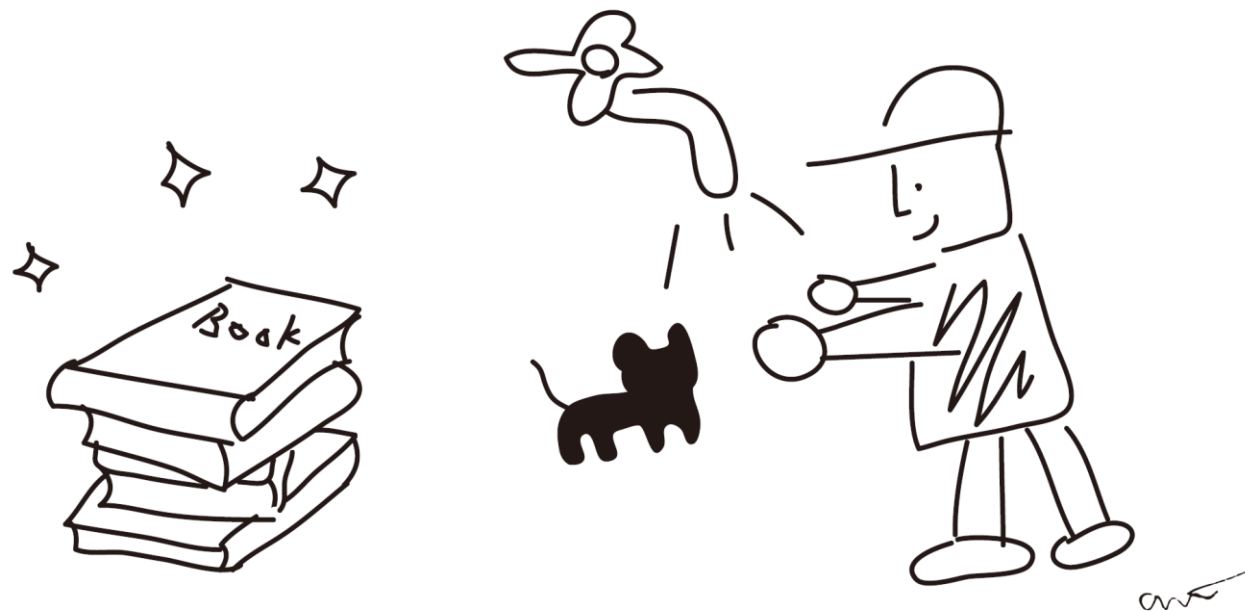


大学図書館と生成AIについて考える

大図研京都ワンディセミナー

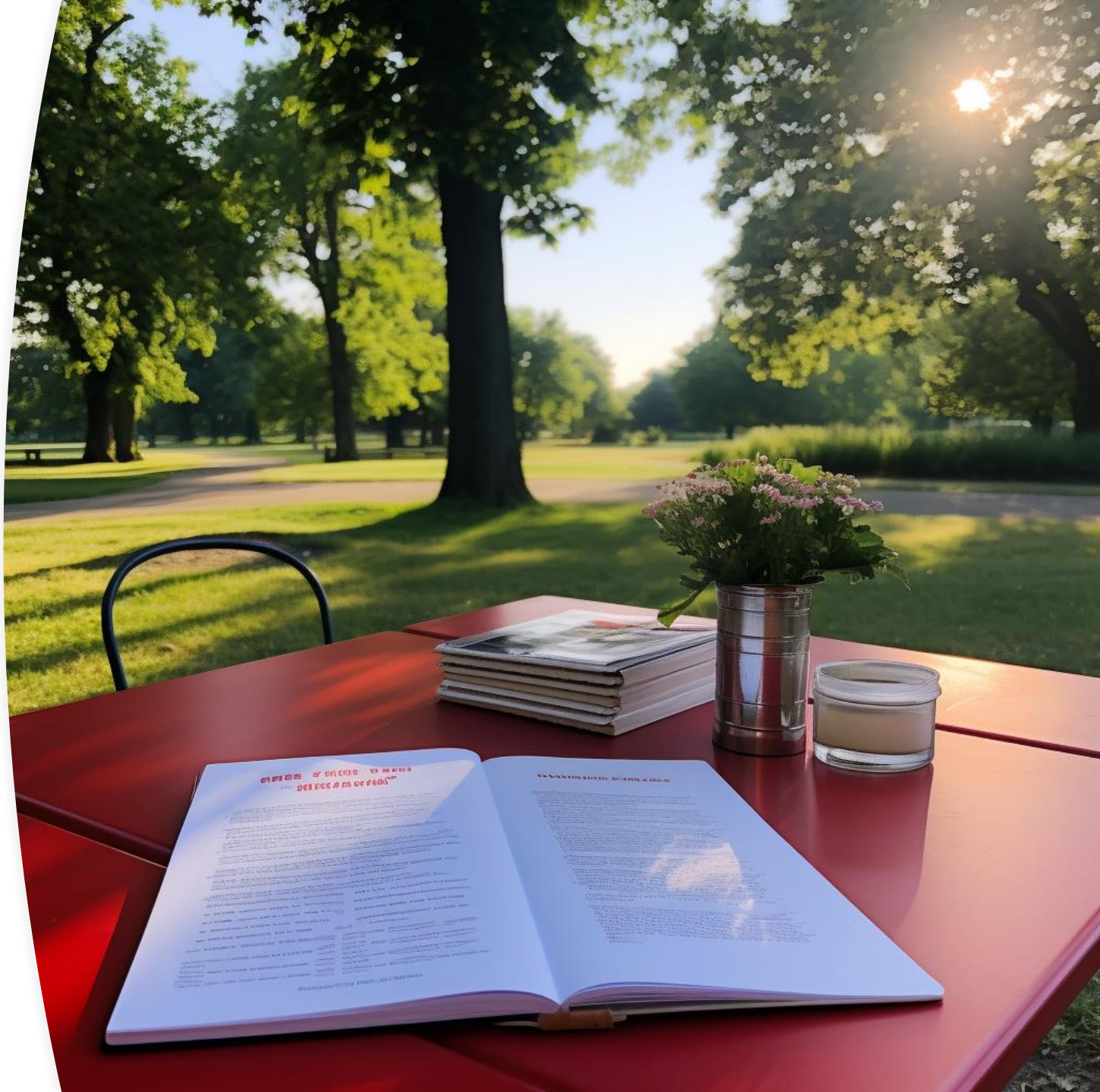
吉本龍司（カーリル）

2023/7/22

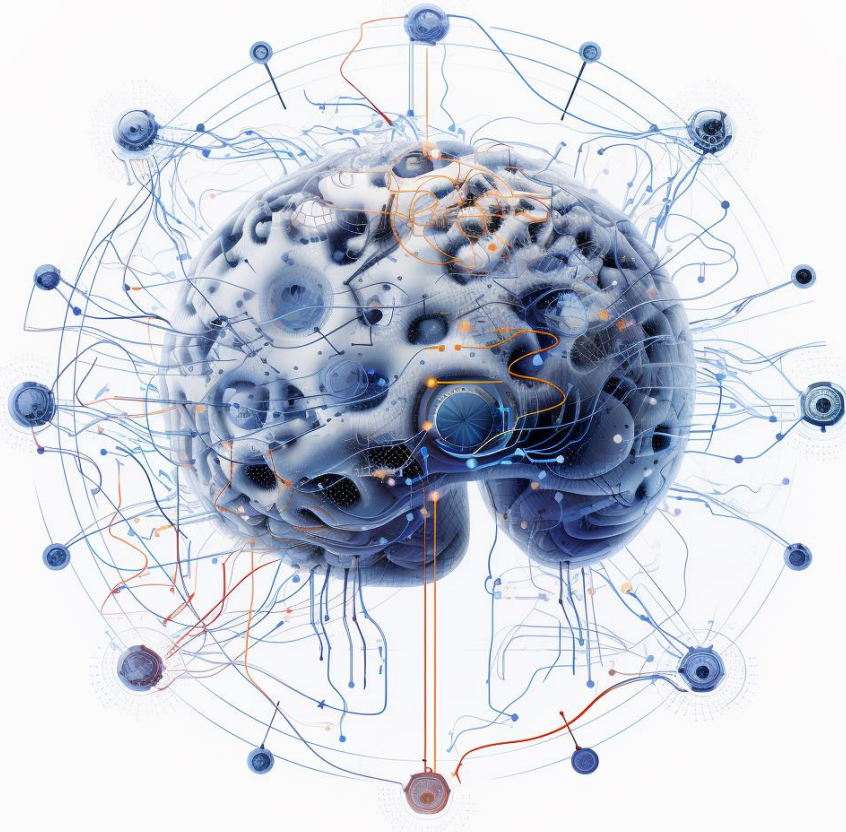


今日の構成

- 自己紹介
- 日常の変化
- APIによる拡張
 - 蔵書検索サポーターの実装
 - できそうなこと
- Chat-GPTと図書館



Chat-GPTの仕組みについては解説しません



- ブラックボックスです
- 大規模言語モデルや生成AIは、ChatGPT以外にも開発が進んでいます
exp. Llama 2 by Meta
- 大量のメモリを積んだ計算機をタイムシェアしています
- 図書館での活用イメージと、新しい図書館の役割を一緒に考えてみる

Everyone is connected via the internet to the AI, which is an image of a line-drawing brain.

約 3,970 件 (0.31 秒)

www.odyssey-com.co.jp > essay > essay91 ▾

リレーエッセイ Vol.91 吉本 龍司さん | 株式会社 オデッセイ ...

