How to get more people onboard with Pharo ?

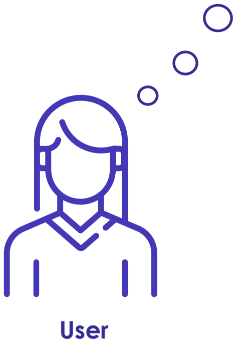
Applying Large Language Models (LLM) as support for the onboarding of new developers.

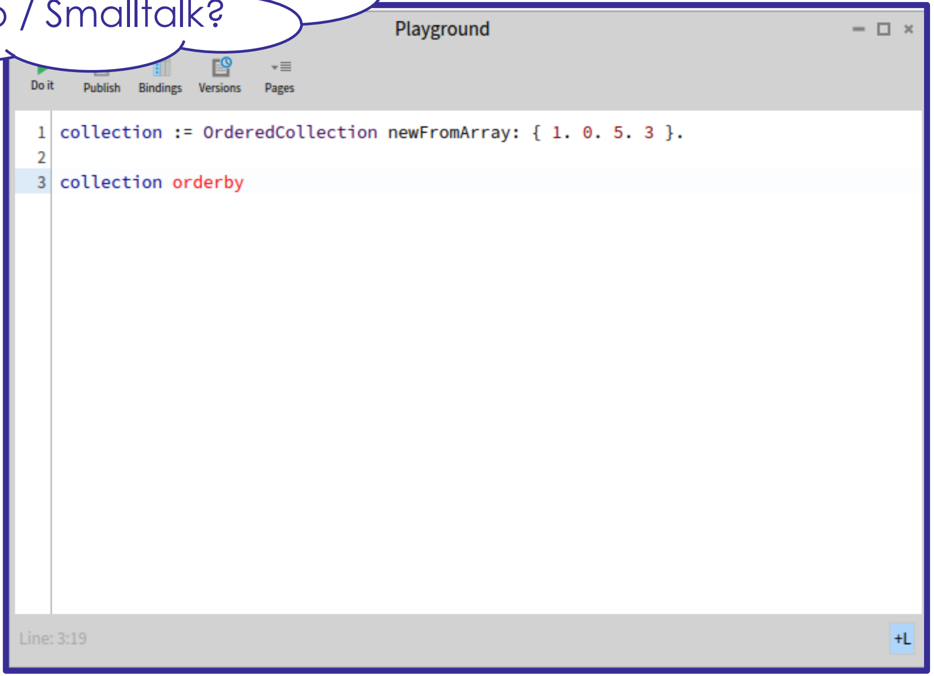
Marius Pingaud
Pascal Zaragoza

ChatGPT3.5

The Typical New User Experience: How do I do this?

How do I do … in Pharo / Smalltalk?



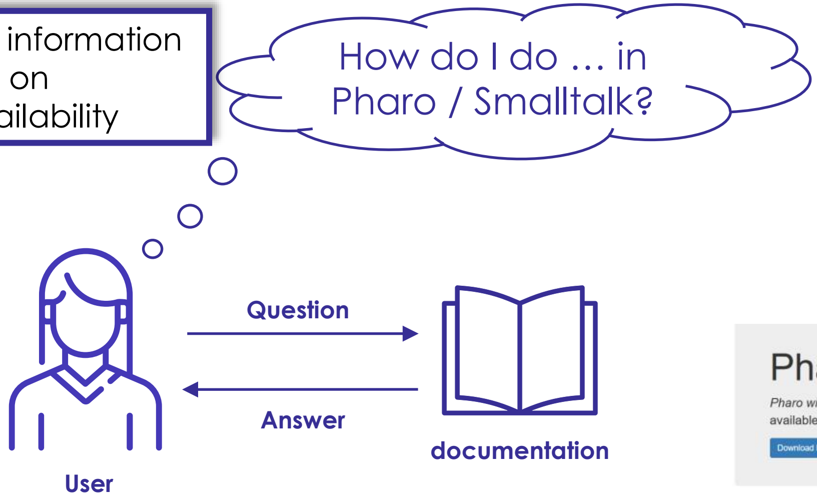


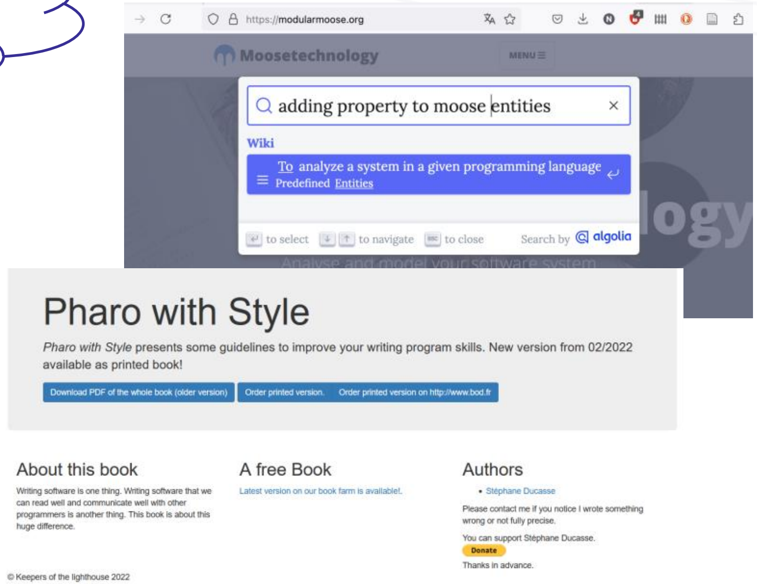
```
collection := OrderedCollection newFromArray: { 1. 0. 5. 3 }.

collection orderby
```

ESUG 2024

The Typical New User Experience: Using documentation

+ highly complete information
- Highly dependent on documentation availability


How do I do … in Pharo / Smalltalk?


https://modularmoose.org
Moosetechnology
MENU ☰

adding property to moose entities

Wiki

To analyze a system in a given programming language
Predefined Entities

to select
to navigate
to close
Search by algolia
Analyse and model your software system

Pharo with Style

Pharo with Style presents some guidelines to improve your writing program skills. New version from 02/2022 available as printed book!

Download PDF of the whole book (older version)
Order printed version.
Order printed version on http://www.bod.fr

About this book

Writing software is one thing. Writing software that we can read well and communicate well with other programmers is another thing. This book is about this huge difference.

A free Book

Latest version on our book farm is available!.

Authors

- Stéphane Ducasse

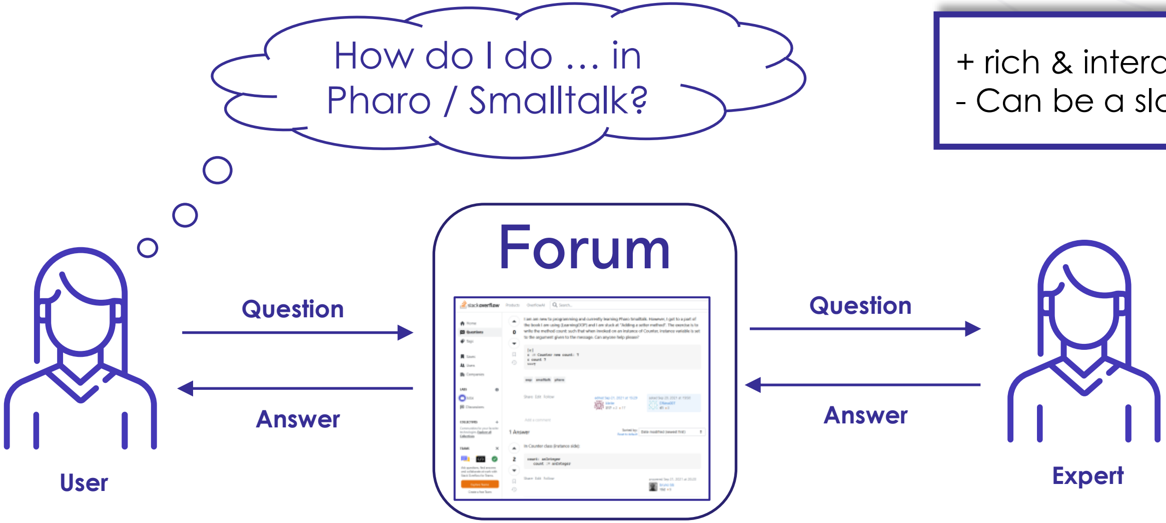
Please contact me if you notice I wrote something wrong or not fully precise.

You can support Stéphane Ducasse.

Thanks in advance.

© Keepers of the lighthouse 2022

The Typical New User Experience: Asking experts



How do I do … in Pharo / Smalltalk?

Forum

I am am new to programming and currently learning Pharo Smalltalk. However, I got to a part of the book I am using (LearningOOP) and I am stuck at "Adding a setter method". The exercise is to write the method count: such that when invoked on an instance of Counter, instance variable is set to the argument given to the message. Can anyone help please?

```
| c |
c := Counter new count: 7
c count ?
>>>7
```

oop smalltalk pharo

1 Answer

In Counter class (instance side):

```
count: anInteger
   count := anInteger
```

Question

Answer

User

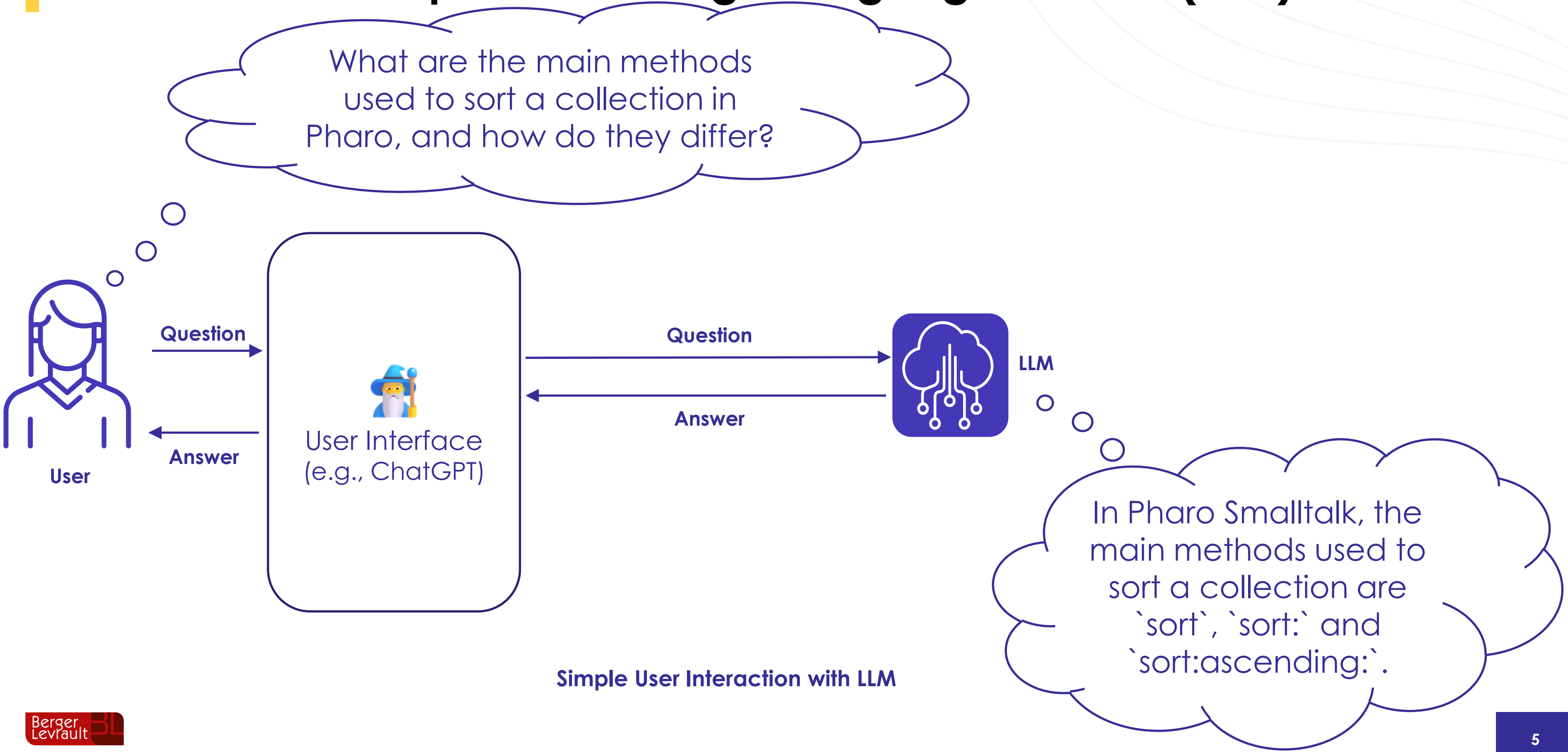
Question

Answer

Expert

+ rich & interactive response
- Can be a slow process

The New User Experience: Large Language Models (LLM)

