



GENERATIVE AI AND INTELLECTUAL PROPERTY PRACTICE:

A Webinar Series on Guardrails, Guidelines, and
Considerations

Before We Get Started...



Recording

A link to the recording and slides will be emailed to all registrants. The slides are posted on our website.



Questions

Type into the Q&A and we will try to get to your question at the end.



Social

Follow us on LinkedIn or go to SLW Institute on slwip.com to see upcoming and on demand webinars.

Polling Question

Do you regularly use generative AI in your job?



Panel



Steve Lundberg

Principal

Schwegman
Lundberg &
Woessner



Andre Marais

Principal

Schwegman
Lundberg &
Woessner



Ana-Maria Popescu

Attorney

Schwegman
Lundberg &
Woessner



Tyler Nasiedlak

Principal

Schwegman
Lundberg &
Woessner



Lucas Hjelle

Principal

Schwegman
Lundberg &
Woessner



Alexis Liistro

Attorney

Schwegman
Lundberg &
Woessner



Anup Suresh

Principal

Schwegman
Lundberg &
Woessner

Overview

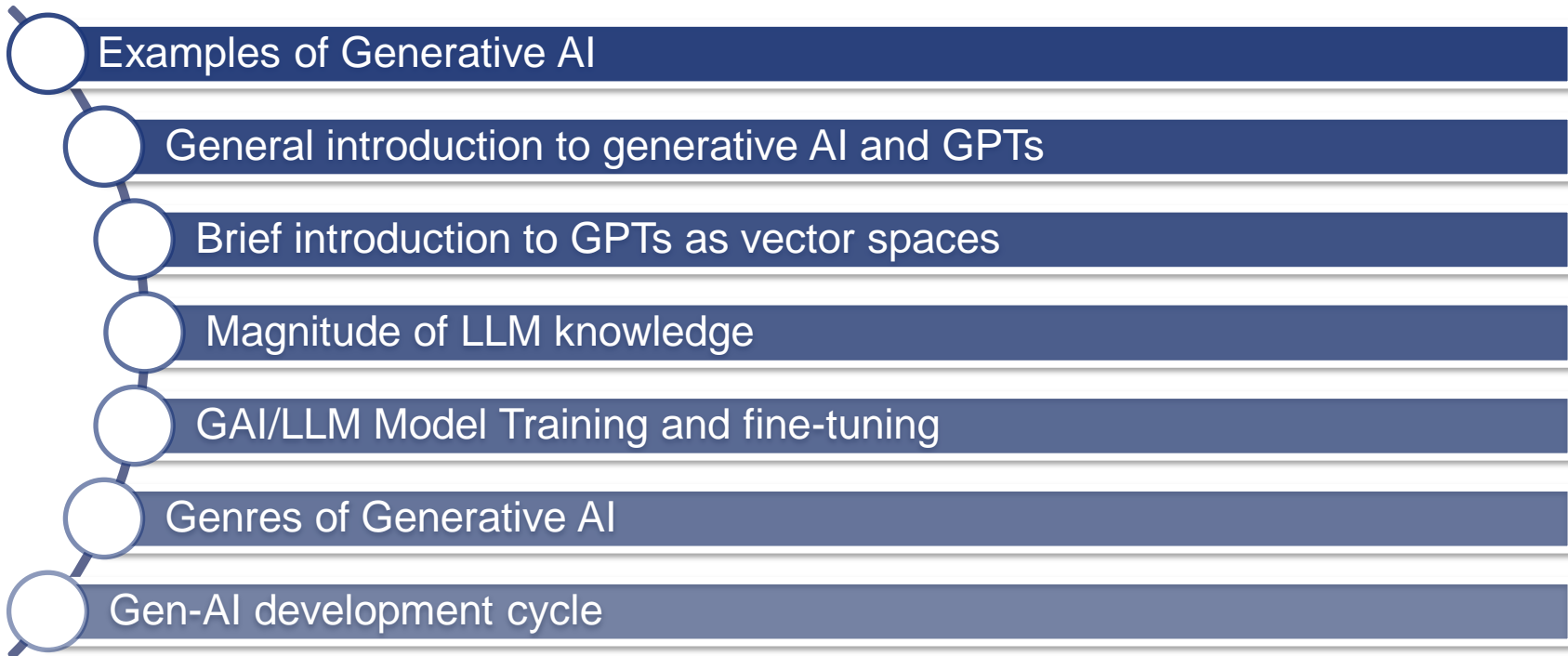
- What is generative artificial intelligence?
- GenAI and Privacy & Security
- Practical: What we see on the ground
- What can lawyers do to safely use GenAI in their practice?
- What can clients do to safely use GAI in their products?