リレーエッセイ Vol.91 吉本 龍司さん (株式会社 カーリル 代表取締役) : 「暮らしと仕事のあいだ。」

吉本龍司 の画像検索結果

府立図書館

internet watch

柳津弘

ソフト

京都府立

カーリル

龍司さん

蔵書

代 >

[→ 吉本龍司で見つかった他の画像](#)[画像を報告](#)

researchmap.jp > ryuuji ▾

吉本 龍司 (Ryuuji Yoshimoto) - マイポータル - researchmap

ヒュー・マクガイア, ブライアン・オレアリ, 浅野紀予, 石松久幸, 堺屋七左衛門, 柴野京子, 高橋征義, 秦 隆司, 宮家あゆみ, 室 大輔, 吉本龍司, yomoyomo(担当:共訳)(原著:ヒュー・マクガイア, ブライアン・オレアリ, アンドリュー・サヴィカス, ...

ja-jp.facebook.com > ryuuji ▾

吉本 龍司 | Facebook

吉本 龍司さんはFacebookを利用しています。Facebookに登録して、吉本 龍司さんや他の知り合いと交流しましょう。Facebookは、人々が簡単に情報をシェアできる、オープンでつながりのある世界の構築をお手伝いします。

www.projectdesign.jp > bigdata ▾

ITで図書館の未来をつくる「蔵書データ」が生んだ新ビジネス ...

