

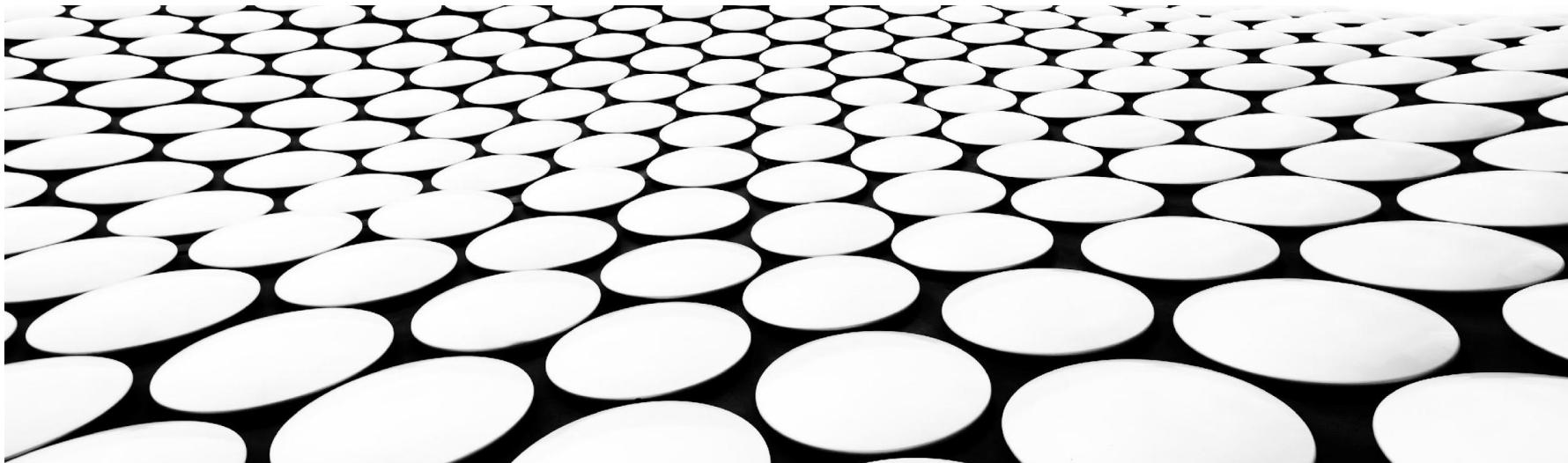
---

---

---

# IGNITE CHANGE: BUILD GENERATIVE AI CHATBOT SOLUTIONS FOR NON-PROFITS

THE WHAT, HOW AND WHY OF GENAI FOR NON-PROFITS



# Tim Spann

Twitter: @PaasDev // Blog: [datainmotion.dev](http://datainmotion.dev)

**Principal Developer Advocate.**

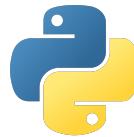
Princeton Future of Data Meetup.

Hightstown Elks Volunteer

ex-Pivotal, ex-Hortonworks, ex-StreamNative,  
ex-PwC, ex-HPE

<https://medium.com/@tspann>

<https://github.com/tspannhw>



DZone REF CARDS TREND REPORTS EXPERTS

## Top IoT Experts

Tim Spann  
Principal Developer Advocate, Cloudera  
<https://github.com/tspannhw/SpeakerProfile/>  
Tim Spann is a Principal Developer Advocate in Data in Motion for Cloudera. He works with Apache NiFi, Apache Pulsar, Apache...



---

# What is Generative AI?

**GenAI** encompasses a wide range of techniques used to create entirely new data – text, code, images, and even music.

While **ChatGPT** is a powerful generative pre-trained transformer model, it represents just a single approach within the broader field of generative AI.

# GenAI Beyond ChatGPT: A Boon for Nonprofits

Here's how it goes beyond **ChatGPT** and benefits nonprofits through **chatbots**:

- **Versatility:** GenAI isn't limited to chat; it can create compelling content – blog posts, social media updates, fundraising materials – tailored to a specific audience.
- **Data Augmentation:** It can generate synthetic data to train other AI models, helping nonprofits with limited resources leverage AI tools.
- **Personalization:** GenAI chatbots can personalize interactions based on user data, providing a more engaging experience.
- **24/7 Availability:** Chatbots provide constant support and answer basic queries, freeing up staff time for more complex tasks.

---

# ChatGPT Says: What is Generative AI?

Generative AI refers to a class of artificial intelligence algorithms and models that are designed to generate new content or data that is similar to the input data it was trained on. These models are particularly good at creating new content such as images, text, audio, and video.

Generative AI works by learning the patterns and features of a dataset and then using that knowledge to generate new examples that resemble the original data. This is often done through deep learning techniques, such as Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs).