What is Generative Artificial Intelligence?



Overview of generative artificial intelligence (GAI) and large language models (LLMs):

- 
- Examples of Generative AI
 - General introduction to generative AI and GPTs
 - Brief introduction to GPTs as vector spaces
 - Magnitude of LLM knowledge
 - GAI/LLM Model Training and fine-tuning
 - Genres of Generative AI
 - Gen-AI development cycle

Examples of Genres of Gen AI:





Text-to-Image (T2I)




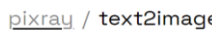
Text-to-Text (T2T)



Text-to-Video (T2V)

Text-to-Audio (T2A)




Text-to-Motion (T2M)




DALL·E 2 - OpenAI    





   




GauGAN2 by NVIDIA  




.....





 Simplified <https://simplified.com/ai>  Jasper <https://www.jasper.ai>  frase.io <https://www.frase.io>

 eleuther.ai <https://www.eleuther.ai>  Copy.ai <https://www.copy.ai>  marketmuse.com <https://www.marketmuse.com>

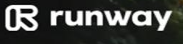
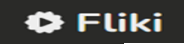
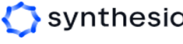
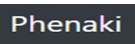
 writesonic.com <https://writesonic.com>  cohere.com <https://cohere.com>  chibi.ai <https://chibi.ai>  copysmith.ai <https://copysmith.ai>



 ideanote.io <https://ideanote.io/idea-generator>  copysmith.ai <https://copysmith.ai>  Flowrite <https://www.flowrite.com>

 sudowrite.com <https://www.sudowrite.com>  textcortex.com <https://textcortex.com>  nichesss.com <https://nichesss.com>





   

.....



   

 Meta <https://ai.meta.com>  Google AI <https://ai.google>

.....

   wellsaidlabs.com <https://wellsaidlabs.com>  descript.com <https://www.descript.com>

.....

Examples of Genres of Gen AI:

Text-to-Code (T2C)

Text-to-NFT (T2N)

Text-to-3D (T2D)


Audio-to-Text (A2T)

Audio-to-Audio (A2A)

Image-to-Text (I2T)

Brain-to-Text (B2T)


Multimodal (TTS, A2T)

 Replit
<https://replit.com/ai>


 GitHub
<https://github.com/features/copilot>

 chaintext.net
<https://chaintext.net>


 lens-ai.com
<https://lens-ai.com>


 dreamfusion3d.github.io
<https://dreamfusion3d.github.io>


 alpha3d.io
<https://www.alpha3d.io>


 the-decoder.com
<https://the-decoder.com>

 descript.com
<https://www.descript.com>

 AssemblyAI
<https://www.assemblyai.com>

 whisper.ai
<https://whisper.ai>

 Voicemod
<https://www.voicemod.net>


 DeepAI
<https://deepai.org>

 VEED.IO
<https://www.veed.io>

 classace.io
<https://www.classace.io>

 FreeConvert
<https://www.freeconvert.com>

 **speech from brain** **non-invasive brain recordings**

 Hugging Face
<https://huggingface.co/speechbrain>

GAI/LLM: General Introduction

- Generative AI refers to a class of AI that can generate new content
 - Can generate text, images, or music without direct human input
 - GAI mimics human creativity and problem-solving by learning statistical patterns from vast datasets
- Generative Pre-trained Transformers (GPTs) are GAI models that use transformer architecture for natural language processing (NLP)
 - GPT models are pre-trained on datasets before fine-tuning for specific tasks
- Large Language Models (LLMs) capture statistical relationships between words within vast amounts of linguistic training data
 - May be used for various generative or transformative language tasks

GAI/LLM: Vector Spaces

- LLMs use vector spaces to store and manipulate words
 - Parameters of an LLM include vectors that represent the numerical relationships between words (e.g., weights and biases in the neural network)
 - Each word or token the model is associated with a unique vector (word embedding)
- Vector spaces provide a mathematical framework and enable mathematical operations
 - Example vector operation: "king" - "man" + "woman" = "queen"
- Similar words have similar vectors, capturing their semantic similarity
 - LLM training adjusts vectors based on the context in which words appear
- LLMs have no concept of grammar or other rules of language
 - LLMs operate as statistical models based on patterns learned from vast training data
 - LLMs generate text that align with statistical regularities observed in training data