プロフィール

吉本龍司 (よしもとりゅうじ)

株式会社カーリル 代表取締役

社長、あるいは、フリーのエンジニア

技術の社会実装

図書館をもっと楽しく。



カーリルの新しい収益源を教えてください



カーリル (Calil) は、日本の図書館検索サービスです。利用者はこのサイトを通じて、地元の図書館で利用できる資料を調べることができます。カーリルが新たな収益源を持つためのいくつかのアイデアは以下の通りです：



- 広告収入:** カーリルのウェブサイトやアプリに広告を掲載することで収益を得る。広告は関連性の高いもの、例えば出版社や書店、オンライン学習プラットフォームなどの広告を掲載することができます。
- 有料プレミアムサービス:** より詳細な検索オプション、待機リストの優先アクセス、新刊通知などの特典を提供する有料メンバーシップを導入します。
- パートナーシップ:** 出版社や書店とのパートナーシップを結び、リンクを通じて本の販売による手数料を得る。
- データ分析サービス:** カーリルが収集する大量のデータを分析し、出版業界や教育機関などのクライアントに販売する。このデータは、人々の読書傾向や需要、地域別の人気作品など、価値ある洞察を提供できます。
- 寄付:** ユーザーからの寄付を募ることも一つの方法です。特に公共の利益を追求するサービスであるカーリルは、この方法を試みることができます。

これらは全て仮設であり、実際にはカーリルのビジネスモデルや経営戦略、法律など多くの要因を考慮する必要があります。また、利用者のプライバシーとデータ保護に関する問題も十分に考慮する必要があります。

🔄 Regenerate response

Send a message





アサヒ光學

@asa_hit

図書館員時代からの友人が、ChatGPTが吐いたウソ書籍情報を元に「この世に存在しない本に対するILL (図書館間相互貸借) 請求」が送られたケースについてFBに投稿していて、そういう「実害」もあるのか、とおもうなどする。実在しない本の非実在を証明するのは悪魔の証明みたいなものだからな.....

午前9:22 · 2023年3月8日 · 26.6万 件の表示



絢李 (じゅんり) @junri3106 · 3月29日

【RT】なんでchatGPTに聞くねん。chatGPT使えるなら、カーリルなり国立国会図書館サーチなり使えるやろ。



1



81



ぬまきち @obenkyounuma · 3月29日

こう言うのは、Chat-GPTを使った後は必ずカーリル使って下さいね、で終わるので比較的問題になりにくい例。



calil.jp

日本最大の図書館蔵書検索サイト「カーリル」

カーリルは全国7,400以上の図書館から書籍とその貸し出し状況を簡単に横断検索できるサービスです。



2



7



936



AIによる日常の変化

- ウェブ翻訳 (Deep-L)
- プログラミングの補助者 (GitHub Copilot)
- 文章作成 (Notion AI)
 - 議事録要約
 - 箇条書きからプレスリリース原稿を草案を作る
 - 文章校正
- 契約書チェック (ChatGPT)
- ちょっとしたことをやってもらう (ChatGPT)

昨日仕事でやったこと



ありそうでない、日本の市町村名を30個考えて



もちろんです。以下に日本の架空の市町村名を30個提案します。これらは架空の名前であり、実際の地名とは関係ありません。



1. 朝霧市 (あさぎりし)
2. 霞浦町 (かすみうらちょう)
3. 鏡川市 (かがみがわし)
4. 蒼海町 (あおみちょう)
5. 翠山市 (みどりやまし)
6. 桜峰町 (さくらみねちょう)
7. 玉瀬市 (たませし)
8. 深川町 (ふかがわちょう)
9. 紫陽花市 (あじさいし)
10. 音羽町 (おとわちょう)

ダミー横断検索デモサイト

[詳細検索](#)

66件見つかりました。

地域で絞り込み

[すべて](#)[公立図書館](#)[大学図書館](#)

外部サイト

[NDLサーチ](#)[CiNii Books](#)

Powered by [カーリル](#)

タイトル

1Q84 a novel BOOK2 7月-9月

23館所蔵

[雪島県立図書館](#)[雪島市](#)[霞浦町](#)[紅葉町](#)[鏡川市](#)[梅花市](#)[深川町](#)[蒼海](#)[蓬萊村](#)[霞浦町](#)[青葉町](#)[高原町](#)[高原町](#)[黄金山市](#)[紫陽花市](#)[雪島大学](#)[雪島国際大学](#)[雪島短期大学](#)[雪島福祉短期大学](#)[雪島高等専門学校](#)

著者名

村上 春樹 / 著

出版者

新潮社

出版年

2009/0

1Q84 a novel BOOK1 4月-6月

村上 春樹 / 著

新潮社

2009/0

1Q84 a novel BOOK3 10月-12月

村上 春樹 / 著

新潮社

2010/0


村上春樹『1Q84』をどう読むか

河出書房新社編集部 / 編 安藤 礼二 / (ほか著) 河出書房新社

2009/0

日常生活の変化

- プログラミングの補助者 (GitHub Copilot)

```
69 def search_isbn(systemid: str, endpoint: str, isbn: str, **_)
70     """
71     ISBNで検索する
72     """
73     _queue = search(queue=UTQueueV1(systemid=systemid, endpo
74     holdings: list[ Holding ] = []
75     for __ in _queue['books']:
76         holdings.append(Holding(id='', library='', place='',
77         url = _queue['books'][0]['url'] if len(_queue['books'])
78
79      if len(holdings) == 0:
80         raise EmptyResultsError()
81     return url, holdings
```

そもそも小学生のころ…

漢字は覚えないと決意→IMEがある！

提出物もだいたいプリンター

テストは全部ひらがな

AIによる日常の変化

- よくわからないことを検索したい（レファレンス）
 - どう検索していいかわからない
 - 対話から関連する情報を引き出す→Googleへ
or (Bing AI)
- 苦手なことを補完できる = アクセシブル

ChatGPTを活用した実証実験を始めたきっかけ

TOPICS

[See more](#)

RESOURCES

r/librarians • 6 mo. ago
by [deleted][Join](#)[Translate](#)

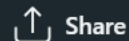
How is ChatGPT going to affect librarianship?

Has anyone fiddled with ChatGPT and thought about how it can drastically disrupt our field?

I've asked ChatGPT some test questions, and find that it is particularly useful when it comes to writing book summaries, brainstorming ideas for programs, and answering simple reference and reader's advisory questions (not with 100% accuracy - though I foresee this will improve in the future). Other users have used it for copywriting (e.g. blog posts, marketing materials, etc.), writing code for websites, writing emails, etc. If it becomes a ubiquitous tool at work in the future, what does it mean for people working in libraries?

