Secure and Private Alternatives to ChatGPT

Stephen Drew, PhD – Chief AI Officer
SoHye Park – Applied AI Scientist
July 23, 2024

Introduction

Stephen Drew — Chief AI Officer and Head of Product Management, RNL
20 years of experience leading the development and support of machine learning and conversational AI systems. Former VP of AI @ Five9, VP of AI @ Uniphore, Global Head of Contact Center Technology @ Cigna. Doctorate, Applied AI, DePaul.

SoHye Park — Applied AI Scientist, RNL
5 years of experience in data science, data analytics, and linguistics research. MA in Applied Linguistics from Ewha University, MS in Quantitative Methodology from University of Maryland, Ph.D. candidate in Applied Linguistics at University of Maryland.

“All models are wrong. Some are useful.”

- George Box

“All models are beautiful. Some are smart.”

- Derek Zoolander

Balancing innovation and responsibility

Responsible AI

Formalize the AI use case intake process with defined requirements, including managing compliance and risks.

Recognition of the unpredictability and difficulty in explaining outcomes from neural network-based models such as large language models and computer vision systems.

AI governance and risk management framework, tooling, and transparency reporting for RNL and our clients.

Innovation

Exploration of generative AI models' integration with existing structured and unstructured data sources.

Leveraging modern AI models and knowledge systems to drive intelligent conversations, distill information quickly, and uncover valuable insights from data.

Controlled releases of innovative new products and services that build on the consulting expertise and predictive analytics capabilities of RNL.

RNL advocates for responsible AI use, focusing on fairness, empathy, accountability, and transparency. We guide institutions in strategically implementing AI while prioritizing ethical considerations.

Avoid cracking walnuts with sledgehammers

The Problem with Large Language Models

“ On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜***”***

❗ They are trained on what is primarily white/euro-centric text.

❗ They are ”stochastic parrots” (i.e., known for haphazardly stitching together sequences of linguistic forms … according to probabilistic information about how they combine, but without any reference to meaning.”)

❗ Therefore, in many cases they can be biased (based on their underlying training data) and incorrect (based on their actual abilities).

ChatGPT 4o

Can you create a picture of a normal American in his 30s?

Here are the images of a normal American man in his 30s as you requested. If you need any adjustments or additional images, feel free to let me know!

ChatGPT 4o

I want you to generate an image of a Korean woman who loves snacking--especially nuts and cereal. She also loves crunching her snacks.

Here is the image of a Korean woman who loves snacking on nuts and cereal. If there's anything else you'd like or any adjustments needed, just let me know!

AI Hallucinations are not surprising

Business Insider
https://www.businessinsider.com › Tech › News

Google's Ad for ChatGPT Rival Bard Shows Inaccurate ...

Feb 8, 2023 — **Google's** new AI chatbot **Bard** showed an inaccurate answer to a question about the discoveries of the James Webb Space Telescope in an online ...

Missing: ~~river~~ | Show results with: river

BY JAKE OFFENHARTZ

Updated 7:11 PM EDT, April 3, 2024

Share

NEW YORK (AP) — An artificial intelligence-powered chatbot created by New York City to help small business owners is under criticism for dispensing bizarre advice that misstates local policies and advises companies to violate the law.

But days after the issues were first reported last week by tech news outlet The Markup, the city has opted to leave the tool on its official government website. Mayor Eric Adams defended the decision this week even as he acknowledged the chatbot's answers were "wrong in some areas."

