BAB 2
TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Infrastruktur Informasi Vital (IIV)

Infrastruktur Informasi Vital (IIV) adalah bagian penting dari sebuah negara. IIV menyangkut setiap organisasi atau lingkup pemerintahan yang sangat krusial bagi kemajuan dan perkembangan negara. IIV mengacu pada sistem elektronik yang menggunakan teknologi informasi atau teknologi operasional dalam proses independen atau saling berhubungan dengan sistem elektronik lain yang mendukung sektor strategis (Biro Hukum dan Komunikasi Publik BSSN, 2021). IIV sendiri memiliki sektor-sektor yang telah ditetapkan dalam Peraturan Presiden (PERPRES) No. 82 Tahun 2022 tentang Perlindungan Infrastruktur Informasi Vital sebagai sektor strategis dan krusial, antara lain Administrasi Pemerintahan, Energi dan Sumber Daya Mineral, Transportasi, Keuangan, Kesehatan, Teknologi Informasi dan Komunikasi, Pangan, dan Pertahanan.

Ancaman Siber menjadi tantangan dalam dunia teknologi saat ini, sebagaimana didefinisikan oleh PERMENHAN No.82 tahun 2014 tentang Pertahanan Siber, ancaman siber adalah suatu bentuk keinginan untuk melakukan kegiatan pelanggaran secara ilegal dan melanggar hukum, norma, atau keamanan informasi dengan tujuan untuk mendapatkan keuntungan manfaat material dan tidak berwujud. Ancaman ini dilakukan oleh aktor negara dan non-negara, termasuk individu, kelompok, organisasi, atau negara.

2.1.2 Keamanan Nasional di Bidang Pertahanan Siber (National Security in Cyber Defense)

Ancaman keamanan siber dilakukan oleh berbagai pelaku ancaman, baik di dalam maupun di luar negeri (Luke Irwin, 2023)..

Negara-negara harus memiliki strategi pertahanan siber multi-tahap untuk melindungi diri mereka sendiri dari ancaman siber. Tingkat pertama adalah deteksi, yang mencakup identifikasi ancaman dan menentukan di mana pertahanan siber yang paling rentan. Tahap kedua adalah perlindungan, yang meliputi penerapan perlindungan yang memadai untuk mengurangi bahaya yang terdeteksi selama tahap

deteksi. Tahap ketiga adalah manajemen, yang meliputi pemantauan dan pengendalian pertahanan siber untuk memastikan keefektifannya. Tahap keempat adalah respons, yaitu menangani peristiwa siber saat terjadi. Tahap kelima adalah pemulihan, yang merupakan proses pemulihan sistem dan data setelah insiden siber (Luke Irwin, 2023).

Sistem Perlindungan Keamanan Siber Nasional (National Cybersecurity Protection System/NCPS) adalah sistem-sistem terintegrasi yang memberikan berbagai kemampuan, seperti deteksi penyusupan, analisis, berbagi informasi, dan pencegahan penyusupan. NCPS menyediakan fondasi teknologi yang memungkinkan Cybersecurity and Infrastructure Security Agency (CISA) untuk mengamankan dan mempertahankan infrastruktur teknologi informasi lembaga-lembaga Federal Civilian Executive Branch (FCEB) dari ancaman siber tingkat lanjut.

Negara-negara harus berinvestasi dalam penangkal ancaman keamanan siber selain memiliki strategi pertahanan siber yang lengkap. Hal ini mencakup penggunaan kata sandi yang aman, membatasi akses ke informasi dan sistem, memasang firewall, menggunakan perangkat lunak keamanan, dan memperbarui perangkat lunak dan perangkat keras secara rutin. Pendidikan karyawan/personel tentang prosedur keamanan siber terbaik sangat penting untuk mencegah rekayasa sosial.

2.1.3 Keamanan Nasional Pertahanan Siber di Indonesia

Perlindungan infrastruktur dan data suatu negara dari ancaman siber sangat bergantung pada keamanan nasional dalam pertahanan siber. Pada tahun 2019, Badan Siber dan Sandi Negara (BSSN) melaporkan 290 juta kasus serangan siber, Indonesia telah mengalami lonjakan serangan siber yang nyata. Indonesia harus memiliki strategi pertahanan siber yang komprehensif yang mencakup beberapa fase, termasuk deteksi, perlindungan, administrasi, reaksi, dan pemulihan, untuk melindungi diri dari serangan siber.

Pada bulan Juli 2023, Presiden Joko Widodo memperkenalkan peraturan baru yang memberlakukan strategi keamanan siber Indonesia, Peraturan Presiden (PERPRES) No. 47 Tahun 2023 tentang Strategi Keamanan Siber Nasional dan Manajemen Krisis Siber, yang memberikan penekanan signifikan pada pendidikan dan pelatihan, mendorong literasi siber dan praktik terbaik, serta menguraikan protokol khusus untuk mengidentifikasi, melaporkan, dan mengurangi krisis siber. Strategi ini juga mencakup investasi yang signifikan dalam infrastruktur keamanan

siber yang canggih. Pemerintah Indonesia telah memprakarsai strategi keamanan siber nasional dan menjalankan program jangka pendek dan jangka panjang, tetapi masih ada kebutuhan akan peraturan yang komprehensif tentang keamanan siber di Indonesia. Parlemen Indonesia telah membahas RUU (Rancangan Undang-Undang) Keamanan Siber yang menyeluruh, tetapi prosesnya tidak melibatkan sektor swasta, sehingga menghasilkan ketentuan yang terlalu rumit dan mahal bagi bisnis.

Hackers memanfaatkan pertahanan keamanan siber Indonesia yang tidak memadai. Koneksi IoT meningkatkan permintaan akan langkah-langkah keamanan siber, tetapi bisnis keamanan siber di Indonesia masih sulit berkembang karena kurangnya tenaga ahli keamanan siber. Pasar keamanan siber di Indonesia tersegmentasi ke dalam tiga area: *Offering (Security Type and Services), Deployment (Cloud, On-premises), dan End User (BFSI, Healthcare, Manufacturing, Government & Defense, IT and Telecommunication)* (Y.Tanuwidjaja, 2023).

2.1.4 Generative AI - Large Language Model (LLM)

GenAI Large Language Model (LLM) adalah cabang kecerdasan buatan yang menggunakan perangkat lunak komputer untuk memahami dan membuat bahasa alami. LLM dibangun dengan menggunakan teknik pembelajaran mesin yang disebut deep learning, dimana model ini dilatih pada sejumlah besar teks bahasa manusia yang telah dikumpulkan dari buku, artikel, dan website (Luknanto, 2023). LLM dapat melakukan berbagai tugas, seperti pemrosesan bahasa alami, penerjemahan mesin, pengenalan suara, dan lain-lain (Luknanto, 2023). Salah satu keunggulan utama LLM adalah kemampuannya untuk mempelajari pola-pola kompleks dalam bahasa alami (Luknanto, 2023). Contoh LLM yang populer adalah GPT-3 (Generative Pre-trained Transformer 3) yang dikembangkan oleh Perplexity, yang memiliki lebih dari 175 miliar parameter dan mampu melakukan berbagai tugas bahasa manusia (Luknanto, 2023).