ChatGPT has also made me think about the skills that we have that are NOT affected by ChatGPT. Programming, communication and instruction are such a core part of librarianship and while ChatGPT can accelerate the planning process (e.g. coming up with content and materials), the actual act of human delivery can never be replaced. These face-to-face interactions (either in-person or online) and community relationships is also what I find a lot of patrons crave and enjoy. Additionally, because of the uncanny valley effect, I highly doubt that there will be enough industry motivation to create robots who can replace us in carrying out these activities in the future. Should the library field, including MLIS programs, put more focus on these soft skills instead with ChatGPT looming in the future? There is also a lot of advice about learning programming to increase chances of employment. Would that still be viable advice 5 years from now?

What does everyone think? There has been some discussion about how ChatGPT is affecting academia (e.g. plagiarism by

[Read more](#)

Reddit 海外図書館員コミュニティの議論から

- How is ChatGPT going to affect librarianship?
 - 常連客の少なくとも半数はコンピューターを怖がっています。したがって、レファレンス図書館員は当面は安全だと思います。
 - ChatGPT が Google の検索エンジンのように主流になったら、その影響は Google の世界と同じになるのではないかと思います。
- The changing role of librarians - are you ready?
 - ChatGPT のような新しい技術でデータベースを検索して引用を引き出すことができるようになったとしても、将来的にはそれらが包括的になったり無料になったりするとは思えません。図書館員だけでなく出版業界も、AI が真に革命的になることを妨げるでしょう。図書館員は引き続き、独自の図書館ごとにデータベースと電子リソースのパッケージを購入する必要があります。そして、完全で正確な情報が確実に得られるように、ギャップを埋める必要があります。

図書館でつかってみる実験

配架場所を案内



あなたは大学図書館の司書です。聞かれたことに対して、その本がある場所を日本十進分類法で答えてください。



もちろん、あなたが探している本のタイトルや著者名、またはその本の具体的な内容を教えていただけますか？それによって、日本十進分類法に従い、その本の場所を特定することが可能になります。



生成AIについて知りたい



生成AI（Generative AI）についての書籍を探しているとのこと、了解しました。これは人工知能や機械学習に関連するトピックですので、日本十進分類法では「情報科学」に関連する区分となります。



具体的には以下のカテゴリーに分類されます：

- 007：コンピュータ科学
- 007.6：人工知能

この情報を基に、あなたの図書館の該当セクションを探してみてください。ただし、具体的な本の位置は図書館の規模や収蔵状況によって異なる可能性があることをご了承ください。また、図書館のスタッフに直接ご質問いただくと、より具体的な情報を提供できるかと思います。

🔄 Regenerate response

NDC Navi

大阪市立大学 村上研究室 浦 芳伸

検索

「007.13 人工知能. パターン認識」に関する情報

分類記号	007.13
事項	人工知能. パターン認識
関連する用語	人工知能 , 図形認識 , パターン認識
本を探す	<input type="button" value="NDL API"/> <input type="button" value="大阪公立大学図書館OPAC"/>

分類表

分類記号	事項
	人工知能. パターン認識



亀を飼いたいと思っています。どうしたらいいですか



亀を飼うことに興味があるとのこと、了解しました。亀の飼育に関する情報は、「動物」や「ペット」に関連するカテゴリで見つけることができます。日本十進分類法では次のようになります。



- 599：爬虫類、両生類
- 597.8：亀類

また、ペットとしての亀の飼育について具体的なガイドラインやアドバイスが含まれている書籍は、「趣味・実用」に関連するカテゴリにも分類されることがあります。

- 796：動物関連の趣味

これらの分類を基に図書館で書籍を探してみてください。ただし、具体的な本の位置は図書館の規模や収蔵状況によって異なる可能性があることをご了承ください。また、図書館のスタッフに直接ご質問いただくと、より具体的な情報を提供できるかと思います。

また、亀を飼う場合は、種類や生息地、飼育環境に関する知識を得るだけでなく、法律や規制についても確認してください。特定の種類の亀は保護されている場合があり、飼育が制限されていることがあります。飼育に必要な許可や手続きについても確認することが重要です。

790	諸芸. 娯楽
791	茶道
792	香道
793	花道
794	撞球
795	囲碁
796	将棋
797	射倅ゲーム
798	室内娯楽
799	ダンス

出典 : <https://www.libnet.pref.okayama.jp/shiryuu/ndc/ndc70.htm>

ぽいけど、ちがう？

覚え違いタイトルを推定



「100万回生きたねこ」というタイトルの本を覚え違った場合のタイトルを10個考えて。



もちろんです、以下のようなバリエーションが考えられます：



1. 「100万年生きたねこ」
2. 「10万回生きたねこ」
3. 「100万回死んだねこ」
4. 「100回生きたねこ」
5. 「100万回蘇ったねこ」
6. 「100万の命を持つねこ」
7. 「100万回生き返ったねこ」
8. 「100万回と生きた猫」
9. 「100万回生まれ変わったねこ」
10. 「ねこが100万回生きた」

検索結果を評価する



以下は「100万回死んだねこ」という検索キーワードに対しての検索結果上位3冊です。
この結果をどう評価しますか？
また、ほかに考えられる検索キーワードがあれば「」でくって教えてください。