GAI/LLM: Magnitude of LLM knowledge

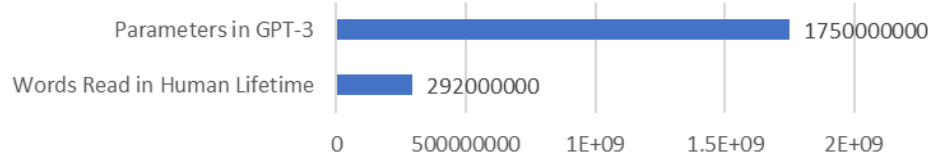
Humans are exposed to about 292 million words in a lifetime

GPT-3 includes 1.75 billion parameters (vectors representing word relationships)

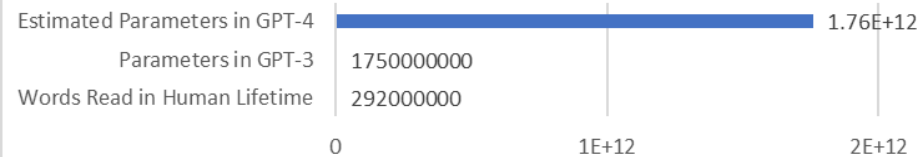
GPT-4 is estimated to include 1.75 trillion parameters

LLMs are trained on about 20x parameters than in final model

Human Word Exposure vs. GPT-3



Word Exposure and GPT-3 vs. GPT-4



GAI/LLM: Model Training and Fine-Tuning

Pre-training

- **Training data** Web text, book corpora, textbooks, code
- **Models** neural network, mixture of experts (network layers)
- **Self-supervised** Given a token sequence, predict next token

Querying LLMs: given prompt, complete answer

Dataset (GPT-3)	Weight (training mix)
Common Crawl	60%
WebText2	22%
Books1 & Books2	16%
Wikipedia	3%

Fine-tuning

- **Partially re-train pre-trained model with new data**
- **Reinforcement Learning from Human Feedback (RLHF)**

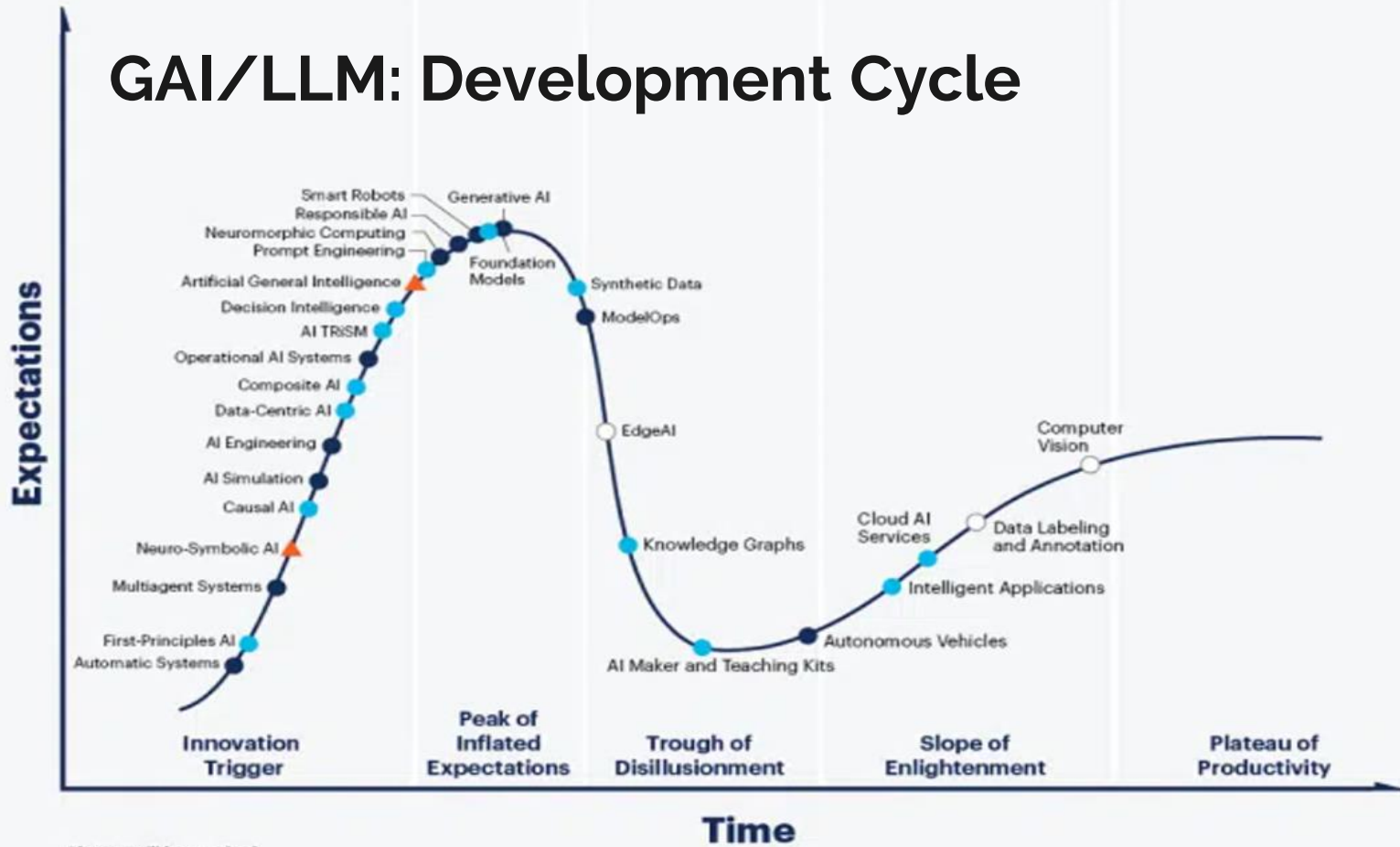
demo of expected behavior ← → ranking of system outputs