Sangatlah penting menguasai pemahaman yang menyeluruh tentang teknologi dan algoritma yang mendasarinya agar dapat memahami sepenuhnya potensi dan batasan alat generatif AI. Untuk membuat data atau materi baru yang meniru konten dibuat oleh manusia, teknologi generatif AI menggunakan ML (Machine Learning) dan NLP (Natural Language Processing). The Brains dari alat bantu menulis generatif AI adalah Large Language Models (LLMs).

2.1.5 Pemahaman dan Sejarah GPT-4

Generative Pre-trained Transformer 4 (GPT-4) adalah model pemrosesan bahasa alami (Natural Language Processing/NLP) yang akan datang yang diharapkan akan dirilis oleh OpenAI dalam waktu dekat. GPT-4 diharapkan menjadi peningkatan yang signifikan dari pendahulunya, GPT-3, yang telah menunjukkan kemampuan yang mengesankan dalam menghasilkan teks yang mirip manusia. Sejarah GPT-4 dapat ditelusuri kembali ke pengembangan model GPT asli oleh OpenAI pada tahun 2018. Model GPT dirancang untuk menghasilkan teks yang mirip manusia dengan melatih Deep Neural Network (DNN) pada data teks yang besar. Sejak saat itu, OpenAI terus menyempurnakan model GPT, yang berpuncak pada peluncuran GPT-3 pada tahun 2020.

GPT-3 telah menunjukkan kemampuan yang mengesankan dalam menghasilkan teks yang mirip dengan manusia, termasuk menulis esai, puisi, dan bahkan kode komputer. Namun, model ini juga dikritik karena kurangnya transparansi dan potensi bias. GPT-4 diharapkan dapat mengatasi beberapa masalah ini dengan menggabungkan teknik dan algoritma baru yang meningkatkan akurasi dan transparansi model.

Gambar 2. 1 Evolution of Generative AI-Language Models GPT
(Sumber: Soumyadarshani Dash, 2023)

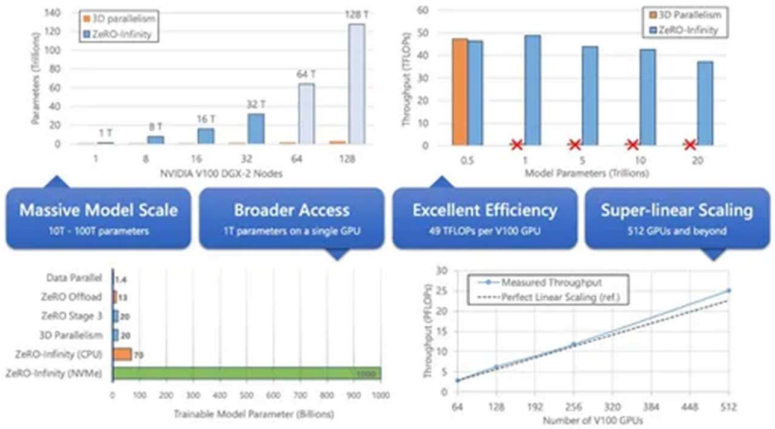
Untuk memahami kekuatan dan keterbatasan GPT-4 dengan baik, diperlukan pemahaman yang menyeluruh tentang teknologi dan algoritma yang mendasarinya. GPT-4 diharapkan didasarkan pada arsitektur transformator, yang merupakan semacam jaringan saraf (Neural Network) yang sangat cocok untuk aplikasi NLP. Model ini juga diharapkan akan mencakup strategi baru untuk penyempurnaan model pada tugas-tugas tertentu, serta metode baru untuk mengatur model's output.

Studi lebih lanjut diperlukan untuk memahami efek model terhadap masyarakat/komunitas dan ekonomi seiring dengan berkembangnya GPT-4. Para

peneliti dan profesional dapat menciptakan strategi yang lebih efisien untuk menerapkan GPT-4 guna mengatasi masalah di dunia nyata dengan memiliki pemahaman yang lebih baik tentang latar belakang dan kemampuannya.

2.1.6 Generative AI Tool berdasarkan GPT-4

Generative AI tool berdasarkan GPT-4 adalah bidang kecerdasan buatan yang sedang berkembang yang melibatkan penggunaan algoritma pembelajaran mesin untuk menghasilkan data atau konten baru yang meniru konten yang dibuat oleh manusia. Model bahasa baru GPT-4, yang dikembangkan oleh OpenAI, adalah sistem multimodal besar yang dapat menerima input dalam bentuk gambar dan teks dan menghasilkan output. Transformer architecture, a class of neural network yang unggul dalam NLP tasks, diantisipasi untuk menjadi fondasi GPT-4. Model ini juga diantisipasi untuk menyertakan pendekatan baru untuk mengatur model's output selain metode baru untuk menyempurnakan model pada tugas-tugas tertentu.



Gambar 2.2 Unveiling Enhanced Generative AI – GPT-4 Revealed
(Sumber: Soumyadarshani Dash, 2023)

Generatif AI berdasarkan GPT-4 memiliki potensi untuk digunakan dalam berbagai aplikasi, termasuk pembuatan konten, desain, penerjemahan bahasa, dan bahkan menghasilkan teks atau gambar asli. GPT-4 diharapkan lebih andal, kreatif, dan mampu menangani instruksi yang lebih bernuansa daripada GPT-3.5. Namun, teknologi ini juga menghadirkan tantangan etika dan hukum, dan ada kebutuhan untuk mempertimbangkan implikasi Generatif AI untuk penelitian, praktik, dan kebijakan.

Seiring dengan kemajuan teknologi Generatif AI, diperlukan lebih banyak penelitian untuk mengembangkan cara yang lebih efektif dalam mendeteksi dan

mencegah Generatif AI, serta menilai dampak Generatif AI terhadap masyarakat dan ekonomi. Para peneliti dan praktisi dapat merancang teknik yang lebih efektif untuk menerapkan teknologi ini ketika mengatasi masalah dunia nyata dengan mengetahui sejarah dan kemampuan Generatif AI Tool berbasis GPT-4.

2.1.7 Prompting

Prompting adalah suatu teknik dalam pemrosesan bahasa alami (NLP) yang digunakan untuk membantu komputer memahami dan membuat bahasa alami.