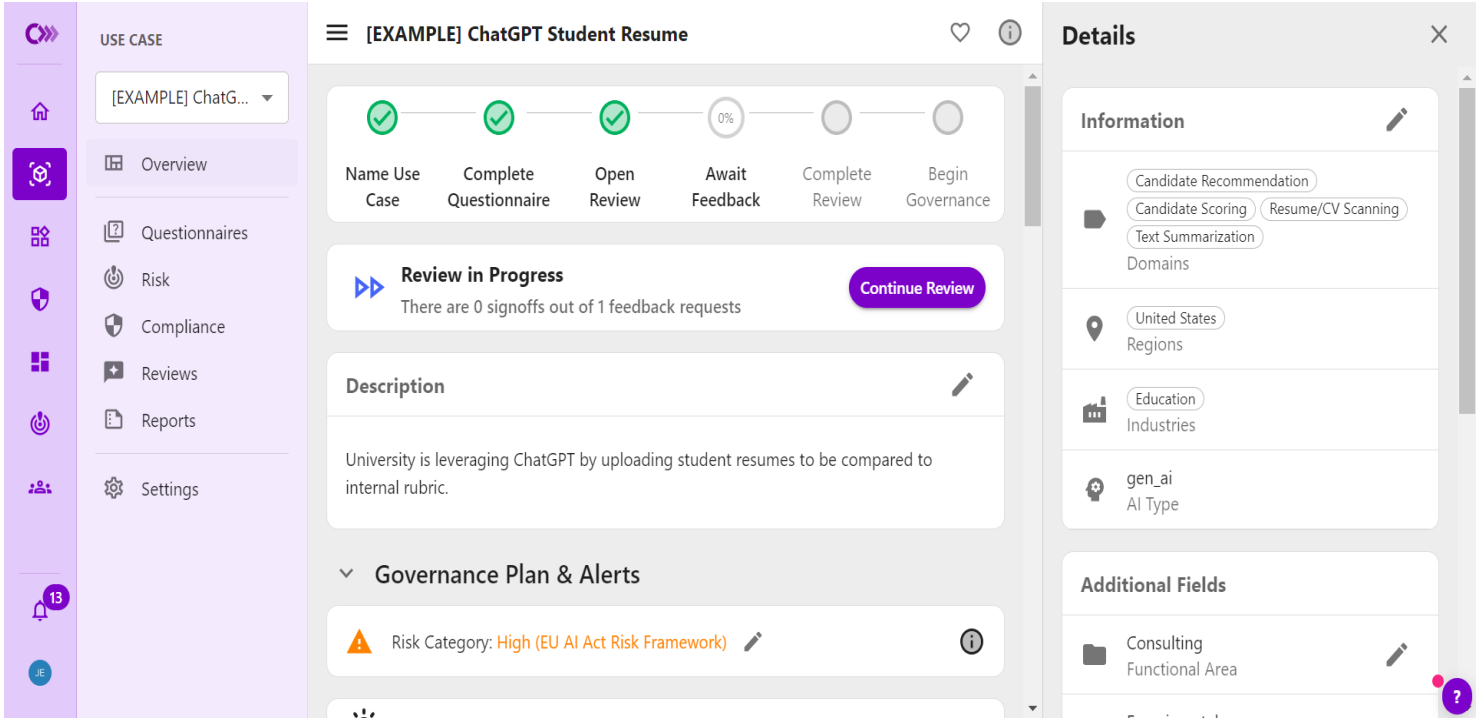
Mashable
https://mashable.com › Tech

Air Canada loses court case after its chatbot hallucinated ...

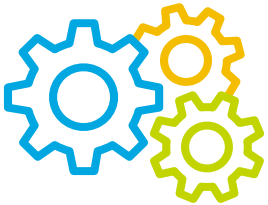
Feb 17, 2024 — **Air Canada was ordered to reimburse a customer in lawsuit over** the airline's chatbot giving out inaccurate policy information.

Remedies for unintentional bias/harm in AI applications

- Establish a process & implement tooling for managing AI use case evaluation, risk management, and governance and adhere to standards such as NIST (https://www.nist.gov/itl/ai-risk-management-framework)
- Establish a cross-functional AI council, train all employees on responsible AI
- Separate initial experimentation/evaluation of use case from engineering function