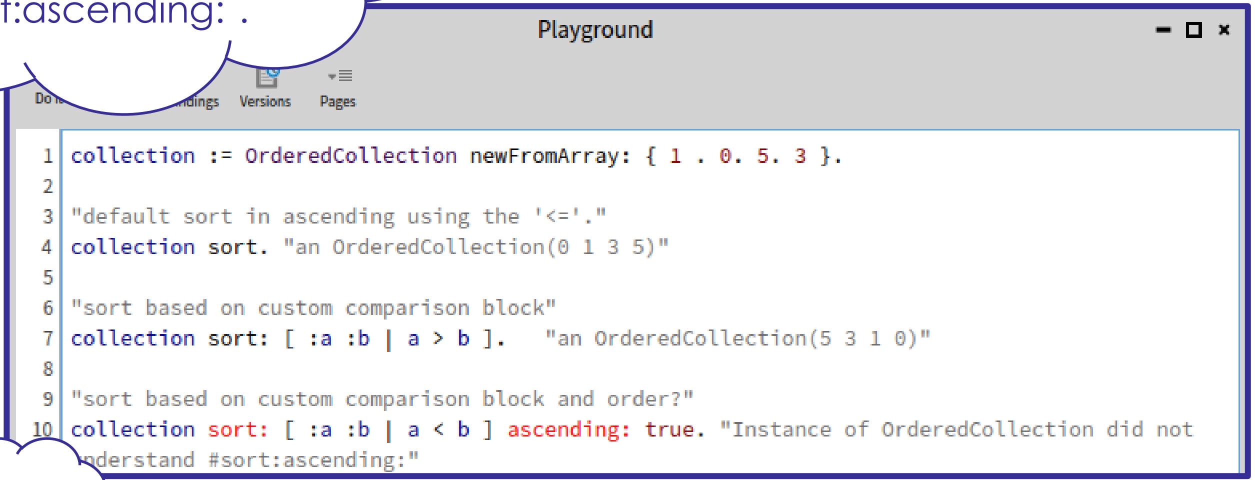

What are the main methods used to sort a collection in Pharo, and how do they differ?
Question
Answer
User
User Interface (e.g., ChatGPT)
Question
Answer
LLM
In Pharo Smalltalk, the main methods used to sort a collection are `sort`, `sort:` and `sort:ascending:`.

Simple User Interaction with LLM

New User Experience using Large Language Models (LLM)

In Pharo Smalltalk, the main methods used to sort a collection are `sort`, `sort:` and `sort:ascending:`.

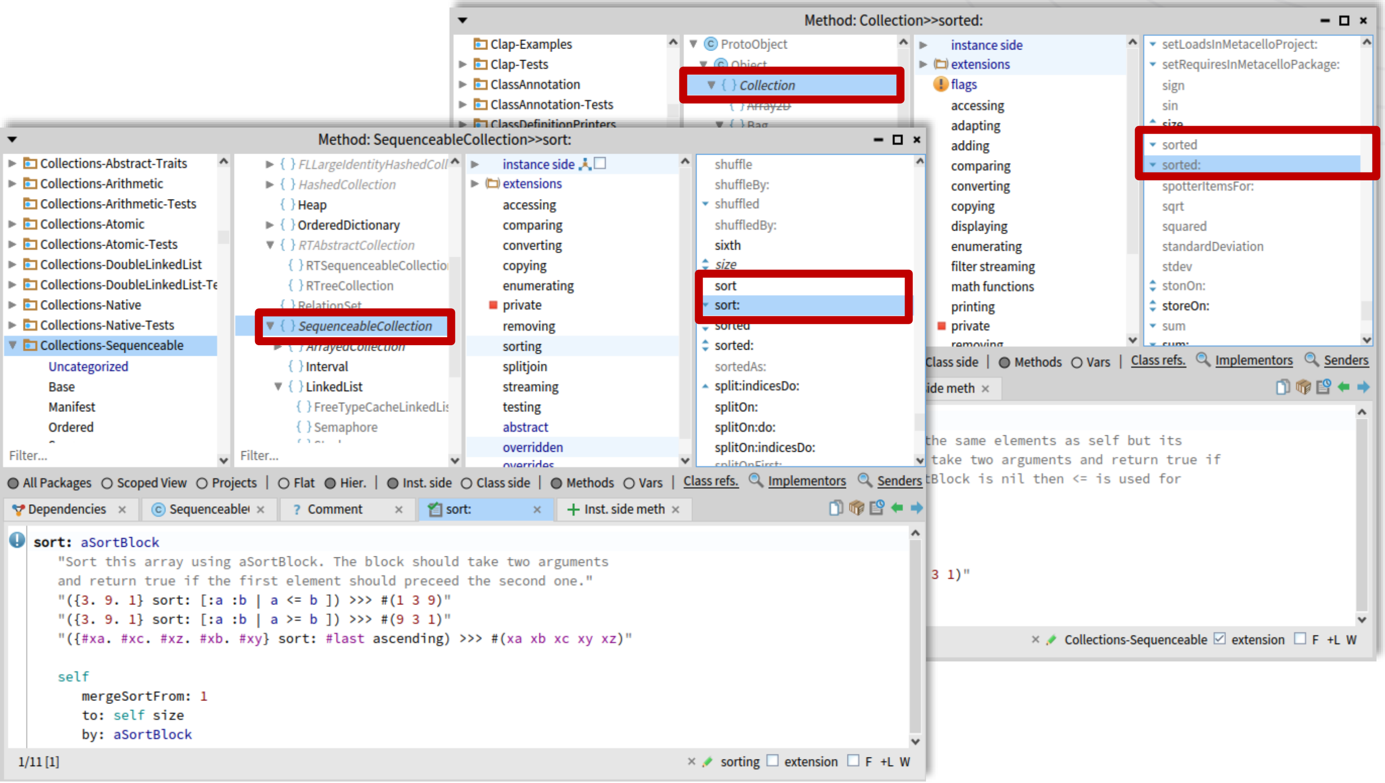
User



```
collection := OrderedCollection newFromArray: { 1 . 0. 5. 3 }.

"default sort in ascending using the '<='."
collection sort. "an OrderedCollection(0 1 3 5)"

"sort based on custom comparison block"
collection sort: [ :a :b | a > b ].   "an OrderedCollection(5 3 1 0)"

"sort based on custom comparison block and order?"
collection sort: [ :a :b | a < b ] ascending: true. "Instance of OrderedCollection did not
nderstand #sort:ascending:"
```

What?!

New User Experience using Large Language Models (LLM)



```
sort: aSortBlock
	"Sort this array using aSortBlock. The block should take two arguments
	and return true if the first element should preceed the second one."
	"({3. 9. 1} sort: [:a :b | a <= b ]) >>> #(1 3 9)"
	"({3. 9. 1} sort: [:a :b | a >= b ]) >>> #(9 3 1)"
	"({#xa. #xc. #xz. #xb. #xy} sort: #last ascending) >>> #(xa xb xc xy xz)"

	self
		mergeSortFrom: 1
		to: self size
		by: aSortBlock
```

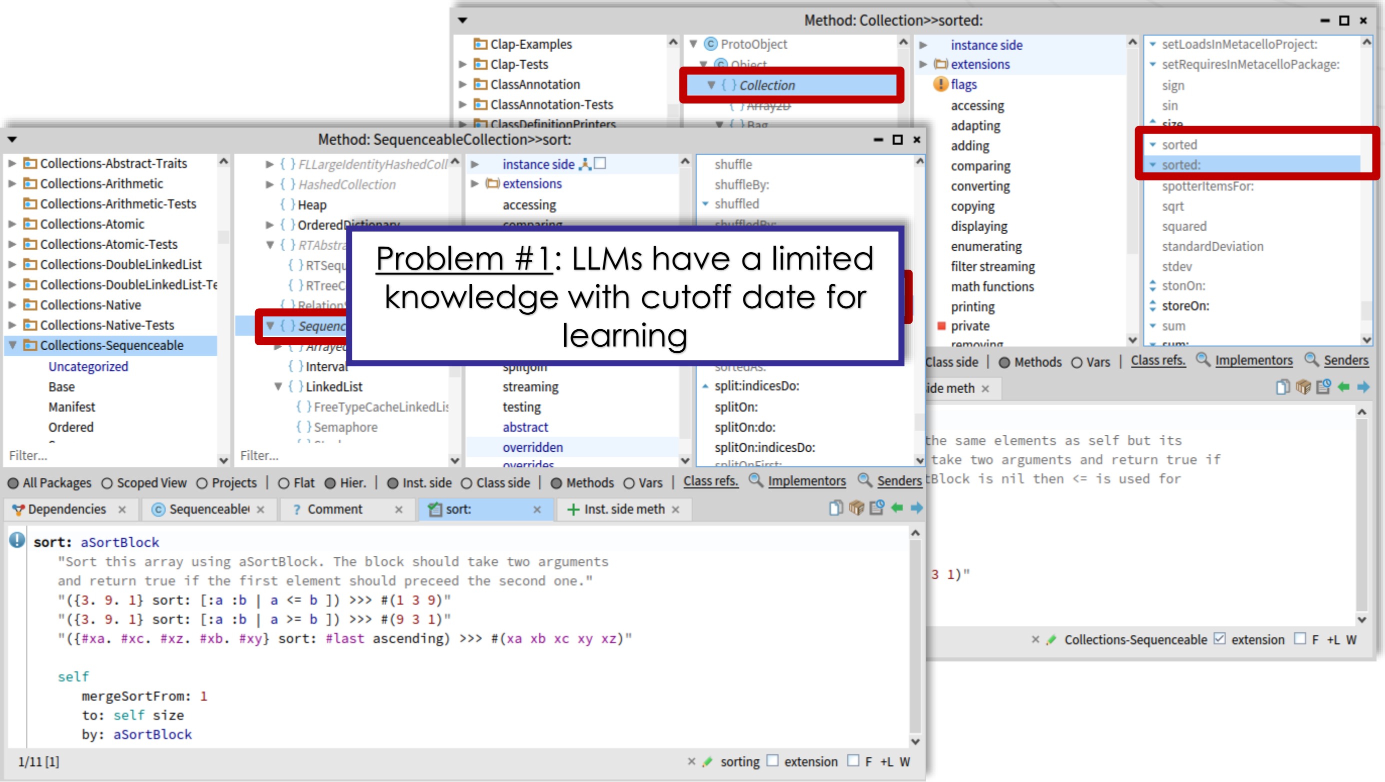
```
the same elements as self but its
take two arguments and return true if
tBlock is nil then <= is used for

3 1)"
```

User

Sorting methods for Collection & Ordered Collection.