Prompting dalam ChatGPT4 dan model GPT secara umum merupakan area yang kaya akan potensi dan kompleksitas. Kemampuannya untuk menginterpretasikan dan merespons berbagai jenis prompt membuatnya menjadi alat yang sangat fleksibel dan kuat. Namun, seperti semua teknologi, efektivitasnya sangat bergantung pada bagaimana ia digunakan dan dipahami. Dengan pemahaman yang tepat tentang prinsip dan teknik prompting, GPT dapat menjadi alat yang sangat berharga di berbagai bidang, dari pendidikan hingga industri kreatif.

Prompting dalam model ChatGPT4 dan GPT telah dieksplorasi dalam beberapa penelitian. Penggunaan dorongan berulang telah menunjukkan harapan dalam meningkatkan efektivitas dan efisiensi ringkasan teks otomatis untuk penilaian efek makanan dalam dokumen Aplikasi Obat Baru (Shi et al., 2023). Namun, telah ditemukan bahwa mendorong GPT-3/4 saja mungkin tidak dapat diandalkan meningkatkan pengalaman pengguna (UX) chatbot, karena dapat menyebabkan kerusakan interaksi dan kurangnya pagar pembatas dalam percakapan. Di sisi lain, mendorong dengan gaya prompt yang kurang ambigu, seperti instruksi kode semu, telah terbukti meningkatkan kinerja model bahasa pra-terlatih, dengan hasil yang lebih baik dalam tugas klasifikasi dan tugas bahasa generatif (Wang & Jin, 2023). Secara keseluruhan, teknik dorongan memiliki potensi untuk meningkatkan kemampuan model GPT, tetapi penelitian lebih lanjut diperlukan untuk mengatasi tantangan dan mengoptimalkan pengalaman pengguna (Mishra et al., 2023).

Prinsip Dasar Prompting pada dasarnya adalah proses komunikasi antara manusia dan AI. Dalam kasus GPT, ini berarti memberikan model serangkaian kata atau kalimat yang memicu respons. Namun, penting untuk memahami bahwa prompting tidak hanya tentang kata-kata itu sendiri; ini juga tentang konteks, nuansa, dan maksud di balik kata-kata tersebut. Dengan GPT, prompt yang efektif sering

memerlukan pemahaman tentang bagaimana model tersebut dilatih, data apa yang ia ketahui, dan batas-batas kemampuannya. Dengan demikian, prinsip dasar dorongan dalam penerapannya di GPT adalah memberikan instruksi atau isyarat khusus untuk memandu proses pembuatan model. Prompting membantu meningkatkan keandalan dan kinerja model GPT dalam berbagai aspek seperti generalisasi, bias sosial, kalibrasi, dan faktualitas (Si et al., 2022; Wang & Jin, 2023). Dengan menggunakan petunjuk yang sesuai, model GPT dapat menggeneralisasi ke data di luar distribusi, mengurangi bias sosial, mengkalibrasi probabilitas keluaran, dan memperbarui pengetahuan faktual dan rantai penalaran. Prompting juga dapat memungkinkan model GPT untuk melakukan perilaku berulang yang diperlukan untuk mengeksekusi program yang melibatkan loop, yang mengarah ke hasil yang akurat bahkan melampaui model yang lebih kuat (Jojic, Wang, & Jojic, 2023). Selain itu, dorongan berulang telah terbukti efektif dalam meningkatkan akurasi ringkasan teks otomatis (Shi et al., 2023).

Prompting dalam ChatGPT4, dengan kemampuan pemrosesan bahasa alami yang lebih canggih, membawa konsep prompting ke tingkat yang lebih tinggi. Ini dapat memahami dan merespons prompt yang lebih kompleks, memungkinkan interaksi yang lebih mendalam dan beragam. Fitur kunci dari ChatGPT4 adalah kemampuannya untuk mempertahankan konteks dalam percakapan yang berlangsung, memahami nuansa bahasa, dan bahkan memperkirakan maksud pengguna yang tidak sepenuhnya diungkapkan.

Untuk menghasilkan petunjuk yang efektif untuk ChatGPT4, penting untuk mengikuti teknik rekayasa yang cepat dan praktik terbaik. Ini termasuk membuat petunjuk dengan kejelasan, batasan eksplisit, dan berbagai jenis pertanyaan (Ekin, 2023). Juga bermanfaat untuk bereksperimen dengan desain yang cepat dan memanfaatkan sumber daya eksternal (C. Liu et al., 2023). Penyempurnaan berulang dan penyeimbangan maksud pengguna direkomendasikan untuk hasil yang optimal. Selain itu, strategi lanjutan seperti kontrol suhu dan token, rantai cepat, adaptasi spesifik domain, dan penanganan input ambigu dapat digunakan. Studi kasus dunia nyata menunjukkan aplikasi praktis rekayasa cepat di berbagai domain, termasuk dukungan pelanggan, pembuatan konten, dan penceritaan interakti. Rekayasa cepat yang efektif secara signifikan berdampak pada kinerja ChatGPT, dan penelitian masa depan harus fokus pada pembinaan kreativitas dan kolaborasi dalam komunitas ChatGPT.

Untuk mendapatkan hasil terbaik dari GPT, penting untuk menggunakan teknik prompting yang efektif. Ini termasuk menggunakan bahasa yang jelas dan langsung, menyediakan informasi yang cukup untuk memandu respons, dan memformulasikan pertanyaan atau pernyataan dengan cara yang memudahkan pemahaman AI. Dalam ChatGPT4, ini juga berarti memanfaatkan kemampuannya untuk memahami prompt yang lebih rumit dan mempertahankan konteks yang lebih luas.

Implikasi Praktis dari Prompting dengan ChatGPT-4 memiliki implikasi praktis di berbagai domain. Ini memungkinkan untuk augmentasi kumpulan data berlabel kecil dengan data sintetis, yang dapat bermanfaat dalam pengaturan sumber daya rendah dan untuk mengidentifikasi kelas langka (Møller, Dalsgaard, Pera, & Aiello, 2023). Selain itu, ChatGPT-4 menunjukkan kinerja *zero-shot* yang kuat di berbagai tugas, menunjukkan keserbagunaan dan kemampuannya untuk menghasilkan respons yang akurat tanpa pelatihan sebelumnya tentang tugas-tugas tertentu. Teknik rekayasa yang cepat memainkan peran penting dalam memaksimalkan potensi ChatGPT-4. Teknik seperti kejelasan, kendala eksplisit, dan memanfaatkan berbagai jenis pertanyaan dapat membantu memperoleh tanggapan yang diinginkan dari model (Ekin, 2023). Dengan merancang prompt dengan hati-hati, kinerja generasi ChatGPT-4 dapat ditingkatkan secara substansif, seperti yang ditunjukkan dalam eksperimen pada tugas pembuatan kode (C. Liu et al., 2023). Temuan ini menyoroti pentingnya rekayasa cepat dan potensi ChatGPT-4 dalam berbagai aplikasi praktis.Dalam praktiknya, efektivitas prompting memiliki implikasi yang luas. Di bidang pendidikan, misalnya, guru dapat menggunakan GPT untuk menghasilkan bahan ajar yang disesuaikan. Dalam bisnis, perusahaan dapat memanfaatkan ChatGPT4 untuk layanan pelanggan yang lebih responsif dan personalisasi. Di bidang kreatif, penulis dan seniman dapat menggunakan prompt untuk menginspirasi karya baru atau menjelajahi ide-ide kreatif.

