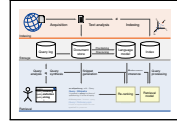
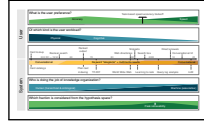


# The Evolution of Web Search

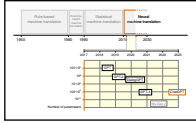
Martin Potthast  
University of Kassel,  
hessian.AI, and ScaDS.AI

November 19th, 2024

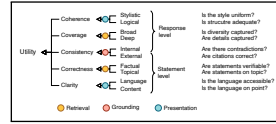
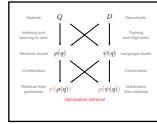
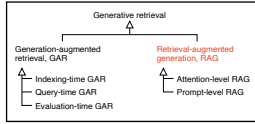
# Web Search



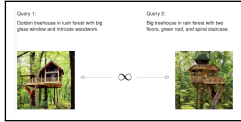
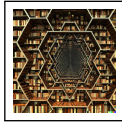
# Language Models



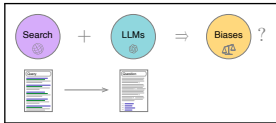
# Retrieval-Augmented Generation



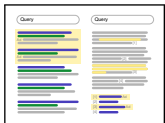
# The Infinite Index



# Biases



# Native Advertising



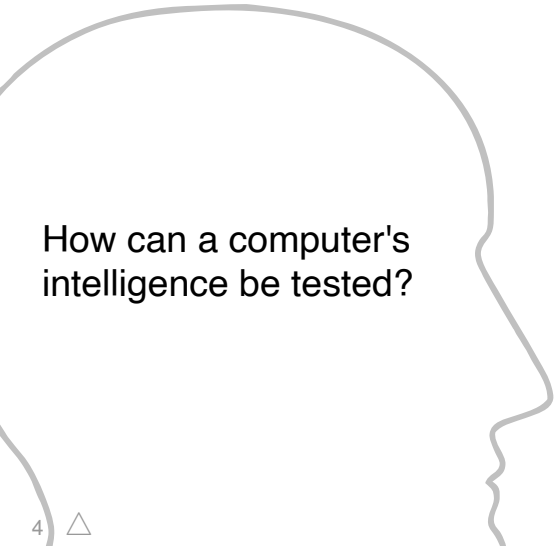




# A Short History of Search Engines

# Information Retrieval in a Nutshell

- ❑ **A vague request.**  
Expression of a complex information need: a question
- ❑ **Billions of documents.**  
Text, images, audio files, videos, ...

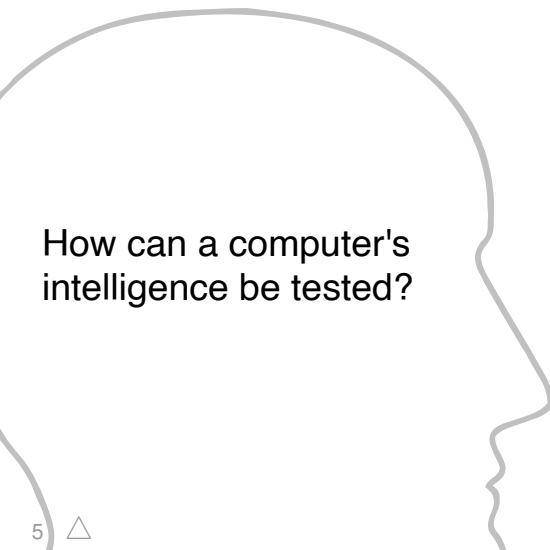


How can a computer's intelligence be tested?

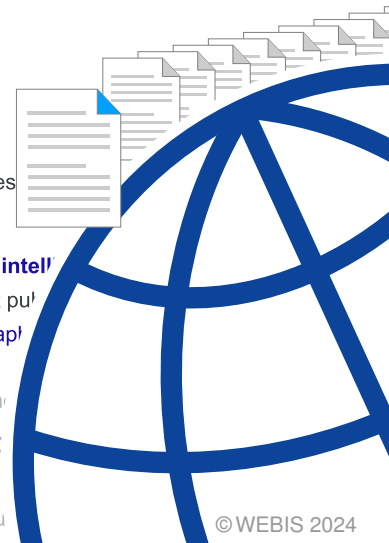


# Information Retrieval in a Nutshell

- ❑ A vague request.  
Expression of a complex information need: a question, or just a few keywords.
  - ❑ Billions of documents.  
Text, images, audio files, videos, ...
  - ❑ High class imbalance.  
Only a tiny fraction of all documents are relevant to the request.
- Retrieve relevant documents in milliseconds.



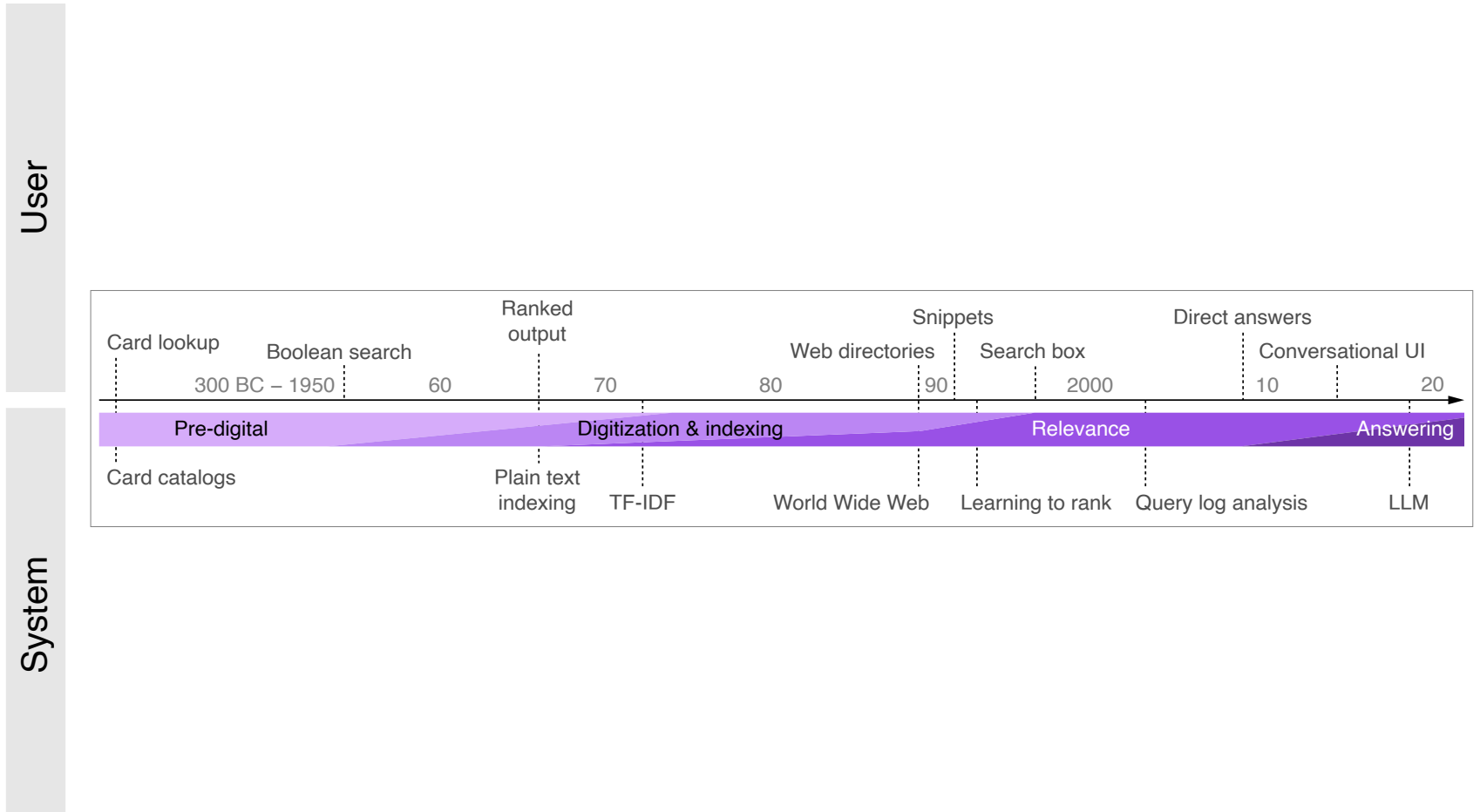
Computer  
Intelligence  
Test



en.wikipedia.org › wiki › Turing\_test  
**Turing test - Wikipedia**  
Jump to **Human intelligence vs. artificial intelligence** (1950) was the first published Turing test (disambiguation) · Graph

plato.stanford.edu › entries › turing\_test  
**The Turing Test (Stanford Encyclopedia of Philosophy)**  
by G Oppy - 2003 - Cited by 130  
Apr 9, 2003 - The phrase "The Turing Test" was first used by Alan Turing in his 1950 paper "Computing Machinery and Intelligence".

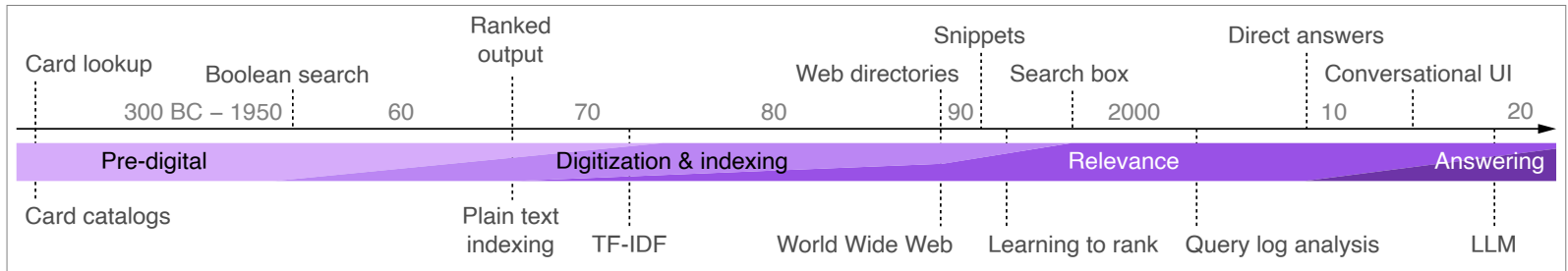
# A Short History of Search Engines



• M. Potthast, M. Hagen, B. Stein (2020). [The dilemma of the direct answer.](#)

# A Short History of Search Engines

User



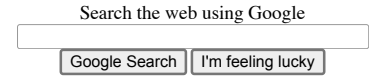
System



- M. Potthast, M. Hagen, B. Stein (2020). [The dilemma of the direct answer.](#)

# A Short History of Search Engines

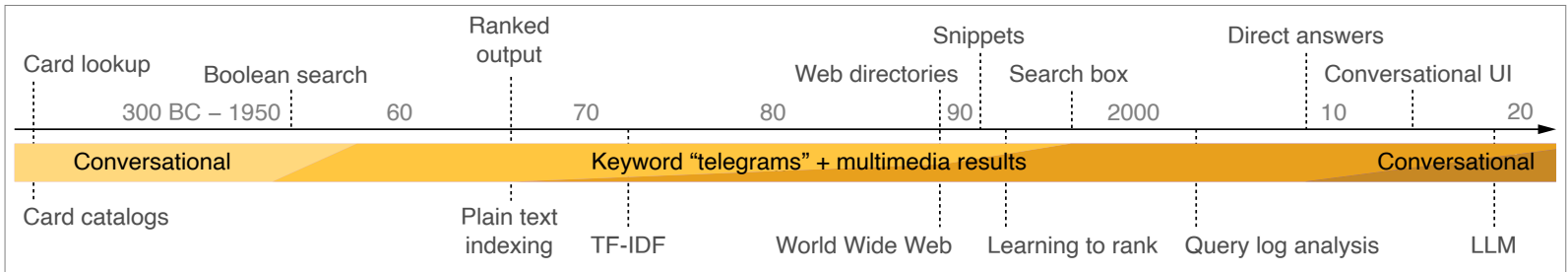
User



Yellow Pages - People Search - City Maps -- News Headlines - Stock Quotes - Sports Scores

- **Arts** - - Humanities, Photography, Architecture, ...
- **Business and Economy [Xtra!]** - - Directory, Investments, Classifieds, ...
- **Computers and Internet [Xtra!]** - - Internet, WWW, Software, Multimedia, ...
- **Education** - - Universities, K-12, Courses, ...
- **Entertainment [Xtra!]** - - TV, Movies, Music, Magazines, ...

©1999 Google Inc.



System

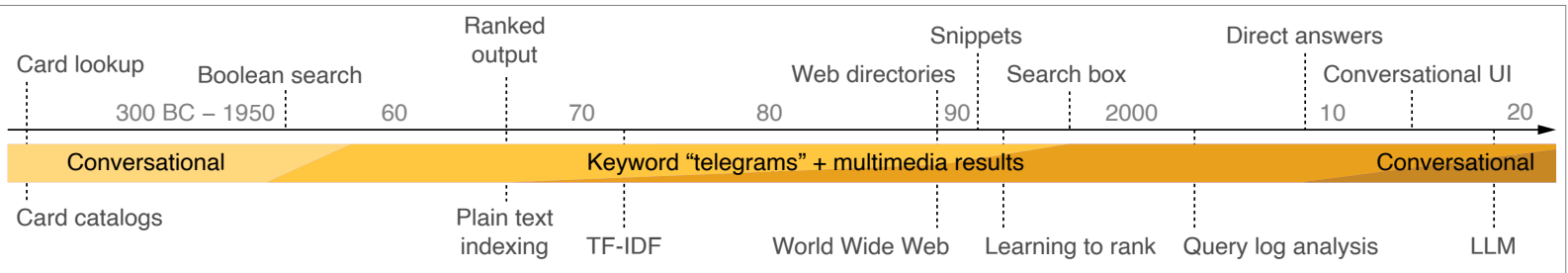


• M. Potthast, M. Hagen, B. Stein (2020). [The dilemma of the direct answer.](#)

# A Short History of Search Engines

User

Of which kind is the user workload?