100万回死んだねこ 覚え違いタイトル集福井県立図書館／著・編集講談社
2021978406525892747

佐野洋子対談集

人生のきほん 佐野洋子／著 西原理恵子／著 リリー・フランキー／著講談社201197840621609020

100万回生きたねこ佐野洋子／著講談社19779784061272743



検索キーワード「100万回死んだねこ」に対する上位3冊の検索結果について、以下のように評価します。



1. 「100万回死んだねこ 覚え違いタイトル集」：この本は、福井県立図書館が編集し、講談社から発行された書籍です。この本のタイトルは、佐野洋子の『100万回生きたねこ』を「100万回死んだねこ」と覚え違いされた事例に由来するようです。検索キーワードと完全に一致しており、適切な検索結果といえます。
2. 「佐野洋子対談集 人生のきほん」：この本は、佐野洋子が著者の一人である対談集です。佐野洋子は『100万回生きたねこ』の著者であり、この作品が検索キーワード「100万回死んだねこ」に関連しているため、佐野洋子の他の作品も検索結果として上がってきたと考えられます。しかし、この本自体は検索キーワードと直接的な関連性はなさそうです。
3. 「100万回生きたねこ」：この本は、佐野洋子が著した作品で、「100万回死んだねこ」という検索キーワードがこの作品を覚え違いしたものであると考えられます。したがって、検索キーワードと深く関連してい

APIで拡張してみよう

[Code](#)[Blame](#)

22 lines (19 loc) · 894 Bytes

[Raw](#)

```
1 import yaml
2 import openai
3 import json
4 from rich.console import Console
5
6 console = Console()
7 with open('../secret.yaml') as file:
8     secret = yaml.safe_load(file)
9
10 openai.organization = secret['OPENAI_ORG']
11 openai.api_key = secret['OPENAI_API_KEY']
12
13 completion = openai.ChatCompletion.create(
14     model='gpt-3.5-turbo',
15     messages=[
16         {'role': 'system', 'content': "あなたは優秀な公共図書館の司書です。" +
17         "質問に沿って最適な検索キーワードを提案してください。" +
18         "キーワードは短い単語で20個提案してください。接続詞と助詞を含まないようにしてください。"},
19         {'role': 'user', 'content': "OCRを開発したい"}
20     ])
21 # print(type(completion.choices[0].message.content))
22 console.print_json(json.dumps(dict(completion)))
```

蔵書検索サポーターの実装

利用イメージを体験するために、カーリルでは学校図書館支援プログラムで提供している学校図書館向けの蔵書検索サービスにChatGPT (OpenAI API)を組み込み、「蔵書検索サポーター」を実装することにしました。例えば、ユーザーが検索キーワードにヒットする本がなかった場合に、別のキーワードを提案したり、自然文の質問から適切な検索キーワードを提案することを想定します。カーリルが提供する検索サービスもAPIでアクセスできるため、OpenAI APIと接続するプロセスは驚くほど簡単でした。

AIはこれからどうなるか

AIの未来について検索していますか。[AIの未来](#)やAIの展望で検索すると見つかるかもしれません。

AIは技術の進歩により、より高度な自然言語処理や画像認識が可能になることが期待されています。さらに、自動運転やロボット技術の発展にも寄与し、社会全体の効率化や新たなサービスの創出が進むでしょう。

質問を送信



× 閉じる

APIのプロンプト

- あなたは図書館の蔵書検索サポーターです。ユーザーから与えられたキーワードをもとに、検索している意図を想定して適切な蔵書検索キーワードの候補を考えて、キーワードは必ず[]で囲った上で、キーワードが難しい漢字の場合、一般的にひらがなでも通用するなら、ひらがなにして、次のフォーマットで応答してください。”〇〇について検索していますか。[キーワード]や[キーワード]で検索すると見つかるかもしれません。”
そのあと、質問に簡潔に答えてください。

2 / 2 >



あなたは図書館の蔵書検索サポーターです。ユーザーから与えられたキーワードをもとに、検索している意図を想定して適切な蔵書検索キーワードの候補を10個以上考えて、キーワードは必ず□で囲った上で、キーワードが難しい漢字の場合、一般的にひらがなでも通用するなら、ひらがなにして、次のフォーマットで応答してください。"○○について検索していますか。[キーワード]や[キーワード]で検索すると見つかるかもしれません。"そのあと、質問に簡潔に答えてください。