Initial Remedies for Hallucinations



Fine Tuning

Adjust the system

Augment Relevant data

Feed extra data source

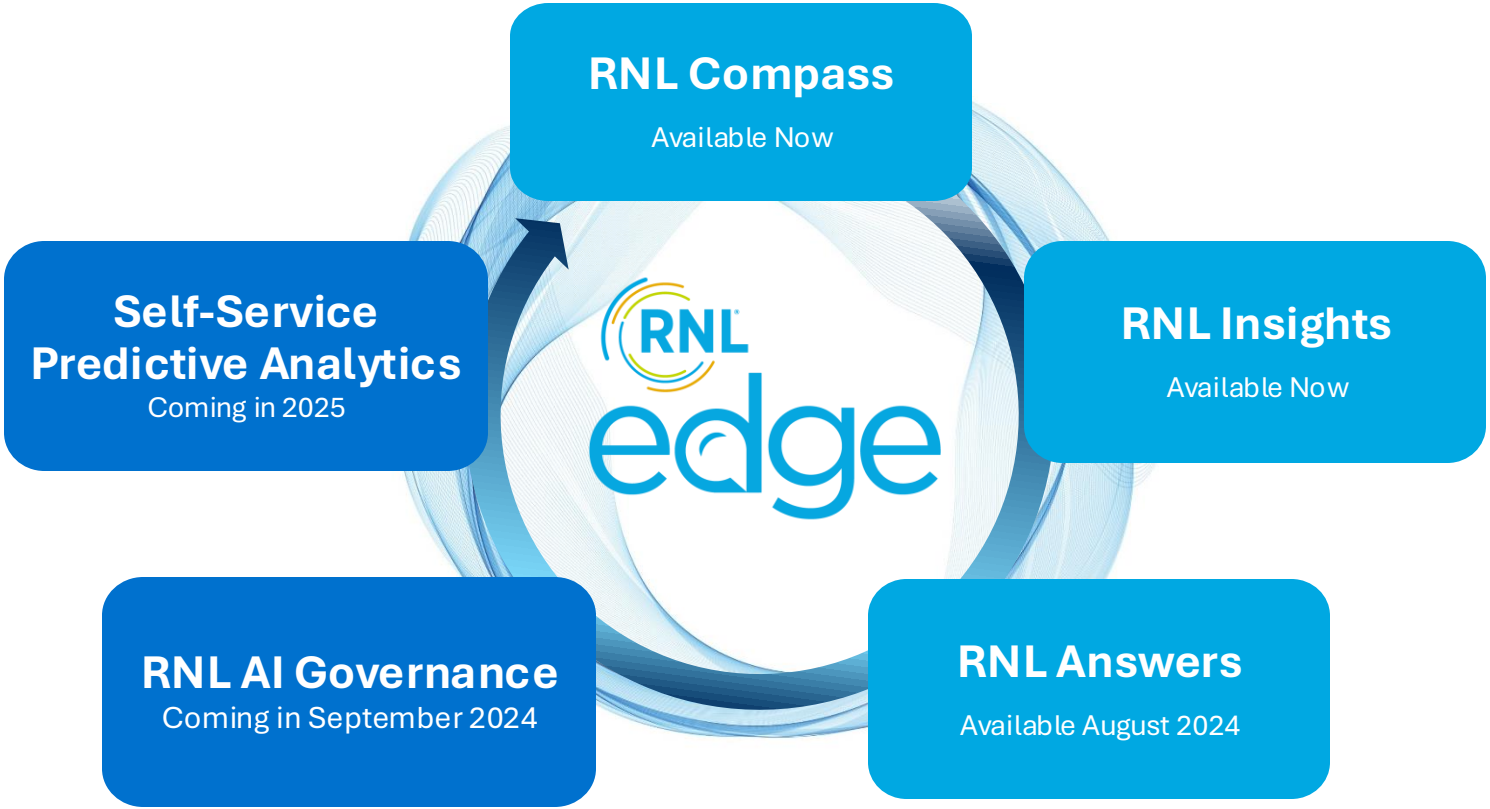
Prompt Engineering

Write a better query

Not perfect but a good start

An Alternative: RNL Answers

Introducing RNL Edge Portfolio



AI @ RNL for 2024

RNL AI solutions leverage both private local language models augmented with contextually relevant data along with selective use of commercial language models. We guarantee complete privacy and security of client-provided data.

RNL Compass

Multi-modal Conversational Assistants for enrollment, student success, and fundraising.

