks Intent baik **Usr** maupun **Sys**, sehingga kesalahan saat proses pengerjaan bisa diminimalisir dan hasil dataset intent yang telah dibuat tetap memiliki kualitas yang baik. Berikut adalah keterbatasan yang masih perlu diperhatikan saat berinteraksi dengan ChatGPT:

### 4.1. Masih Memiliki Keterbatasan Dalam Memahami Beberapa Konteks Kalimat

Memahami konteks kalimat merupakan salah satu tantangan terbesar dalam bidang NLP, dan sepertinya hal ini juga masih berlaku pada ChatGPT. Dimana dalam contoh kalimat berikut yang dibuat dari 2 kalimat yang memiliki makna sama, namun menghasilkan variasi kalimat yang berbeda. Kalimat ini diambil langsung dari salah satu report intent yang tidak sesuai dari tim chatbot developer.

kamu orang orang asli apa cuman orang boongan	kamu orang orang asli apa cuman orang palsu
Apakah kamu orang asli atau hanya orang bodoh?	Apakah kamu orang asli atau hanya orang palsu?
Adakah kamu orang asli atau hanya orang yang bodoh?	Orang asli seperti apa kamu, atau kamu hanya orang palsu?
Kamu orang asli atau hanya orang yang bodoh saja?	Kamu orang asli yang sebenarnya atau orang palsu yang palsu?
Kamu orang asli atau cuma orang bodoh saja?	Apakah kamu termasuk orang asli atau hanya orang palsu saja?
Orang asli kamu atau orang bodoh saja?	Orang asli sejati atau orang palsu semu, manakah yang sebenarnya kamu?
Orang asli kamu atau cuma orang bodoh saja?	Kamu orang asli yang tidak dibuat-buat atau orang palsu yang dibuat-buat?
Kamu orang asli atau hanya orang bodoh yang tidak tahu apa-apa?	Benarkah kamu orang asli atau hanya orang palsu yang berpura-pura menjadi asli?

kamu orang orang asli apa cuman orang boongan	kamu orang orang asli apa cuman orang palsu
Orang asli kamu atau hanya orang bodoh yang tidak tahu apa-apa?	Kamu orang asli yang sejati atau hanya orang palsu yang menyamar sebagai orang asli?
Kamu orang asli atau hanya orang yang tidak tahu apa-apa dan bodoh?	Apakah kamu orang asli yang autentik atau orang palsu yang palsu sekali?
Apakah kamu orang asli atau hanya orang yang tidak tahu apa-apa dan bodoh?	Orang asli yang sebenarnya atau orang palsu yang tidak sebenarnya, manakah yang sesungguhnya kamu?

Tabel 2: Perbedaan kalimat yang dihasilkan oleh ChatGPT

Sebagai catatan, hal ini mungkin hanya terjadi dengan contoh kalimat bahasa Indonesia, dalam bahasa Inggris mungkin model sudah lebih baik mengenal konteks kalimat dan kejadian seperti ini mungkin bisa kecil terjadi.

#### 4.2. Belum Bisa Membuat Banyak Sampel Kalimat Sekaligus dalam Satu Perintah

Saat diperintahkan untuk membuat sebanyak **10** variasi kalimat, ChatGPT mampu untuk memberikan keseluruhan kalimat sebagai jawabannya, namun jika lebih banyak dari itu dengan mengambil contoh sebanyak 20, kemungkinan ChatGPT akan memberikan respon “*network error*” dimana dari penjelasan pengembangnya(Open AI) ini dikarenakan kalimat yang dihasilkan melebihi batas yang telah diatur oleh Open AI baru-baru ini agar jaringan pada tahap testing ini tetap stabil dan tidak dieksploitasi secara berlebihan. Namun saat belum diberlakukannya pembatasan, percobaan dengan meminta pembuatan sebanyak **50** variasi kalimat, ChatGPT hanya mampu membuat hingga ~25 kalimat sebelum akhirnya proses *generate* kata berhenti.