Namun, ada juga tantangan dalam prompting. Salah satunya adalah kemungkinan bias dalam respons yang dihasilkan, yang dapat dipengaruhi oleh cara prompt diformulasikan atau data latih yang digunakan oleh model. Selain itu, terdapat risiko bahwa respons yang dihasilkan mungkin tidak selalu akurat atau relevan, terutama jika prompt tidak jelas atau terlalu ambigu. Dengan kata lain, tantangan utama dalam menggunakan prompt dengan ChatGPT4 termasuk kendala konten dan potensi penyalahgunaan, serta kebutuhan untuk pembuatan prompt yang kuat dan pencegahan jailbreaking model bahasa (Y. Liu et al., 2023). Selain itu, ada kebutuhan akan petunjuk yang lebih kompleks untuk secara konsisten melampaui kumpulan data

yang dihasilkan manusia dalam hal kinerja. Selanjutnya, hasil ChatGPT kadang-kadang dapat menyajikan keacakan dalam tanggapan, informasi yang terlalu disederhanakan atau diabaikan, yang dapat dikurangi dengan menggunakan petunjuk yang lebih rinci (Møller et al., 2023). Secara keseluruhan, tantangannya terletak pada memastikan bahwa petunjuk secara efektif memandu respons model bahasa sambil menghindari penghindaran kendala dan mempertahankan akurasi dan keandalan dalam output yang dihasilkan.

.

Contoh :

Contoh Coding dan Output menggunakan prinsip-prinsip prompting yang menunjukkan pembuatan teks menggunakan GPT-3.5 dan model hipotetis GPT-4.

Code:

```
# Import necessary libraries
from transformers import GPT2LMHeadModel, GPT2Tokenizer

# Load GPT-3.5 model and tokenizer
gpt3_model = GPT2LMHeadModel.from_pretrained("Use your API Key")
gpt3_tokenizer = GPT2Tokenizer.from_pretrained("Use your API Key")

# Load hypothetical GPT-4 model and tokenizer
gpt4_model = GPT2LMHeadModel.from_pretrained("Use your API Key")  # Replace "gpt2"
with the actual GPT-4 model name
gpt4_tokenizer = GPT2Tokenizer.from_pretrained("Use your API Key")  # Replace
"gpt2" with the actual GPT-4 tokenizer name

# Define a prompt for both models
prompt = "Once upon a time"

# Generate text with GPT-3.5
gpt3_input_ids = gpt3_tokenizer.encode(prompt, return_tensors="pt")
gpt3_output = gpt3_model.generate(gpt3_input_ids, max_length=100,
num_return_sequences=1)
gpt3_text = gpt3_tokenizer.decode(gpt3_output[0], skip_special_tokens=True)

# Generate text with hypothetical GPT-4
gpt4_input_ids = gpt4_tokenizer.encode(prompt, return_tensors="pt")
gpt4_output = gpt4_model.generate(gpt4_input_ids, max_length=100,
num_return_sequences=1)
gpt4_text = gpt4_tokenizer.decode(gpt4_output[0], skip_special_tokens=True)

# Print generated text
print("Generated text with GPT-3.5:")
print(gpt3_text)
```

```
print("\nGenerated text with GPT-4:")
print(gpt4_text)
```

Output:

```
Generated text with GPT-3.5:
Once upon a time, in a land far away, there lived a brave knight named Sir Arthur.
He was known throughout the kingdom for his courage and honor.
One day, a fearsome dragon attacked the village, threatening to destroy everything
in its path.
Sir Arthur knew he had to act quickly to save his people. With his trusty sword in
hand, he rode out to face the dragon and protect his home.
The battle was fierce, but Sir Arthur's determination and skill prevailed.
He slayed the dragon and returned to the village as a hero, celebrated by all.


Generated text with GPT-4:
Once upon a time, in a world filled with magic and mystery, there existed a hidden
realm known as Eldoria.
This realm was inhabited by creatures of ancient lore, from graceful unicorns to
mischievous sprites.
The balance of power in Eldoria was maintained by the Guardians of the Elements,
individuals gifted with the ability to control fire, water, earth, and air.
But one fateful day, a dark force began to encroach upon Eldoria, threatening to
disrupt the harmony that had prevailed for centuries.
As the skies darkened and the land trembled, a young orphan named Elysia
discovered an ancient prophecy that foretold of a chosen one who would rise to
challenge the darkness and restore balance to the realm.
With a heart full of courage and a determination to fulfill her destiny, Elysia
embarked on a quest that would test her limits, forge unexpected alliances, and
unveil the true power of her own spirit.
```

Code Explanation (Penjelasan Coding):

- Kode ini menggunakan pustaka Transformers untuk bekerja dengan model bahasa terlatih, khususnya GPT-2.
- Import modul yang diperlukan: GPT2LMHeadModel dan GPT2Tokenizer.
- Model dan tokenizer GPT-3.5 (gpt3_model dan gpt3_tokenizer) dimuat menggunakan nama model “gpt2”.
- Model dan tokenizer hipotetis GPT-4 (gpt4_model dan gpt4_tokenizer) dimuat menggunakan nama model “gpt2” (diganti dengan nama model sebenarnya).
- Perintah untuk membuat teks didefinisikan sebagai “Pada suatu ketika”.

Generated Text with GPT-3.5:

- Outputnya dimulai dengan ungkapan cerita klasik: “Pada suatu ketika.”

- Ini memperkenalkan seorang ksatria pemberani bernama Sir Arthur, yang dikenal karena keberanian dan kehormatannya.
- Seekor naga mengancam desa, dan Sir Arthur bersiap menghadapinya.
- Pertempuran sengit terjadi kemudian, dan Sir Arthur muncul sebagai pemenang, menjadi pahlawan.

Generated Text with GPT-4:

- Outputnya dimulai dengan “Pada suatu ketika” di dunia sihir dan misteri.
- Ini memperkenalkan dunia tersembunyi Eldoria dan penghuninya, termasuk makhluk Ajaib.
- Penjaga Elemen menjaga keseimbangan, namun kekuatan gelap mengancamnya.
- Seorang anak yatim piatu bernama Elysia menemukan sebuah ramalan dan memulai sebuah pencarian.
- Elysia menghadapi tantangan, membentuk aliansi, dan mengungkap kekuatan sejatinya untuk memulihkan keseimbangan.