もちろんですが、ユーザーからのキーワードが必要です。どのトピックを検索しようとしていますか？



プログラマーになりたい



プログラマーになりたいという意図であれば、次のようなキーワードで図書館の蔵書を検索すると良いでしょう。



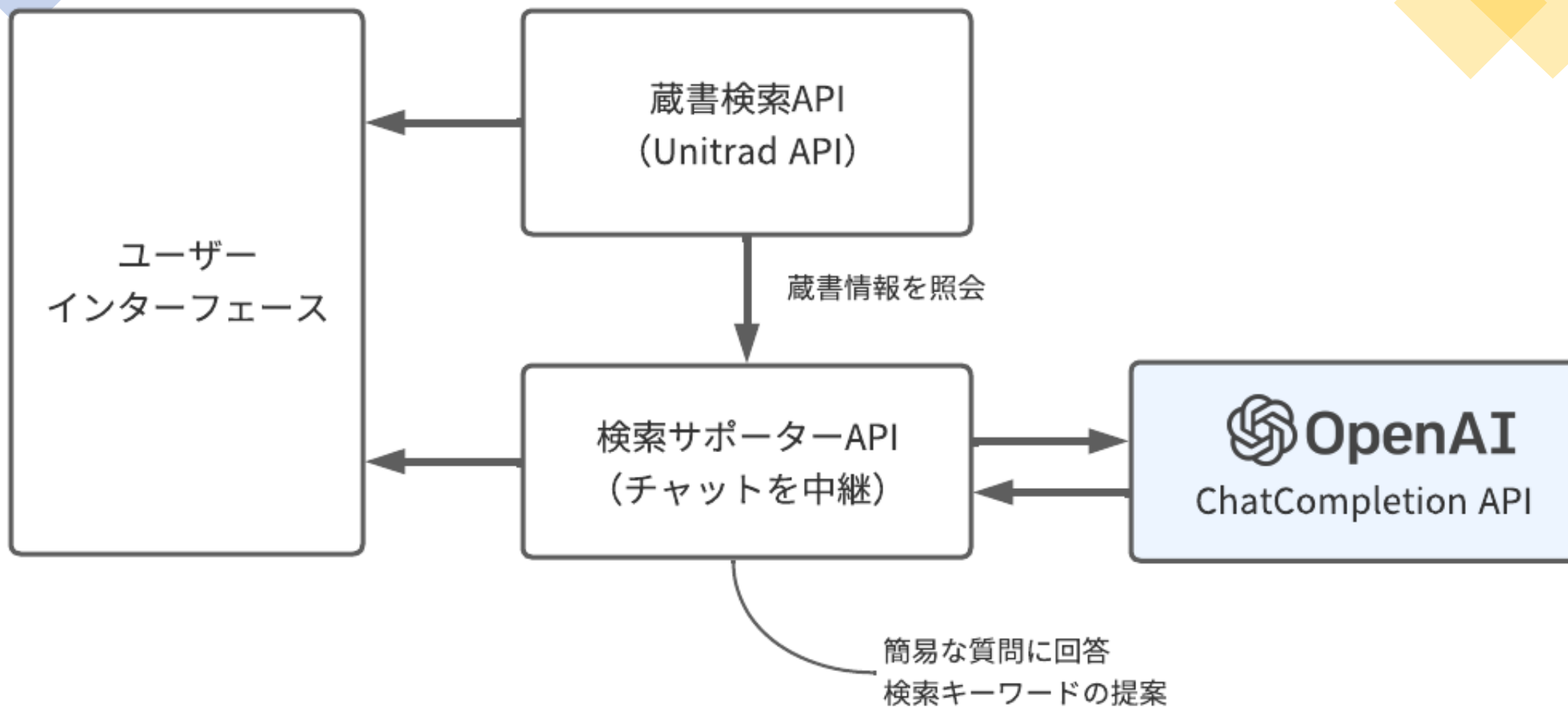
OPACで検索する

"プログラマーになりたいということで、[プログラミング入門]や[コーディング学習]、[プログラム言語]、[ソフトウェア開発]、[アルゴリズム]、[データ構造]、[オブジェクト指向プログラミング]、[フロントエンド開発]、[バックエンド開発]、[フルスタック開発]、[コンピュータサイエンス]で検索すると見つかるかもしれません。"

また、特定のプログラミング言語について学びたい場合、その言語の名前（例えば[Python]、[Java]、[JavaScript]など）でも検索できます。

ChatGPTの応答を蔵書検索サービスに照会

- 10件程度並行処理
- 検索結果があるキーワードのみを応答に含める
- 事前学習なしで図書館の蔵書にあわせた応答になる





亀を育てたい



亀の飼育について検索していますか。[亀の飼い方]や[亀の種類]、[亀の飼育環境]、[亀のエサ]、[亀の健康管理]、[亀の寿命]、[ペットとしての亀]、[亀の行動]、[亀の繁殖]、[亀の病気]、[亀の手入れ方法]で検索すると見つかるかもしれません。



亀を飼う際には、その種類や性格、寿命、必要な飼育環境などをしっかりと理解してから始めることが重要です。また、亀は驚くほど長寿で、適切なケアを行えば数十年以上にわたり飼うことができます。そのため、飼い始める前には、長期的なコミットメントが必要であることを理解しておくべきです。また、亀は比較的手間のかからないペットではありますが、適切な飼育環境を提供し、栄養バランスの取れた食事を与えることが求められます。亀の飼育に関する情報をしっかりと調べてから飼い始めることをお勧めします。

実際にはプログラムから呼んでいますが、基本的には同じ。



提案いただいたキーワードではヒットしませんでした。もう少し上位概念のキーワードを提案いただけないでしょうか。



それならば、より広範なキーワードで検索を試みる方が良いかもしれません。亀の飼育について検索していただけますか。[ペットの飼い方]、[爬虫類の飼育]、[動物の飼い方]、[生物学]、[生態学]、[動物行動学]、[動物保護]、[動物の健康]、[動物病理学]、[家庭でのペットの飼い方]で検索すると見つかるかもしれません。



なお、特定の情報を見つけるのに困っている場合、その情報がどのような分野や概念に関連しているかを考えることで、より効果的な検索キーワードを思いつくことがあります。亀の飼育は生物学、動物学、動物行動学などの広範な分野に関連していますので、そのような分野についての一般的な情報を検索することで、求めている情報を見つける手がかりを得られるかもしれません。