RNL Insights

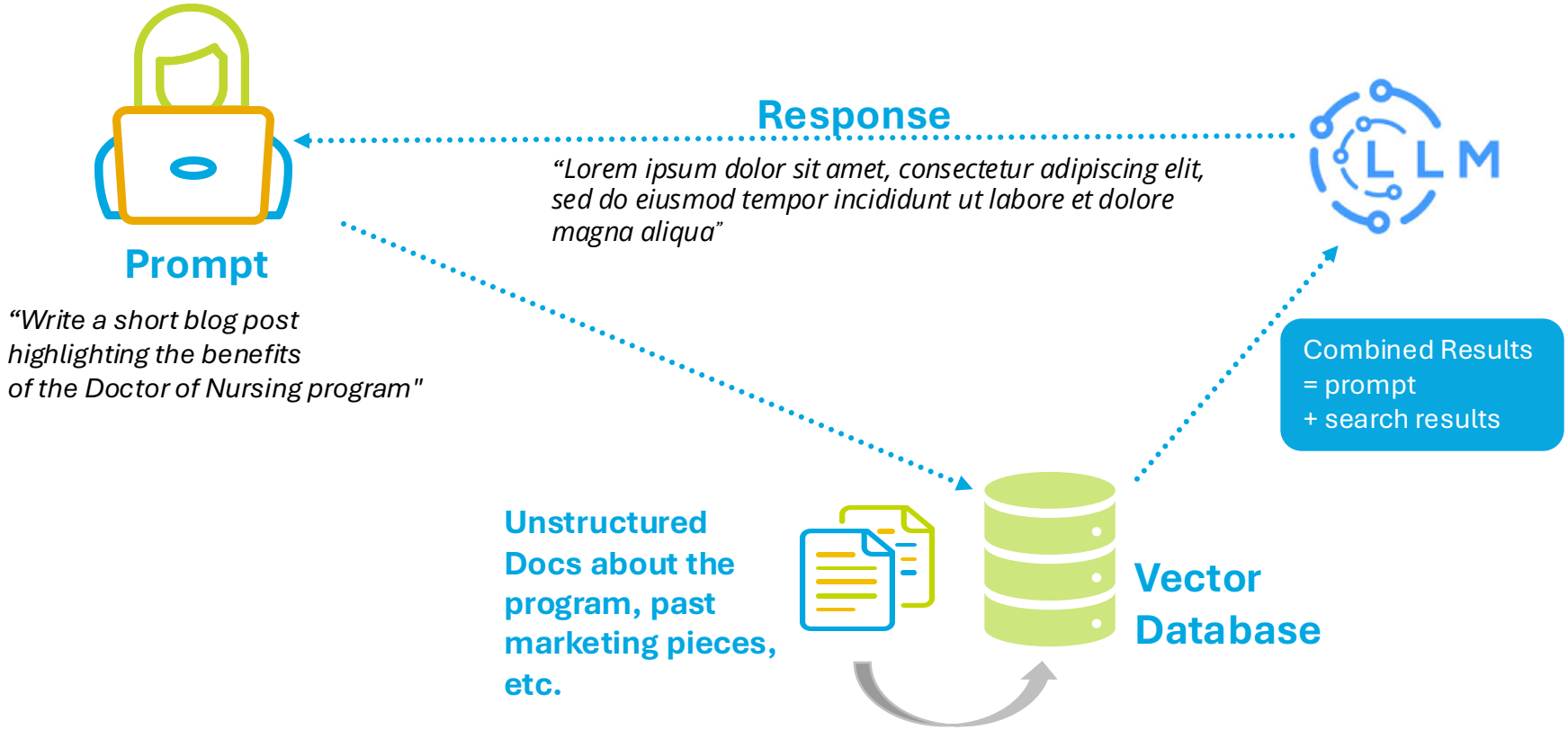
A conversational interface (digital analyst) using client data (CRM, SIS, etc.) and RNL-generated data to produce strategic insights.