(Source: https://www.analyticsvidhya.com/blog/author/soumyadarshani5884821)

2.1.8 Grounded Theory

Grounded Theory (Teori Berbasis Data) adalah pendekatan penelitian yang dikembangkan oleh dua sosiolog, Barney G. Glaser dan Anselm L. Strauss pada tahun 1960-an. Metode ini awalnya digunakan dalam bidang sosiologi, tetapi seiring berjalannya waktu, grounded theory telah diterapkan luas dalam berbagai disiplin ilmu, termasuk Teknologi Informasi (TI), Sistem Informasi (SI), dan Teknik Informatika (TIK). Pendekatan ini memiliki nilai besar dalam memahami fenomena kompleks dan berkembang di bidang-bidang tersebut.

Menurut (Dr. Susan G & Dr. Jim W, 2013) Dalam bidang Teknologi Informasi, Sistem Informasi, dan Teknik Informatika, grounded theory dapat digunakan untuk memahami bagaimana penggunaan teknologi informasi mempengaruhi perilaku pengguna dan organisasi, serta bagaimana sistem informasi dapat ditingkatkan untuk memenuhi kebutuhan pengguna.

Menurut (Strauss & Corbin, 1990) mengarah pada praktik riset dimana data sampling, analisis data dan pengembangan teori tidak dilihat berbeda dan terpisah, tetapi sebagai langkah yang berbeda harus diulang sampai menggambarkan dan

menjelaskan fenomena yang diteliti. Yang paling membedakan grounded theory dari banyak metode riset kualitatif lainnya adalah bahwa hal itu secara eksplisit muncul, dimana metode grounded theory tidak menguji hipotesis, namun menetapkan untuk menemukan teori yang bagaimana untuk situasi riset seperti itu.

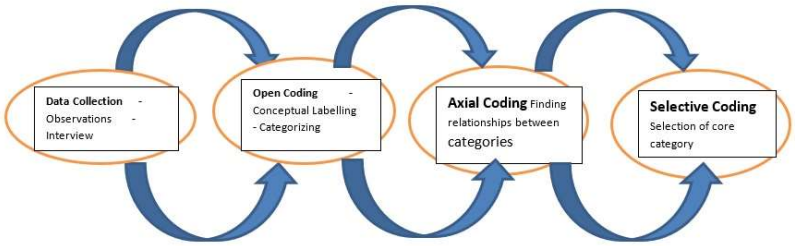
Grounded Theory berfokus pada pemahaman mendalam tentang fenomena yang sedang diteliti, dengan mengidentifikasi pola, hubungan, dan konsep yang muncul dari data. Cara kerja grounded theory melibatkan tiga tahap utama: pembandingan terus-menerus, pengkodean, dan pengembangan teori (Juhro, Solikin, 2021). Menurut (Muhadjir, 2002) riset Grounded theory dimaksudkan untuk membuat spesifikasi-spesifikasi terhadap: 1) kondisi yang menjadi sebab terjadinya suatu fenomena, 2) tindakan atau interaksi yang merupakan respon terhadap kondisi tersebut, dan 3) konsekuensi-konsekuensi yang timbul dari tindakan atau interaksi tersebut.

Dalam konteks penelitian yang berbeda, Grounded theory dapat diterapkan dengan beberapa variasi, seperti Open Coding, Axial Coding, dan Selective Coding. Open Coding melibatkan identifikasi konsep dan kategori yang muncul dari data secara umum. Axial Coding melibatkan pengorganisasian konsep dan kategori ke dalam kerangka yang lebih terstruktur. Selective Coding melibatkan pengembangan teori yang lebih terperinci berdasarkan konsep dan kategori yang telah diidentifikasi.

Grounded Theory adalah metodologi terkenal yang digunakan dalam banyak penelitian (Chun Tie et al, 2019). Teknik pengumpulan data kualitatif dan kuantitatif dapat digunakan dalam studi grounded theory. Grounded theory bertujuan untuk menemukan atau membangun teori dari data dan mengembangkan teori baru berdasarkan data yang dikumpulkan, yang diperoleh secara sistematis dan dianalisis menggunakan analisis komparatif. Meskipun grounded theory pada dasarnya fleksibel, namun merupakan metodologi yang kompleks. Grounded theory didasarkan pada konsep bahwa peneliti harus membangun teori dari data yang diperoleh dari lapangan (grounded in data). Gasson Susan & Waters James, 2013, pendekatan ini berfokus pada pengumpulan dan analisis data lapangan, dan tujuannya adalah mengembangkan teori yang muncul dari temuan yang sebenarnya, bukan dari teori yang telah ada sebelumnya.

Beberapa aspek penting Grounded Theory dan aplikasinya dalam TI, SI, dan TIK seperti, Pengumpulan dan Analisis Data, Pengembangan Sistem Informasi, dan

Metode ini memungkinkan pemahaman yang mendalam dan komprehensif tentang fenomena yang diteliti.



Gambar 2.3 Pendekatan Grounded Theory
(Sumber: Chetty, 2020)

2.1.9 MISUSE Threat Modelling (Misuse Theory)

MISUSE digunakan untuk mengidentifikasi kemungkinan niat jahat dari pelaku penyalahgunaan yang difasilitasi teknologi. Laporan ini menyoroti enam dimensi ancaman, yang merangkum potensi tujuan pemanfaatan teknologi secara jahat terhadap orang yang rentan.

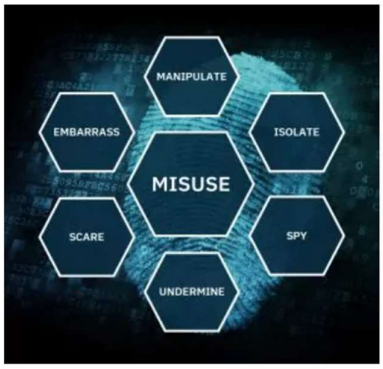
MANIPULATE – Steering, controlling, or influencing vulnerable individuals.

ISOLATE – Controlling contact to cut vulnerable individuals off from their support system.

SPY – Monitoring and tracking activities, conversations, and whereabouts.

UNDERMINE – Wearing down a vulnerable individual's self-esteem or lessening how they are perceived by others.

SCARE – Unnerving, worrying or frightening vulnerable individuals.



EMBARRASS – Causing a vulnerable individual to feel self-conscious, anxious, or ashamed.

Gambar 2.4 MISUSE framework
(Sumber: IBM, 2022)

Dengan menjadikan enam dimensi ancaman ini sebagai inti dari pemodelan ancaman MISUSE, para ahli teknologi dapat memperoleh wawasan tentang bagaimana kreasi mereka dapat digunakan untuk tujuan yang merugikan. Dengan pemahaman ini, mereka dapat berupaya memitigasi niat jahat tersebut dengan meningkatkan keamanan, privasi, dan kegunaan teknologi mereka.