ESUG 2024

New User Experience using Large Language Models (LLM)



```
the same elements as self but its
take two arguments and return true if
tBlock is nil then <= is used for

3 1)"
```

```
sort: aSortBlock
	"Sort this array using aSortBlock. The block should take two arguments
	and return true if the first element should preceed the second one."
	"({3. 9. 1} sort: [:a :b | a <= b ]) >>> #(1 3 9)"
	"({3. 9. 1} sort: [:a :b | a >= b ]) >>> #(9 3 1)"
	"({#xa. #xc. #xz. #xb. #xy} sort: #last ascending) >>> #(xa xb xc xy xz)"

	self
		mergeSortFrom: 1
		to: self size
		by: aSortBlock
```

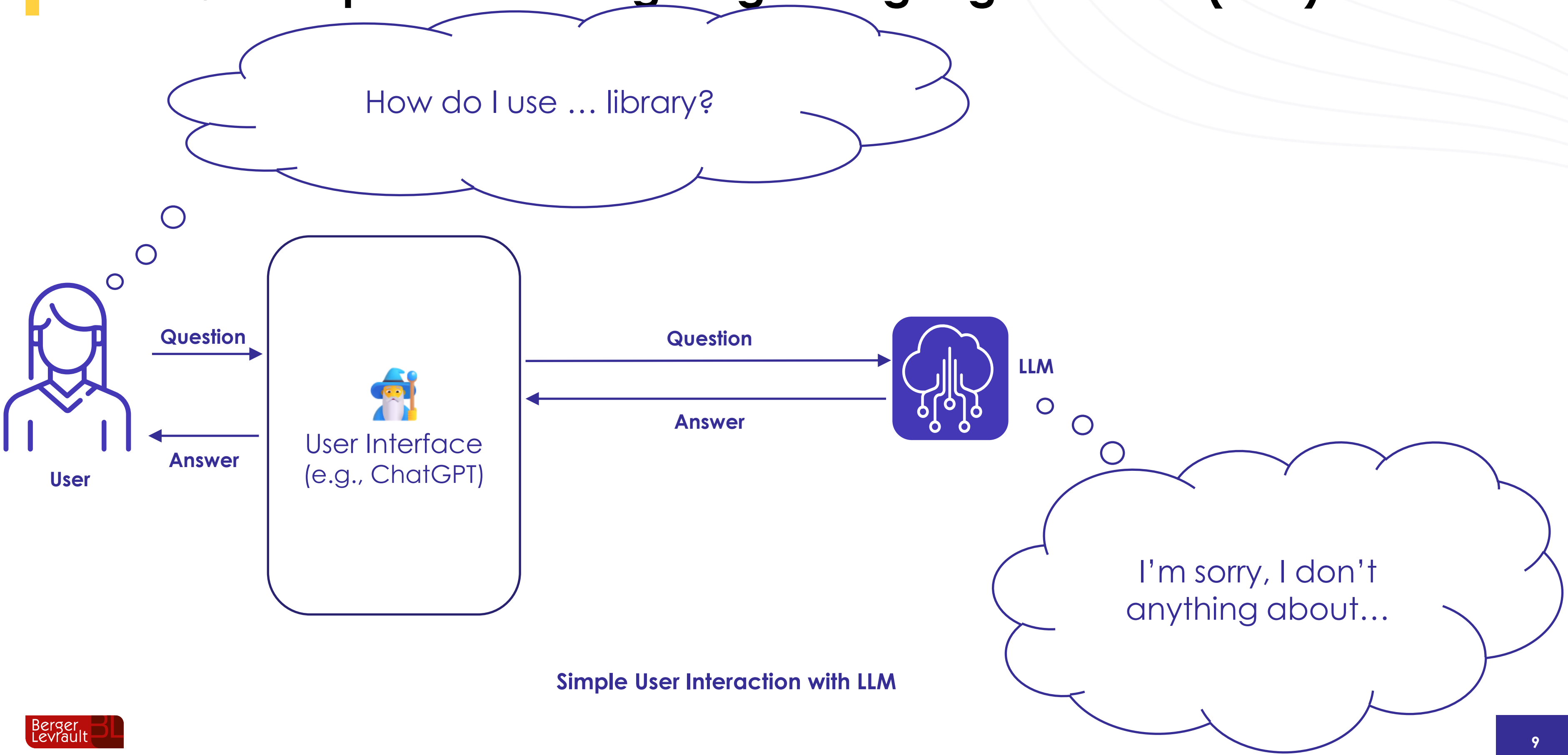
Problem #1: LLMs have a limited knowledge with cutoff date for learning

User

Sorting methods for Collection & Ordered Collection.

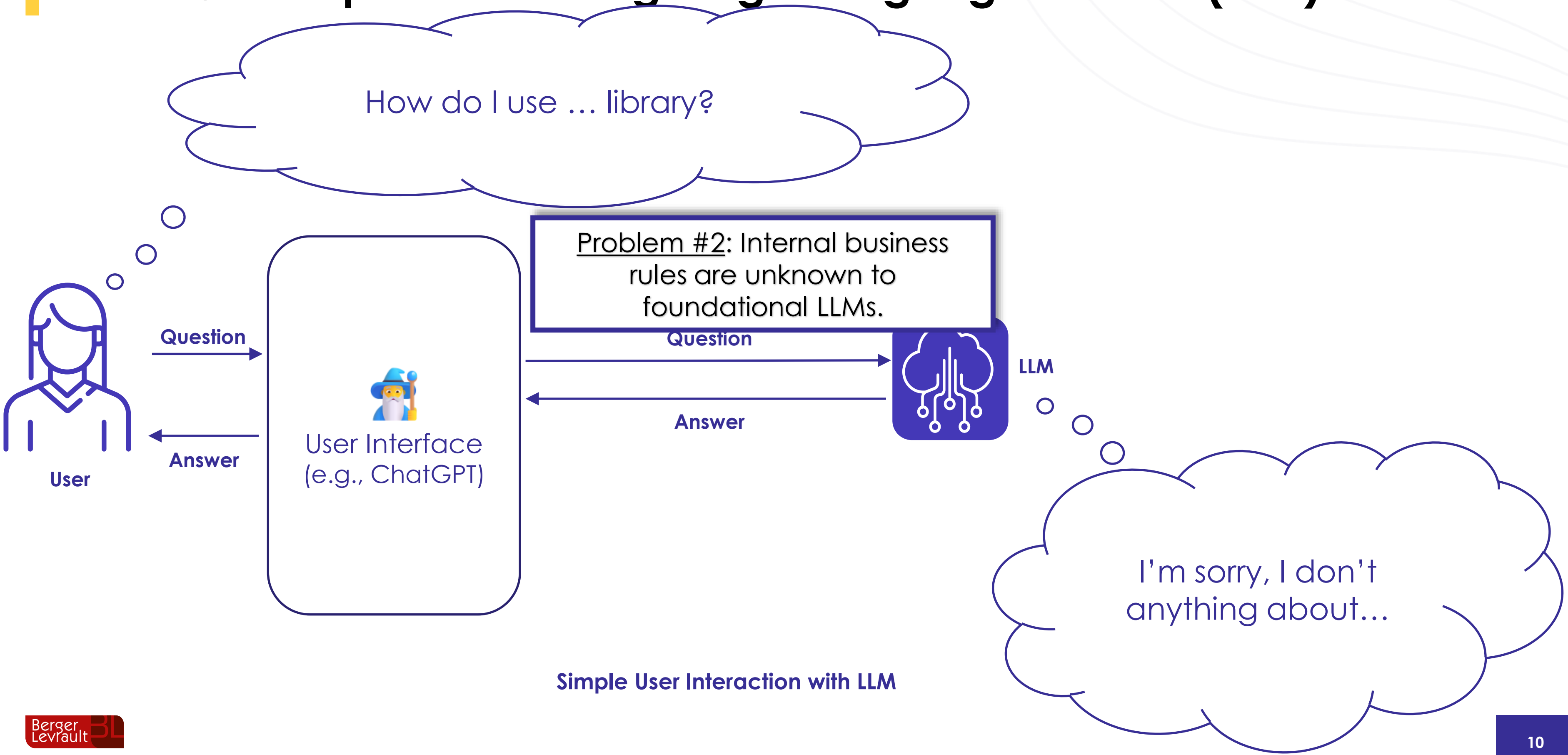
ESUG 2024

New User Experience using Large Language Models (LLM)



Simple User Interaction with LLM

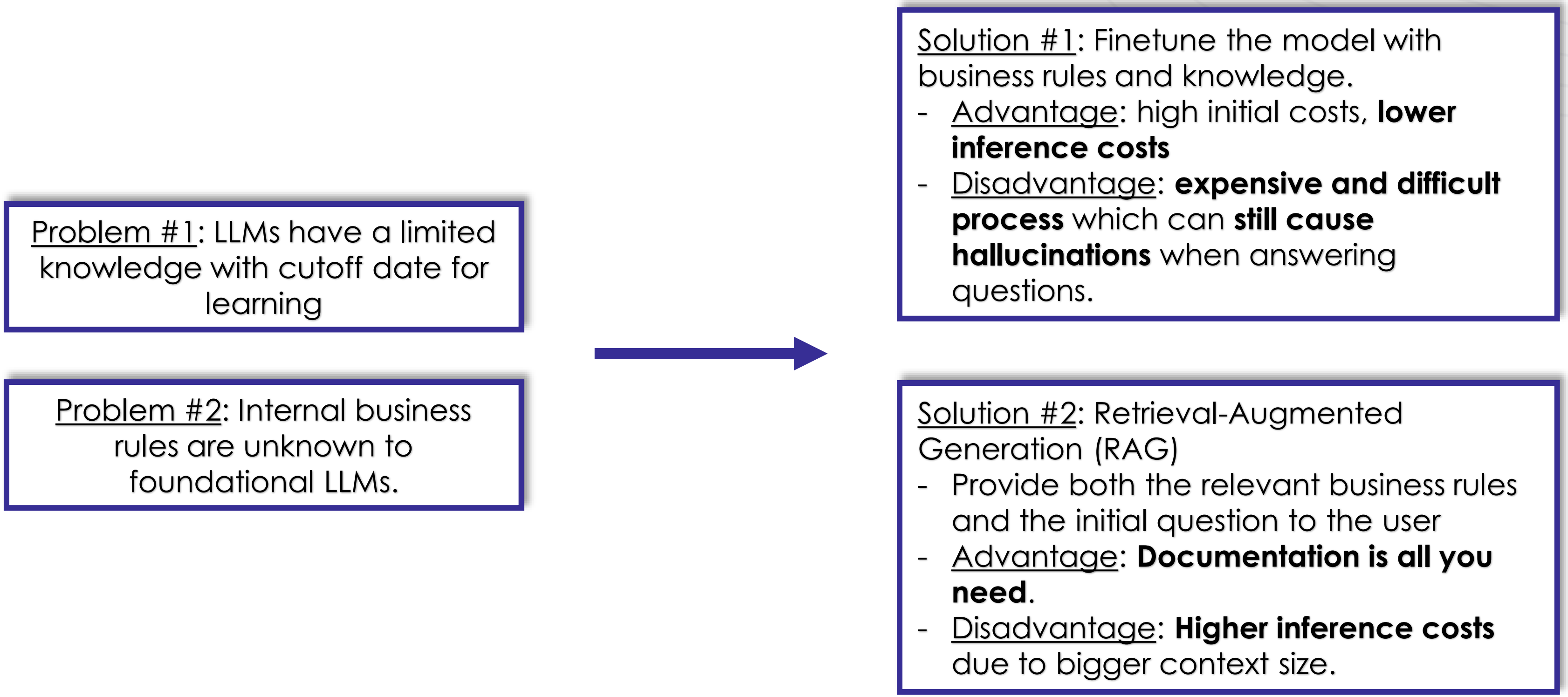
New User Experience using Large Language Models (LLM)



Problem #2: Internal business rules are unknown to foundational LLMs.