Build reward model that predicts average labelers' preference or given output

Use reward model to guide GPT-4 training

Reward refusals to answer some prompts, good answers to sensitive prompts: medical, legal

GAI/LLM: Development Cycle



2023 Gartner AI Hype Cycle

Plateau will be reached:

- less than 2 years
- 2 to 5 years
- 5 to 10 years
- ▲ more than 10 years
- ⊗ obsolete before plateau

As of July 2023

Trough of Disillusionment



Courtesy of DALL-E 3

GAI/LLM: LLM Strengths and Weaknesses

- **Strengths**

- Multi-task, off-the-shelf natural language processing functionality

- Coherent, realistic generative output

- **Weaknesses**

- Hallucinations: invented, incorrect, non-factual content

- Unsupported generated claims**
 - Unsupported summaries**
 - Invented case law**

- Limited evidence of "reasoning" or consistent "self-correction"

- Inconsistent argument or claims**
 - Inconsistent post-revision quality**

- Harmful content: offensive or harassing content

- instructions for committing illegal acts

- Inadvertent retrieval and disclosure of protected or sensitive information

GenAI & Privacy, Security, Data and Protection



Security and Privacy Considerations

Data Security

Retention of IP Ownership of Inputs and Outputs

Foundation Model Used

Personally Identifiable Information (PII) and Compliance

Ethical Obligations

Internal Policies

Keeping Up-to-Date

Existing Tools Providers Continue to Innovate/Rely on Generative AI

Security and Privacy Considerations (Non-AI specific)



4th Party Concerns

Insurance

Access Control

Practical: What We See on the Ground



Practical Law Firm Use Cases:



Courtesy of DALL-E 3


Practical Law Firm Use Cases:



POLICY:

- Enforces **appropriate usage** within our legal and ethical frameworks.
- Continuous **human oversight (HITL)**, is needed to realize the benefits of AI while safeguarding quality, accountability, and trust.