2.2 Hasil Penelitian Terdahulu

Penelitian dalam bidang kecerdasan buatan dan penggunaan Generative Pre-trained Transformers (GPT) terus berkembang, dengan berbagai studi yang mengkaji aspek yang berbeda-beda. Mengamati penelitian terdahulu dan membandingkannya dengan rencana penelitian yang sedang dijalankan, memberikan wawasan berharga tentang perkembangan di bidang ini.

Sebagaimana Tabel 2.1, penelitian terdahulu dan studi yang sedang dilaksanakan ini memiliki persamaan dan perbedaan. Dari tinjauan literatur sistematis, ada tujuh penelitian yang memiliki relevansi untuk menjadi acuan studi ini, yakni:

1. Poldrack et al. (2023) mengkaji penggunaan GPT-4 dalam alat pengkodean AI. Mereka menyoroti pentingnya partisipasi manusia dalam proses pengkodean, sebuah aspek yang sering terlewatkan dalam penelitian AI generatif.
2. Peng et al. (2023) berfokus pada penyetelan instruksi menggunakan GPT-4, mengeksplorasi bagaimana GPT-4 dapat digunakan untuk meningkatkan kinerja zero-shot dalam tugas-tugas baru.
3. Savelka et al. (2023) meneliti kemampuan GPT-4 dalam menganalisis data tekstual yang membutuhkan keahlian domain khusus, seperti hukum, menunjukkan bahwa GPT-4 dapat menjadi alat yang berguna dalam analisis semacam ini.
4. Katz et al. (2023) menemukan bahwa GPT-4 dapat mengungguli manusia dan model sebelumnya dalam Ujian Pengacara, menunjukkan potensi model bahasa besar dalam mendukung layanan hukum.
5. Zheng et al. (2023) menguji kemampuan GPT-4 dalam mencari arsitektur neural, sebuah pendekatan inovatif yang menyoroti fleksibilitas dan kegunaan GPT-4 di luar bidang bahasa.
6. Tran et al. (2022) fokus pada penggunaan BERT dan GPT dalam deteksi intrusi cyber, menekankan pentingnya kurasi data dan jaminan kualitas.
7. Clark (2020) membahas tentang metodologi pra-pelatihan yang efisien dalam model BERT, sebuah konsep yang penting dalam pemahaman dasar algoritma AI generatif.

Studi ini direncanakan berfokus pada efektivitas Generative AI Tool berbasis GPT-4 dalam mensimulasikan serangan phishing, mencakup pembuatan email

phishing yang realistis, pengembangan formula prompting optimal, dan evaluasi efektivitas email phishing yang dihasilkan. Semua penelitian, termasuk rencana penelitian ini, berbagi minat umum dalam memahami dan memanfaatkan kemampuan model GPT-4. Fokus utamanya adalah pada bagaimana GPT-4 dapat diaplikasikan dalam berbagai skenario praktis, dari coding hingga keamanan siber. Perbedaan dari penelitian ini dan yang terdahulu adalah dalam hal luas cakupan aplikasi GPT-4, dimana rencana penelitian ini secara khusus mengkaji penerapan GPT-4 dalam konteks keamanan siber, khususnya dalam simulasi serangan phishing. Penelitian ini tidak hanya menguji kecanggihan teks yang dihasilkan oleh GPT-4, tetapi juga mendalami aspek formula prompting dan evaluasi efektivitasnya dalam konteks yang lebih spesifik.

Tabel 2.1 Hasil Penelitian terdahulu, persamaan dan Perbedaan

No	Peneliti, Tahun, Judul	Tujuan Penelitian	Ringkasan Penelitian	
			Persamaan	Perbedaan
1	Russell A. Poldrack, Thomas Lu, Gašper Beguš. (April, 2023). AI-assisted coding: Experiments with GPT-4. IEEE. arXiv:2304.13187. https://www.doi.org/10.48550/arXiv.2304.13187	Memahami kemampuan dan keterbatasan alat pengkodean AI berbasis GPT-4 dan menyoroti pentingnya partisipasi manusia dalam proses pengkodean.	Penggunaan Genertaive AI tool berbasis GPT-4 dalam penelitian.	Mengeksplorasi penggunaan Genertaive AI tool berbasis GPT-4 berdasarkan model bahasa besar, untuk pembuatan kode komputer dan refactoring.
2	Baolin Peng, Chunyuan Li, Pengcheng He, M Galley, Jianfeng Gao. (April, 2023). Instruction Tuning with GPT-4. arXiv.org-Vol. abs/2304.03277. https://www.doi.org/10.48550/arXiv.2304.03277	Penggunaan GPT-4 untuk menghasilkan data mengikuti instruksi untuk menyempurnakan model bahasa besar, yang mengarah ke kinerja zero-shot yang unggul pada tugas-tugas baru.	Menggunakan GPT-4, model bahasa, untuk menghasilkan data mengikuti instruksi	Penelitian menyoroti potensi penggunaan GPT-4 untuk penyetelan instruksi dan kemampuannya untuk meningkatkan kemampuan tembakan nol LLM.
3	Jaromir Savelka, Kevin D. Ashley, Morgan Gray, Hannes Westermann, Huihui Xu. (June, 2023). Can GPT-4 Support Analysis of Textual Data in Tasks Requiring Highly Specialized Domain Expertise?. arXiv.org-Vol. abs/2306.13906. https://www.doi.org/10.48550/arXiv.2306.13906	Bisakah GPT-4 Mendukung Analisis Data Tekstual dalam Tugas yang Membutuhkan Keahlian Domain yang Sangat Terspesialisasi? Ya, GPT-4 berkinerja baik dalam menganalisis data tekstual dalam tugas-tugas yang membutuhkan keahlian	Generative AI tool GPT-4 mendukung analisis data tekstual dalam tugas yang membutuhkan keahlian terspesialisasi dan GPT-4 untuk membantu keamanan siber.	Penelitian ini secara khusus, fokusnya adalah menganalisis pendapat pengadilan untuk menafsirkan konsep hukum dan keamanan.