Simple User Interaction with LLM

The problem with foundational LLMs



Problem #1: LLMs have a limited knowledge with cutoff date for learning

Problem #2: Internal business rules are unknown to foundational LLMs.

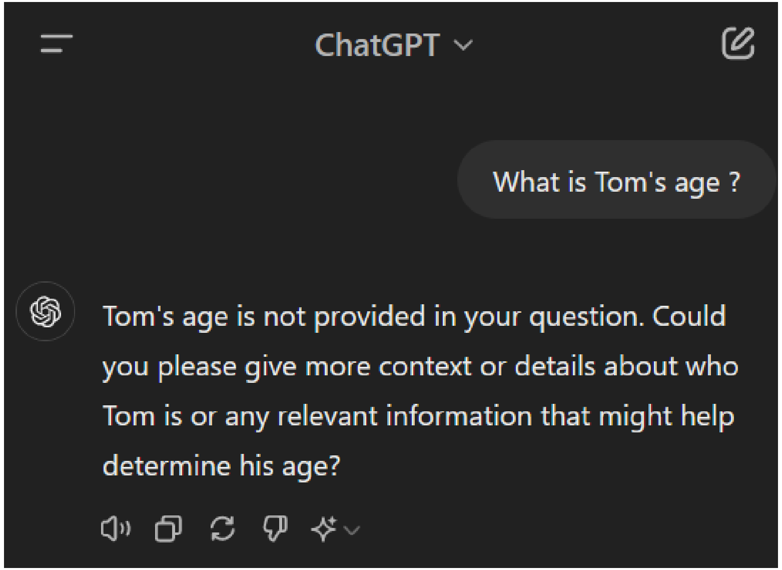
Solution #1: Finetune the model with business rules and knowledge.

- Advantage: high initial costs, **lower inference costs**
- Disadvantage: **expensive and difficult process** which can **still cause hallucinations** when answering questions.

Solution #2: Retrieval-Augmented Generation (RAG)

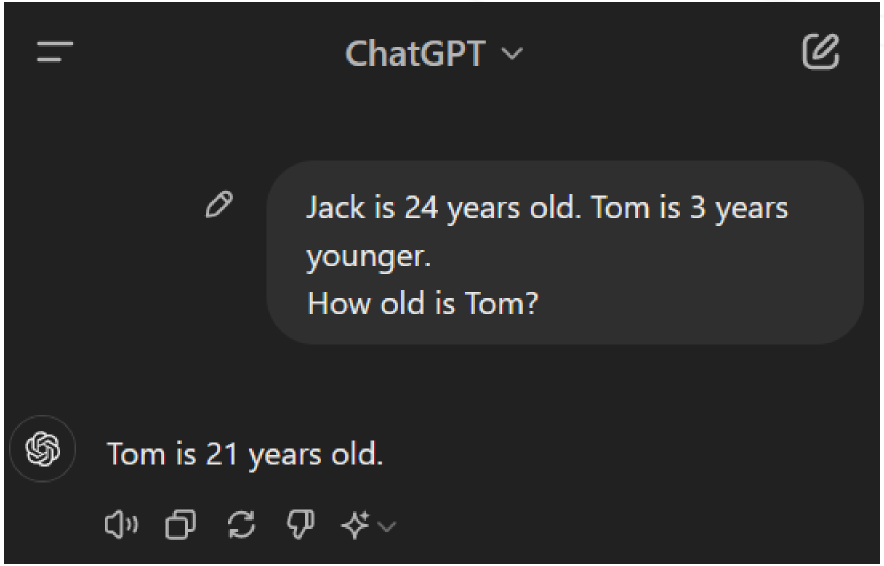
- Provide both the relevant business rules and the initial question to the user
- Advantage: **Documentation is all you need**.
- Disadvantage: **Higher inference costs** due to bigger context size.

The problem with foundational LLMs



What is Tom's age ?

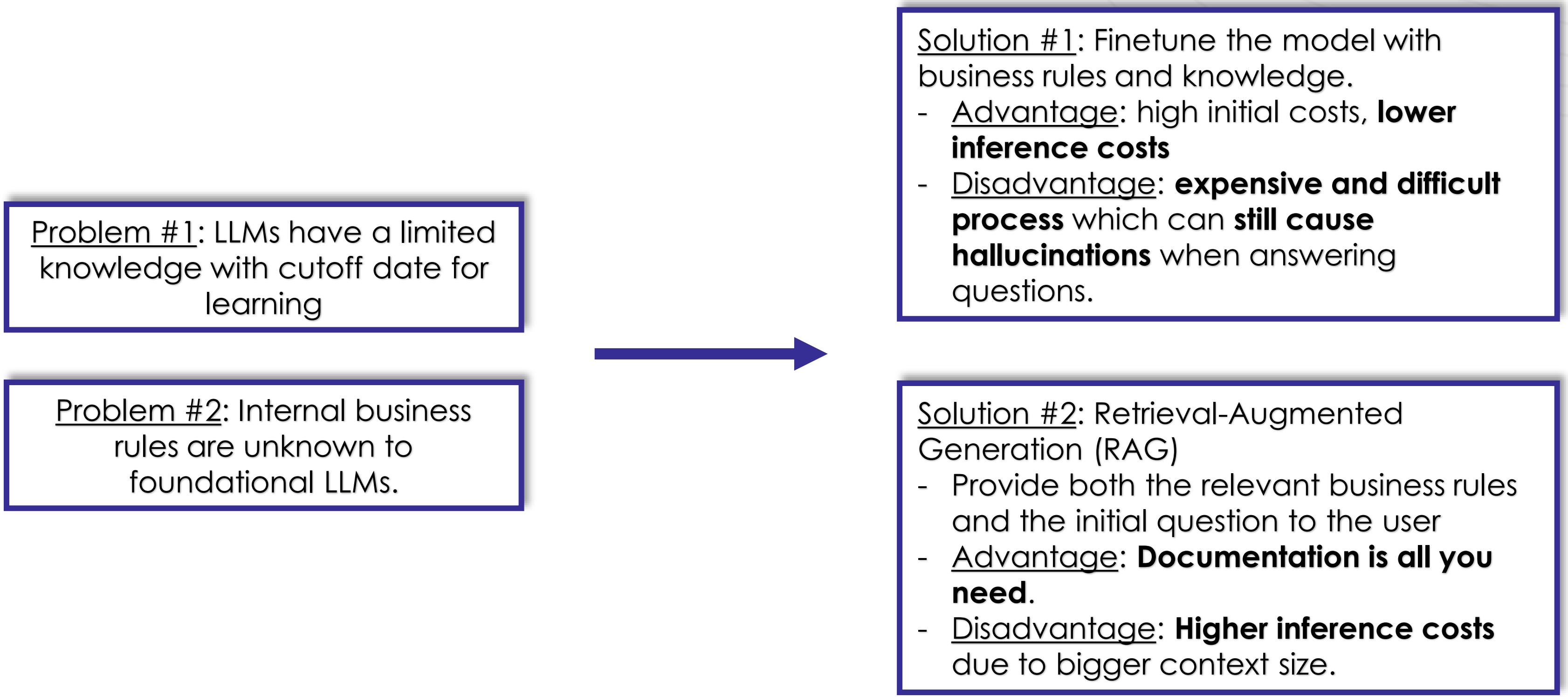
Tom's age is not provided in your question. Could you please give more context or details about who Tom is or any relevant information that might help determine his age?



Jack is 24 years old. Tom is 3 years younger.
How old is Tom?

Tom is 21 years old.

The problem with foundational LLMs



Problem #1: LLMs have a limited knowledge with cutoff date for learning

Problem #2: Internal business rules are unknown to foundational LLMs.

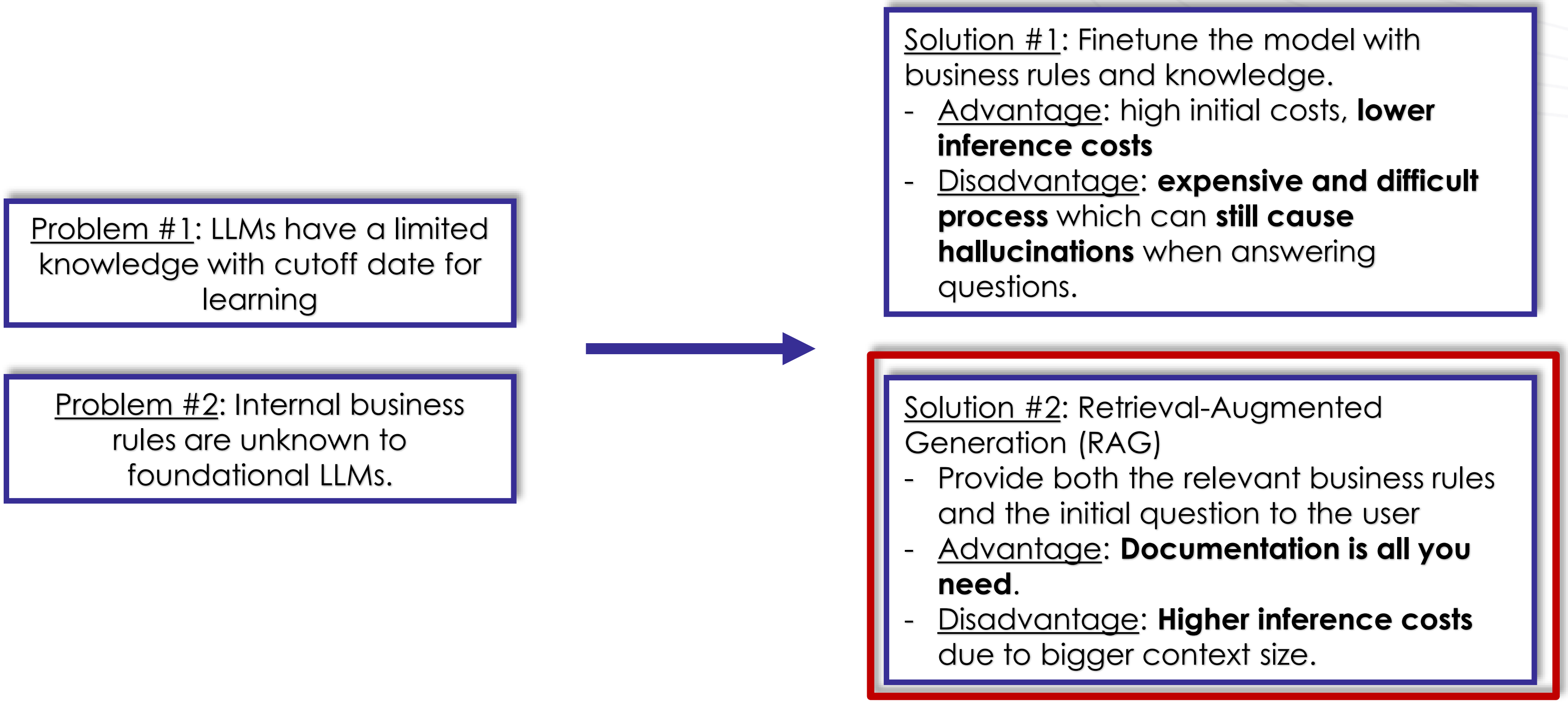
Solution #1: Finetune the model with business rules and knowledge.