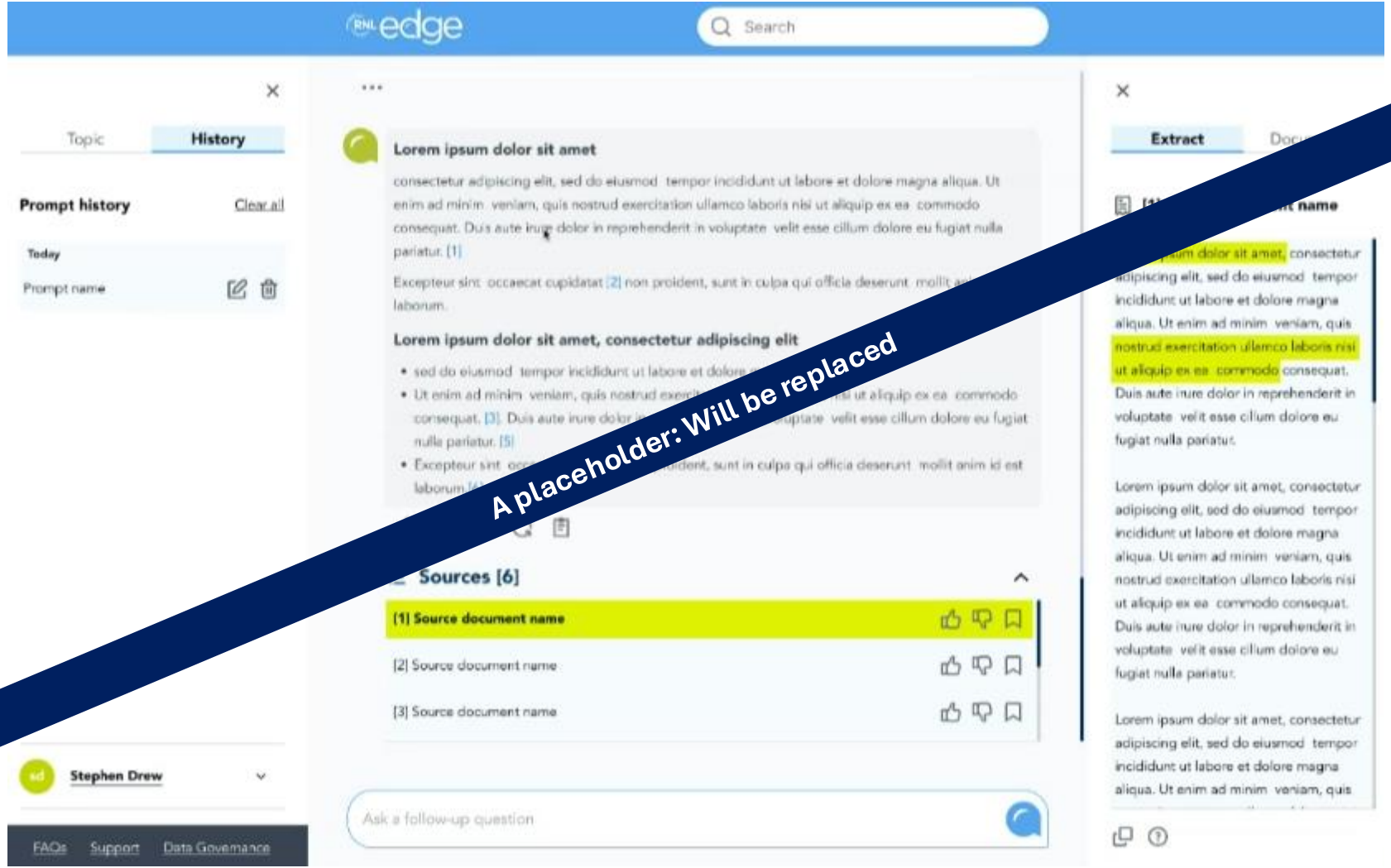
RNL Answers (August 2024)

A private and secure ChatGPT-like experience for RNL clients and generative AI features for RNL SaaS platforms.