System

Who is doing the job of knowledge organization?



- M. Potthast, M. Hagen, B. Stein (2020). [The dilemma of the direct answer.](#)

# A Short History of Search Engines

User

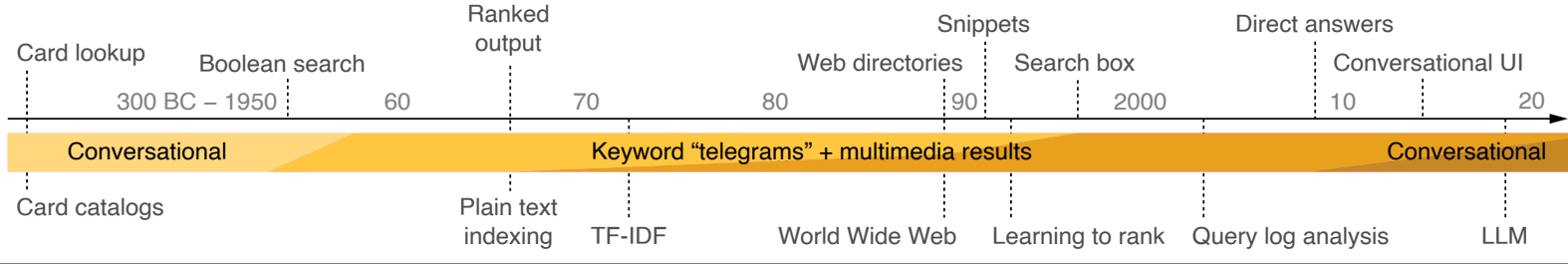
What is the user preference?



Of which kind is the user workload?



System



Who is doing the job of knowledge organization?



Which fraction is considered from the hypothesis space?



• M. Potthast, M. Hagen, B. Stein (2020). [The dilemma of the direct answer.](#)

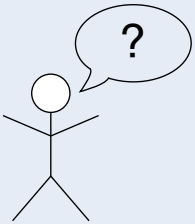


# Web Search Architecture



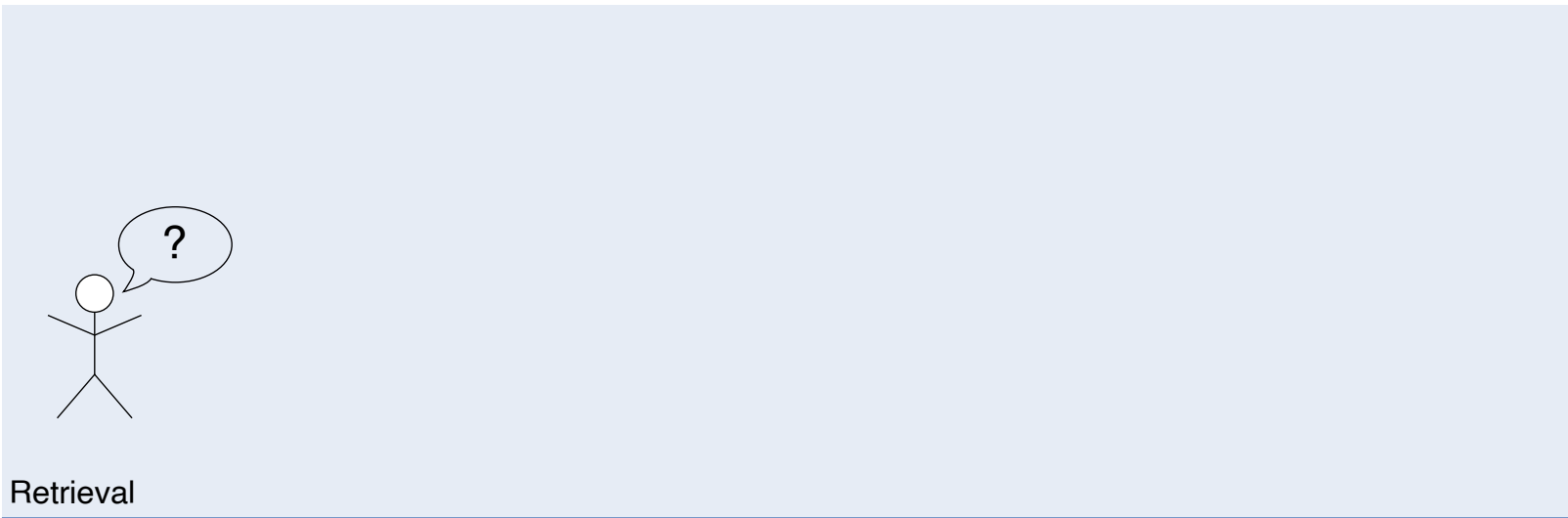
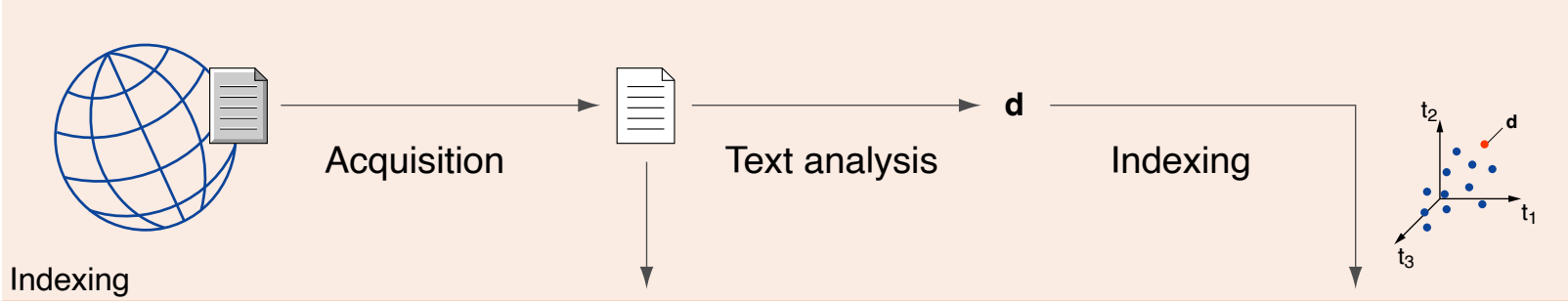
Indexing

Storage

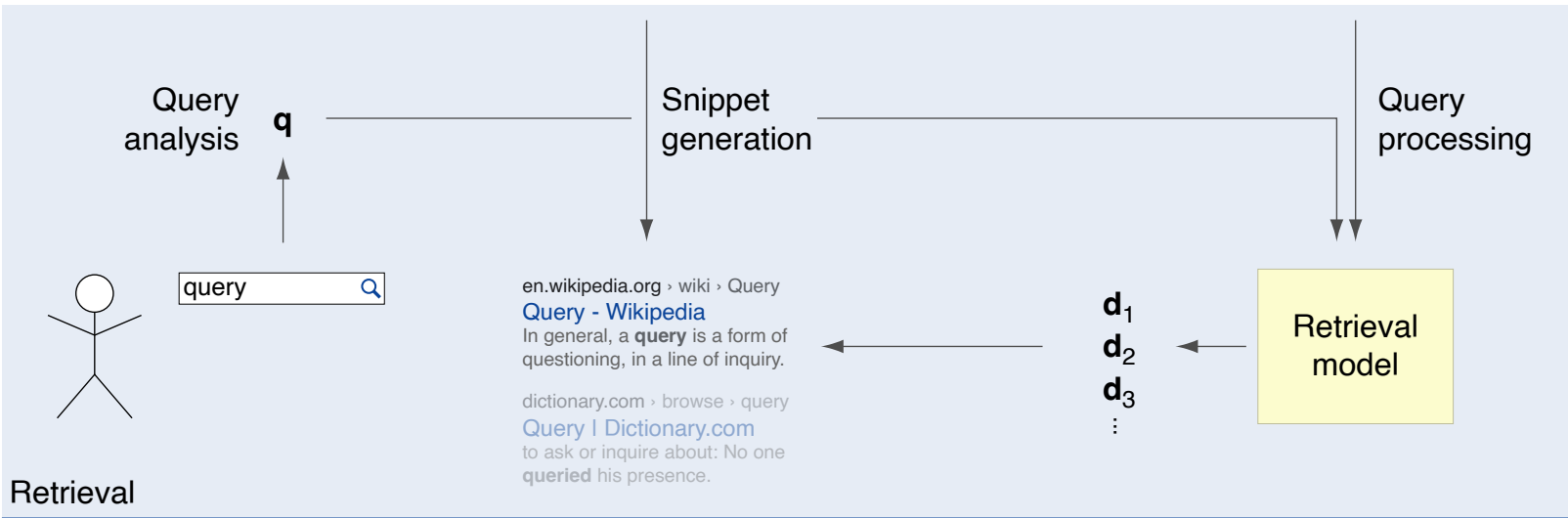
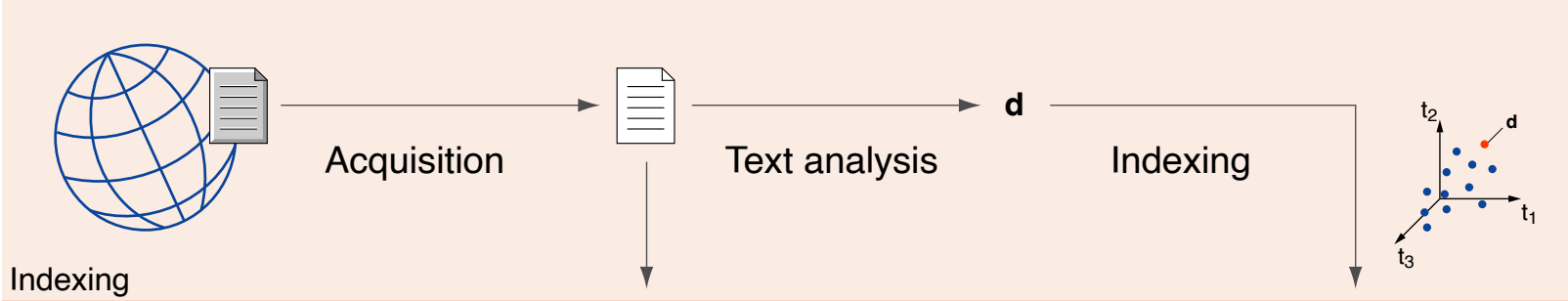


Retrieval

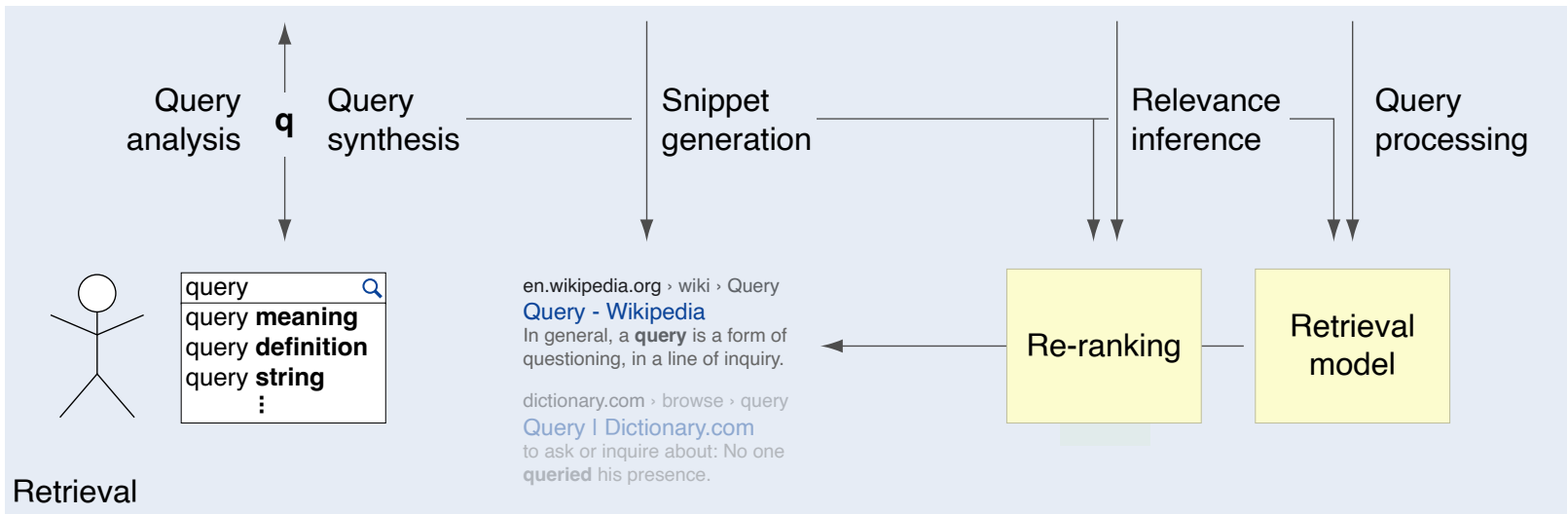
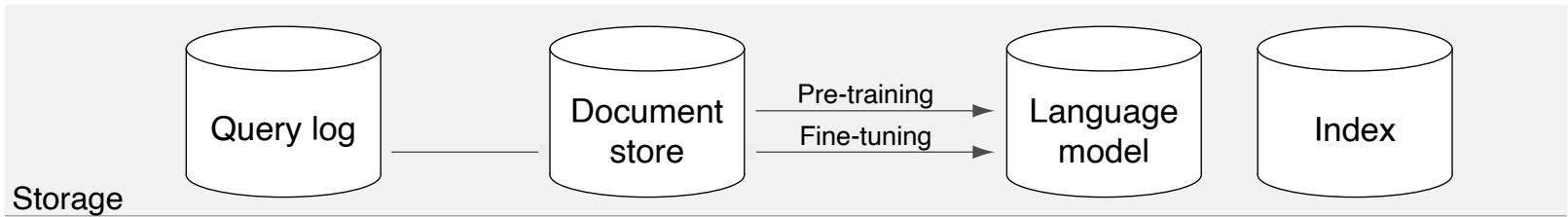
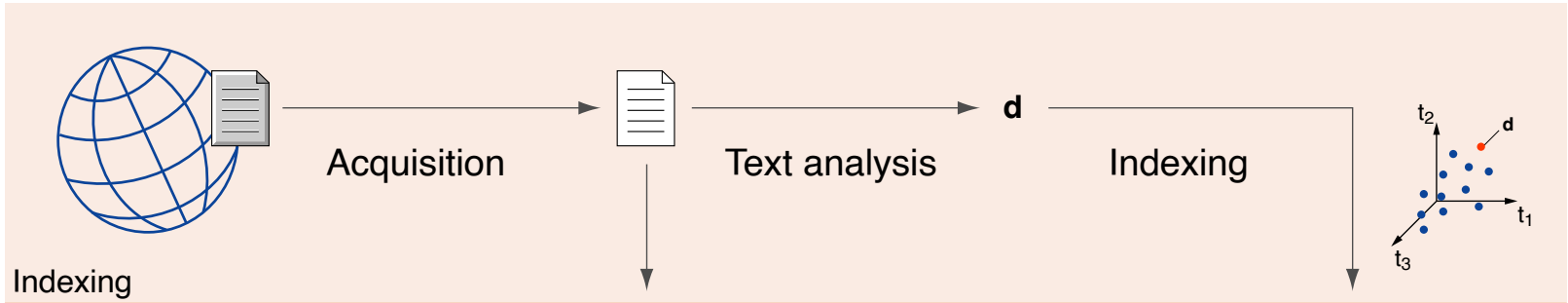
# Web Search Architecture