Here are some common applications of generative AI:

- 1. Image Generation:** Generative models can create new images that resemble real photographs. This has applications in creating art, generating realistic images for games and movies, and even designing new products.
- 2. Text Generation:** Models like OpenAI's GPT (Generative Pre-trained Transformer) can generate coherent and contextually relevant text. This can be used for tasks like language translation, content creation, and chatbots.

---

## ChatGPT Says 2: What is Generative AI?

3. **Music and Audio Generation:** Generative models can create new music compositions, mimic the style of a particular artist, or even generate realistic speech.
4. **Video Synthesis:** There are models that can generate new video content, such as deepfakes which can create realistic videos of people saying or doing things they never actually did.
5. **Data Augmentation:** Generative models can also be used to augment datasets for training other AI models. For example, generating new images to add to a dataset for image recognition tasks.

Generative AI has incredible potential for creativity and innovation but also raises ethical considerations, especially regarding the generation of realistic but fake content (deepfakes) and potential misuse.

# Benefits of GenAI Chatbots for Nonprofits

## Fundraising & Engagement

## Information & Awareness

## Efficiency & Operations

Targeted outreach  
Event promotion  
Automated email campaigns

Educational content  
Multilingual communication

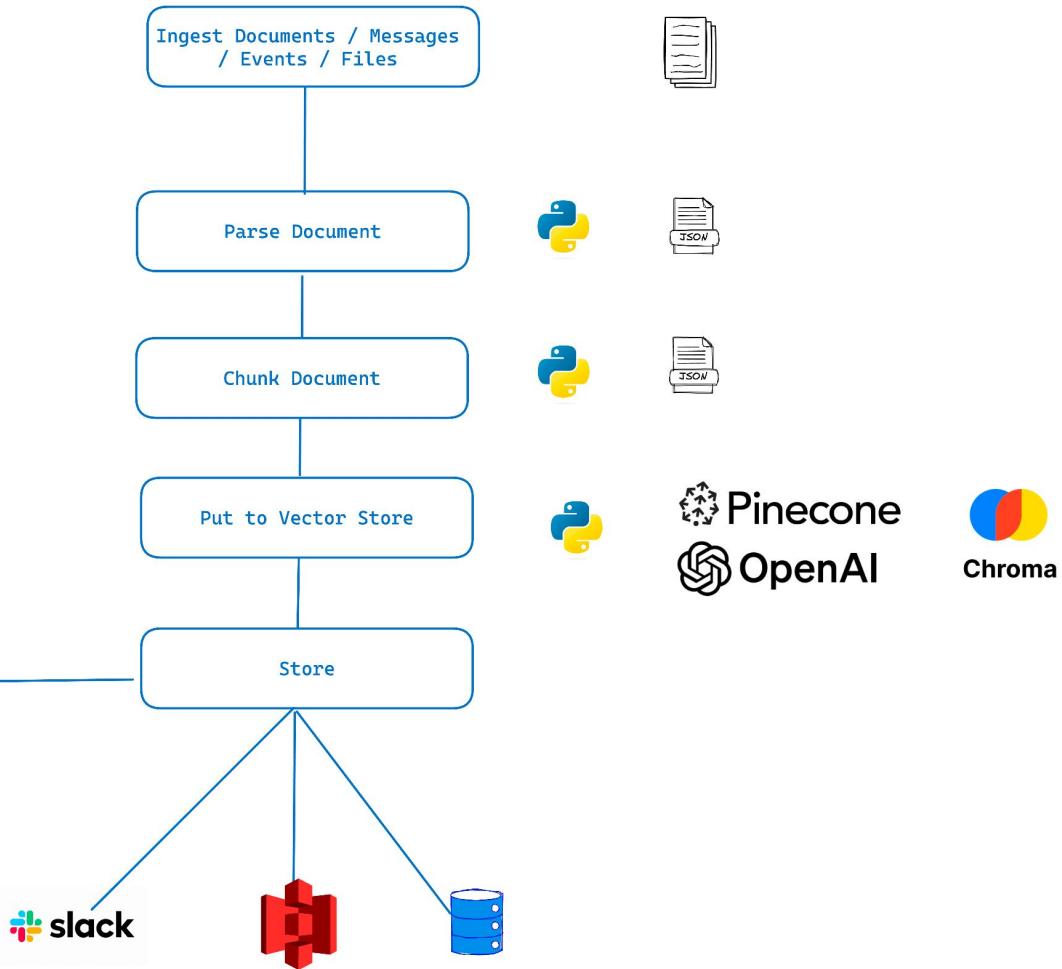
Data entry automation  
Feedback collection  
Cost reduction

24/7 donor support  
Volunteer recruitment

Personalized support  
Accessibility

Schedule management  
Report generation  
Reduced staff workload





Pinecone  
OpenAI





# Real-Time Pipelines

## External Context Ingest

Ingesting, routing, clean, enrich, transforming, parsing, chunking and vectorizing structured, unstructured, semistructured, binary data and documents