		domain yang sangat khusus, seperti menganalisis pendapat pengadilan untuk menafsirkan konsep hukum. Temuan penelitian ini dapat bermanfaat bagi peneliti dan praktisi yang terlibat dalam anotasi teks semantik/pragmatis dalam tugas-tugas yang membutuhkan keahlian domain yang sangat khusus.		
4	Daniel Martin Katz, Michael James Bommarito, Shan Gao, Pablo D. Arredondo. (March, 2023). GPT-4 Passes the Bar Exam. SSRN: https://ssrn.com/abstract=4389233. https://www.doi.org/10.2139/ssrn.4389233	GPT-4 mengungguli manusia dan model sebelumnya pada Ujian Pengacara, menunjukkan potensi model bahasa untuk mendukung layanan hukum.	Penggunaan Genertaive AI tool berbasis GPT-4 dalam penelitian, menunjukan potensi model bahasa dan mendukung keamanan siber	GPT-4 mengungguli peserta tes manusia dan model sebelumnya pada Multistate Bar Examination (MBE) pilihan ganda, dengan peningkatan 26% dibandingkan ChatGPT dan mengalahkan manusia di lima dari tujuh bidang studi. Kemajuan yang cepat dan luar biasa dari kinerja model bahasa yang besar dan menyoroti potensi

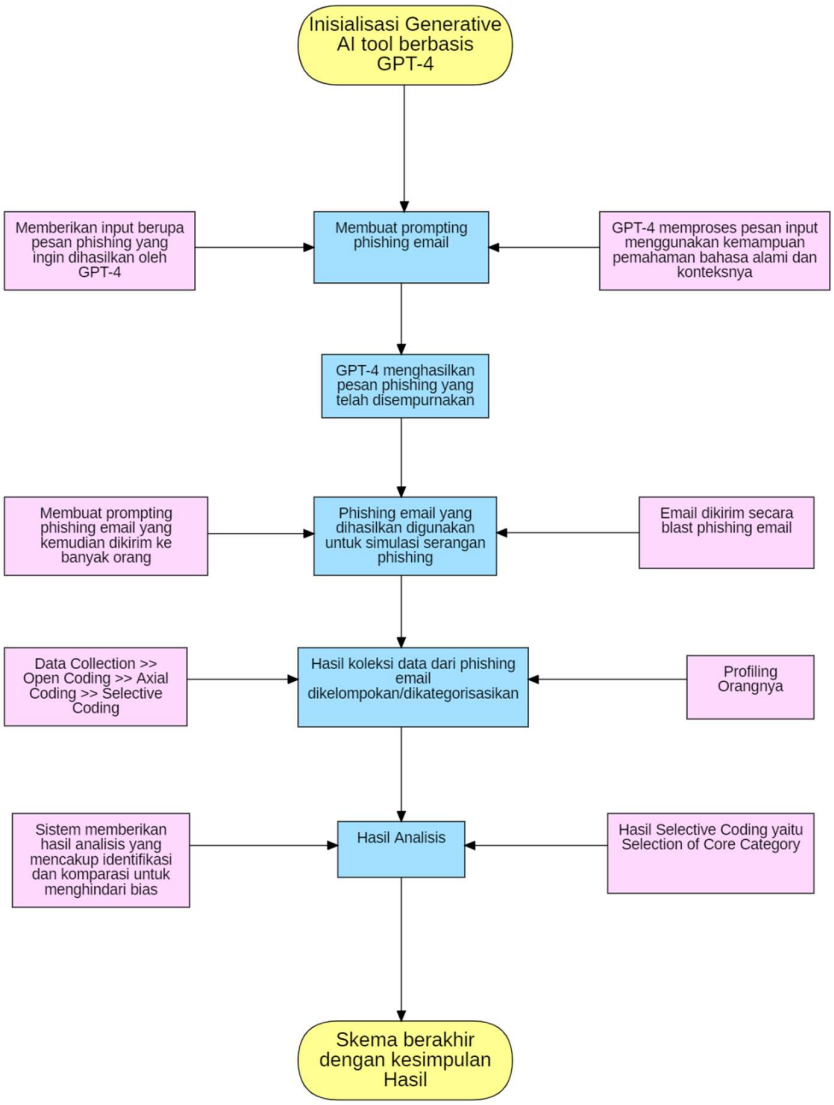
				model tersebut untuk mendukung penyampaian layanan hukum di masyarakat.
5	Min Zheng, Xiu-ping Su, Shan You, Fei Wang, Chen Qian, Chang Xu, Samuel Albanie. (April, 2023). "Can GPT-4 Perform Neural Architecture Search?". Cornell University, Ithaca, New York. arXiv:2304.10970. https://www.doi.org/10.48550/arXiv.2304.10970	Penelitian ini untuk menyoroti potensi GPT-4 untuk membantu penelitian tentang masalah teknis yang menantang melalui skema pendorong sederhana yang membutuhkan keahlian domain terbatas. Makalah ini menunjukkan bahwa penelitian di masa depan dapat memanfaatkan model bahasa tujuan umum untuk tugas pengoptimalan yang beragam dan menyoroti batasan penting untuk penelitian ini, serta implikasi untuk keselamatan AI.	Memanfaatkan penggunaan GPT-4 dalam penelitian. Generatif GPT-4 menunjukkan efektivitasnya dalam menavigasi ruang pencarian arsitektur dan meningkatkan kinerja.	Penelitian mengkaji efektivitas GENIUS (Metode Genius Learning/Holistik Learning) dalam menavigasi ruang pencarian arsitektur dengan cepat, menunjukkan dengan tepat kandidat yang menjanjikan, dan secara berulang menyempurnakannya untuk meningkatkan kinerja. Studi ini membandingkan GENIUS dengan teknik NAS (Neural Architecture Search) yang memanfaatkan kemampuan generatif GPT-4.

6	N. Tran, H. Chen, J. Bhuyan and J. Ding. (October, 2022). "Data Curation and Quality Evaluation for Machine Learning-Based Cyber Intrusion Detection," in IEEE Access, vol. 10, pp. 121900-121923, 2022, https://ieeexplore.ieee.org/document/9907008. doi: 10.1109/ACCESS.2022.3211313. Source Scopus. or https://arxiv.org/abs/2105.10041v1. https://doi.org/10.48550/arXiv.2105.10041	Penelitian ini bertujuan untuk memulai perspektif kualitas data bagi peneliti dan praktisi untuk meningkatkan kinerja deteksi intrusi berbasis pembelajaran mesin.	BERT dan GPT ditemukan sebagai algoritma terbaik pada semua kumpulan data.	Penelitian berfokus pada pentingnya kurasi data dan jaminan kualitas untuk deteksi intrusi cyber berbasis pembelajaran mesin.
7.	K. Clark. (March, 2020). ELECTRA: PRE-TRAINING TEXT ENCODERS AS DISCRIMINATORS RATHER THAN GENERATORS. 8th International Conference on Learning Representations, ICLR 2020, cited by 945 (315.00 per year). arXiv:2003.10555. https://doi.org/10.48550/arXiv.2003.10555	Metode pra-pelatihan pemodelan bahasa bertopeng (MLM) seperti BERT merusak input dengan mengganti beberapa token dengan [MASK] dan kemudian melatih model untuk merekonstruksi token asli.	Menggunakan algoritma BERT	Menunjukkan bahwa tugas pra-pelatihan yang baru lebih efisien daripada MLM karena tugas ini didefinisikan atas semua token input daripada hanya subset kecil yang disamarkan. Hasilnya, representasi kontekstual yang dipelajari oleh pendekatan secara substansial mengungguli representasi yang dipelajari oleh BERT dengan ukuran model, data, dan komputasi yang sama.