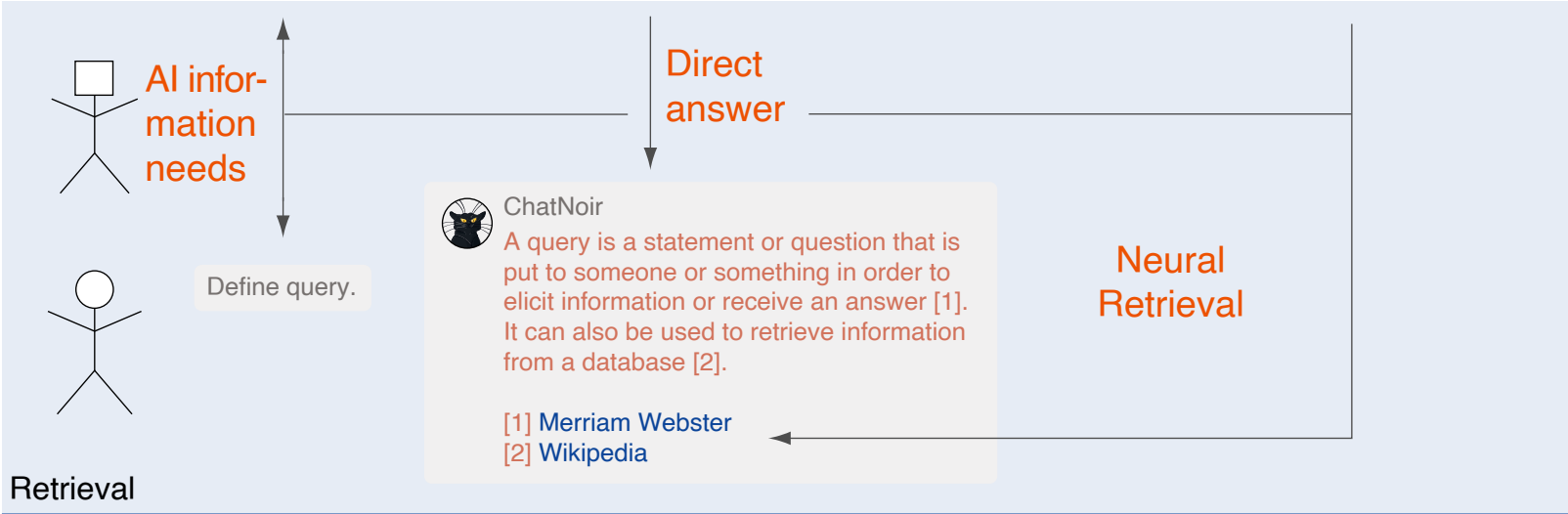
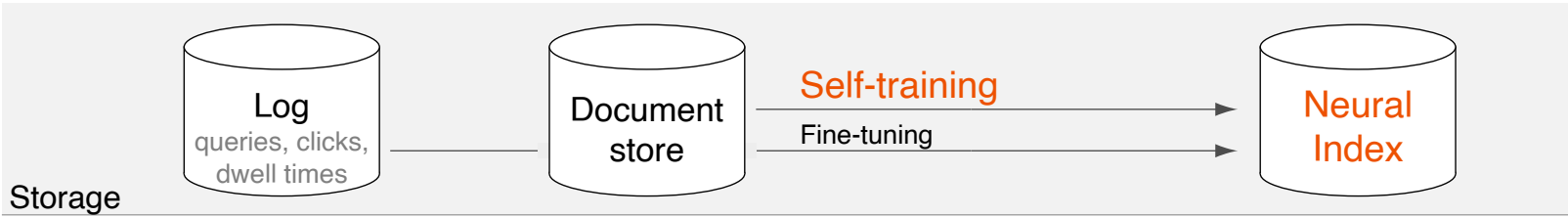
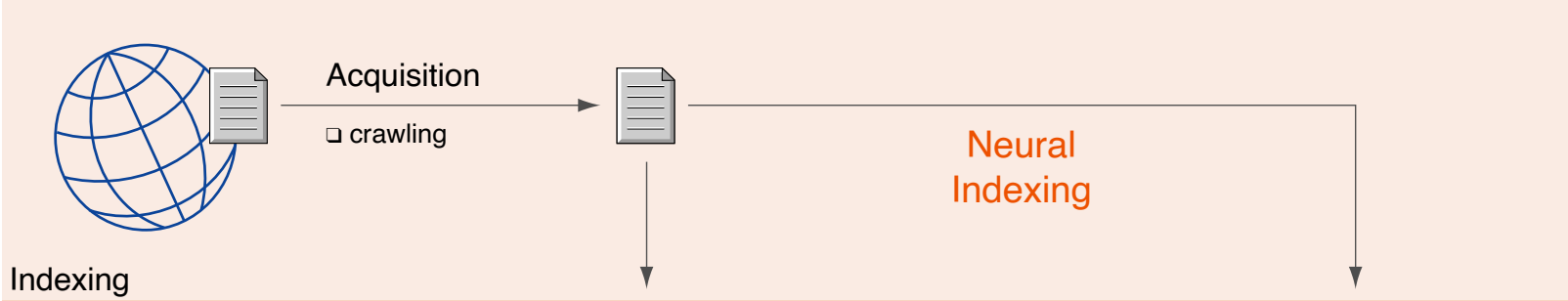
# Web Search Architecture



# Web Search Architecture



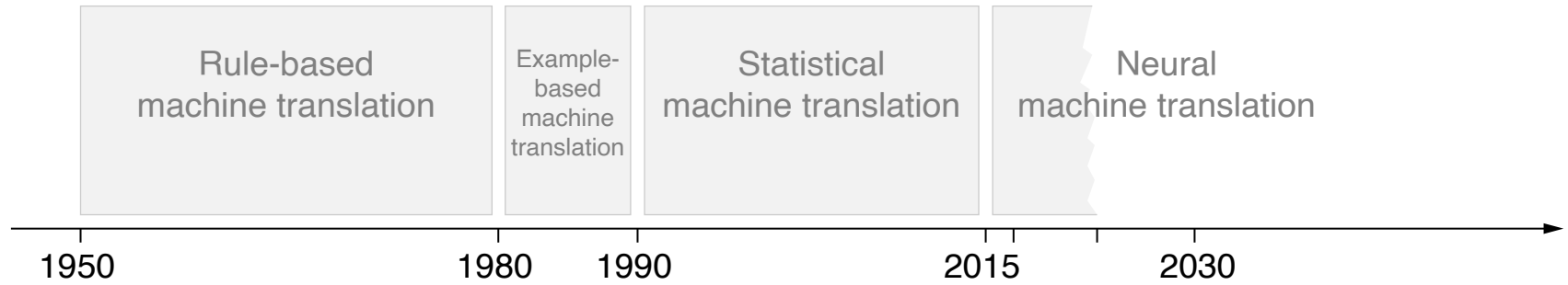
# Web Search Architecture



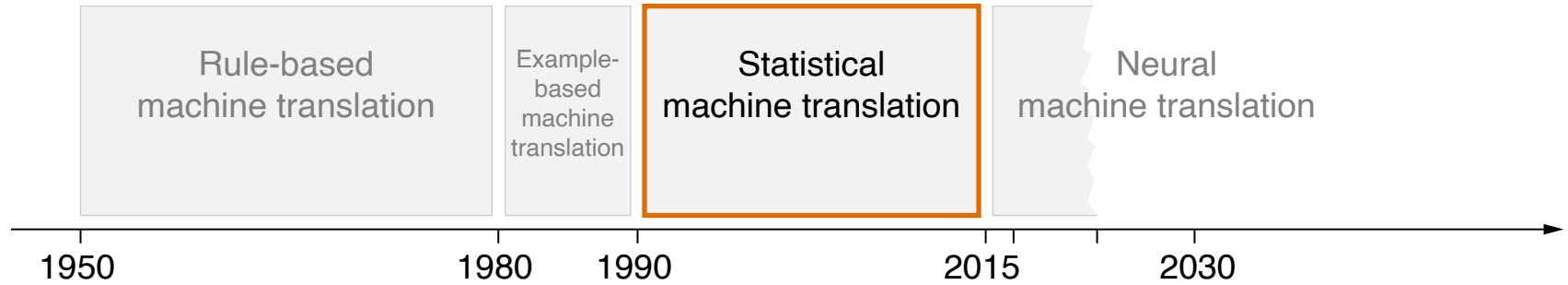


## A Short History of Language Models

# A Short History of Language Models



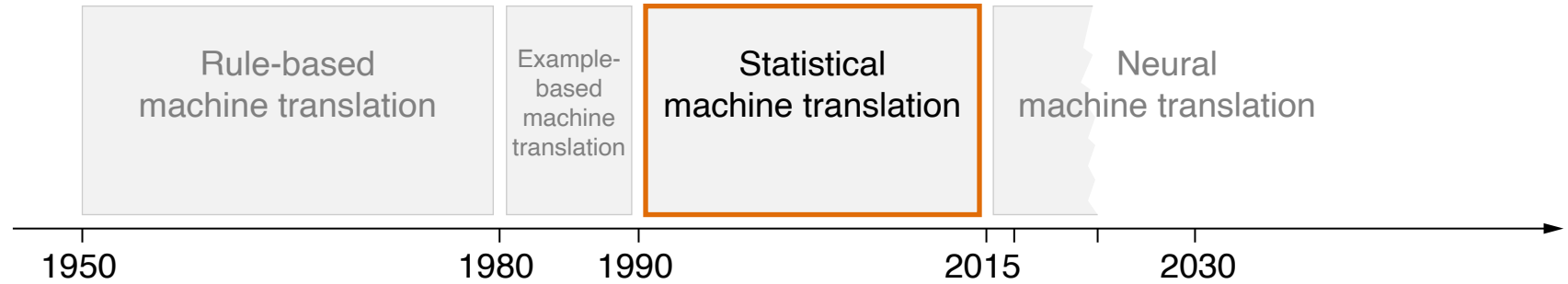
# A Short History of Language Models



A statistical language model  
is a probability distribution over all possible texts.



# A Short History of Language Models



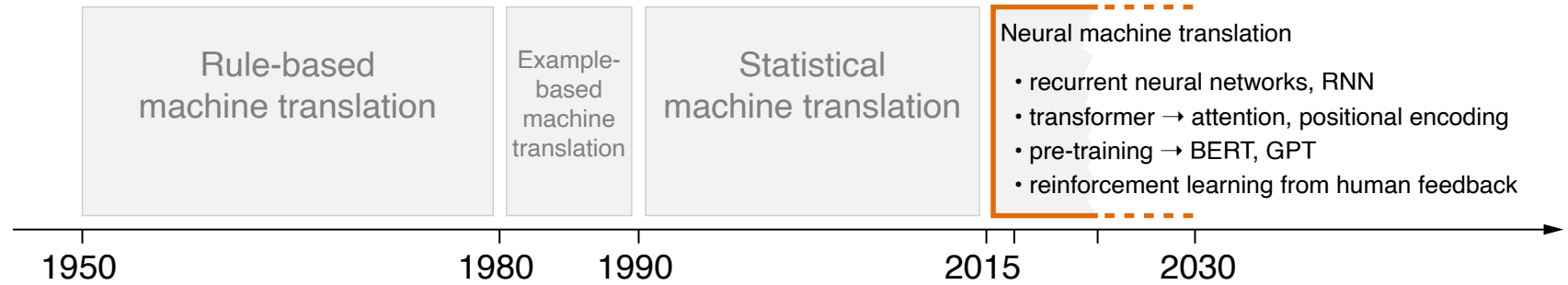
A statistical language model  
is a probability distribution over all possible texts.