## Prompt engineering

Crafting and structuring queries to optimize LLM responses

## Context Retrieval

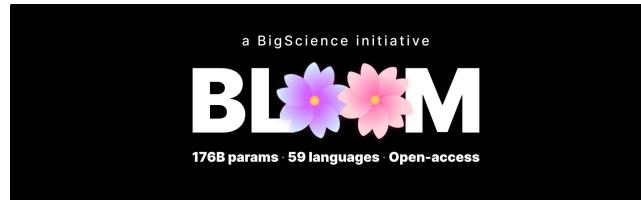
Enhancing LLM with external context such as Retrieval Augmented Generation (RAG)

## Roundtrip Interface

Act as a Discord, REST, Kafka, SQL, Slack bot to roundtrip discussions

# Think SMALLER

<https://github.com/tspannhw/FLaNK-HuggingFace-DistilBert-SentimentAnalysis>



watsonx.ai

**CLOUDERA**  
Machine Learning

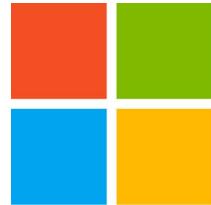
<https://github.com/tspannhw/FLaNK-LLM>



# Small Language Models

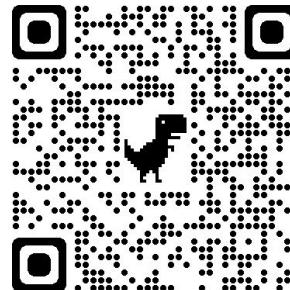
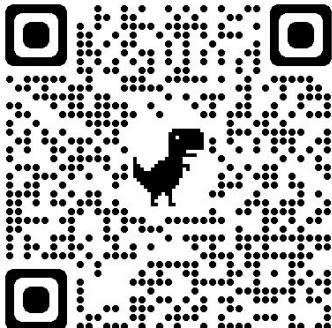


TinyLlama 1.1B Llama model



Microsoft

Microsoft PHI-2 2.7B model



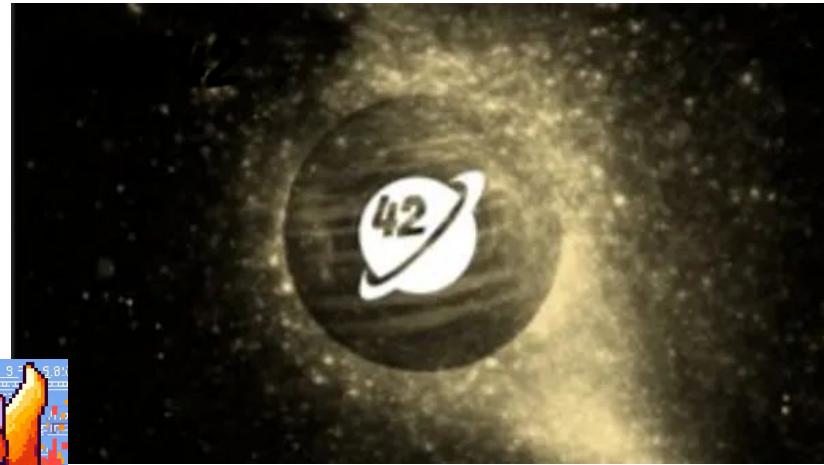
# UNSTRUCTURED DATA WITH NIFI

- **Archives** - tar, gzipped, zipped, ...
- **Images** - PNG, JPG, GIF, BMP, ...
- **Documents** - HTML, Markdown, RSS, PDF, Doc, RTF, Plain Text, ...
- **Videos** - MP4, Clips, Mov, Youtube URL...
- **Sound** - MP3, ...
- **Social / Chat** - Slack, Discord, Twitter, REST, Email, ...
- **Identify Mime Types, Chunk Documents, Store to Vector Database**
- **Parse Documents** - HTML, Markdown, PDF, Word, Excel, Powerpoint





**Startup Grind Princeton & NJ Big  
Data Alliance Generative AI  
Reverse Pitch - March 27**



star\*tree | CLOUDERA

IN-PERSON MEETUP

**Discover Data Delights: A Slice of  
Real-Time Analytics and GenAI!**

March 28 | 05:30 PM EST | NYC

**Startup Grind hosts AI Max  
Summit - April 12**



# Future of Data - NYC + NJ + Philly + Virtual



<https://www.meetup.com/futureofdata-princeton/>

<https://www.meetup.com/futureofdata-newyork/>

From GenAI to Big Data to Edge to Streaming to Containers to Cloud to Analytics to Cloud Storage to Fast Data to Machine Learning to Microservices to ...



@PaasDev

# FLaNK Stack Weekly by Tim Spann



<https://bit.ly/32dAJft>

<https://www.meetup.com/futureofdata-princeton/>

This week in Apache NiFi, Apache Flink, Apache Kafka, ML, AI, Apache Spark, Apache Iceberg, Python, Java, LLM, GenAI, Vector DB and Open Source friends.

