DOSANJH USA

3/27/2025

AI Model Considerations

Clarity, Conciseness, Security and Trust

Sunny “Dos” Dosanjh
DOSANJH USA (DUSA)

AI Model Considerations

Clarity, Conciseness, Security and Trust

It was a simple ask. An editor and publisher requested that we summarize and provide an abstract for an upcoming book "*Permission to Walk in Peace*." The abstract was going to be distributed nationwide to media, bookstores and other outlets. Context and content related to how US Military Veterans healed from trauma, who helped them and how they recovered to reintegrate into society. We compared the DUSA AI Platform output to five other publicly available AI Prompt and Language Model services and the results are summarized as follows:

Gemini AI

- The paragraph is concise but lacks depth and emotional resonance.
- It doesn't provide a clear understanding of the book's themes or unique aspects.
- The tone is neutral, which might not grab the reader's attention.

Perplexity AI

- The paragraph is a bit longer, but it still feels somewhat generic and doesn't fully capture the book's essence.
- It mentions the book's themes and structure, but the language is not as engaging as the DUSA AI paragraph.
- The tone is informative, but it might not be as compelling to readers seeking a more personal connection.

Cohere AI

- The paragraph feels a bit disjointed, with abrupt transitions between sentences.
- It mentions the book's dedication and structure, but the language is not as vivid or engaging as the DUSA AI paragraph.
- The tone is somewhat dry, which might not appeal to readers seeking a more emotive experience

Mistral AI

- The paragraph is well-structured and provides a clear overview of the book's content.
- However, it might be a bit too focused on the Vietnam War aspect, which might limit its appeal to readers interested in other conflicts.
- The tone is somewhat somber, which might not fully convey the book's themes of hope and resilience.

CHAT GPT

- Lack of specific details around structure and themes of Experience, Handrails and Mending Fences
- Less emphasis of the transformative power of sharing and artistic expression as DUSA AI
- More generic language and less sense of scope about how it will impact a broader audience

AI services like Microsoft CoPilot, Meta AI and Claude were also reviewed but had limitations for their freemium services. These included limitations of file type, file size and an inherent bias towards language that did not meet their filtering criteria. Furthermore, relying on a single vendor for such queries creates a lock-in effect, necessitating multi-year subscription licenses. Meanwhile, Large Language Models (LLMs) that aren't regularly updated or maintained will yield outdated responses and become vulnerable to cybersecurity risks.

DUSA AI Output

The *Permission to Walk in Peace* book was uploaded to a directory in Google Drive, and a query was entered into the customized *Veterans Lifeline* prompt: "*Create a summary of Permission to Walk in Peace of approximately 150 words in length to be used as an annotation on the back of the book*" The book has approximately 300 pages and uploaded as a .pdf file.

Figure 1: DUSA AI Platform Prompt Query Output

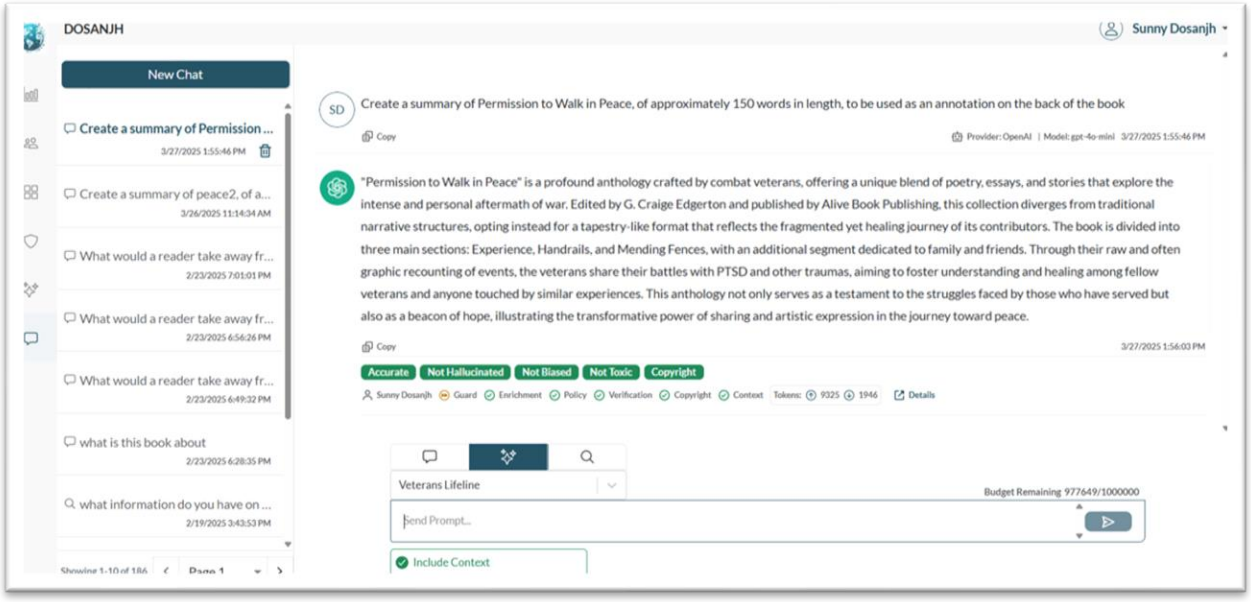


Figure 1 displays the generated output that was selected by the editor & publisher and will be included when the book is released in the next few weeks ~ mid-April 2025. Key considerations were copyright check, context and how the inference policies were implemented.