- Advantage: high initial costs, **lower inference costs**
- Disadvantage: **expensive and difficult process** which can **still cause hallucinations** when answering questions.

Solution #2: Retrieval-Augmented Generation (RAG)

- Provide both the relevant business rules and the initial question to the user
- Advantage: **Documentation is all you need**.
- Disadvantage: **Higher inference costs** due to bigger context size.

The problem with foundational LLMs



Problem #1: LLMs have a limited knowledge with cutoff date for learning

Problem #2: Internal business rules are unknown to foundational LLMs.

Solution #1: Finetune the model with business rules and knowledge.

- Advantage: high initial costs, **lower inference costs**
- Disadvantage: **expensive and difficult process** which can **still cause hallucinations** when answering questions.

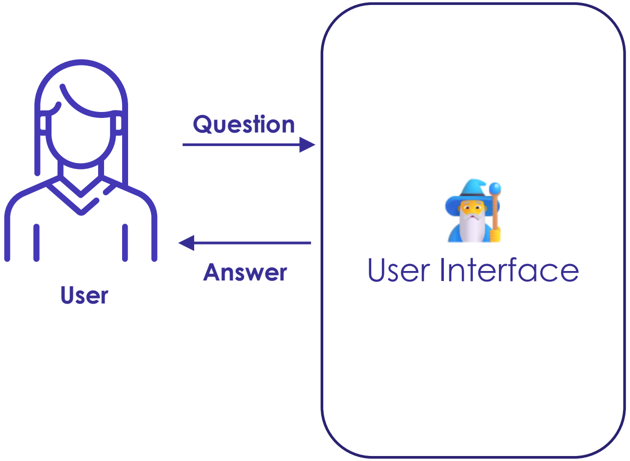
Solution #2: Retrieval-Augmented Generation (RAG)

- Provide both the relevant business rules and the initial question to the user
- Advantage: **Documentation is all you need**.
- Disadvantage: **Higher inference costs** due to bigger context size.

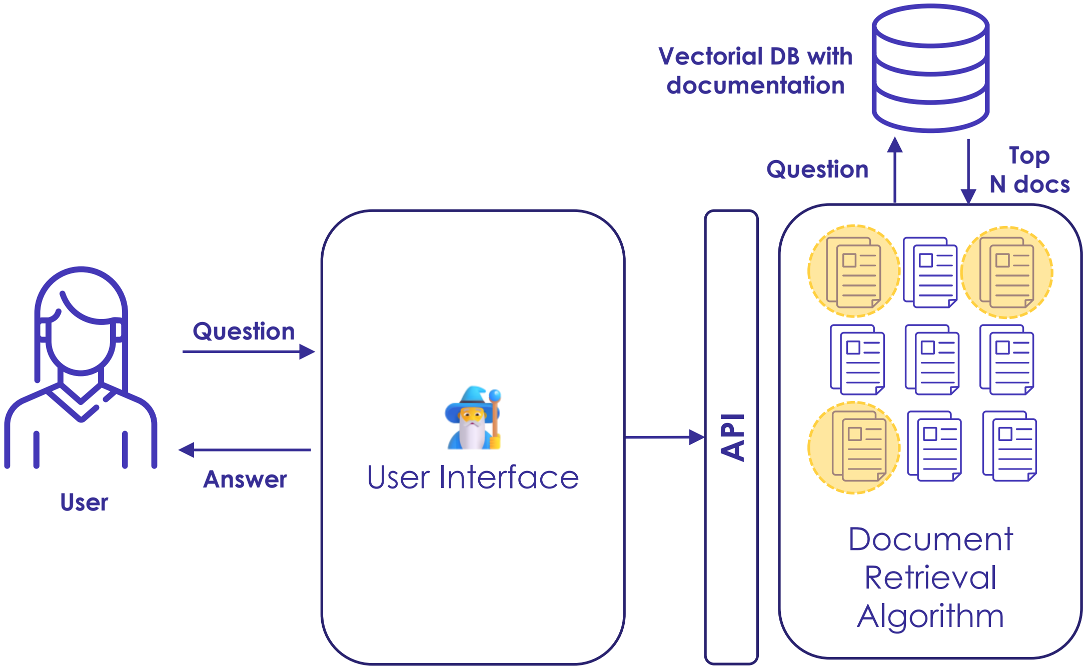
Retrieval Augmented Generation (RAG)

User

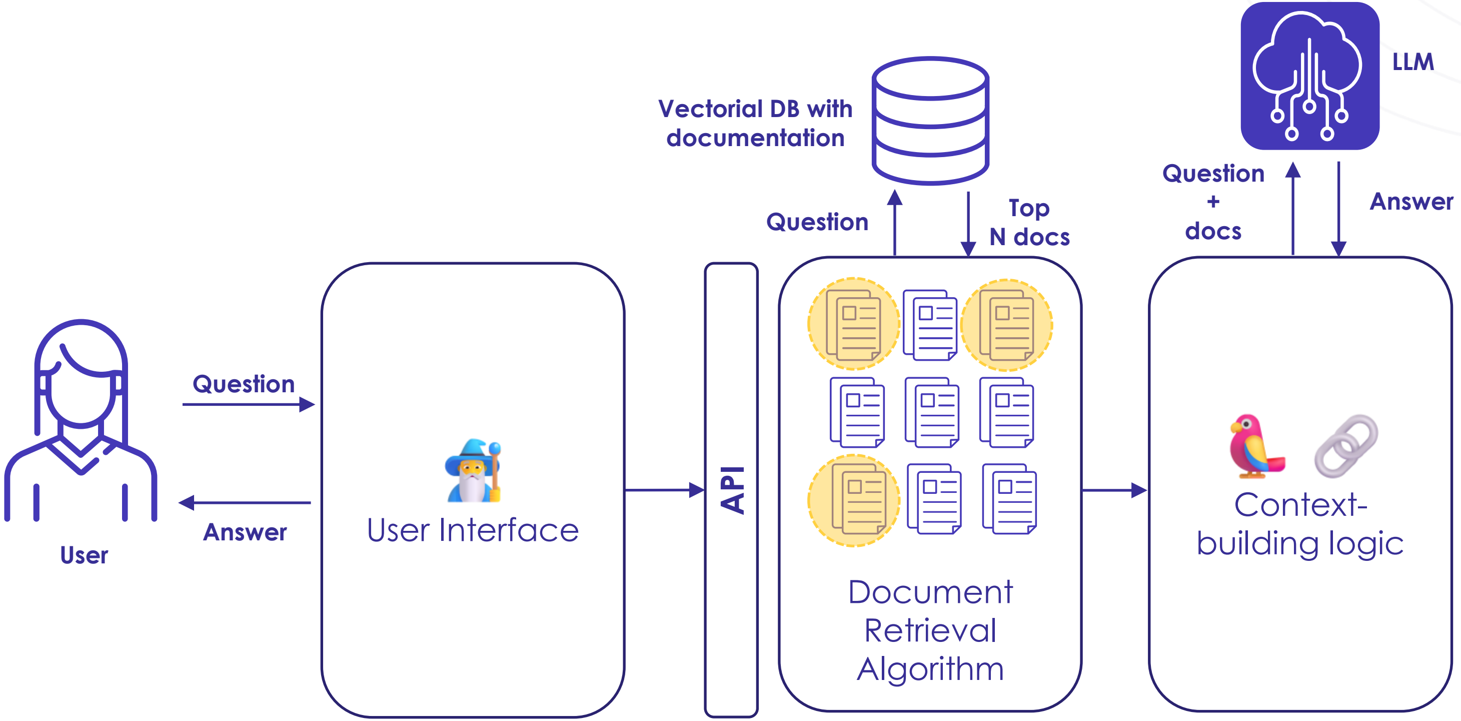
Retrieval Augmented Generation (RAG)



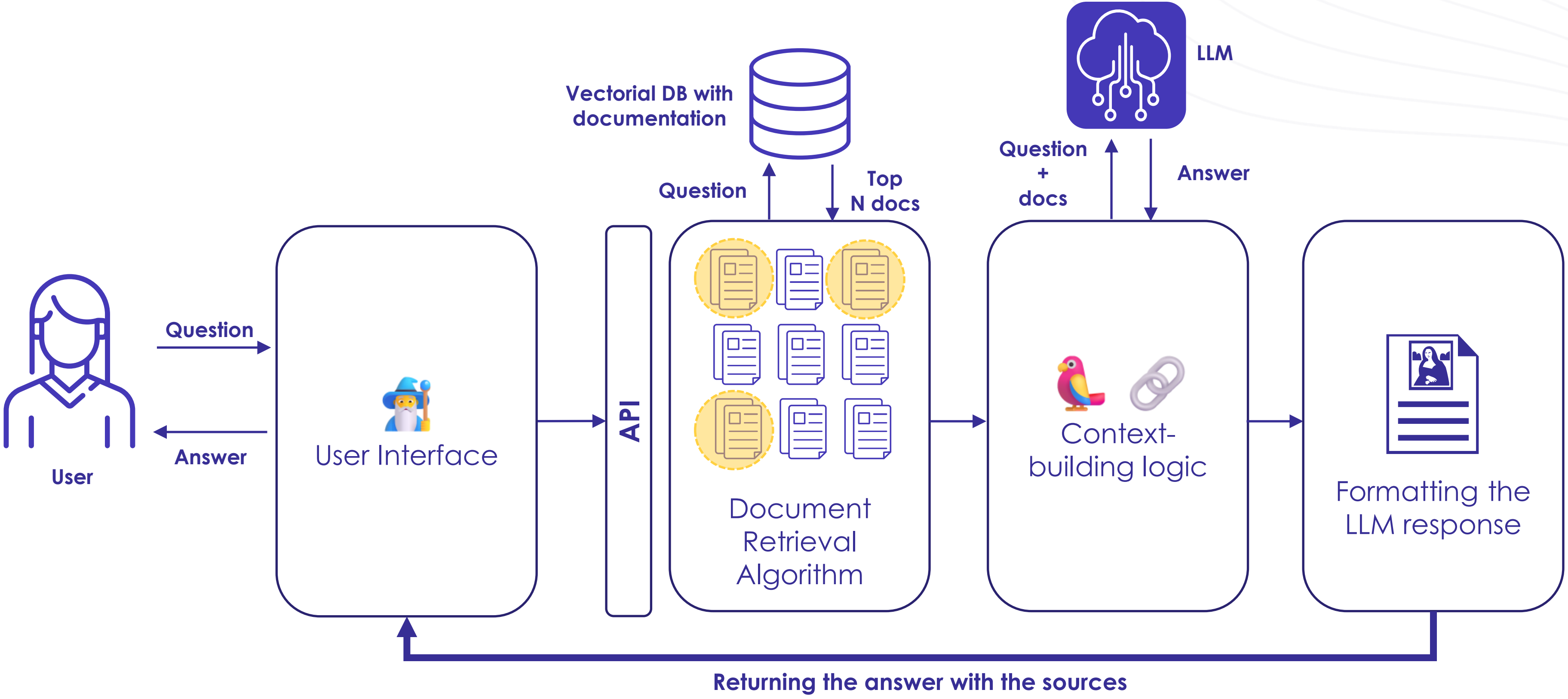
Retrieval Augmented Generation (RAG)