実際にはプログラムから呼んでいますが、基本的には同じ。

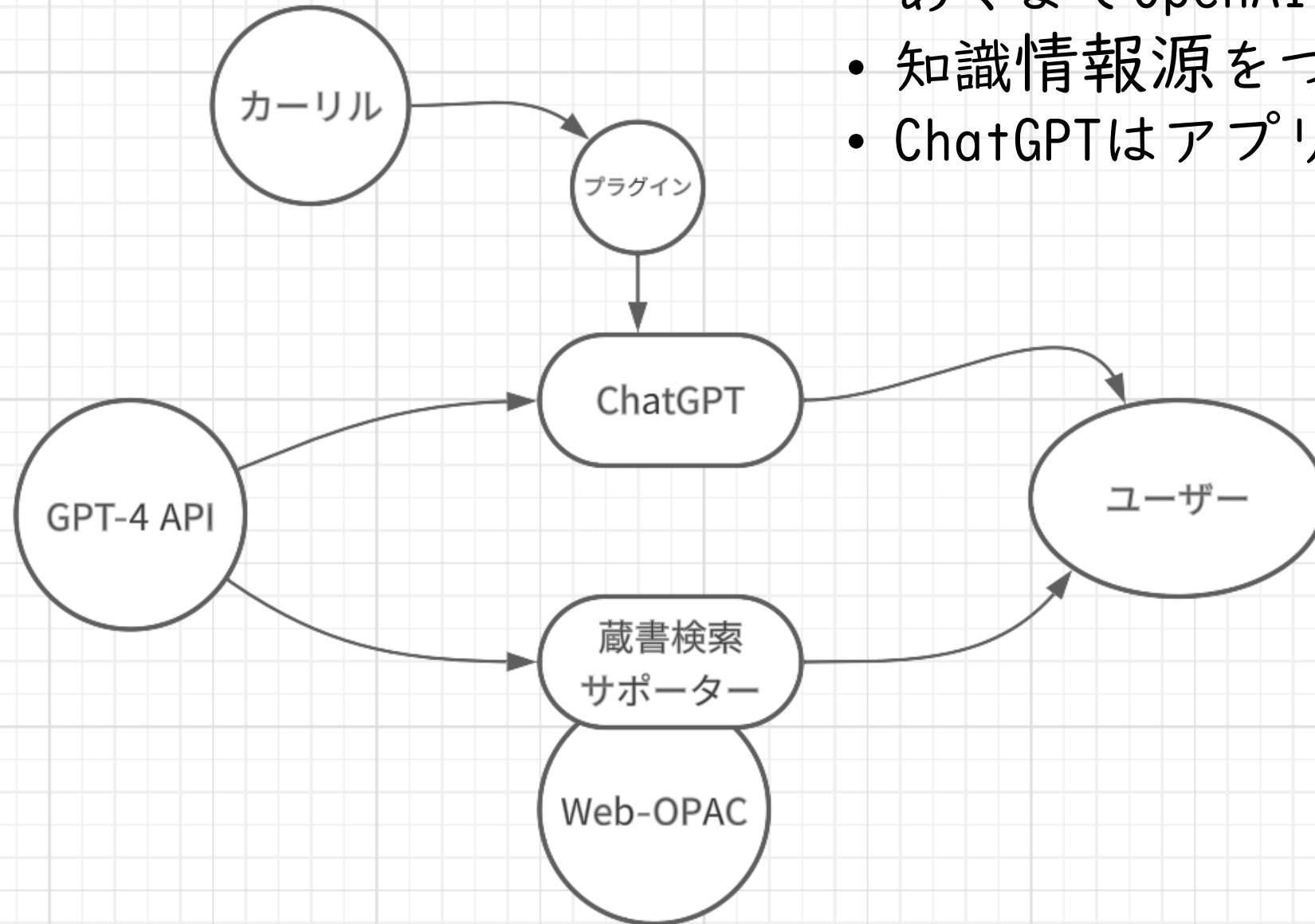
APIを使うと、知識は拡張できる

あるいは、正しい知識を注入できる
並列で大量にリクエストして、その結果を集約したりもできる。

例えば蔵書検索でサポートできそうなこと

- 関連キーワードのサジェスト
- 上位概念や下位概念の整理
- 類似書誌のグループ化、絞り込みの支援
- データベースなど他の情報源への誘導
- これまで個別に開発されてきた様々な要素が汎用的に実現
- 高校生でも読めそうな本。とか聞きたいけど…

- あくまでOpenAIの技術戦略
- 知識情報源をつなぐ仕組み
- ChatGPTはアプリケーション例



ChatGPTで遊んでみて

- 知識情報源ではない
(正確な情報は、検索したほうが速い)
- 自然言語⇔構造化データの変換が簡単になった
- 違う構造のデータをつないだり、加工したりできる
- 図書館メタデータとの相性はよさそう
情報をつなぐことで効率的にデータ処理できる

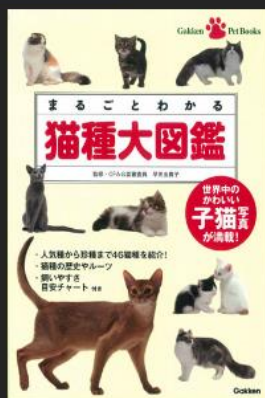
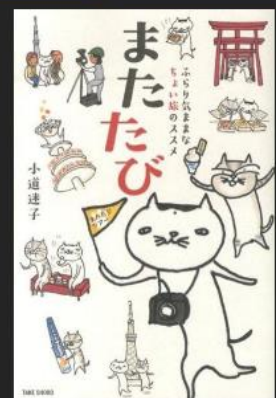
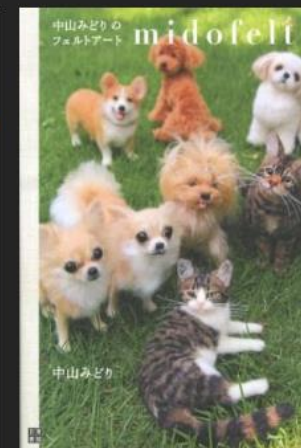
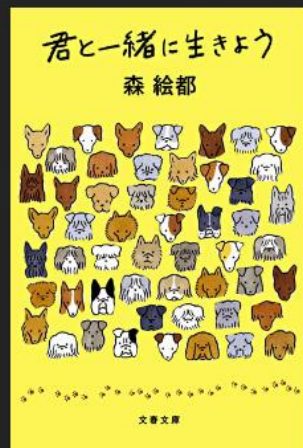