"Permission to Walk in Peace" is a profound anthology crafted by combat veterans, offering a unique blend of poetry, essays, and stories that explore the intense and personal aftermath of war. Edited by G. Craige Edgerton and published by Alive Book Publishing, this collection diverges from traditional narrative structures, opting instead for a tapestry-like format that reflects the fragmented yet healing journey of its contributors.

The book is divided into three main sections: Experience, Handrails, and Mending Fences, with an additional segment dedicated to family and friends. Through their raw and often graphic recounting of events, the veterans share their battles with PTSD and other traumas, aiming to foster understanding and healing among fellow veterans and anyone touched by similar experiences. This anthology not only serves as a testament to the struggles faced by those who have served but also as a beacon of hope, illustrating the transformative power of sharing and artistic expression in the journey toward peace.

AI Security and Trust Considerations

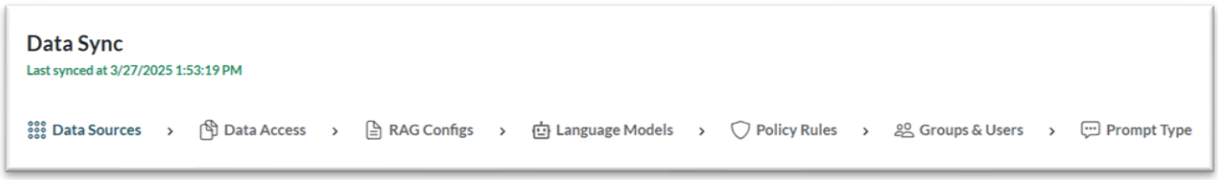
Identity, data protection, observability dashboards and reports are core components of the DUSA AI Platform cybersecurity considerations. Advanced capabilities are integrated via APIs to address updated quantum-resilience requirements or enhanced LLM inspection as related to the following:

- **LLM Evasion**: Address model deception, jailbreaking or prompt injection
- **Data Poisoning:** Detect injection of malicious or misleading data in the training dataset
- **Post-Quantum Cryptography:** Integrate advanced quantum resilience algorithms via API access
- **Excessive Agency:** Policies that are configurable by department, group and user level access

AI Prompt Configuration

A custom prompt, "*Veterans Lifeline*" was configured with the following constructs outlined in Figure 2.

Figure 2: DUSA AI Platform Custom Prompt Data Sync Configuration



These constructs utilize Inference policies to determine:

- **Data Sources & Access:** Location of local or cloud data sources and directory structures
- **RAG Configs:** Selection of Vector DB, Embedding Model, Similarity Algorithm, Chunk Size
- **Language Models:** Authorized LLMs whether "*Bring Your Own Model*" or cloud provisioned
- **Policy Rules:** PII, PHI, Software Attributes, Protected Finance & Legal Information and the option to:
  - Block
  - Redact
  - Anonymize
  - Allow
  - Custom Policy
- **Groups & Users:** Integrate Enterprise IDPs for access controls based upon organization hierarchies
- **Prompt Type:** Dynamic text defined or classic pre-defined prompts