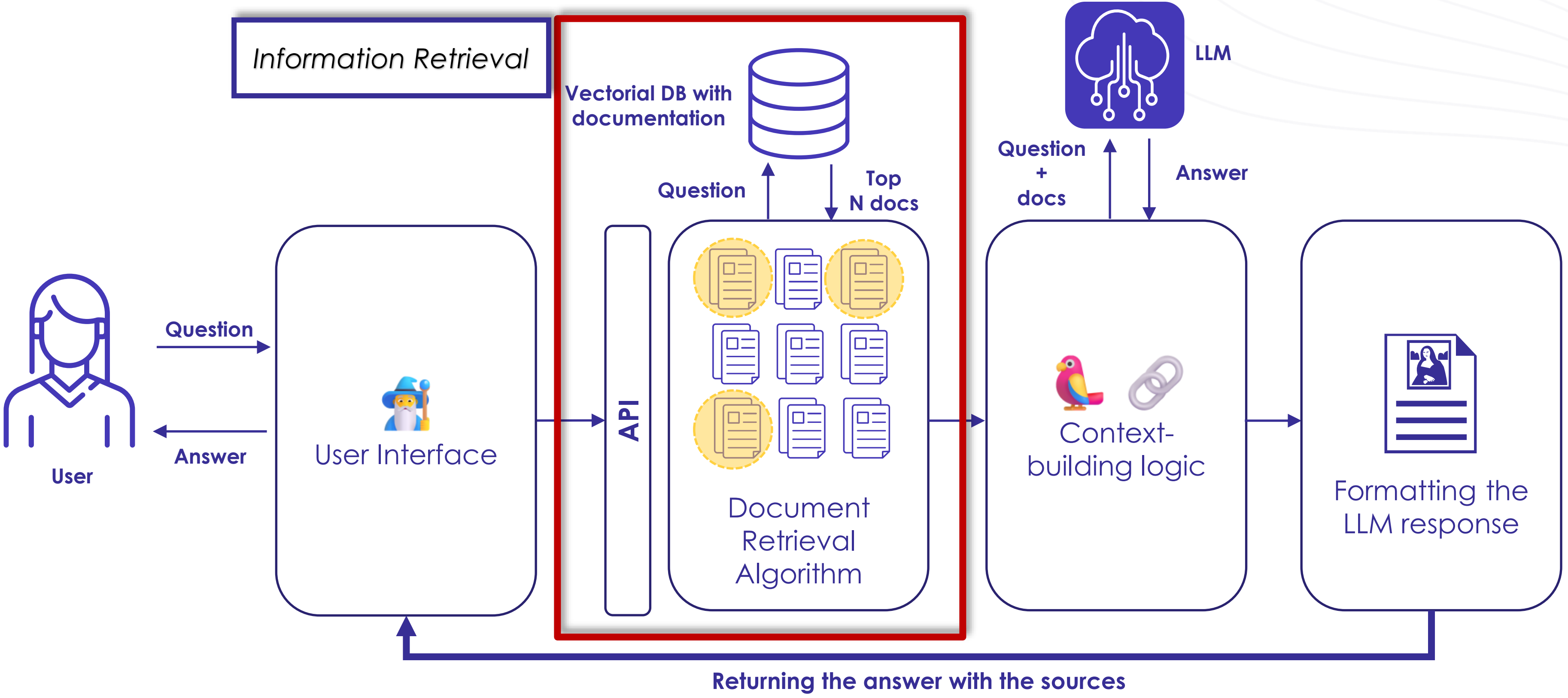
Retrieval Augmented Generation (RAG)



Retrieval Augmented Generation (RAG)

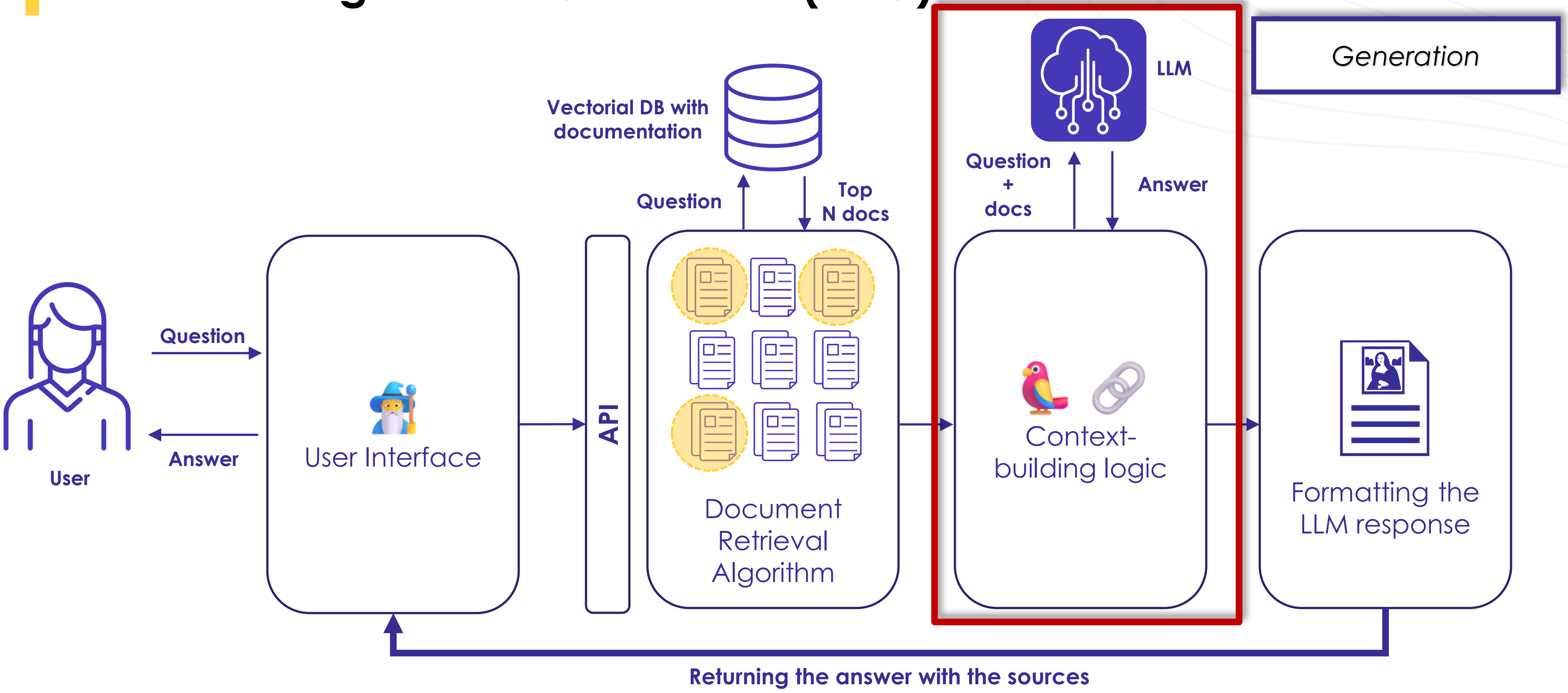

Vectorial DB with documentation
LLM
Question
Top N docs
Question + docs
Answer
Question
Answer
User
User Interface
API
Document Retrieval Algorithm
Context-building logic
Formatting the LLM response
Returning the answer with the sources

Retrieval Augmented Generation (RAG)


Information Retrieval
Vectorial DB with documentation
Question
Top N docs
LLM
Question + docs
Answer
Question
Answer
User
User Interface
API
Document Retrieval Algorithm
Context-building logic
Formatting the LLM response
Returning the answer with the sources

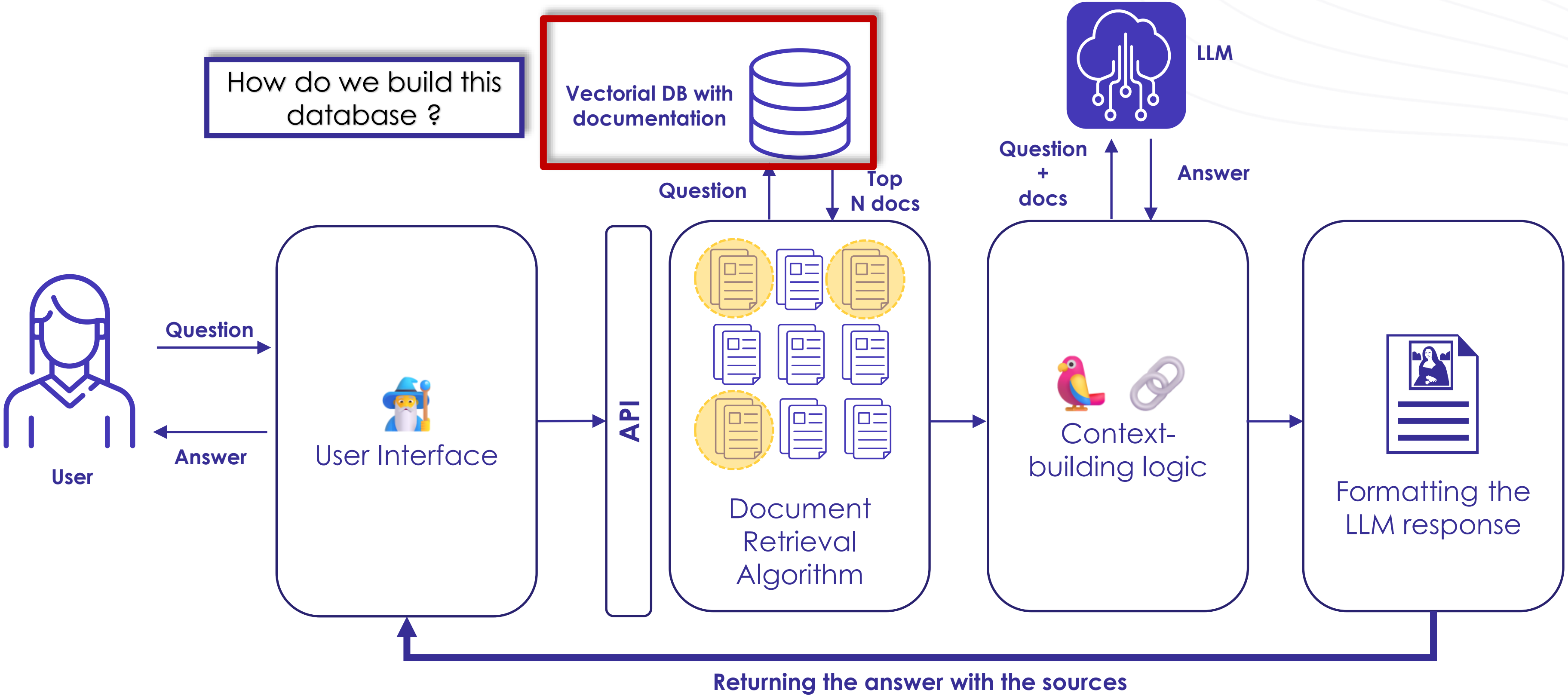
ESUG 2024

Retrieval Augmented Generation (RAG)