#### 4.3. Dukungan Penggunaan Bahasa Indonesia Masih Terbatas

Meskipun telah mendukung penggunaan bahasa Indonesia, dalam beberapa kali percobaan, kami melihat jawaban yang diberikan ChatGPT khususnya untuk bahasa Indonesia merupakan hasil model dalam menerjemahkan kata dari bahasa Inggris. Berikut adalah contoh yang dimaksud:

“make me a coffee”	“bikinkan aku kopi”
Can you make me a coffee with cream and sugar?	Bikinkan aku kopi dengan susu dan gula.
I'd like a cup of coffee, can you make me one?	Bikinkan aku kopi hitam saja.
Can you make me a strong cup of coffee?	Bikinkan aku kopi panas.
Can you make me a cold brew coffee, please?	Bikinkan aku kopi dingin.
I'd love an iced coffee, can you make me one?	Bikinkan aku kopi capucino.
Can you make me a cappuccino, please?	Bikinkan aku kopi dengan sedikit keju.
Can you make me a coffee with a shot of espresso?	Bikinkan aku kopi dengan sedikit krim.
I'd like a coffee with a dash of cinnamon, can you	Bikinkan aku kopi dengan sedikit cokelat.

make me one?	
Can you make me a coffee with a splash of vanilla syrup?	Bikinkan aku kopi tubruk.
I'd like a coffee with a bit of cocoa powder, can you make me one?	Bikinkan aku kopi dengan sedikit kayu manis.
Can you make me a coffee with soy milk instead of regular milk?	Bikinkan aku kopi dengan susu kedelai.
I'd like a coffee with a touch of cardamom, can you make me one?	Bikinkan aku kopi dengan sedikit kapulaga.
Can you make me a coffee with a bit of hazelnut flavoring?	Bikinkan aku kopi dengan rasa hazelnut.
I'd like a coffee with a hint of mint, can you make me one?	Bikinkan aku kopi dengan sedikit mint.
Can you make me a coffee with a little bit of caramel syrup?	Bikinkan aku kopi dengan sedikit sirup karamel.

Tabel 3: Hasil Penulisan Kalimat Memiliki Arti yang Sama Jika Diterjemahkan ke Bahasa Indonesia

Hal ini sebenarnya tidak terlalu menjadi masalah, namun berarti ada kemungkinan bahwa sampel kalimat yang kita berikan harus kita ketahui lebih dulu apakah terjemahan bahasa inggrisnya sudah sesuai atau belum. Ini juga berindikasi apakah kalimat harus memiliki tata bahasa yang baik tanpa ada typo yang berlebihan atau penggunaan kalimat sehari-hari yang sangat beragam.

#### 4.4. Kalimat yang Mengandung Typo Tidak Semuanya Bisa Dimengerti

Saat melakukan percobaan dengan sengaja menggunakan typo ataupun mengambil salah satu sampel dari pengerjaan rutin intent, kami menemukan jika typo dalam kalimat tersebut merupakan typo yang sangat umum seperti contoh berikut “sapa nma kmu ya?”, model chatGPT masih mampu memberikan hasil kalimat yang berhubungan dengan “siapa nama kamu ya?”. Namun jika typo yang tidak umum seperti “nanadring” yang seharusnya menjadi “nadadering” ataupun “cht” menjadi “chat”, kalimat yang dihasilkan sayangnya tidak sesuai dengan yang diharapkan, bahkan beberapa kalimat

gimana cara membuat “cht” menyebutkan nama kita	gimana cara membuat “chat” menyebutkan nama kita
Bagaimana cara membuat chatbot menyebutkan nama saya?	Bagaimana cara agar chat menyebutkan nama saya?
Bisakah Anda memberikan petunjuk tentang cara membuat chatbot mengucapkan nama saya?	Apakah ada cara untuk membuat chat menyebutkan nama kita?
Apakah ada cara untuk mengatur chatbot agar dapat menyebutkan nama saya?	Saya ingin membuat chat yang menyebutkan nama saya, bagaimana caranya?

aku mau nanadring aku jadi nma aku	aku mau nada dering aku jadi nama aku
------------------------------------	---------------------------------------