Preparation Use Cases

- 
- 1. Disclosure Preparation:**
 - a. Process invention disclosure documents
 - b. Attorney education
 - c. Generate questions for inventors
 - d. Establish state-of-the-art (?) - general overview – but not specific references
 - 2. Specification drafting assistance/copilot:**
 - a. NOT good for single input, output
 - b. BETTER for iterative process to:
 - i. Adding extra "meat to the bones"
 - ii. Suggestions for identifying technical problems and solutions in disclosure materials
 - iii. Covert between claim types
 - iv. Suggest dependent claims
 - v. Enablement: Expand certain technical descriptions with known techniques and alternatives
 - 3. Review draft applications for quality and completeness**

Prosecution Use Cases



Summarizing
office actions

Suggesting
arguments

Suggesting
claim
amendments

Analyzing Prior Art and Portfolios

Current **foundation models** with general training struggle to:

Deliver usable patent literature search results

Digest and analyze large volumes of data in a context window

Specifically trained models show great potential for searching and portfolios to identify strengths, weaknesses, and opportunities

Models with **large context windows** or **prompt-chaining solutions** more helpful for reviewing larger bodies of documents

EoU and SEP Patent Identification

Identifying EoU for allowed or issued patents

- **Good at:** identifying potential 3rd party products to investigate
- **Not good at:** claim charts and publication dates

Analyzing claims to determine if a patent is potentially standards essential

- **Good at:** flagging claims potentially applicable to standard specifications
- **Not good at:** claim charts (unless stands specifications in context window)

Administrative Workflows

Automated
docketing

Automated
reporting

Invoice
review

Benefits of GAI

Key benefits of using AI in patent prosecution:

Increased efficiency and productivity

Reduced costs and improved margins

Higher application quality and better prosecution outcomes

Enables attorneys to focus on high-value strategic work

IP Contracts Use Cases

- Drafting Contracts
 - Proofing/grammar – poor with grammar; good with missing/improper definitions
 - Content generation and suggestions
 - Quick and numerous (will bullet point a list of 6-12 suggestions in little time)
 - Most suggestions (75%) are poorly reasoned (e.g., may recommend a change so that IP is jointly owned, and then may summarize that is critical that one party own all IP)
 - Some (25%) are potentially insightful and worth additional review
 - Useful for first draft of unusual contract terms; less helpful for customary terms,
- Analyzing Contracts
 - Can quickly finds relevant sections
 - Can offer some suggestions to pursue (like above, most poorly reasoned or unreliable but some are helpful or worth additional consideration)
- Limited test samples; training will improve results
- Adds value for its quick and numerous suggestions, but cannot be fully relied upon

GenAI Guardrails for Firms and Companies: Practical Patent & Copyright Safeguards



Risk Mitigation Strategies and Safeguards for GenAI



Practical Patent Considerations

- Implement human review of any GAI output
- Clear Terms of Use/Terms of Service
- Watermark/fingerprint/metadata all GAI output
- Maintain ongoing detailed documentation
- Disclaimers/warranties
- Plagiarism detection tools to scan GAI output
- Quality control of GAI output

Risk Mitigation Strategies and Safeguards for GenAI



Practical Copyright Considerations

- Provide attribution
- DMCA Notice and Takedown
- Obtain license/indemnity for copyrighted training data
- Obscurity: use federated (collaborative) learning to obscure source training datasets
- Screen for open-source code
- Limit on employee access to input datasets
- Implement audits
- Use LLM output as starting point

What Are The Rules Now? Where Are They Coming From?

1) GENAI AND CASE LAW/USPTO GUIDELINES

2) USPTO held public listening sessions and request for comments on questions:

- Questions for comment: <https://www.jdsupra.com/legalnews/uspto-is-holding-public-listening-1317823/>
- Listen here: <https://www.uspto.gov/about-us/events/ai-inventorship-listening-session-west-coast> October USPTO Talk

3) Case Law:

- There are primarily copyright cases out as of yet, but we should focus one webinar on applicable case law, USPTO discussions, and listening sessions.
- Future legislative changes will change the way LLMs are trained, likely with the idea of opt-outs requiring completely new training datasets.

4) Relevant Government Discussions:

- The “AI Bill of Rights” <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>
- The “AI Risk Management Framework”
- <https://doi.org/10.6028/NIST.AI.100-1>
- https://airc.nist.gov/AI_RMFKnowledgeBase/Playbook

[FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence | The White House](#)



Thank you for your interest.

What Questions Do You Have?



INSTITUTE
Building Extraordinary IP Assets

Schwegman Lundberg & Woessner | slwip.com



These materials are for general informational purposes only. They are not intended to be legal advice, and should not be taken as legal advice. They do not establish an attorney-client relationship.



INSTITUTE
Building Extraordinary IP Assets

Schwegman Lundberg & Woessner | slwip.com