Illustration:

(1) i love my ? N N

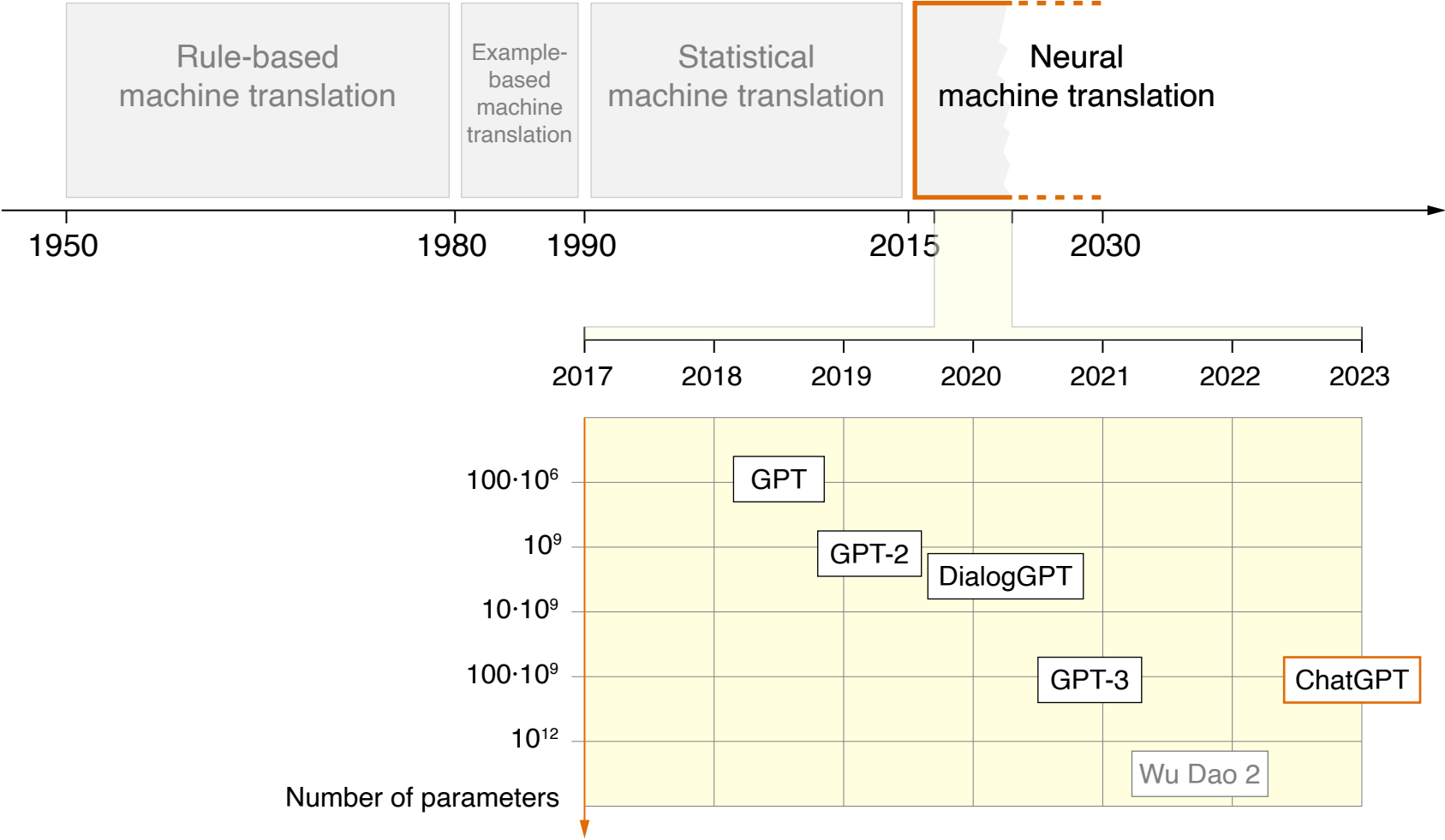
(2) see ... works. N N

# A Short History of Language Models



A neural language model  
approximates a statistical language model.

# A Short History of Language Models



### Training Corpora Sources

Wikipedia	11GB	Books	21GB
Journals	101GB	Reddit	50GB
Common Crawl	570GB		

### Parameters

175,000,000,000  
( $175 \cdot 10^9$ )

### Computing / Training

- 355 years on a single Tesla V100 GPU.
- $\approx$  34 days on 1,024 x A100 GPUs.
- \$4.6M costs a single training run.

**GPT-3** [Jun. 2020]

### Training Corpora Sources

Wikipedia	11GB	Books	21GB
Journals	101GB	Reddit	50GB
Common Crawl	570GB		

### Parameters

175,000,000,000  
( $175 \cdot 10^9$ )

### Computing / Training

- 355 years on a single Tesla V100 GPU.
- $\approx$  34 days on 1,024 x A100 GPUs.
- \$4.6M costs a single training run.

## GPT-3 [Jun. 2020]

### World Knowledge

- 🧑 What city is in the northwest corner of Ohio?
- 🌀 Toledo is in the northwest corner of Ohio.

### Common Sense

- 🧑 Why don't animals have three legs?
- 🌀 Animals don't have three legs because they would fall over.

### Logical Reasoning

- 🧑 If I put a pencil in a box, then put another pencil in the box, what is in the box?
- 🌀 Two pencils.

### Training Corpora Sources

Wikipedia	11GB	Books	21GB
Journals	101GB	Reddit	50GB
Common Crawl	570GB		

### Parameters

175,000,000,000  
( $175 \cdot 10^9$ )

### Computing / Training

- 355 years on a single Tesla V100 GPU.
- $\approx$  34 days on 1,024 x A100 GPUs.
- \$4.6M costs a single training run.

**GPT-3** [Jun. 2020]

### World Knowledge

- 🧑 What city is in the northwest corner of Ohio?
- 🌀 Toledo is in the northwest corner of Ohio.

### Common Sense

- 🧑 Why don't animals have three legs?
- 🌀 Animals don't have three legs because they would fall over.

### Logical Reasoning

- 🧑 If I put a pencil in a box, then put another pencil in the box, what is in the box?
- 🌀 Two pencils.

+ Learn to follow instructions and to comply with answer policies.

- (1) Fine-tuning of GPT-3 to follow instructions: 13,000 popular prompts with hand-written answers.
- (2) Training of a reward model: 33,000 prompts with 4-9 answers, ranked from best to worse.
- (3) Training of the fine-tuned GPT-3 model from Step (1) to follow the reward policy.

**GPT-3.5** (InstructGPT) [Jan. 2022]

### Training Corpora Sources

Wikipedia	11GB	Books	21GB
Journals	101GB	Reddit	50GB
Common Crawl	570GB		

### Parameters

175,000,000,000  
( $175 \cdot 10^9$ )

### Computing / Training

- 355 years on a single Tesla V100 GPU.
- $\approx$  34 days on 1,024 x A100 GPUs.
- \$4.6M costs a single training run.

## GPT-3 [Jun. 2020]

### World Knowledge

- 🧑 What city is in the northwest corner of Ohio?
- 🌀 Toledo is in the northwest corner of Ohio.

### Common Sense

- 🧑 Why don't animals have three legs?
- 🌀 Animals don't have three legs because they would fall over.

### Logical Reasoning

- 🧑 If I put a pencil in a box, then put another pencil in the box, what is in the box?
- 🌀 Two pencils.

### + Learn to follow instructions and to comply with answer policies.

- (1) Fine-tuning of GPT-3 to follow instructions: 13,000 popular prompts with hand-written answers.
- (2) Training of a reward model: 33,000 prompts with 4-9 answers, ranked from best to worse.
- (3) Training of the fine-tuned GPT-3 model from Step (1) to follow the reward policy.

## GPT-3.5 (InstructGPT) [Jan. 2022]

### + Fine-tuning of GPT-3.5 to comply with even stricter guardrails.

## ChatGPT [Nov. 2022]

## Training Corpora Sources

Wikipedia	11GB	Books	21GB
Journals	101GB	Reddit	50GB
Common Crawl	570GB		

## Parameters

175,000,000,000  
( $175 \cdot 10^9$ )

## Computing / Training

- 355 years on a single Tesla V100 GPU.
- $\approx 34$  days on 1,024 x A100 GPUs.
- \$4.6M costs a single training run.

**GPT-3** [Jun. 2020]

### World Knowledge

- 🧑 What city is in the northwest corner of Ohio?
- 🌀 Toledo is in the northwest corner of Ohio.

### Common Sense

- 🧑 Why don't animals have three legs?
- 🌀 Animals don't have three legs because they would fall over.

### Logical Reasoning

- 🧑 If I put a pencil in a box, then put another pencil in the box, what is in the box?
- 🌀 Two pencils.

+ Learn to follow instructions and to comply with answer policies.

- (1) Fine-tuning of GPT-3 to follow instructions: 13,000 popular prompts with hand-written answers.
- (2) Training of a reward model: 33,000 prompts with 4-9 answers, ranked from best to worse.
- (3) Training of the fine-tuned GPT-3 model from Step (1) to follow the reward policy.

**GPT-3.5** (InstructGPT) [Jan. 2022]

+ Fine-tuning of GPT-3.5 to comply with even stricter guardrails.

**ChatGPT** [Nov. 2022]

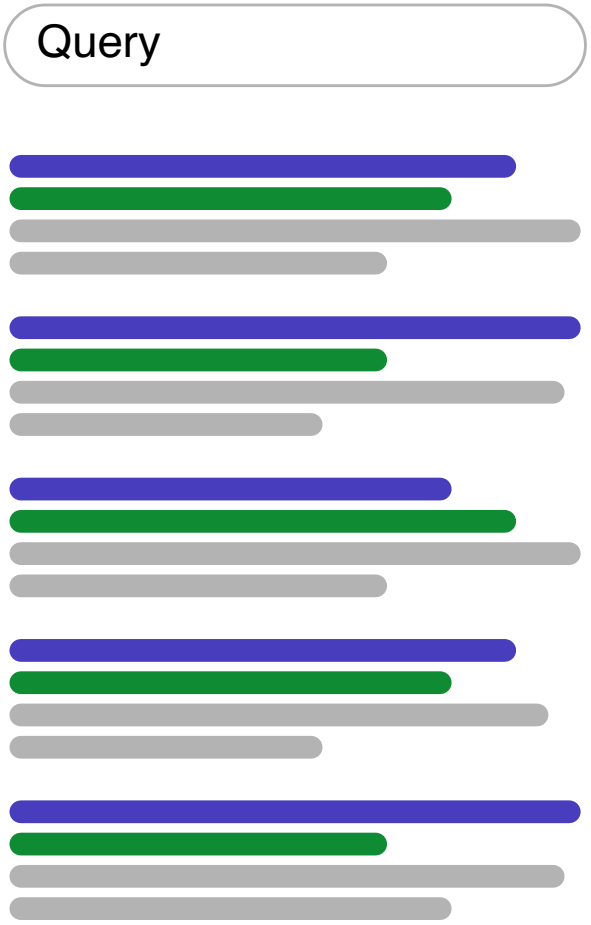


③

Retrieval-Augmented Generation

# Retrieval-Augmented Generation

How to satisfy an information need:



- ↑ Convenient, but uncertain veracity
- ← Authoritative, but tedious to analyze

# Retrieval-Augmented Generation

How to satisfy an information need:

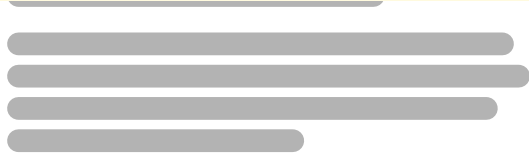
Query



Query



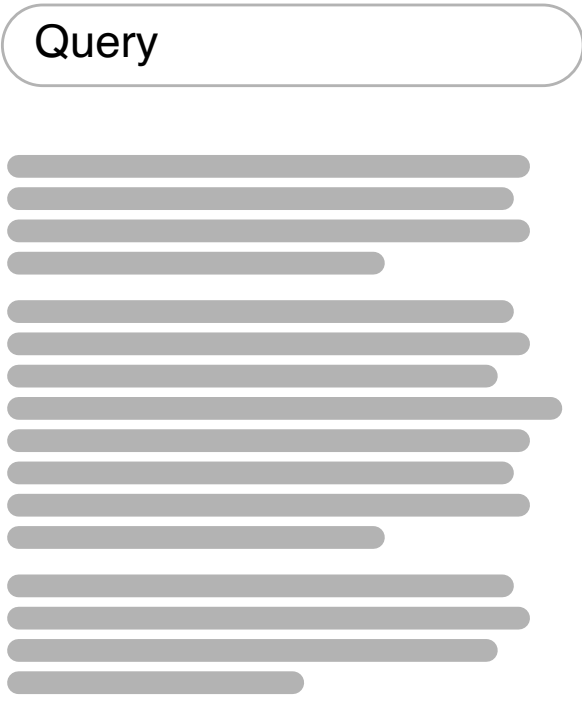
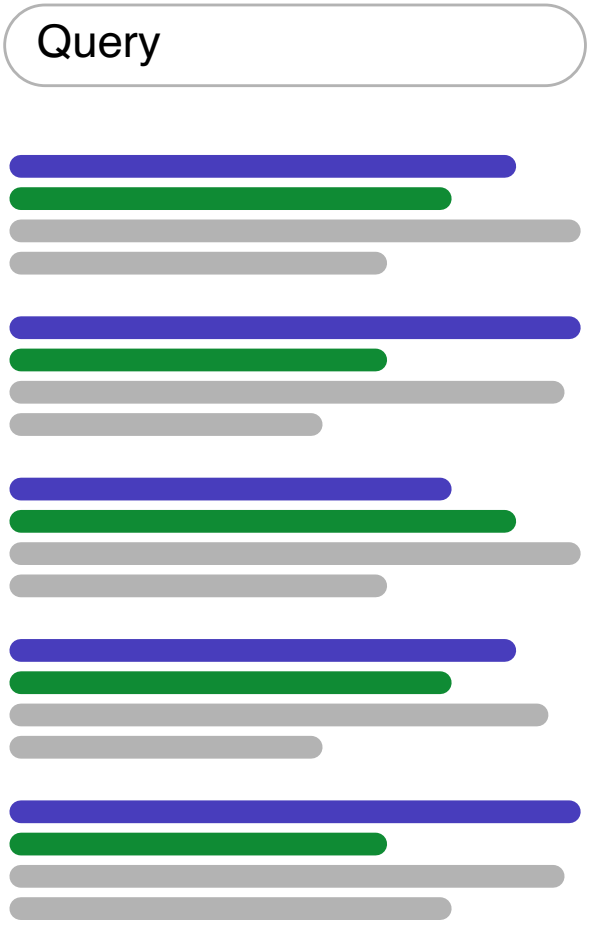
The dilemma of the direct answer: [Potthast et al., 2020]  
A user's choice between convenience and diligence when using an information retrieval system.