Generation
Vectorial DB with documentation
Question
Top N docs
LLM
Question + docs
Answer
User
Question
Answer
User Interface
API
Document Retrieval Algorithm
Context-building logic
Formatting the LLM response
Returning the answer with the sources

ESUG 2024

Retrieval Augmented Generation (RAG)



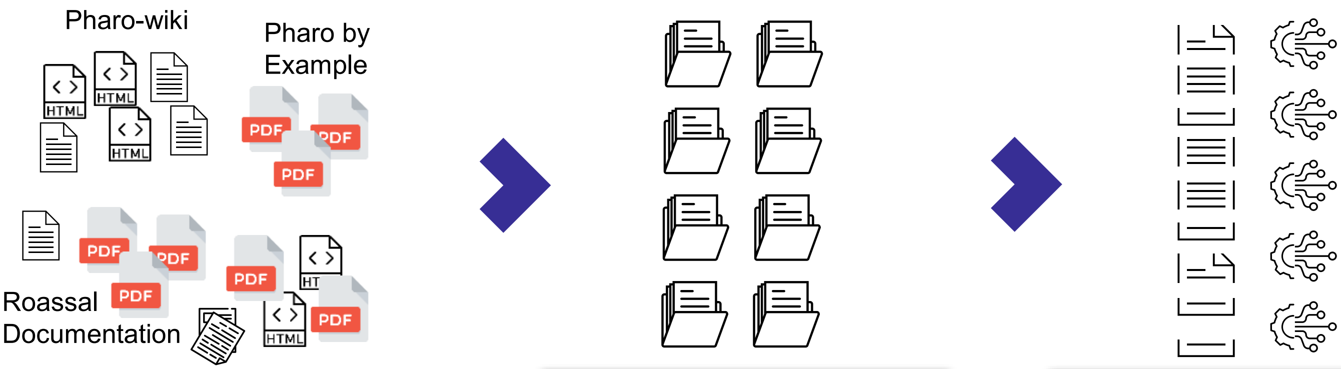
How do we build this database ?
Vectorial DB with documentation
LLM
Question
Top N docs
Question + docs
Answer
Question
Answer
User
User Interface
API
Document Retrieval Algorithm
Context-building logic
Formatting the LLM response
Returning the answer with the sources

Preparing the data source

Collect Pharo documentation

Clean and parse each document

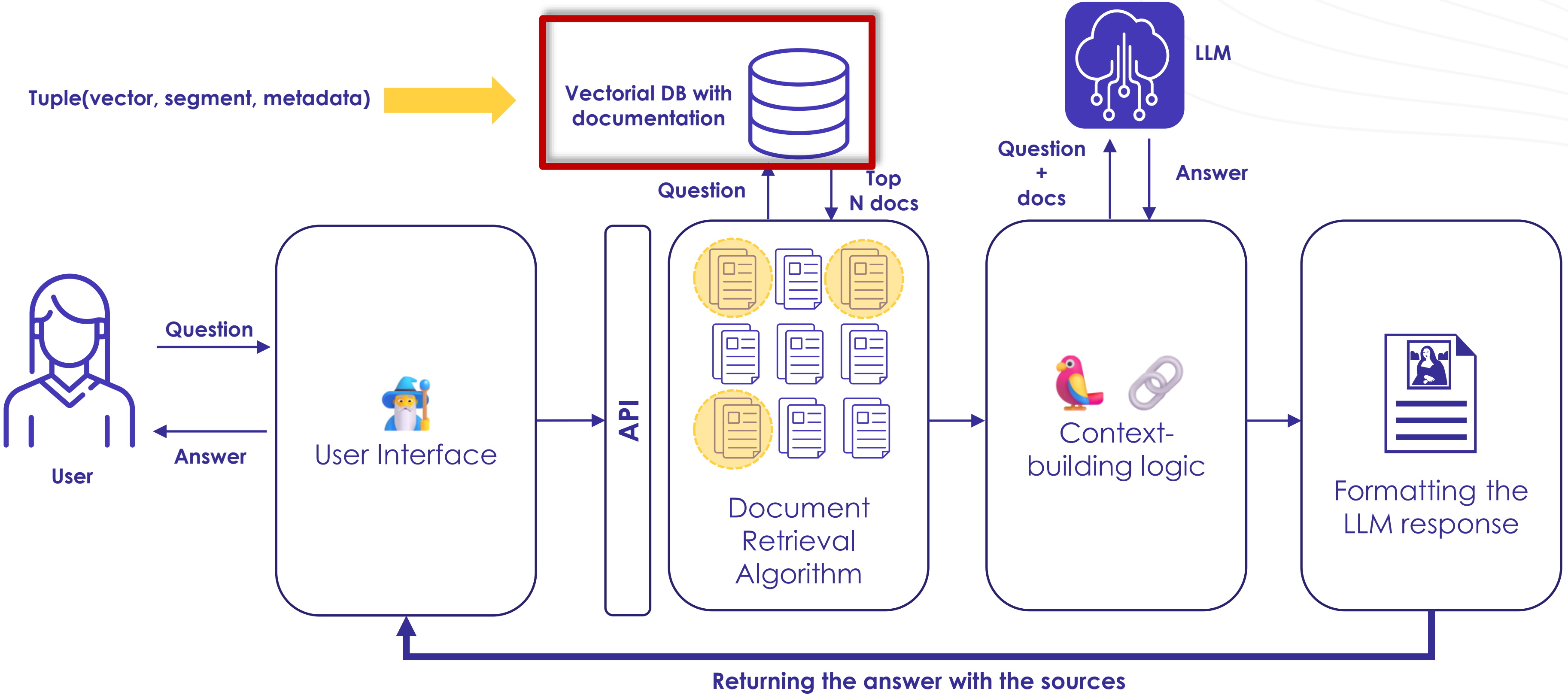
Segment document and embed them



Extracts the data as a readable text

Associate a N-dimension vector to the semantic value of each document

Retrieval Augmented Generation (RAG)


Tuple(vector, segment, metadata)
Vectorial DB with documentation
LLM
Question
Top N docs
Question + docs
Answer
User
Question
Answer
User Interface
API
Document Retrieval Algorithm
Context-building logic
Formatting the LLM response
Returning the answer with the sources

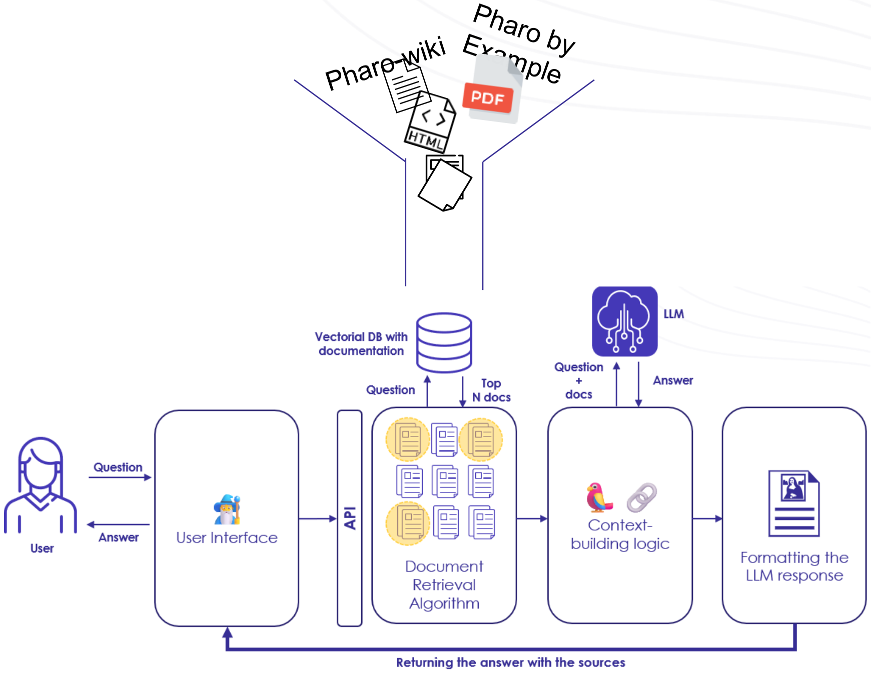
Tooling & Evaluation

ESUG 2024

Tooling

- Set out to create the **initial tooling in Python**

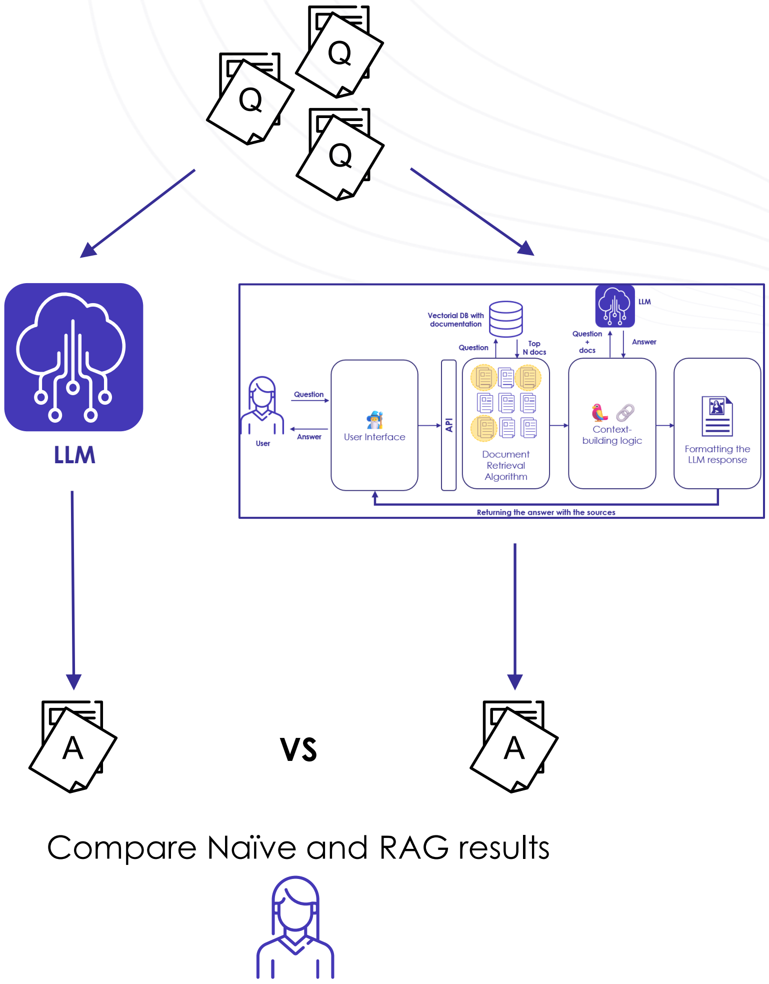
- **Identify** and **process** a set of Pharo documentation:
  - https://github.com/pharo-open-documentation/pharo-wiki.git
  - https://github.com/SquareBracketAssociates/PharoByExample9
  - https://github.com/SquareBracketAssociates/LearningOOPWithPharo
  - https://github.com/SquareBracketAssociates/BuildingApplicationWithSpec2
  - https://github.com/pharo-graphics/RoassalDocumentation.git