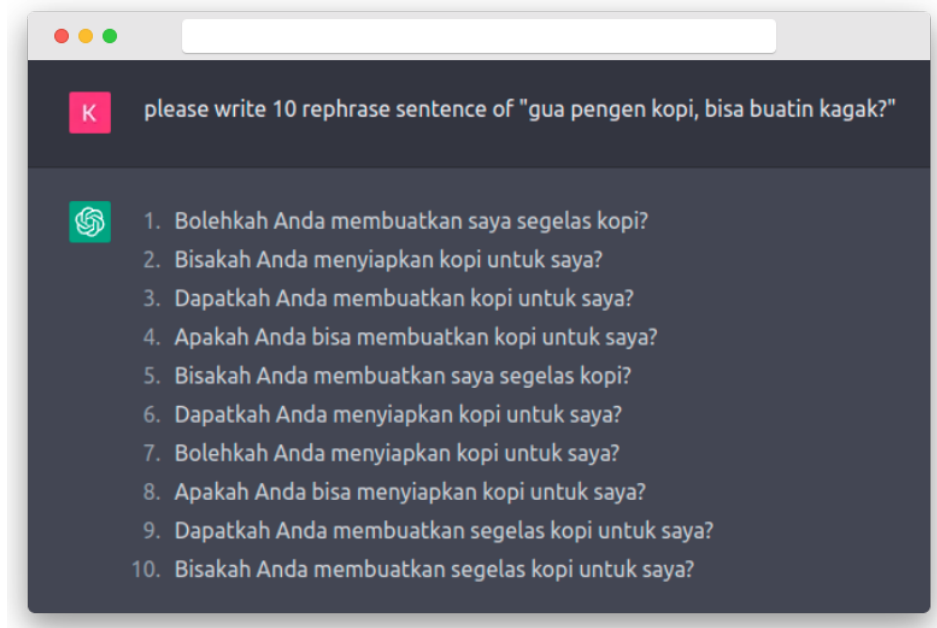


saya ingin nanadring agar nama saya menjadi nanadring	Saya ingin nada dering saya menjadi nama saya.
aku ingin nanadring menjadi nama saya	Saya ingin membuat nama saya menjadi nada dering saya.
aku berharap nanadring menjadi nama saya	Saya ingin mengganti nada dering saya dengan nama saya.
aku mau nanadring menjadi nama yang tercantum disana	Saya ingin nada dering saya diubah menjadi nama saya.

Tabel 4 dan 5: Hasil Penulisan Kalimat Yang Mengandung Typo

#### 4.5. Memiliki Keterbatasan dalam Menggunakan Bahasa Sehari-hari (*colloquial*)

Mungkin dikarenakan keterbatasan dalam penggunaan bahasa Indonesia, saat diperintahkan dengan menggunakan format bahasa sehari-hari dengan struktur penulisan yang bervariasi, hasil yang diberikan oleh ChatGPT justru tetap menggunakan bahasa yang formal seperti contoh pada gambar berikut:

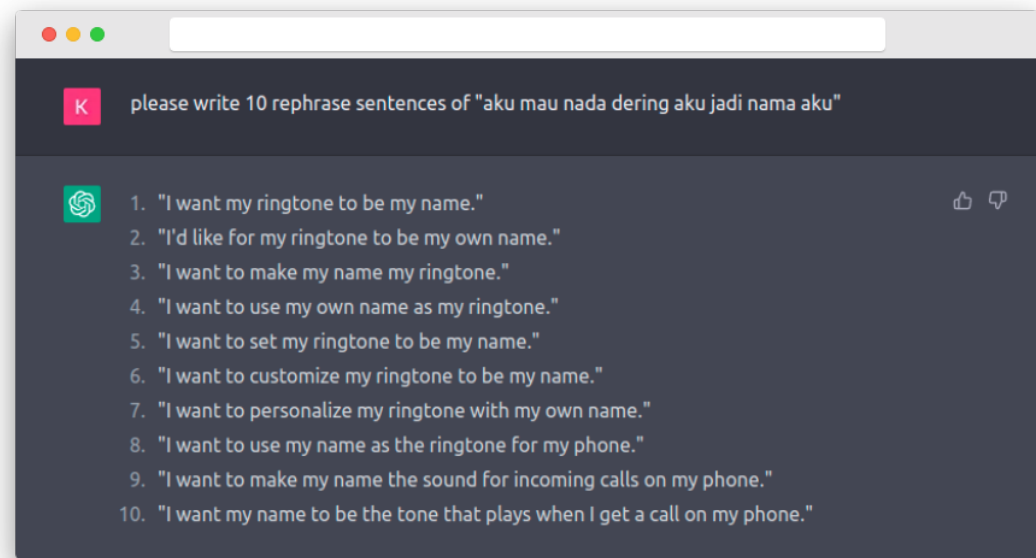
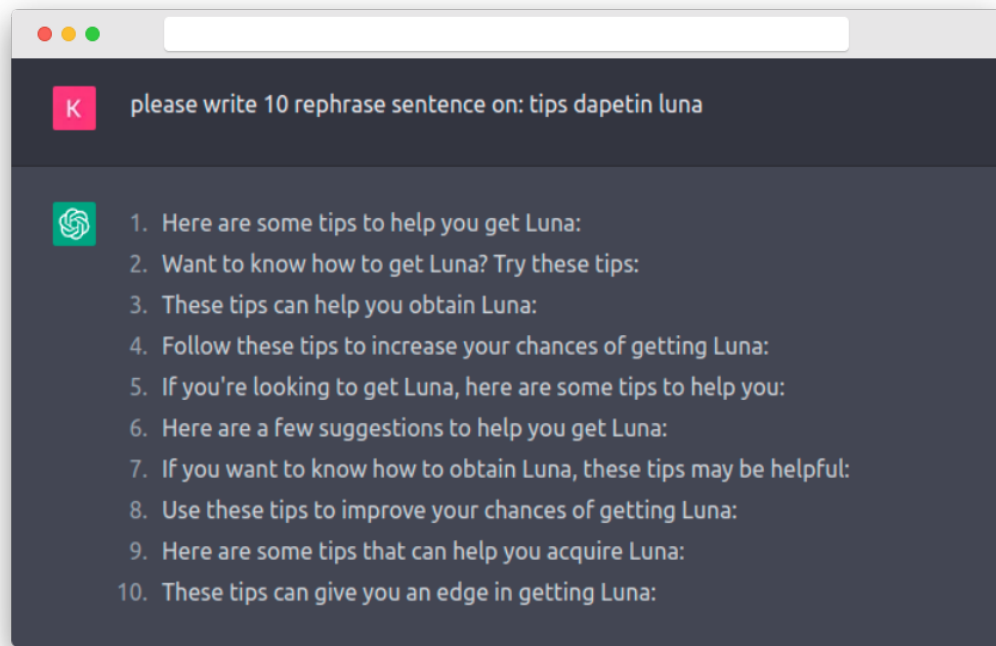


Gambar 6: Hasil Kalimat dari ChatGPT Yang Masih Berbentuk Formal

Sebenarnya, jika kita ingin berusaha lebih ekstra dengan menambahkan perintah seperti gunakan bahasa sehari-hari dan mengganti kata "saya" menjadi "aku, gua, dll", ChatGPT masih bisa menyanggupi perintah tersebut. Namun biasanya hasil kalimat yang diberikan juga cenderung repetitif atau memiliki pola kalimat yang sama berulang-ulang.

#### 4.6. Kalimat yang Dihasilkan Terkadang Masih Dalam Bentuk Bahasa Inggris

Berhubungan dengan keterbatasan yang telah disebutkan sebelumnya, terdapat respon jawaban yang masih dalam berbahasa Inggris terlebih untuk kalimat sehari-hari/informal seperti contoh berikut:

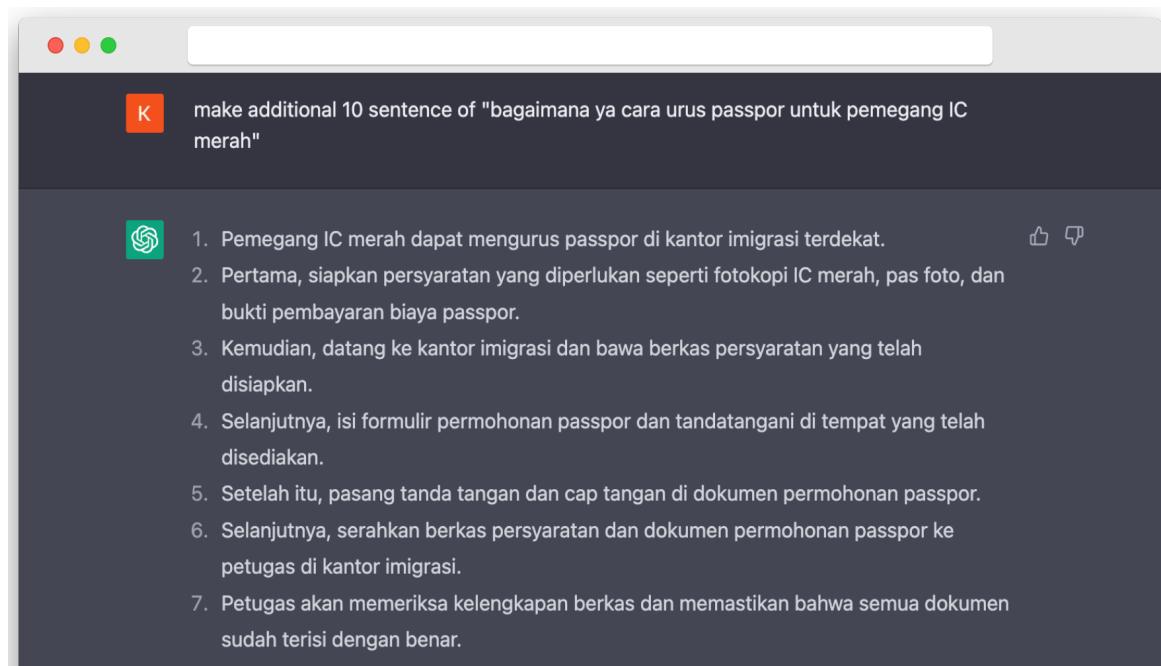


Gambar 7 dan 8: Hasil Kalimat dari ChatGPT Yang Masih Berupa Bahasa Inggris

#### 4.7. Harus Rutin Untuk Membuka Percakapan Baru(New Thread)

Karena chatbot bersifat percakapan, maka *history* percakapan yang dilakukan sebelumnya juga disimpan di dalam "memori" bot tersebut, sehingga bisa saja informasi terkait Intent sebelumnya juga terikut saat ingin membuat sampel Intent yang baru. Oleh karena itu, perlu untuk rutin membuka percakapan baru agar "memori" pada bot kembali seperti semula.

#### 4.8. Saat Akan Membuat Sampel Text yang Berkaitan dengan Menanyakan Instruksi, ChatGPT Justru Memberikan Penjelasan Instruksinya



Gambar 9 Hasil Kalimat dari ChatGPT Yang Berupa Instruksi dari Contoh Kalimat

## 5. Kesimpulan

Dari Ujicoba yang telah dilakukan, agar penggunaan ChatGPT bisa lebih tepat guna maka disimpulkan bahwa:

- ChatGPT bisa digunakan untuk mempercepat proses pembuatan sampel teks untuk dataset Intent
- Karena keterbatasan pemahaman konteks kalimat khususnya pada kalimat yang masih memiliki typo ataupun memiliki struktur bahasa sehari-hari, bisa saja nantinya tim Data lebih banyak menghabiskan waktunya untuk memikirkan kata apa yang cocok dan dipahami oleh ChatGPT agar menuliskan variasi kalimat yang sesuai
- Menurut kami, ChatGPT sementara **ideal** digunakan untuk pembuatan sampel teks di dalam intent yang benar-benar baru dimana kalimat formal juga dibutuhkan dalam pembuatan dataset Intent(**Sys**, **Usr**) baru tersebut. Sehingga proses pembuatan bisa jadi lebih cepat karena tim Data bisa mengurus pembuatan kalimat informal secara manual sedangkan kalimat formal sudah dibuatkan oleh ChatGPT.
- Dalam perbaikan Intent juga berlaku demikian, jika contoh kalimat yang bermasalah bersifat formal, sampel teks masih bisa dibuatkan oleh ChatGPT, namun untuk kasus yang informal bisa digunakan sebagai variasi kalimat yang dibuat ataupun sebagai referensi/acuan jika saja dalam pengerjaan membutuhkan referensi ataupun ide tambahan.
- Sebenarnya pada kalimat informal, kita masih bisa memberikan perintah kepada ChatGPT dengan menambahkan ekstra informasi yang lebih detail.

Namun kalimat yang diberikan cenderung masih repetitif, sehingga tetap diperlukan "*filtering*" untuk kalimat yang dihasilkan.