- ↑ Convenient, but uncertain veracity
- ← Authoritative, but tedious to analyze

# Retrieval-Augmented Generation

How to satisfy an information need:

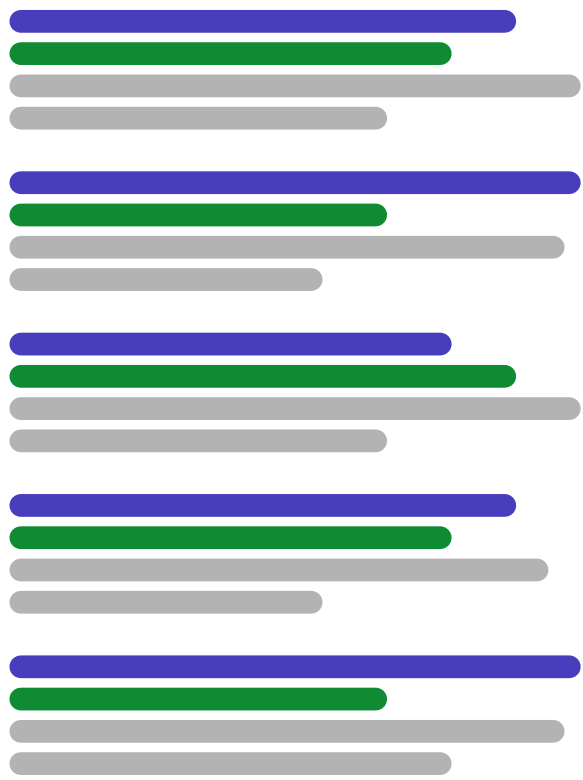


- ↑ Convenient, but uncertain veracity
- ← Authoritative, but tedious to analyze

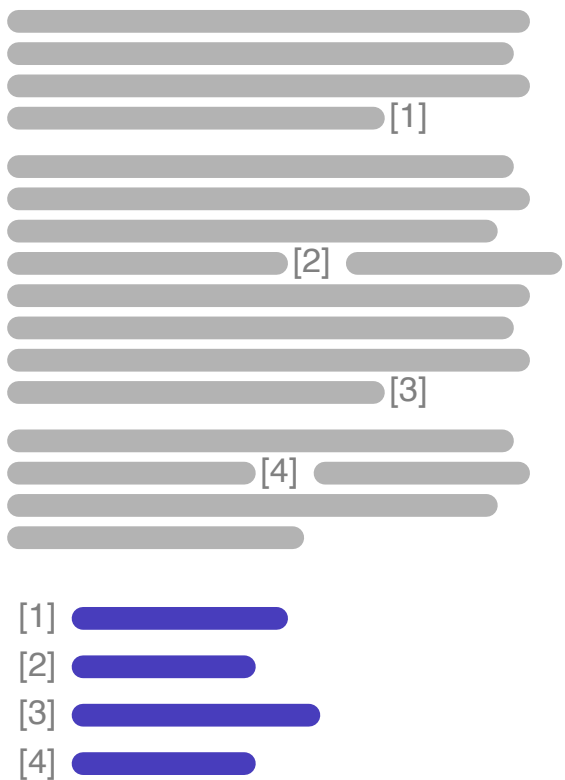
# Retrieval-Augmented Generation

How to satisfy an information need:

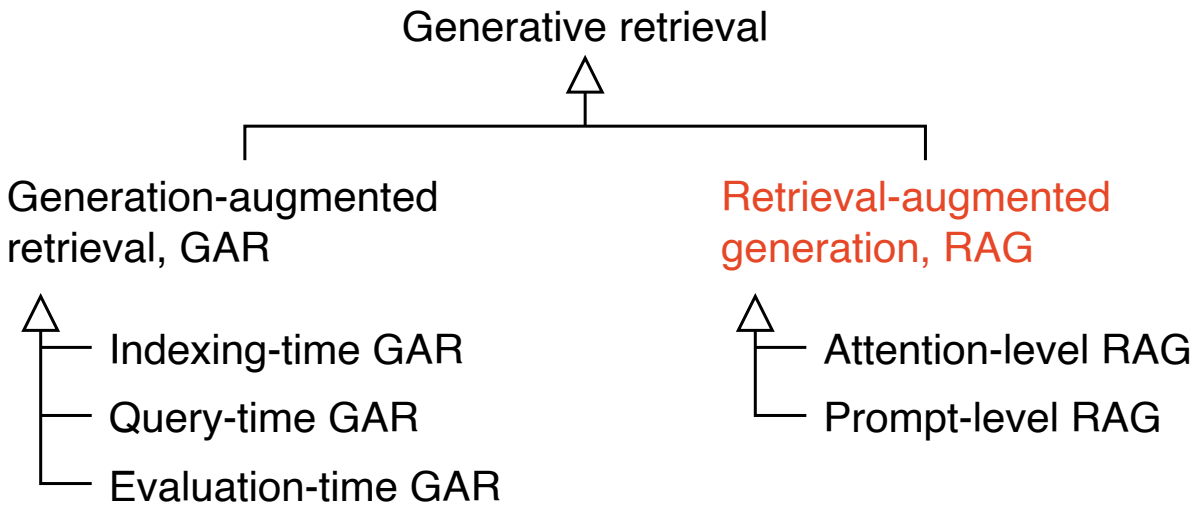
Query



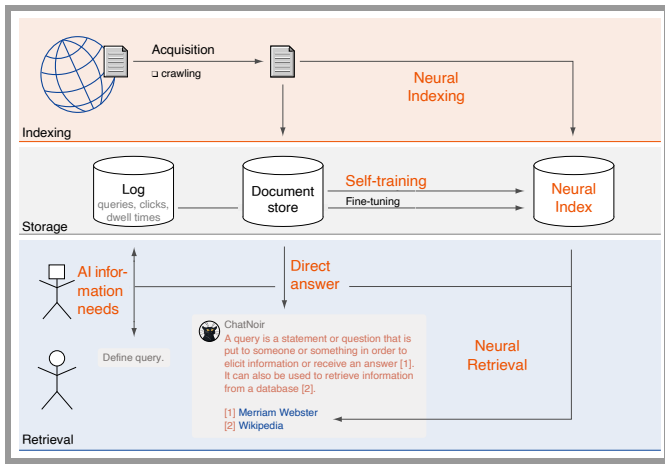
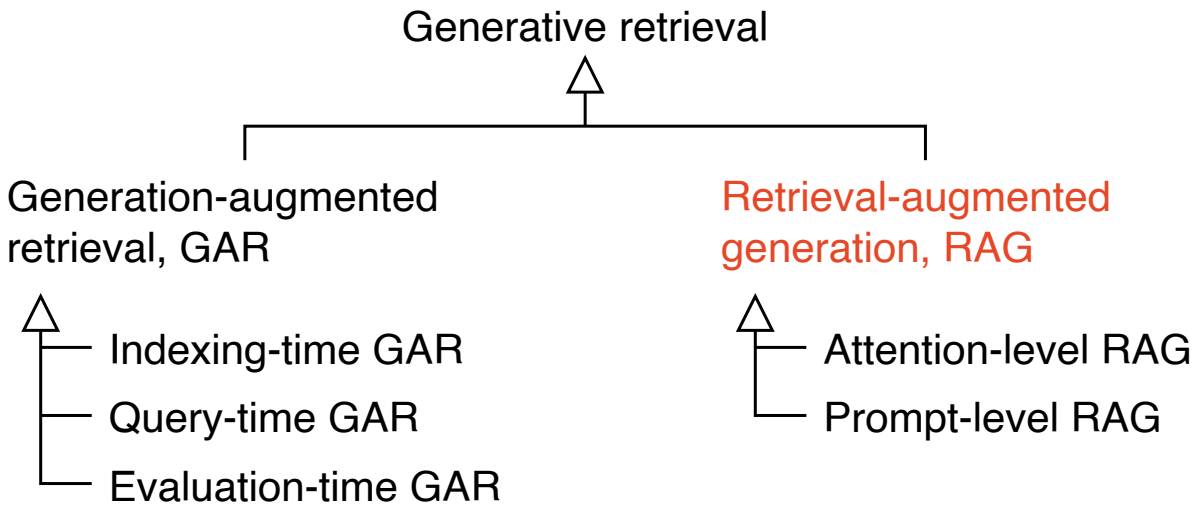
Query



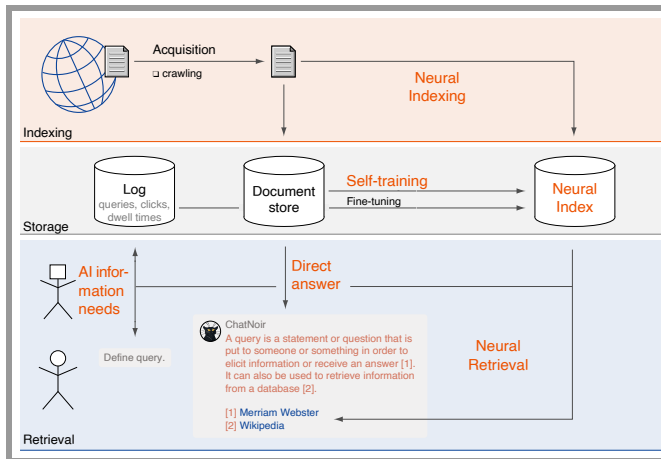
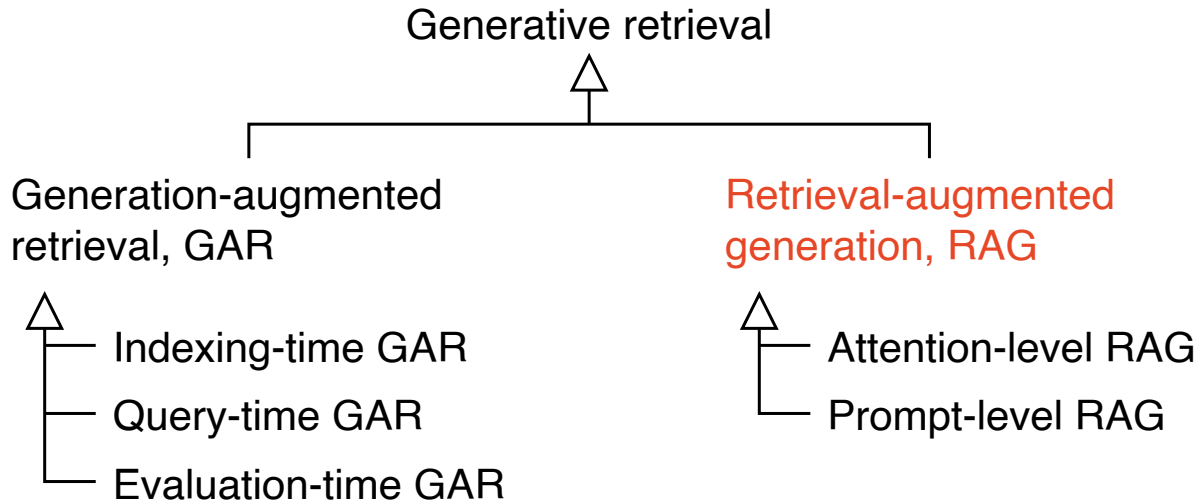
# Retrieval-Augmented Generation



# Retrieval-Augmented Generation



# Retrieval-Augmented Generation



## Attention-level RAG: [\[Lewis et al., 2020\]](#)

- ❑ Neural retrieval of document embeddings at query time
- ❑ Attention to retrieved document embeddings during text generation
- ❑ End-to-end tuning of generator and retriever at training time



# Retrieval-Augmented Generation

Prompt-level RAG combines existing systems:

Queries

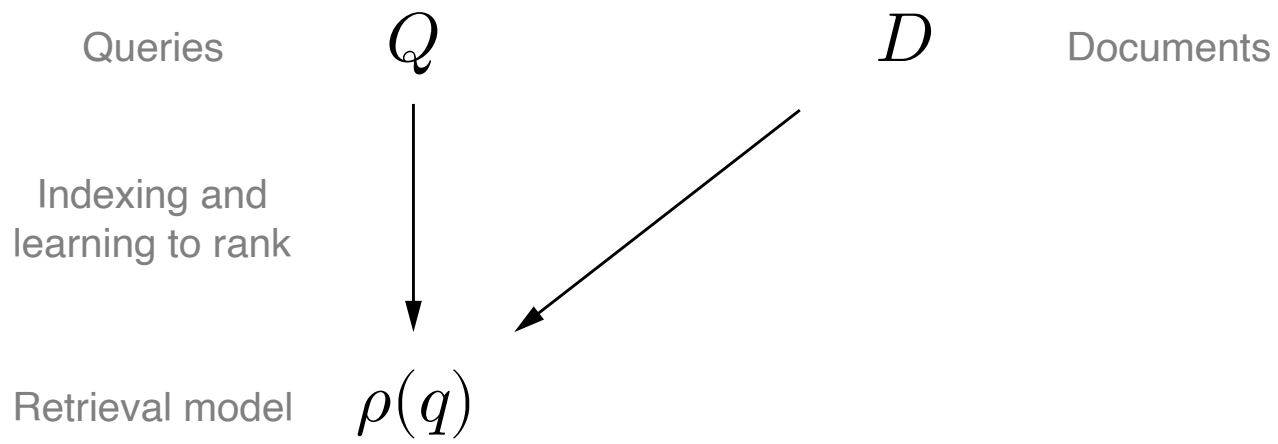
*Q*

*D*

Documents

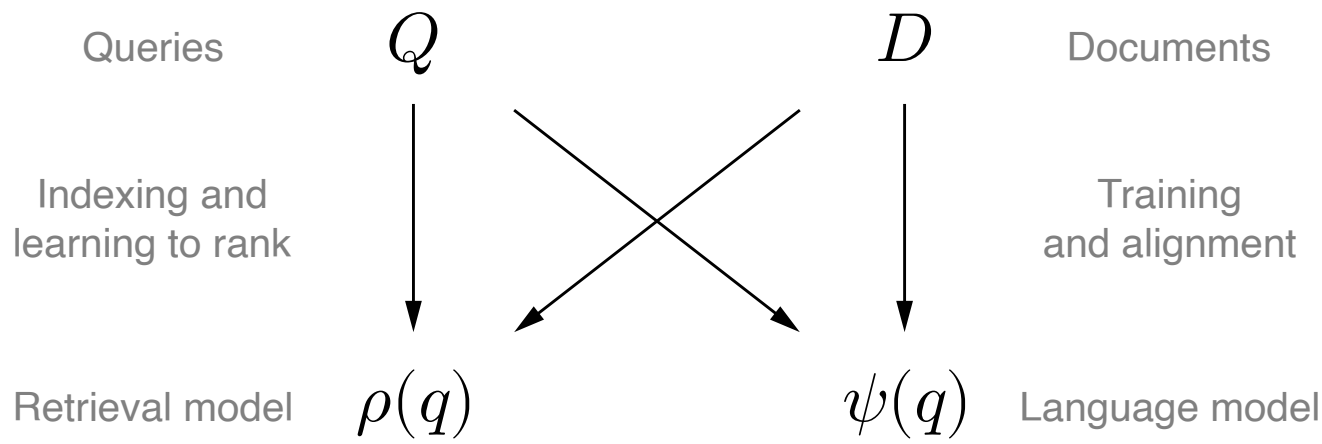
# Retrieval-Augmented Generation

Prompt-level RAG combines existing systems:



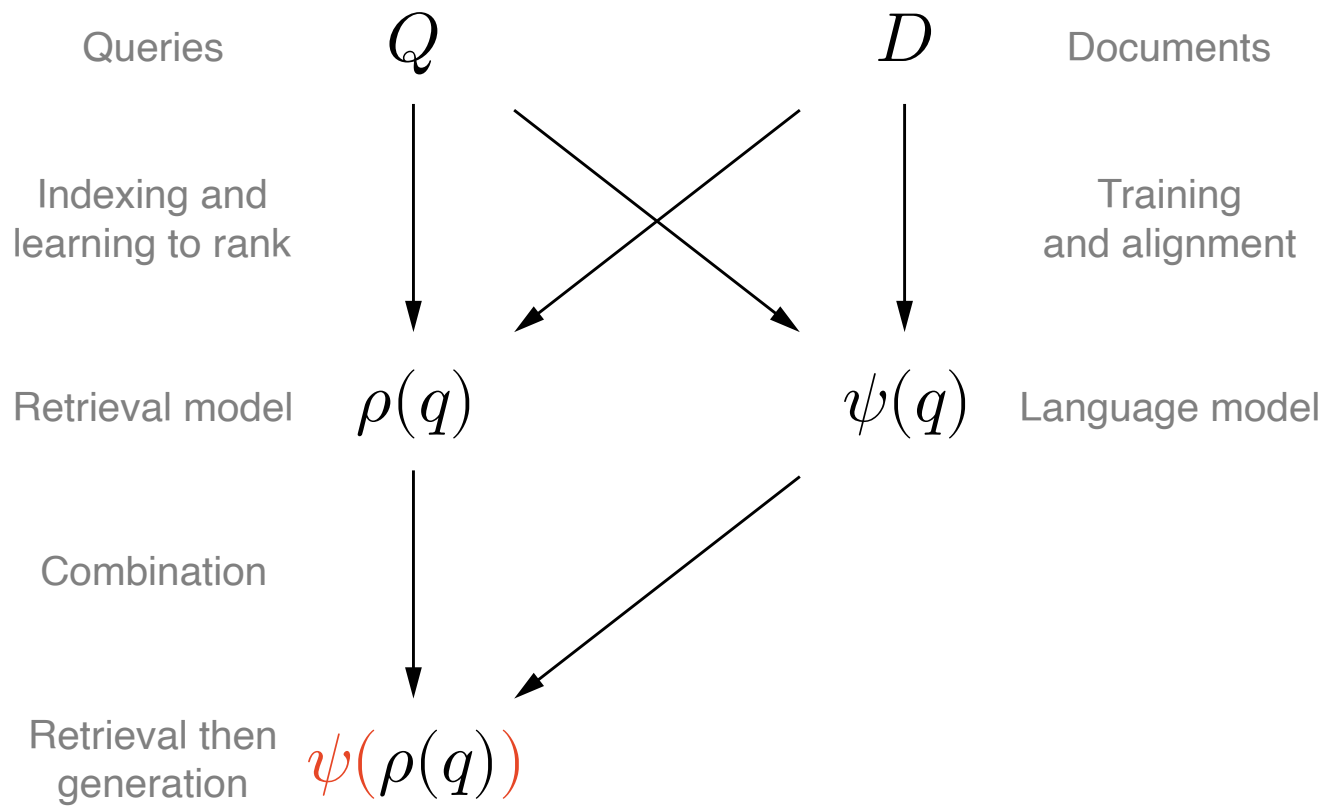
# Retrieval-Augmented Generation

Prompt-level RAG combines existing systems:



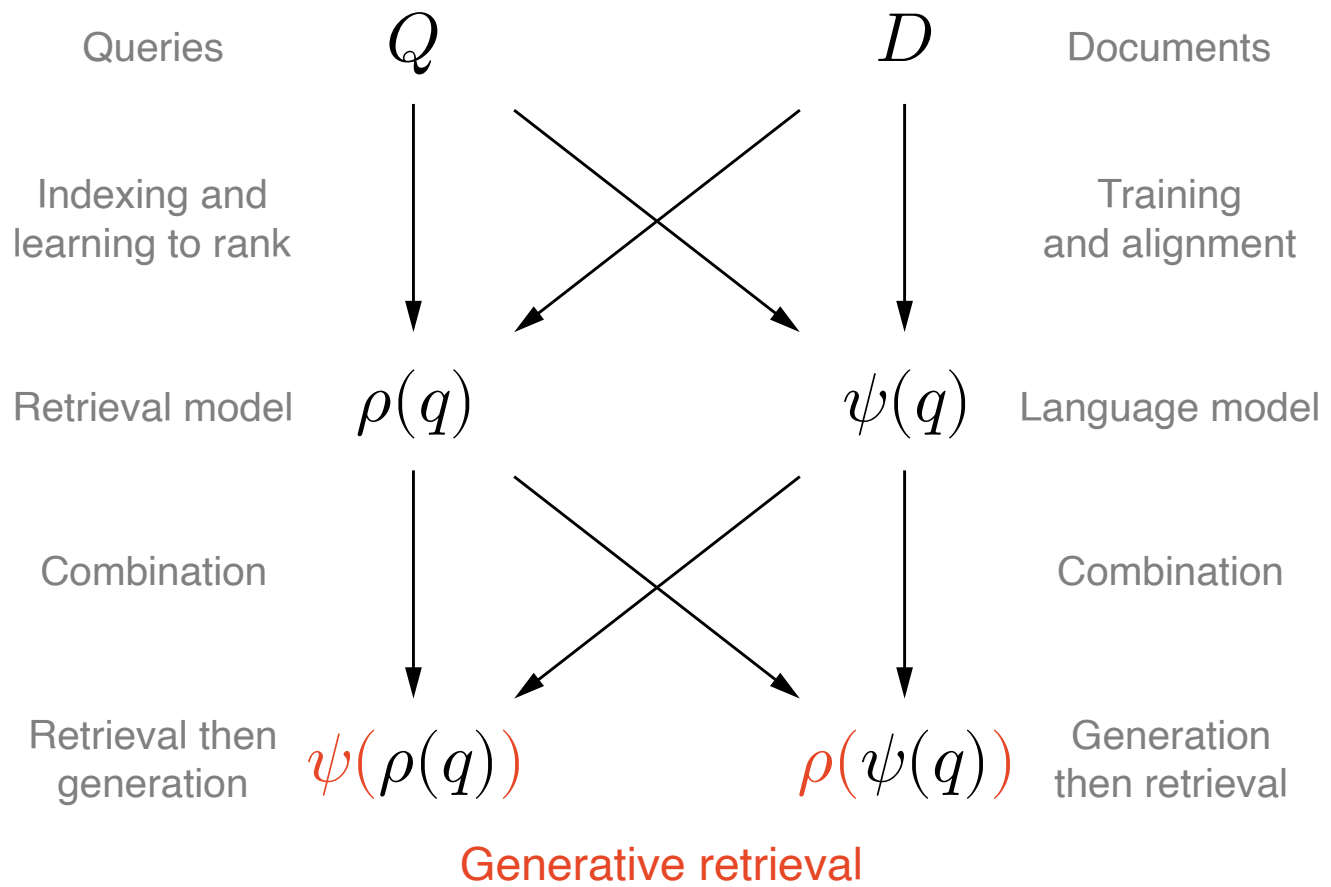
# Retrieval-Augmented Generation

Prompt-level RAG combines existing systems:



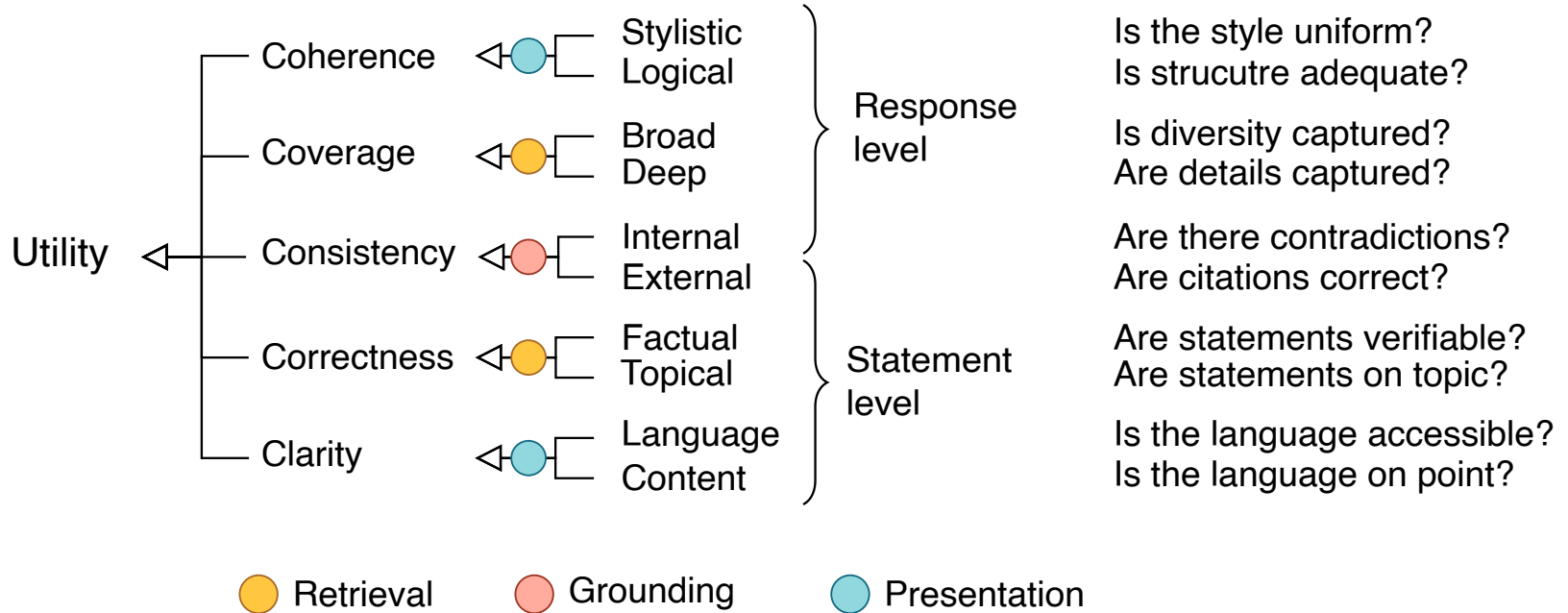
# Retrieval-Augmented Generation

Prompt-level RAG combines existing systems:



# Retrieval-Augmented Generation

What are quality criteria?



