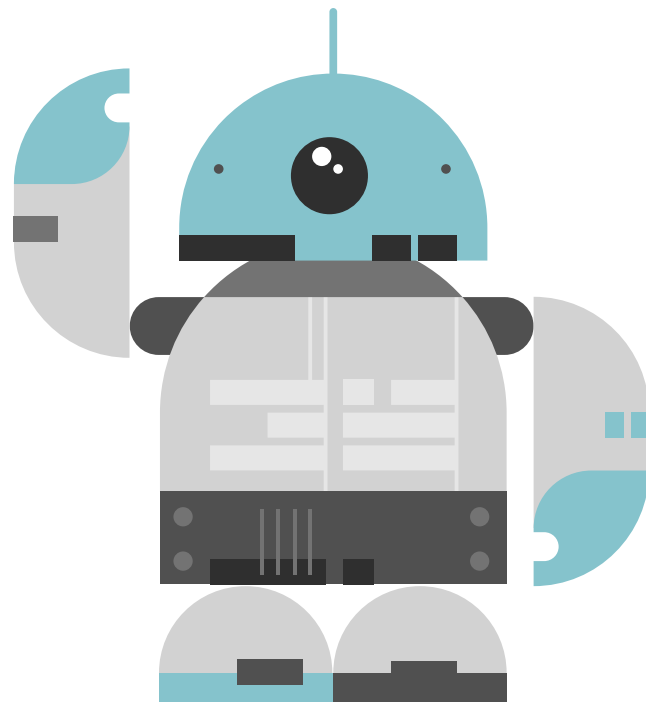
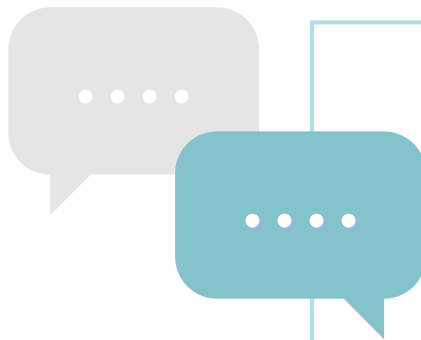
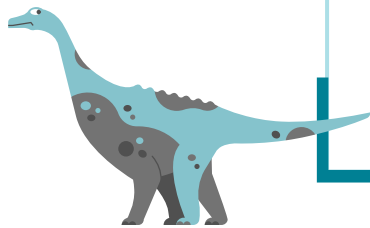


How to build ...



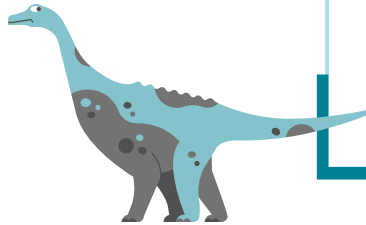
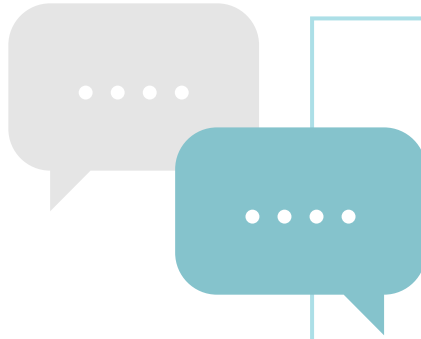
Wim Horinga

ExLibris  innovative
Part of Clarivate



... the next discovery experience

Dad? How heavy is a stegosaurus?



How heavy is a stegosaurus?



Wuerhosaurus homheni, a broad-bodied stegosaur, is estimated to have reached 7 meters (23 ft) in length and 4 metric tons (4.4 short tons) in body mass. Its dorsal plates were at first thought to have been much rounder or flatter than other stegosaurids, but Maidment established this was an illusion caused by breakage: their actual form is unknown.

Sources:

- wikipedia.org/Paranthodon
- wikipedia.org/Dacentrurus

Write your question!

ask >

What can we do in 30 minutes?...

- About the Conversational Discovery Prototype
- Application logic
- Creating the jurassic.ai prototype
- Inspiration to continue building

... with some preparation quite a lot!

About the Conversational Discovery Prototype

What?

- Question & answer, sourced with academic data
- 500k full text documents from PQ One Literature
- In cooperation with AI21 Labs

Why?