Onboarding tool demo

Evaluation protocol

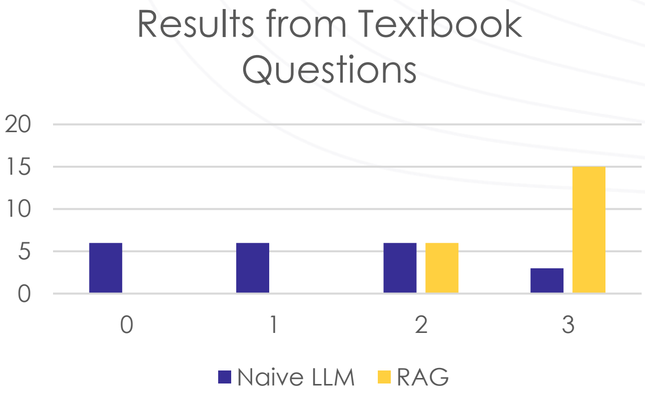
- Objective : **Compare** the LLM's ability to answer Pharo-related questions with a RAG-enabled LLM
  - Generate a **set of textbook questions** (21 questions)
  - Extract a **set of questions from Discord** and **Stack Overflow** (19 questions)
- Manual evaluation of each answer on a scale of 0 to 3
  - 0 : terrible
  - 1 : not ok
  - 2 : ok
  - 3 : perfect


Q
Q
Q
LLM
Vectorial DB with documentation
LLM
Question
Top N docs
Question + docs
Answer
Question
Answer
User
User Interface
API
Document Retrieval Algorithm
Context-building logic
Formatting the LLM response
Returning the answer with the sources
A
VS
A

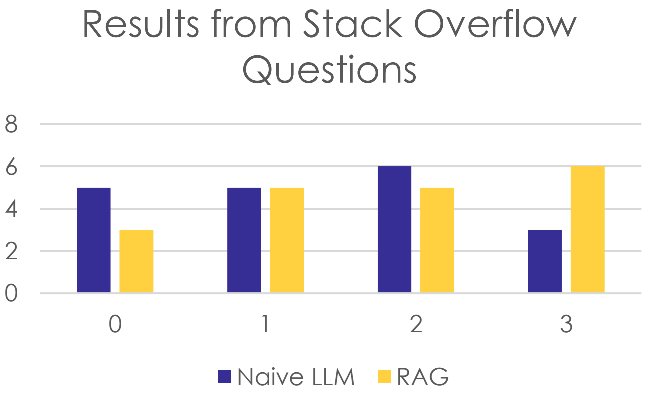
Compare Naïve and RAG results

Initial Results

- Textbook questions (21 questions):
  - When asking basic questions that exists in the documentation, the RAG technique gives near perfect answer
- Stack Overflow questions (19 questions):
  - Initial results are closer
  - There is a shift towards higher answer quality



Results from Textbook Questions

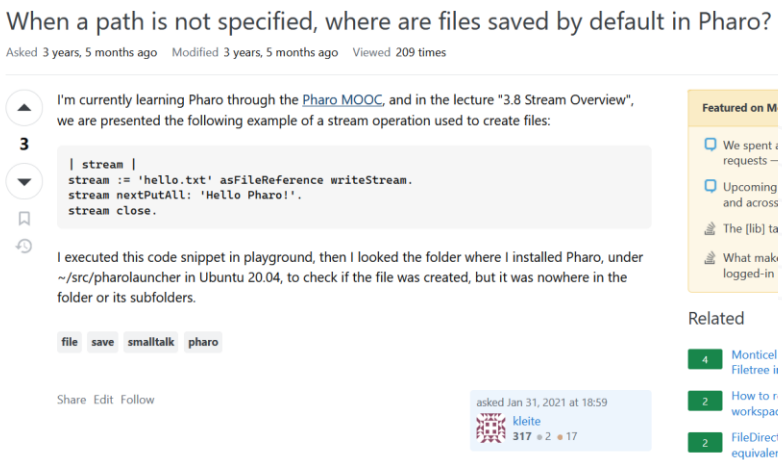


Results from Stack Overflow Questions

Conclusion & Perspectives

Conclusion

- Textbook-based results are extremely positive
- Stack Overflow-based results shows a lack of significant impact for a RAG-based onboarding LLM.
- Why ? What happened ?
  - Answers where the RAG-based technique was better, the documents provided with the questions did **contain an element of the answer needed**.
  - Answers where the RAG-based technique was **not** better, the documents provided with the question did **not** contain any elements of the answer required.
- What could this mean ?
  - Not enough documents to answer some of these questions
  - the question is too complex to answer with a simple request



When a path is not specified, where are files saved by default in Pharo?

Asked 3 years, 5 months ago Modified 3 years, 5 months ago Viewed 209 times

3

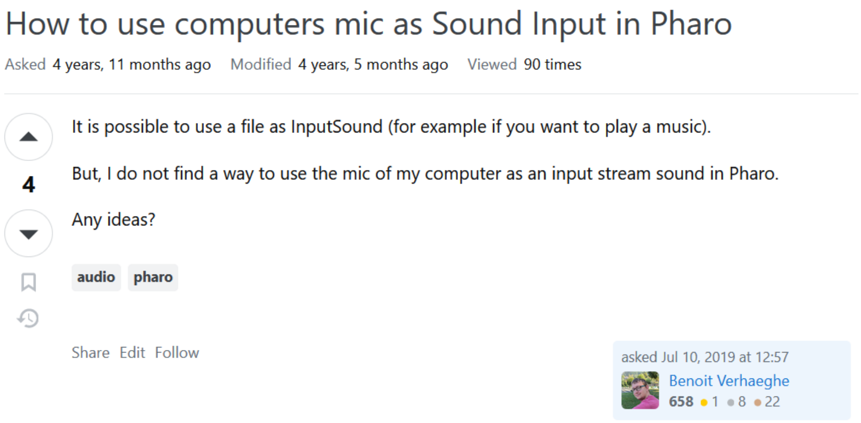
I'm currently learning Pharo through the Pharo MOOC, and in the lecture "3.8 Stream Overview", we are presented the following example of a stream operation used to create files:

```
| stream |
stream := 'hello.txt' asFileReference writeStream.
stream nextPutAll: 'Hello Pharo!'.
stream close.
```

I executed this code snippet in playground, then I looked the folder where I installed Pharo, under ~/src/pharolauncher in Ubuntu 20.04, to check if the file was created, but it was nowhere in the folder or its subfolders.

file save smalltalk pharo

asked Jan 31, 2021 at 18:59
kleite
317 ● 2 ● 17



How to use computers mic as Sound Input in Pharo

Asked 4 years, 11 months ago Modified 4 years, 5 months ago Viewed 90 times

4

It is possible to use a file as InputSound (for example if you want to play a music).

But, I do not find a way to use the mic of my computer as an input stream sound in Pharo.

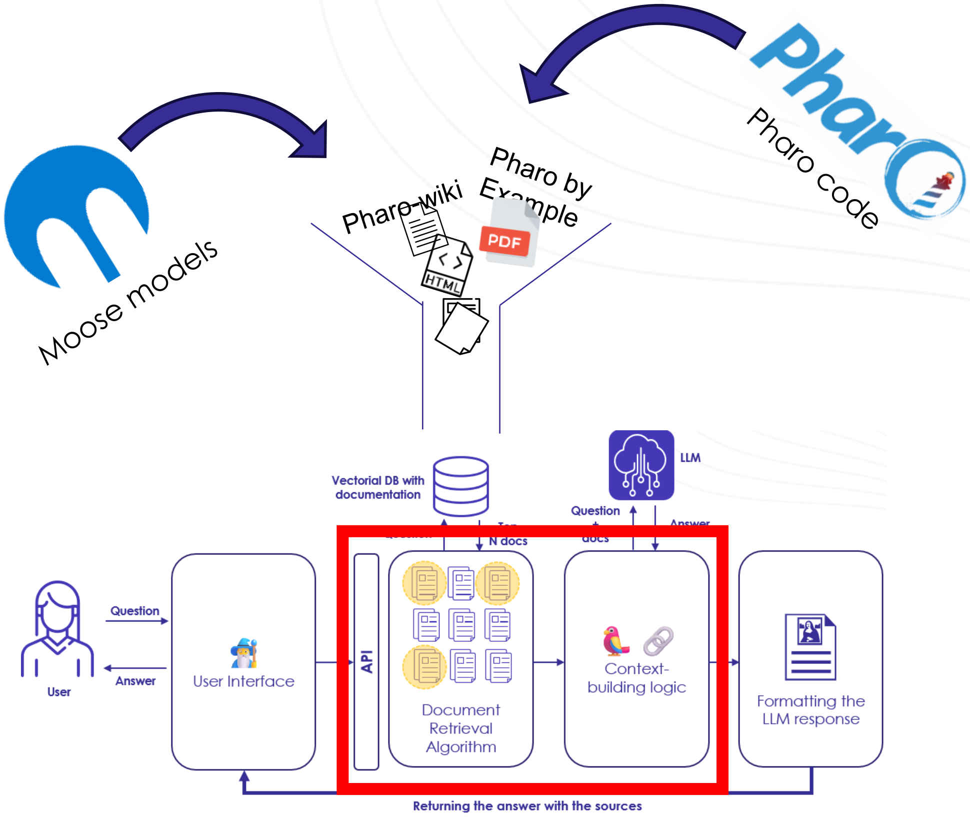
Any ideas?

audio pharo

asked Jul 10, 2019 at 12:57
Benoit Verhaeghe
658 ● 1 ● 8 ● 22

Perspectives

- Increase the amount of documentation we parse.
- Include **other data sources** (e.g., moose models, Pharo code).
  - Allows use to ask questions about a specific code base (e.g., what is the super class of …)
- Use **agent-based** programming to **iterate** over more complex questions.
- User experimentation (launching experimentation within the company so that our users can try the tool -> **collect Q&A with evaluation**)



onboarding.pharo.research-bl.com

Thank you for your attention.

Any Questions?

pascal.zaragoza@berger-levrault.com
marius.pingaud@berger-levrault.com