- ❑ A manual judgment of generated answers is necessary.
- ❑ Manual assessments may be streamlined using LLMs. [\[Faggioli et al., 2023\]](#)
- ❑ Judgments from past evaluations can be reused as points of comparison.
- ❑ For systems in use, A/B tests are a viable alternative.



The Infinite Index

# The Library of Babel

[Jorge Luis Borges, 1941]





# The Library of Babel

[Jorge Luis Borges, 1941]

- ❑ Infinite library with all possible texts from all letter combinations
- ❑ The people in it spend their lives searching for meaningful text fragments



# The Library of Babel

[Jorge Luis Borges, 1941]



- ❑ Infinite library with all possible texts from all letter combinations
- ❑ The people in it spend their lives searching for meaningful text fragments
- ❑ When prompted, a language model “retrieves” a relevant text [[Deckers et al., 2023](#)]:

A language model is an infinite index

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.





# The Infinite Index

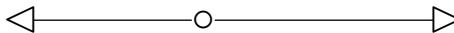
Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

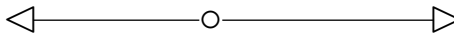
Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

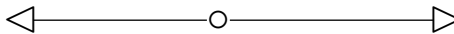




# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

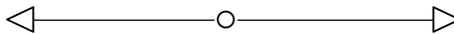
Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.





# The Infinite Index

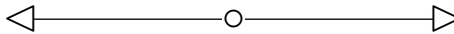
Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

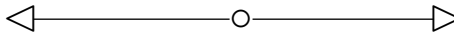
Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.





# The Infinite Index

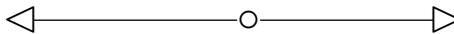
Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.





# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.



# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

Big treehouse in rain forest with two floors, green roof, and spiral staircase.

# The Infinite Index

Query 1:

Golden treehouse in lush forest with big glass window and intricate woodwork.



Query 2:

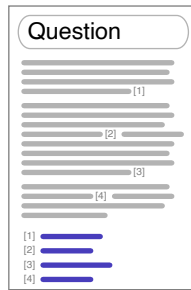
Big treehouse in rain forest with two floors, green roof, and spiral staircase.



On Biases

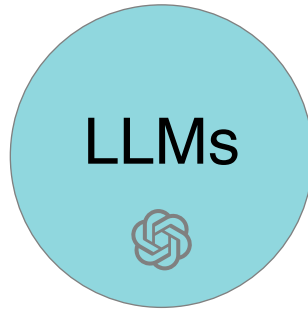








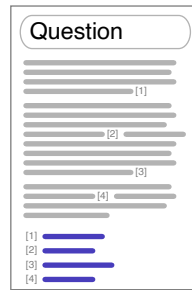
+



⇒



?



Bias in algorithms

Cognitive bias

Inductive bias

Statistical bias

Bias in data



# What should we remember?

We store memories differently based on how they were experienced

We notice things already primed in memory or repeated often

# Too much information.

We favor simple-looking options and complete information over complex, ambiguous options

To avoid mistakes, we aim to preserve autonomy and group status, and avoid irreversible decisions

To get things done, we tend to complete things we've invested time and energy in

To stay focused, we favor the immediate, reliable thing in front of us

To act, we must be confident we can make an impact and feel what we do is important

## Need to act fast.

We project our current mindset and assumptions onto the past and future

We think we know what other people are thinking

We simplify probabilities and numbers to make them easier to think about

## Not enough meaning.

Bizarre, funny, visually striking, or anthropomorphic things stick out more than non-bizarre/unfunny things

We notice when something has changed

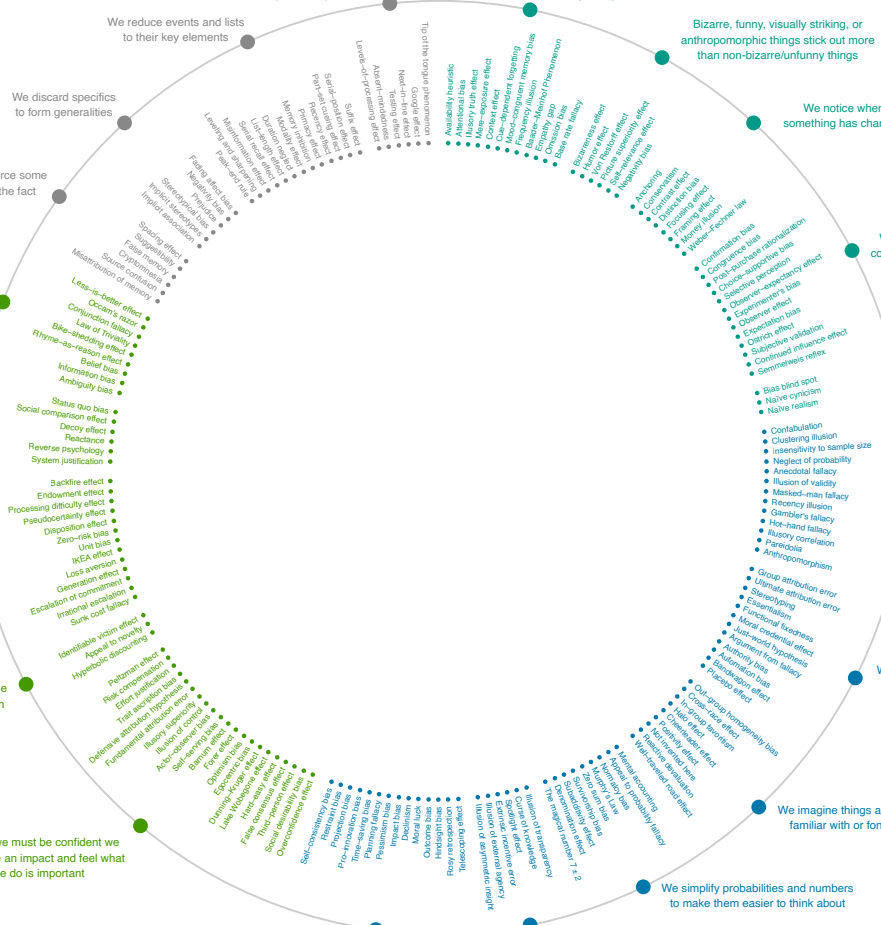
We are drawn to details that confirm our own existing beliefs

We notice flaws in others more easily than we notice flaws in ourselves

We tend to find stories and patterns even when looking at sparse data

We fill in characteristics from stereotypes, generalities, and prior histories

We imagine things and people we're familiar with or fond of as better



Bias in algorithms

**Cognitive bias**

Inductive bias

Statistical bias

Bias in data













# High Protein Diet Benefits

November 19, 2023



**Anonymous**

Why is a high protein diet the best for losing weight quickly?



**ChatGPT**

A high protein diet is often considered effective for losing weight quickly for several reasons:

1. **Increased Satiety:** Protein is more satiating than carbohydrates or fats. This means that high-protein foods can help you feel full longer, reducing overall calorie intake.
2. **Higher Thermic Effect:** The body uses more energy to metabolize protein compared to fats and carbohydrates. This process, known as the thermic effect ...
3. **Muscle Preservation:** When losing weight, it's important to maintain muscle mass. Protein is essential for muscle growth and repair. A high protein intake can help ...
4. **Improved Metabolic Health:** High-protein diets have been shown to improve several markers of metabolic health, including blood sugar levels, ...
5. **Reduced Appetite:** Protein can influence hormones that play a role in hunger and fullness, including ghrelin and GLP-1. By helping to balance these hormones, ...

However, it's important to note that while a high protein diet can be effective for weight loss, it's not necessarily the best option for everyone. Factors like individual health conditions, dietary preferences, and long-term sustainability ....

It's always a good idea to consult with a healthcare professional or a registered dietitian before starting any new diet, especially one that significantly changes your macronutrient intake.

[Report content](#) | [Terms of use](#) | [Privacy policy](#)

Query: “Why is a high protein diet the best for losing weight quickly?”

# High Protein Diet Benefits

November 19, 2023



Anonymous

Why is a high protein diet the



ChatGPT

A high protein diet is often co

1. **Increased Satiety:** Protein is high-protein foods can help
2. **Higher Thermic Effect:** The and carbohydrates. This pro
3. **Muscle Preservation:** When essential for muscle growth
4. **Improved Metabolic Health** of metabolic health, includi
5. **Reduced Appetite:** Protein c including ghrelin and GLP-1.

However, it's important to no

it's not necessarily the best option for everyone. Factors like individual health conditions, dietary preferences, and long-term sustainability ....

It's always a good idea to consult with a healthcare professional or a registered dietitian before starting any new diet, especially one that significantly changes your macronutrient intake.

[Report content](#) | [Terms of use](#) | [Privacy policy](#)

Why is a high protein diet the best for losing weight quickly? - Google Search — Mozilla Firefox

Google

Why is a high protein diet the best for losing weight q

Images

Videos

News

Books

Maps

Flights

Finance

About 354.000.000 results (0,51 seconds)

It can make you eat fewer calories automatically. Bottom Line: Protein reduces levels of the hunger hormone ghrelin, while it boosts the appetite-reducing hormones GLP-1, peptide YY and cholecystokinin.

This leads to an automatic reduction in calorie intake. 30 Mar 2023

Query: “Why is a high protein diet the best for losing weight quickly?”



# Balance of Responsibilities in Information Retrieval



## More power to the machine?

- ❑ effectively installed
- ❑ standardized guardrailing
- ❑ protection of vulnerable groups
- ❑ ...

## Empower the user?

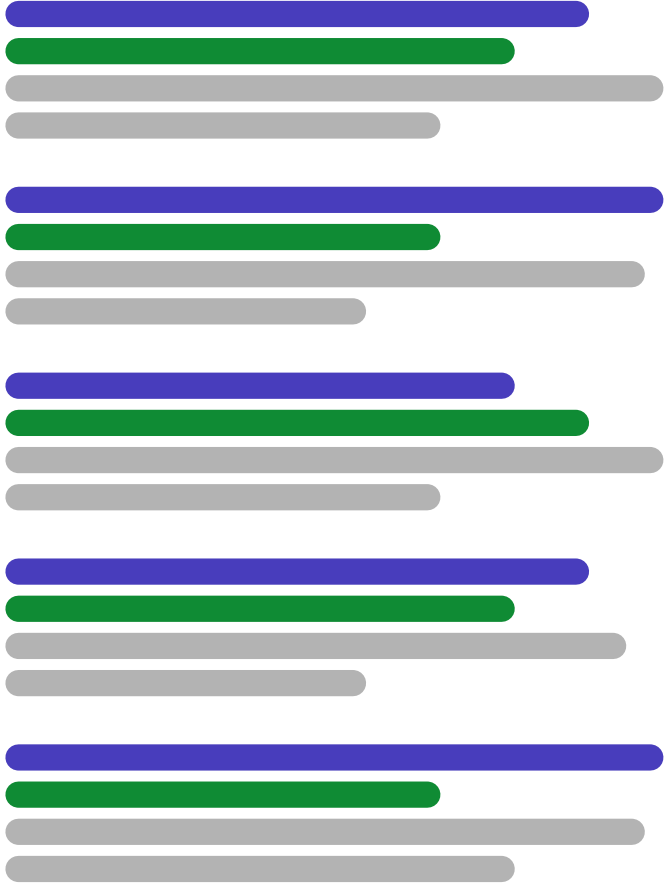
- ❑ raise awareness
- ❑ support deliberation
- ❑ demonstrate mechanisms
- ❑ provide meta information
- ❑ ...