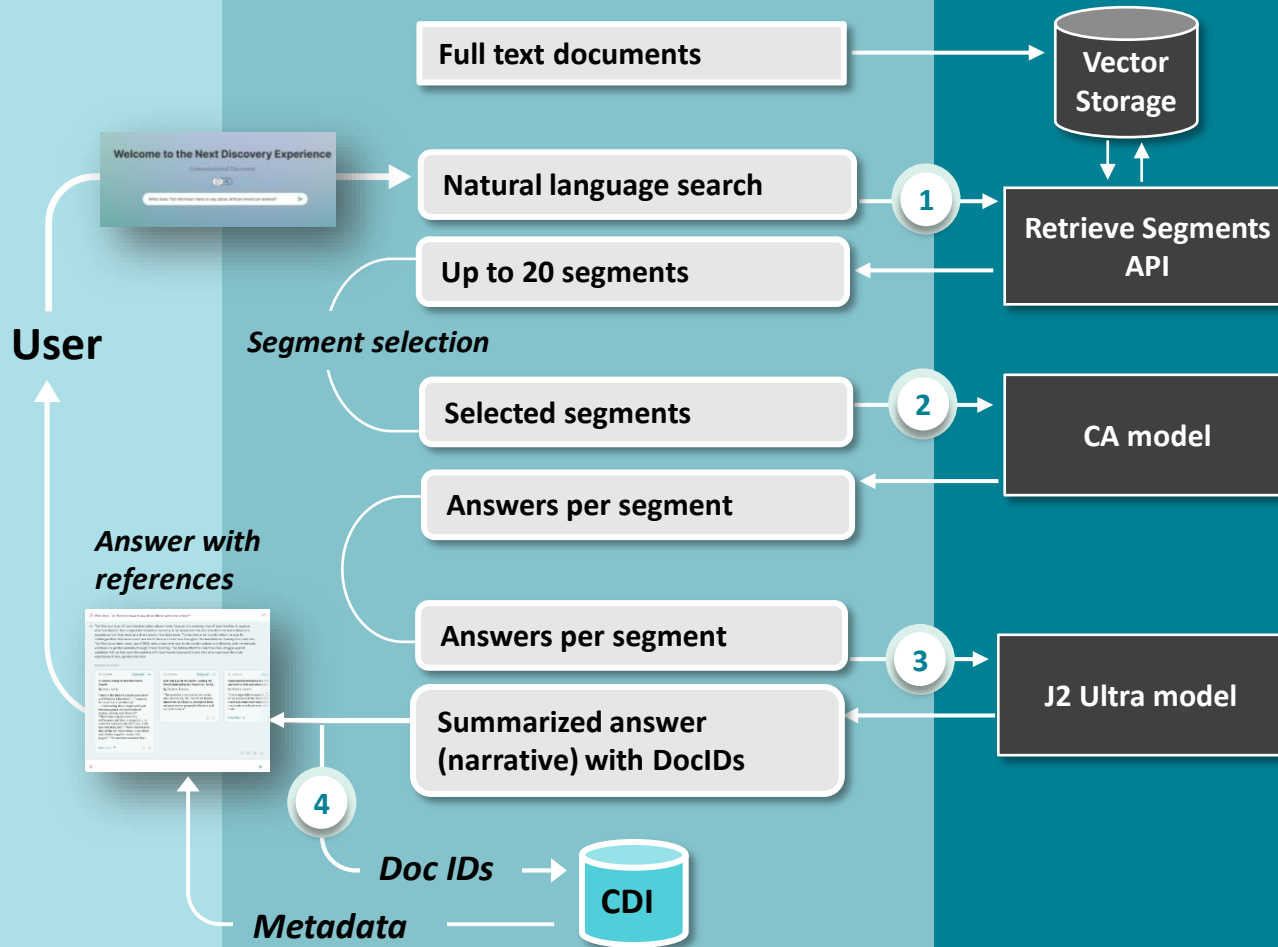
- Gain understanding about using Large Language Models (LLM)
- Experiment with 'grounded' (factual) chat-based discovery of academic data

How?

- I'll explain!