8.	Satnam Narang. (March, 2023). OpenAI's ChatGPT and GPT-4 Used as Lure in Phishing Email, Twitter Scams to Promote Fake OpenAI Tokens. https://www.tenable.com/blog/openai-chatgpt-and-gpt-4-used-as-lure-in-phishing-scams-to-promote-fake-token-airdrop	ChatGPT dan GPT-4 digunakan sebagai umpan dalam serangan phishing untuk mempromosikan token palsu OpenAI.	Penggunaan Genertaive AI tool berbasis GPT-4 melakukan serangan phishing dalam penelitian.	GPT-4 digunakan untuk membuat malware. GPT-4 memiliki potensi untuk digunakan kejahatan cyber lainnya. Oleh karena itu, pentingnya memahami dan mengatasi potensi risiko yang terkait dengan penggunaan GPT-4 dalam konteks keamanan cyber.
9.	Rudolph, J. (April, 2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. Scopus. Journal of Applied Learning and Teaching, 6(1), 364-389, ISSN 2591-801X, <https://doi.org/10.37074/jalt.2023.6.1.23>	Melakukan pemeriksaan mendalam terhadap berbagai chatbot inovatif dalam bahasa Inggris dan Mandarin, menyajikan latar belakang dan sejarah mereka, dengan tujuan mengeksplorasi teknik perbandingan yang digunakan dan melakukan tes komprehensif di ranah pendidikan tinggi.	Penggunaan Genertaive AI tool berbasis GPT-4 dalam penelitian.	Hasil penelitian menunjukkan bahwa meskipun ada kekhawatiran tentang tingkat kecerdasan yang dicapai oleh chatbot AI, GPT-4 dan pendahulunya tampil terbaik, sementara Bing Chat dan Bard tampil serupa dengan siswa berisiko rendah dengan nilai F rata-rata.

10.	Rahman, M.M. (2023). ChatGPT for Education and Research: Opportunities, Threats, and Strategies. Scopus. Applied Sciences (Switzerland), 13(9), ISSN 2076-3417, <https://doi.org/10.3390/app13095783>	Penelitian ini bertujuan untuk menyajikan hasil survei dan analisisnya guna memahami seberapa besar dampak ChatGPT dalam mendukung pembelajaran pemrograman serta bagaimana teknologi ini dapat membantu pengembangan pendidikan secara lebih luas.	Penggunaan Genertaive AI tool berbasis GPT model bahasa besar yang kuat yang dikembangkan oleh OpenAI dalam penelitian.	Terlepas dari kekhawatiran tentang risiko yang ditimbulkan oleh ChatGPT dalam pengaturan pendidikan, termasuk kemungkinan kecurangan dan tantangan dalam mengevaluasi informasi yang dihasilkan, penelitian ini terutama meneliti keuntungan dan potensi teknologi kecerdasan buatan, khususnya ChatGPT, di bidang pendidikan dan pendidikan pemrograman.

2.3 Kerangka Pemikiran



Gambar 2. 5 Kerangka Pemikiran

Gambar 2.4 adalah kerangka konseptual yang merupakan jenis kerangka pemikiran yang menawarkan struktur logis dari konsep-konsep yang saling berhubungan yang membantu memberikan gambaran atau tampilan visual tentang bagaimana ide-ide dalam sebuah penelitian berhubungan satu sama lain. Kerangka pemikiran digunakan untuk membatasi ruang lingkup data yang relevan dengan memfokuskan pada variabel-variabel tertentu dan mendefinisikan sudut pandang (kerangka) tertentu yang akan diambil oleh peneliti dalam menganalisis dan menginterpretasikan data yang akan dikumpulkan. Kerangka pemikiran membantu

memberikan arah dan fokus untuk penelitian, dan berfungsi sebagai peta jalan untuk penelitian, memberikan arah dan fokus untuk penelitian.

Kerangka berpikir membantu peneliti dalam mendapatkan konsep yang matang yang kemudian digunakan untuk menjelaskan setiap permasalahan dalam penelitian. Bagan kerangka berpikir sebagaimana Gambar 2.4 menunjukkan rencana kerangka kerja dan proses penelitian dalam mewujudkan tujuan penelitian dan mempelajari perilaku pengguna dalam menanggapi serangan phishing email dalam rangka menjawab pertanyaan penelitian. Secara lebih rinci, proses penelitian dapat dijabarkan sebagai berikut:

1. Inisialisasi Generative AI tool berbasis GPT-4

   Pada tahap ini, alat Generative AI berbasis GPT-4 diinisialisasi. Alat ini kemudian diberi input berupa pesan phishing yang ingin dihasilkan.
2. Membuat prompting phishing email

   Setelah alat Generative AI berbasis GPT-4 diinisialisasi, langkah selanjutnya adalah membuat prompting phishing email. Prompting phishing email ini digunakan untuk mengarahkan alat Generative AI berbasis GPT-4 dalam menghasilkan pesan phishing email yang realistis dan menarik.
3. GPT-4 memproses pesan input

   GPT-4 kemudian memproses pesan input yang telah diberi prompting. GPT-4 menggunakan kemampuan pemahaman bahasa alami dan konteksnya untuk menghasilkan pesan phishing yang telah disempurnakan.
4. Membuat prompting phishing email yang telah disempurnakan

   GPT-4 kemudian menghasilkan pesan phishing email yang telah disempurnakan. Pesan phishing email ini kemudian dikirim secara blast ke banyak orang untuk simulasi serangan phishing.
5. Data Collection >> Profiling

   Hasil koleksi data dari phishing email kemudian dikelompokan/ dikategorisasikan menggunakan teknik analisis data kualitatif. Teknik analisis data kualitatif yang digunakan adalah teknik open coding, axial coding, dan selective coding.
6. Open Coding

   Pada tahap open coding, data yang telah dikumpulkan dikelompokkan berdasarkan tema-tema umum.

7. Axial Coding

   Pada tahap axial coding, tema-tema umum yang telah diidentifikasi pada tahap open coding kemudian dihubungkan satu sama lain.

8. Selective Coding

   Pada tahap selective coding, tema-tema yang paling penting diidentifikasi dan disimpulkan.

9. Selective Coding

   Pada tahap selective coding, tema-tema yang paling penting diidentifikasi dan disimpulkan.

10. Selective Coding

    Pada tahap selective coding, tema-tema yang paling penting diidentifikasi dan disimpulkan.