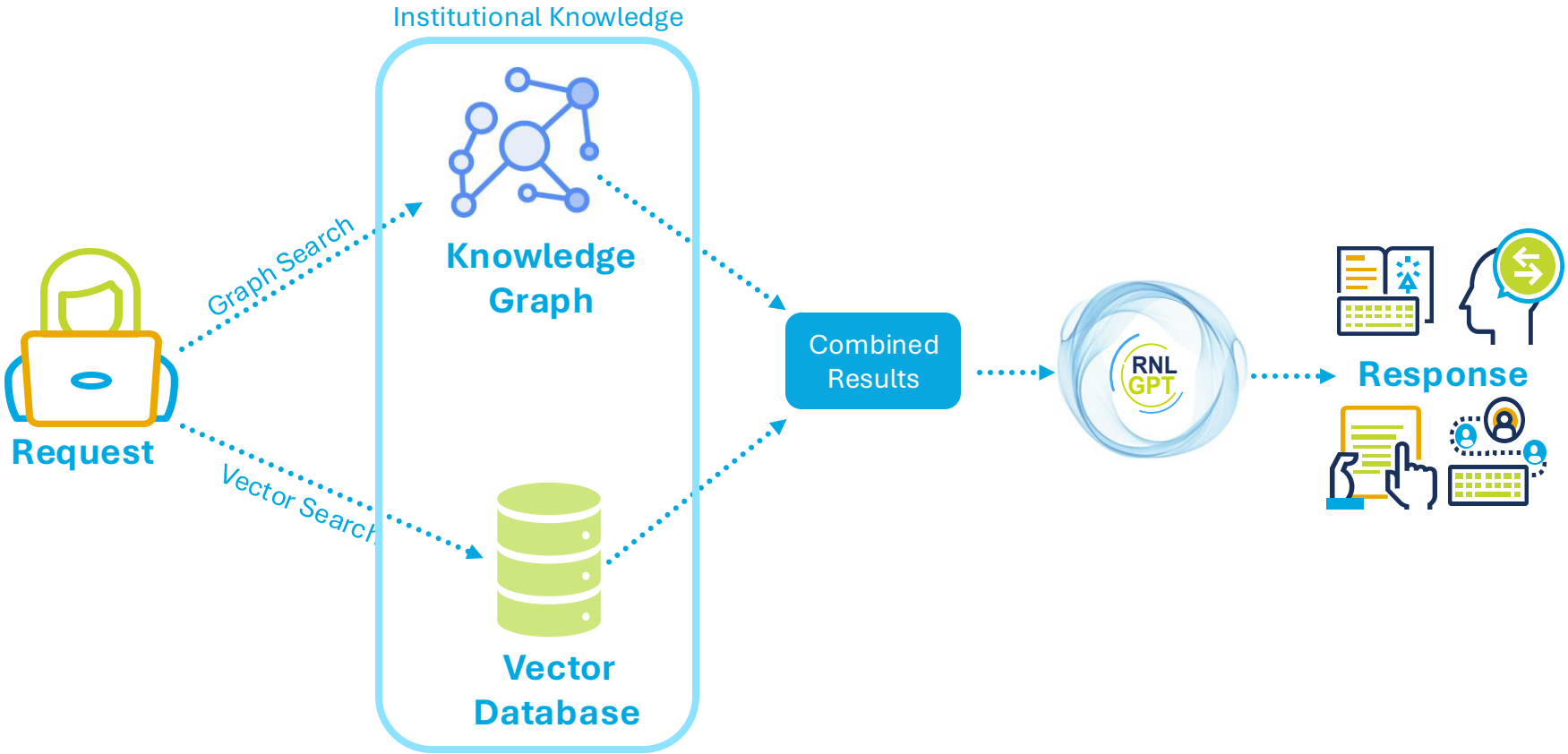
RNL Answers

Retrieval Augmented Generation (RAG)




edge
Search
Topic
History
Prompt history
Clear all
Today
Prompt name
Lorem ipsum dolor sit amet
consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. [1]
Excepteur sint occaecat cupidatat [2] non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.
Lorem ipsum dolor sit amet, consectetur adipiscing elit
sed do eiusmod tempor incididunt ut labore et dolore
Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. [3]. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. [5]
Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.
Sources [6]
[1] Source document name
[2] Source document name
[3] Source document name
Ask a follow-up question
Extract
name
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis
Stephen Drew
FAQs
Support
Data Governance
A placeholder: Will be replaced

Retrieval Augmented Generation (RAG)