Clarivate

AI21



Task: Store the data (embeddings)
API: Multi-Docs API to remove/upload a TXT file

Task: Retrieve segments relevant to user input
API: CA API /designated API

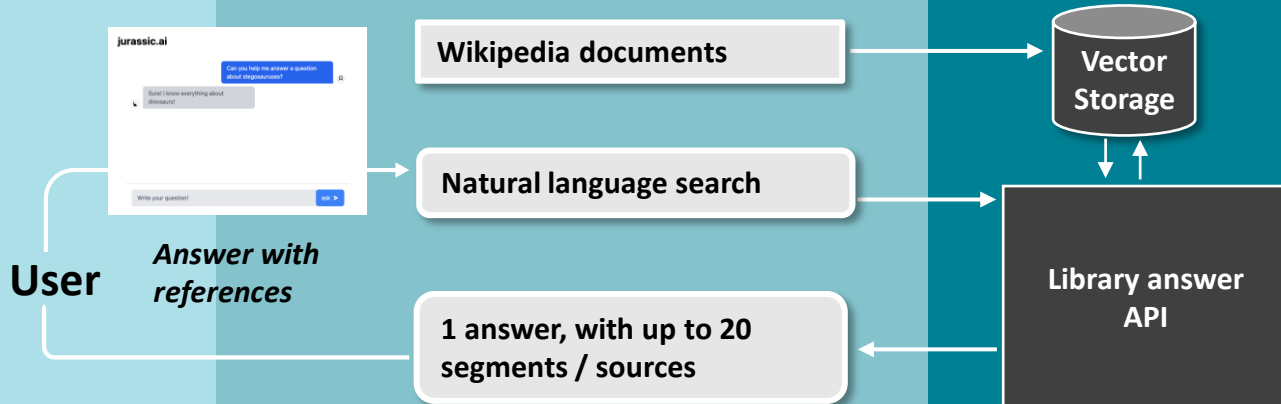
Task: Answer user question from segment
API: CA API call for every segment with user input
Benefits: Grounding/truthfulness

Task: Rephrase & summarize answers
API: Foundation model API
Input: Generated answers
Output: Final answer for the user
Benefits: Detailed answers, number of aspects, summary structure

Conversational Discovery UI

demo

Our demo



AI21

Task: Store the data (embeddings)
API: Multi-Docs API to remove/upload a TXT file

Task: Retrieve segments relevant to user input, and use it to answer the question
API: library/answer API

building jurassic.ai



Follow along: github.com/Wimpje/next.discovery

- ExpressJS backend
- Angular / TailwindCSS frontend
- Using AI21 API

Steps

1. Get a dataset
2. Upload data to AI vector search endpoint
3. Create the UI for asking questions
4. Create a server that connects to the (AI21) API
5. Integrate UI and server code
6. Deploy the application
7. Get all the dinosaur-knowledge we need!

Step 1. Getting a data set

- Website content
- Organizational wiki
- Knowledge base
- Blogposts
- Create your own about your specific topic for Q&A



Step 1. Getting a data set

ChatGPT... can you create a script to get dinosaur info from Wikipedia?



Image by brgfx on Freepik

Default (GPT-3.5)



Certainly! Wikipedia provides an export functionality that allows you to retrieve articles in XML format. Here's an example of how you can use JavaScript to fetch dinosaur-related articles using the Wikipedia API export feature:



1. Create Your Script:

Create a file called `export.js` in your project directory and add the following code:

javascript

Copy code

```
const axios = require('axios');
const fs = require('fs');

const apiUrl = 'https://en.wikipedia.org/w/api.php';

async function fetchDinosaurArticles() {
  try {
    const response = await axios.get(apiUrl, {
      params: {
        action: 'query',
        list: 'categorymembers',
        cmtitle: 'Category:Dinosaurs',
        format: 'json',
        cmlimit: 'max',
      },
    });
  }

  const articles = response.data.query.categorymembers;

  for (let i = 0; i < articles.length; i++) {
    const article = articles[i];
```

Step 2. Make data searchable

1. Use existing (keyword) search engine
 - + don't need to install something new
 - not great at returning results for conversational queries
2. Use semantic search (vector search)
 - + great at 'understanding' conversational questions
 - + it can even handle different languages
 - most organizations need to set it up, or use a cloud service