LLM Clarity and Conciseness

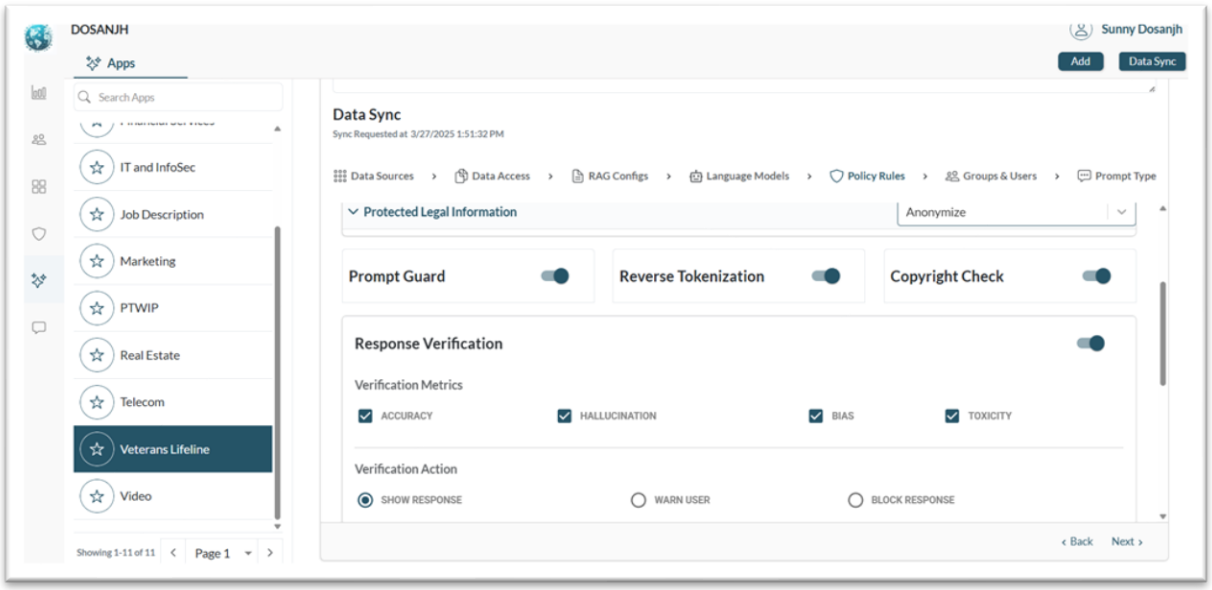
The DUSA AI Platform provided the editor and publisher with an output that they favored. The platform provides the ability to select an LLM and apply both Inference and cybersecurity policies on the entered prompt. An example is illustrated in Figure 3 whereby the platform provides the option to select various content options such as:

- **Prompt Guard:** Protection against inappropriate queries (Jailbreaking / Space Breaker Prompt Injector)
- **Reverse Tokenization:** Level of inspection on the return response
- **Copyright Check:** Verify if the content or "copy" violates any public copyright protections
- **Response Verification Metrics:** Determine the inspection level for the prompt response

The following verification metrics are available on the administration page to promote clear messaging and concise vocabulary. Controls are configured based upon the organizations InfoSec and content use policies to show a response, warn users or block the response.

- **Accuracy:** Measures how often an AI model's predictions match the actual outcomes, reflecting an ability to make correct decisions.
- **Hallucination:** Occurs when an AI model generates or predicts information that isn't based on actual data, resulting in false or misleading results.
- **Bias:** Refers to the unfair or discriminatory tendencies in an AI model's decisions, reflecting prejudices or imbalances in the training data.
- **Toxicity:** Involves the generation or dissemination of hateful, abusive, or harmful content by an AI model, posing risks to individuals or communities.

Figure 3: DUSA AI Platform Prompt Configuration

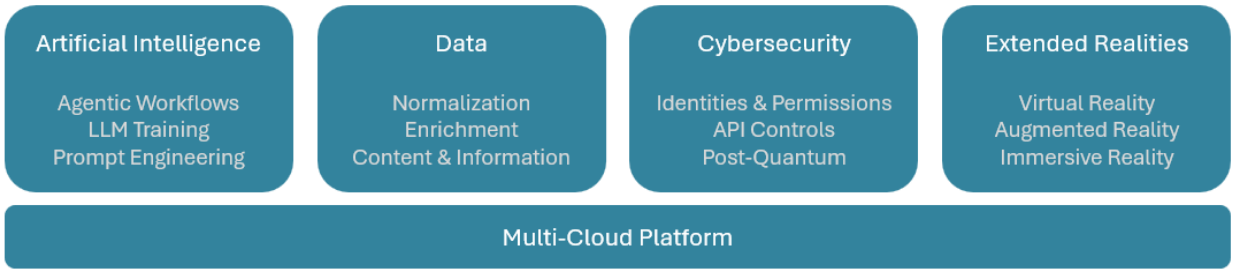


Conclusion

The DUSA AI Platform was selected for the book abstract due to its superior conciseness, clarity, and security. Compared to other AI models, DUSA AI provided a well-structured and emotionally resonant summary that effectively captured the book's themes while maintaining a compelling tone. Its ability to generate a precise and engaging annotation set it apart from competing platforms that lacked depth or coherence.

Additionally, the platform's robust cybersecurity measures—including protections against data poisoning, LLM evasion, and post-quantum cryptography—ensured the integrity and trustworthiness of the output. By offering customizable inference policies, content verification metrics, and advanced access controls, DUSA AI not only delivered an accurate and impactful book summary but also ensured compliance with security and copyright standards. These factors made it the preferred choice for the editor and publisher in preparing the nationwide book release.

ABOUT DUSA:
DUSA is a Silicon Valley Corporation that is invested in the future of technology that benefits societies. Our DUSA AI Platform leverages distributed multi-cloud capabilities that enables organizations to rapidly deploy and expand products and services. We specialize in how current information systems can be prepared for upgrades, new technologies and a comprehensive runbook to manage the operation.



Contact us and we can provide a path forward leveraging these latest technology developments to further expand your business.

CONTACT INFORMATION
info@dosanjhusa.com
https://www.dosanjhusa.com
+1.408.430.DUSA (3872)