CLIP: Connecting text and images

書影をベクトル化・類似検索

- 20万冊の書影データをベクトル化
- クエリ文字をベクトル化して、距離計算
- 書誌情報は一切与えない

many cats

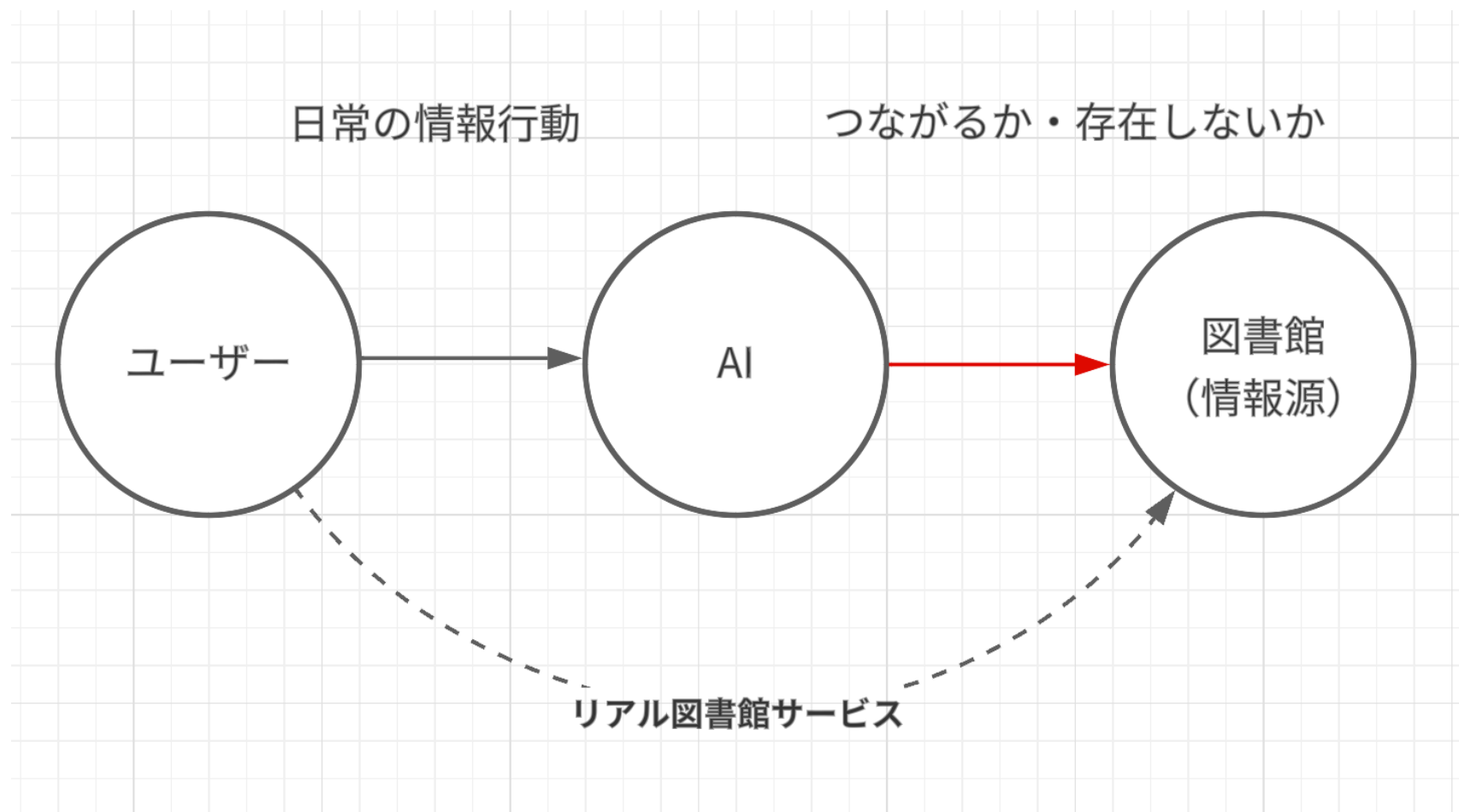


さいごに

Chat-GPTと大学図書館

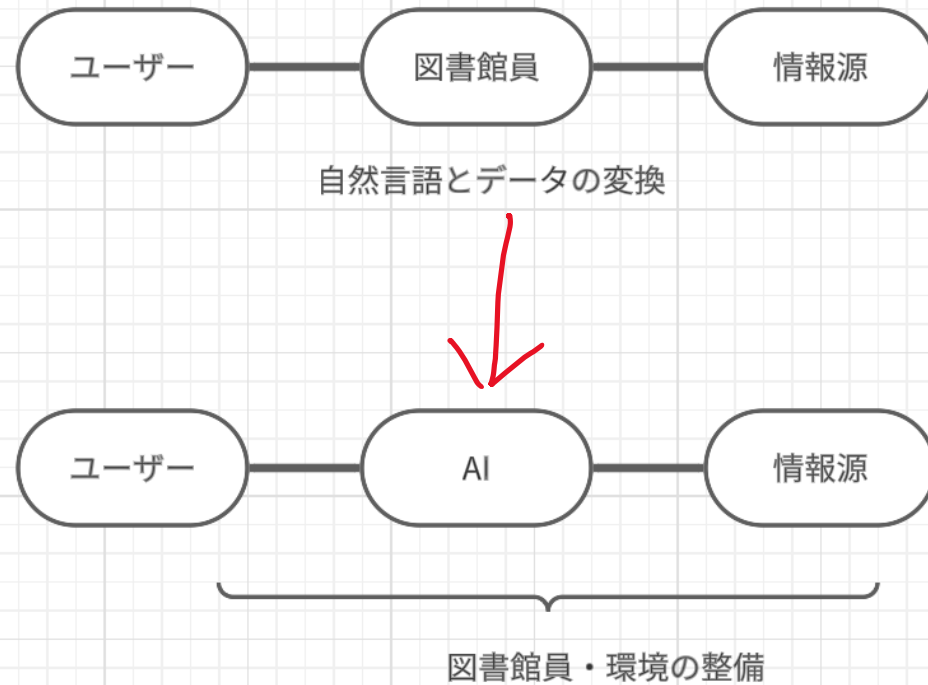
- 図書館業務での活用
- 図書館サービスでの活用

ユーザーの図書館への要求の変化



図書館をどうつないでいくか

図書館員のしごと



目指すべき未来を共有する

“図書館がやります”と宣言できるか、できないか。

Chat-GPTと大学図書館

- 知識情報源としての図書館、
あるいはAI学習ソースの公平性・民主化
- データレイクの費用負担（いまはネット情報が学習源だが…）
 - 情報の信頼性維持
- うまく使える人と使えない人の格差をどう縮めるか