... a quick detour

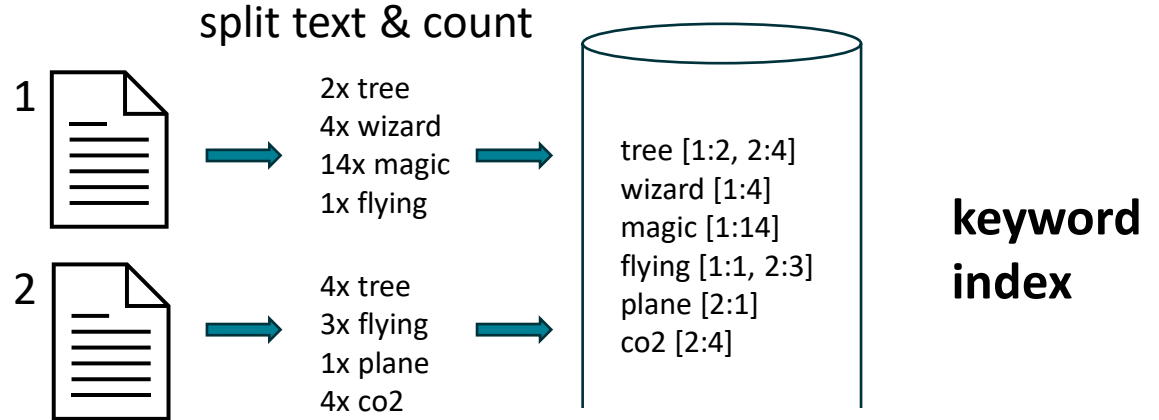
differences between vector and keyword search



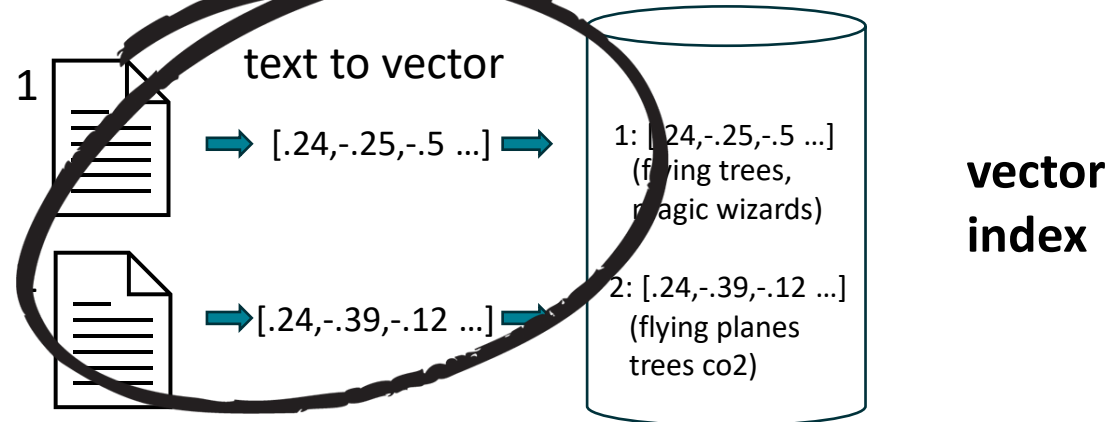
1. fantasy story



2. environmental article



AI provides context



... a quick detour

differences between vector and keyword search



1. fantasy story



2. environmental article

Now search for:

Can you find a **book** about **flying trees**??



split text & count

book AND flying AND trees

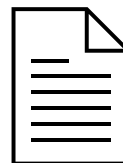
2x tree
4x wizard
14x magic
1x flying

4x tree
3x flying
1x plane
4x co2

keyword search



1. fantasy story



2. environmental article



... a quick detour

differences between vector and keyword search

Now search for:
Can you find a book about flying trees?



text to vector

[.24,-.25,-.5, .12, .47, .8 ...]

1: [.24,-.25,-.5 ...]

2: [.24,-.39,-.12 ...]

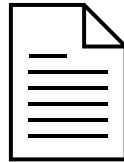
semantic search



1. fantasy story



2. environmental article



1. fantasy story



2. environmental article

AI provides context

Step 2. Make data searchable

Upload online...

studio.ai21.com/documents

Document Library











Documents here provide the basis of context for Contextual Answers API.

Documentation ↗

+ Upload documents

Q Search...