Use Cases & Demo

Generated Models

Augmented with Institutional Knowledge

Academic Support/Advising

Content generation, evaluation and brainstorming based on

- Catalog search
- Policy search
- HR document search

Brainstorming

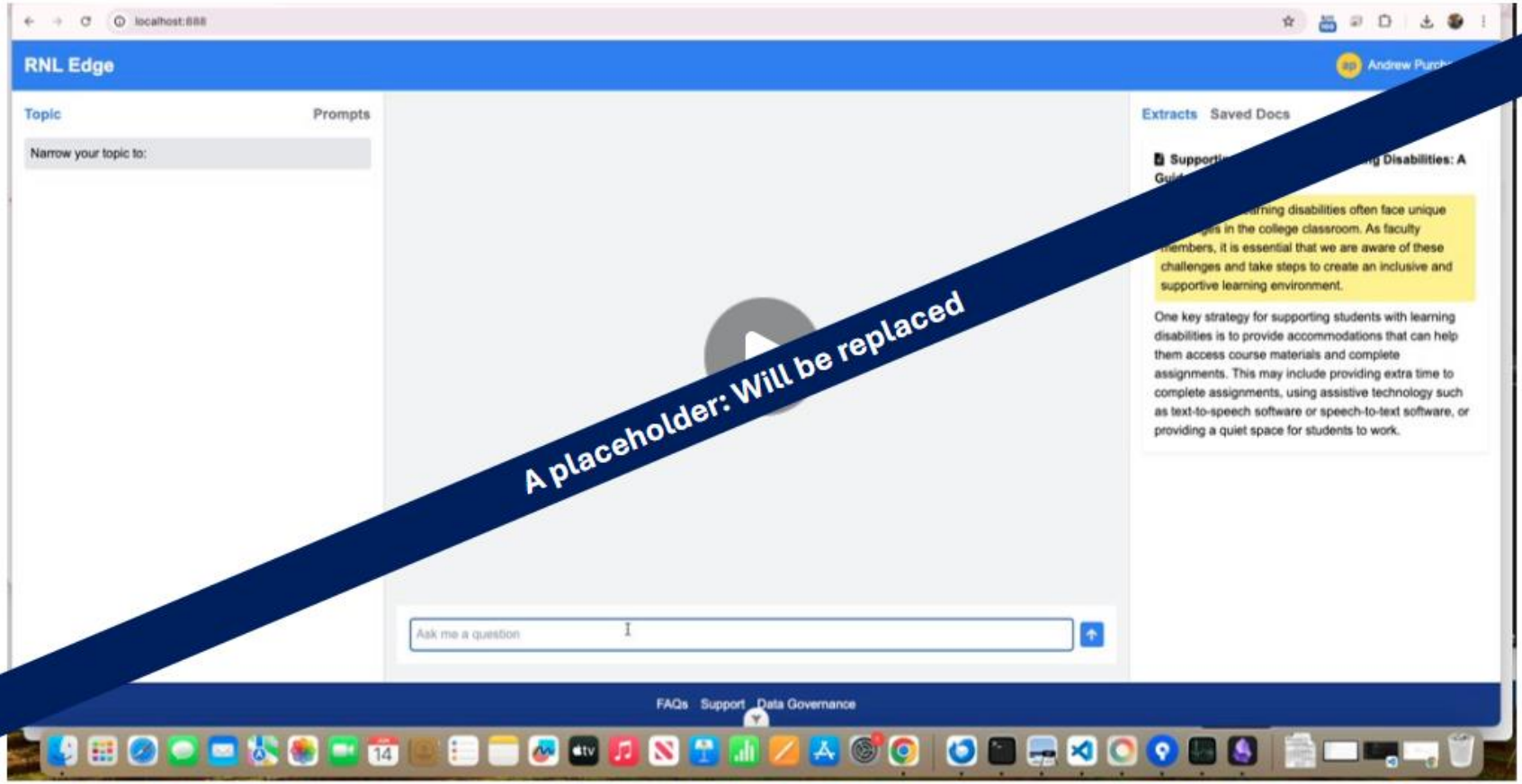
- Marketing ideas
- Outreach strategies
- Brand positioning
- Key benefits
- Best practices

Content Creation

- Emails
- Reports
- Landing pages
- Blog posts
- Social media content
- SEO keyword blueprints
- Re-purposing existing content
- Headlines & captions

Content Evaluation

- Summary
- Writing assistant
- Essay evaluation
- Resume evaluation
- Transcript evaluation


localhost:888
RNL Edge
Andrew Purch
Topic
Prompts
Narrow your topic to:
Extracts
Saved Docs
Disabilities: A
learning disabilities often face unique
in the college classroom. As faculty
members, it is essential that we are aware of these
challenges and take steps to create an inclusive and
supportive learning environment.
One key strategy for supporting students with learning
disabilities is to provide accommodations that can help
them access course materials and complete
assignments. This may include providing extra time to
complete assignments, using assistive technology such
as text-to-speech software or speech-to-text software, or
providing a quiet space for students to work.
Ask me a question
FAQs
Support
Data Governance
A placeholder: Will be replaced

Q&A

Thank you

Stephen Drew – Chief AI Officer
SoHye Park – Applied AI Scientist