</> Ask all documents

Document Name ↑	File Size	ID Number	Date Created	Uploader	Status	
 Aardonyx.txt	7.89KB	f98aa5c...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮
 Abdarainurus.txt	1.87KB	9ccb29...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮
 Abrictosaurus.txt	5.51KB	6c0f68a...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮
 Abydosaurus.txt	2.32KB	764aeaf...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮
 Achillobator.txt	22.29KB	e455c9...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮
 Acrocanthosaurus.txt	21.39KB	f1dd156f...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮
 Adasaurus.txt	13.5KB	28f70f3...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮
 Adeopapposaurus.txt	1005B	9d8a5f9...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	<input type="text" value="Ask a question"/> ⋮
 Aegyptosaurus.txt	1.67KB	4d1725c...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮
 Aegyptosaurus.txt	1.67KB	1b55b7b...	Aug 10th, 2023	wim.horinga@exlibrisgroup.com	✓ Ready	⋮

1-10 of 503 < >

... or with the API

docs.ai21.com/reference/manage-library-ref

Step 3. Create the UI

Angular (angular.io/guide/setup-local)

```
npm install -g @angular/cli  
ng new jurassic.ai
```



Tailwind CSS (tailwindcss.com/docs/installation)

```
npm install -D tailwindcss postcss autoprefixer  
npx tailwindcss init
```



Now get creative!

...or not: tailwindcomponents.com/component/chat

Step 4. Create the server

ExpressJS (expressjs.com)

Express

```
npm install express body-parser node-fetch connect-history-api-fallback
```

Now get coding!

- Connect to API21
- Handle client requests

Later:

- Add authentication ?
- Add persistent storage (mongo / sql) ?



Source: deviantart- bloodysamoan

Step 5. Integrate UI and server

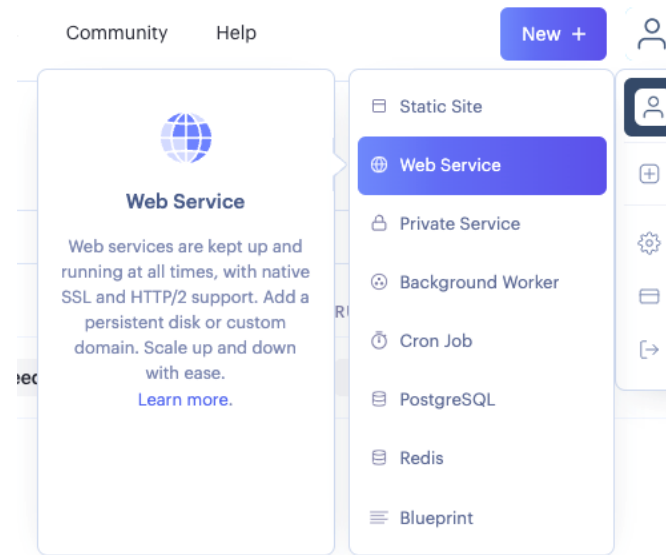


AI21 labs

Step 6. Deploy

render.com (render.com/docs)

- Create web service
- Connect to your Github repo
- Set up environment variables
- Settings
 - Build Command: `npm install; npm run build`
 - Start command: `npm run start`
- Wait for the build
- It's up! jurassic-ai.onrender.com



... how to accelerate development

Use AI to code (!)

- Copilot
- ChatGPT
- phind.com

Use frameworks

- Frontend
 - Angular, Vue, Svelte, React
- Backend
 - .NET core, Spring, ExpressJS, Django, Laravel
- CSS
 - Bootstrap, Tailwind, Bulma, Pure

Use component libraries

- Tailwind-elements, Bootstrap, Flowbite, Material Design

... how to
accelerate development

Use hosted AI solutions

- AI21
- OpenAI
- Azure
- Hugging Face

Deploy to cloud hosting

- Render.com
- Google cloud
- Azure
- AWS

inspiration

Tools that simplify working with LLMs


- langchain.com
- llamaindex.ai
- streamlit.io

Vector databases / semantic search

- Weaviate (cloud / self)
- Chroma (self)
- Milvus (self)

Using AI Models / services

- Hugging face (train / use your own)
- OpenAI / AI21 / Claude (APIs)



I hope you're inspired!
Enjoy coding something cool!