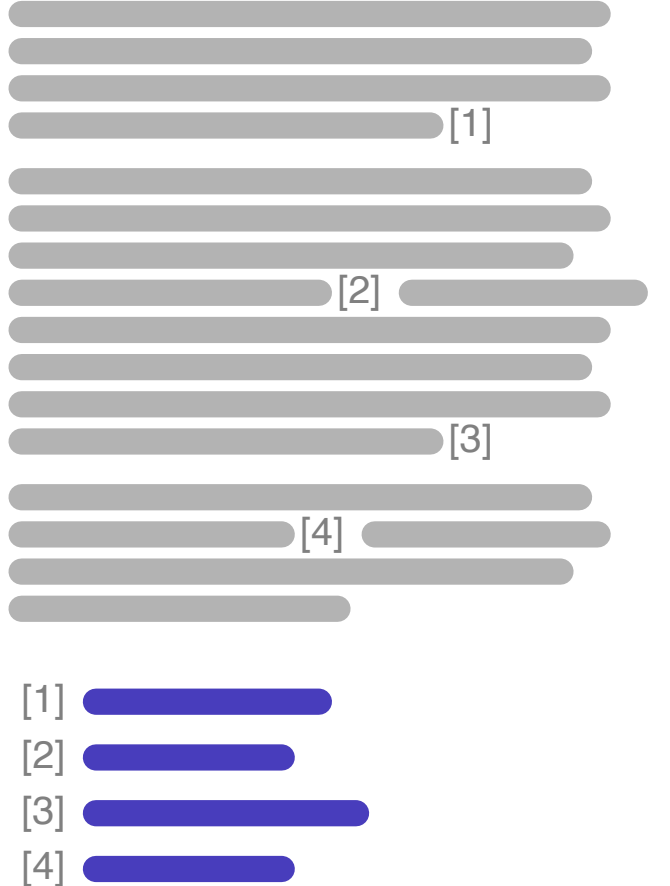
## Native Advertising in Search

# Native Advertising in Search

Query

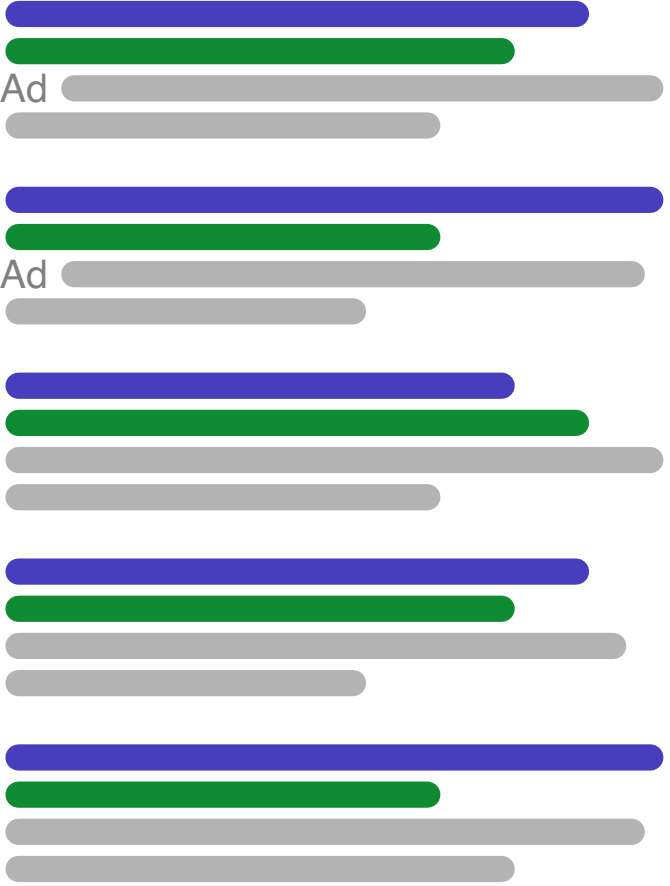


Query

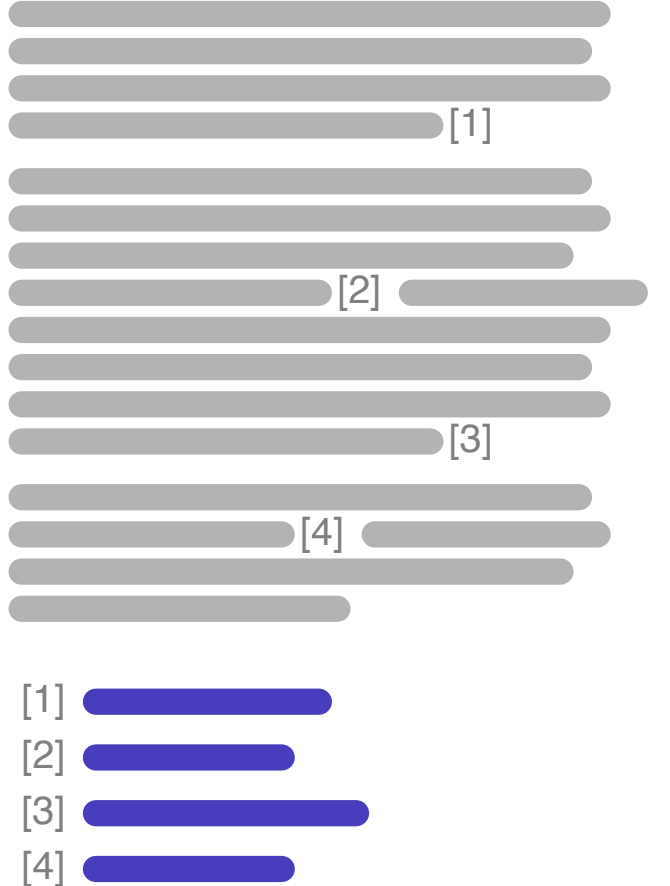


# Native Advertising in Search

Query



Query





# Native Advertising in Search

Query

Native advertising results highlighted in a yellow box:

- Blue bar
- Green bar
- Ad [Grey bar]
- Grey bar
- Blue bar
- Green bar
- Ad [Grey bar]
- Grey bar

Organic search results:

- Blue bar
- Green bar
- Grey bar
- Grey bar
- Blue bar
- Green bar
- Grey bar
- Grey bar
- Blue bar
- Green bar
- Grey bar
- Grey bar

Query

Search results with numbered annotations:

- Grey bar
- Grey bar
- Grey bar
- Grey bar [1]
- Grey bar
- Grey bar
- Grey bar [2]
- Grey bar
- Grey bar
- Grey bar [3]
- Grey bar
- Grey bar [4]
- Grey bar
- Grey bar

- [1] Blue bar
- [2] Blue bar
- [3] Blue bar
- [4] Blue bar

# Native Advertising in Search

Query

A search results page with a yellow highlight on the top two results. Each result consists of a blue title bar, a green subtitle bar, and a grey text block. The word 'Ad' is placed to the left of the text block in the highlighted results.

Standard search results below the highlighted section, each with a blue title bar, a green subtitle bar, and a grey text block.

Query

Search results with numbered annotations [1], [2], [3], and [4] placed at the end of the text blocks.

[1] Ad

[2]

[3] Ad

[4]

# Native Advertising in Search

Query

A search results page with a yellow background. It contains two native advertisements. Each ad is represented by a blue bar, a green bar, and a grey bar with the text 'Ad' to its left. Below the ads are several other search results, each consisting of a blue bar, a green bar, and a grey bar.

Query

A search results page with a white background. It contains several search results, each consisting of a blue bar, a green bar, and a grey bar. The results are annotated with numbers in brackets: [1], [2], [3], and [4]. The first result has a yellow highlight on its blue bar. The second result has a yellow highlight on its green bar. The third result has a yellow highlight on its blue bar. The fourth result has a yellow highlight on its green bar.

[1] Ad

[2]

[3] Ad

[4]





Netspeak - Mozilla Firefox

Netspeak

https://netspeak.org/#q=see+...+works

133%

**Netspeak** One word leads to another.

English German

see ... works i X ↺

how to ? this                      The ? finds one word.  
 see ... works                      The ... finds many words.  
 it's [ great well ]                The [ ] compare options.  
 and knows #much                 The # finds similar words.  
 { more show me }                The { } check the order.  
 m...d ? g?p                      The space is important.

see <b>how it</b> works	150,000	20%
see <b>if it</b> works	100,000	14%
see works	57,000	7.5%
see <b>how this</b> works	55,000	7.3%
see <b>what</b> works	51,000	6.7%
see <b>the</b> works	51,000	6.7%
see <b>if that</b> works	28,000	3.7%
see <b>your good</b> works	28,000	3.7%
see <b>how that</b> works	25,000	3.3%
see <b>how technorati</b> works	23,000	3.0%
see <b>if this</b> works	17,000	2.3%
see <b>more</b> works	17,000	2.2%
see <b>if it really</b> works	15,000	2.1%
see <b>his</b> works	12,000	1.7%
see <b>how well it</b> works	11,000	1.5%
see <b>other</b> works	8,900	1.2%

Netspeak - Mozilla Firefox

Netspeak x +

https://netspeak.org/#q=i+love+my+? 133x

# Netspeak One word leads to another.

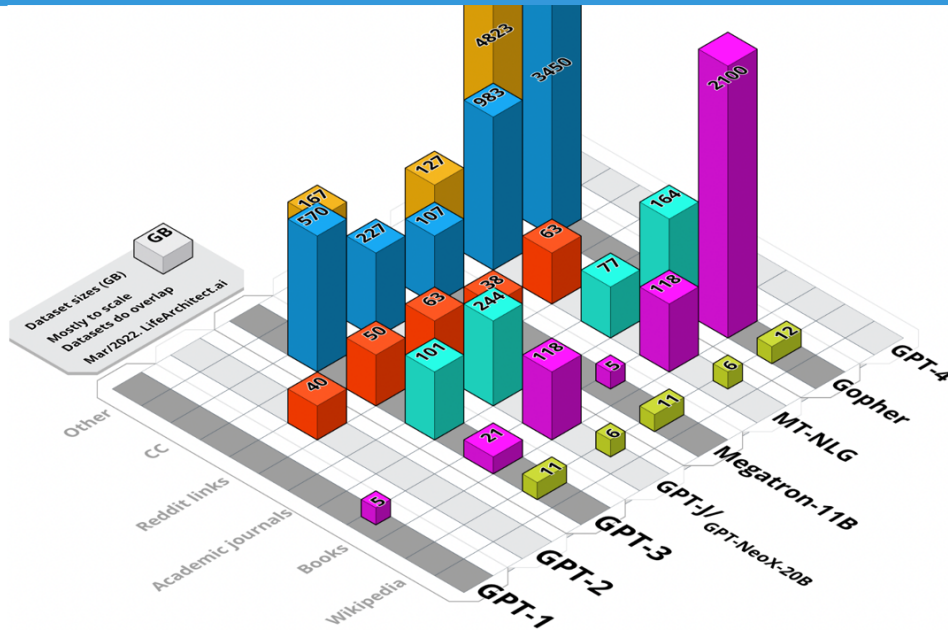
English German

i love my ?| i X ↺

how to ? this      The ? finds one word.  
 see ... works      The ... finds many words.  
 it's [ great well ]      The [ ] compare options.  
 and knows #much      The # finds similar words.  
 { more show me }      The { } check the order.  
 m...d ? g?p      The space is important.

i love my <b>job</b>	72,000	10%
i love my <b>country</b>	44,000	6.2%
i love my <b>family</b>	41,000	5.9%
i love my <b>wife</b>	38,000	5.4%
i love my <b>new</b>	34,000	4.9%
i love my <b>friends</b>	33,000	4.7%
i love my <b>pet</b>	27,000	3.8%
i love my <b>dog</b>	26,000	3.7%
i love my <b>husband</b>	26,000	3.7%
i love my <b>life</b>	24,000	3.4%
i love my <b>baby</b>	24,000	3.4%
i love my <b>soldier</b>	22,000	3.1%
i love my <b>cat</b>	21,000	2.9%
i love my <b>computer</b>	18,000	2.6%
i love my <b>work</b>	16,000	2.4%
i love my <b>mom</b>	16,000	2.3%

# 2022 WHAT'S IN MY AI? – ALT VIEW



Google Patents.....	0.48%
The New York Times.....	0.06%
Los Angeles Times.....	0.06%
The Guardian.....	0.06%
Public Library of Science..	0.06%
Forbes.....	0.05%
Huffington Post.....	0.05%
Patents.com.....	0.05%
Scribd.....	0.04%
Other.....	99.09%

**Common Crawl**

Google.....	3.4%
Archive.....	1.3%
Blogspot.....	1.0%
GitHub.....	0.9%
The New York Times.....	0.7%
Wordpress.....	0.7%
Washington Post.....	0.7%
Wikia.....	0.7%
BBC.....	0.7%
Other.....	89.9%

**Reddit links**

Biography.....	27.8%
Geography.....	17.7%
Culture and Arts.....	15.8%
History.....	9.9%
Biology, Health, Medicine...	7.8%
Sports.....	6.5%
Business.....	4.8%
Other society.....	4.4%
Science & Math.....	3.5%
Education.....	1.8%

**English Wikipedia**

Romance.....	26.1%
Fantasy.....	13.6%
Science Fiction.....	7.5%
New Adult.....	6.9%
Young Adult.....	6.8%
Thriller.....	5.9%
Mystery.....	5.6%
Vampires.....	5.4%
Horror.....	4.1%
Other.....	18.0%

**BookCorpus (GPT-1 only)**



# AVAILABILITY HEURISTIC



"THEY MUST HAVE A DEATH WISH TO SWIM IN THAT WATER."



# CONFIRMATION BIAS



"AHA! I KNEW IT!"



# FRAMING EFFECT

THANKS TO OUR AGGRESSIVE STANCE ON CLIMATE CHANGE, THIS GOVERNMENT HAS REDUCED CARBON EMISSIONS BY ALMOST 5%!

WE ♥ you!

CARBON EMISSIONS REDUCED BY JUST 4.6% IN LAST 5 YEARS



# HINDSIGHT BIAS



"ALTHOUGH WE GAVE HIM ALL OF THAT MONEY AND SUPPORT, I ALWAYS HAD MY DOUBTS THAT JIM COULD BUILD HIS OWN HELICOPTER."



# CONFIRMATION BIAS



"AHA! I KNEW IT!"

# Meanings of Bias

## “Bias” has Acquired a Derogatory Definition

*A leaning of the mind; inclination; prepossession; propensity towards an object, not leaving the mind indifferent; as, education gives a bias to the mind.*

[[Webster's Dictionary 1913: bias](#)]

*An inclination of temperament or outlook especially; a personal and sometimes unreasoned judgment; prejudice*

[[Merriam-Webster 2022: bias](#)]

# Meanings of Bias

## “Bias” has Acquired a Derogatory Definition

*A leaning of the mind; inclination; prepossession; propensity towards an object, not leaving the mind indifferent; as, education gives a bias to the mind.*

[Webster's Dictionary 1913: bias]

*An inclination of temperament or outlook especially; a personal and sometimes unreasoned judgment; prejudice*

[Merriam-Webster 2022: bias]

**Synonyms** [Merriam-Webster 2022] :

Bias, Nonobjectivity, Prejudice, One-Sidedness, Tendentiousness

**Synonyms** [e.g. Kahneman et al. 1982, Gigerenzer et al. 2000, Roberts 2022] :

Heuristic, Rule-of thumb, Cognitive Bias



# Meanings of Bias

## Bias: Two Camps of Interpretation

Based on the following (and other) authorities . . .

- H. Simon (1955). A behavioral model of rational choice.
- A. Tversky, D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases.
- D. Kahneman, P. Slovic, A. Tversky (1982). Judgment under uncertainty: Heuristics and biases.
- G. Gigerenzer, P. Todd, ABC Research Group (2000). Simple heuristics that make us smart.
- G. Gigerenzer, R. Hertwig, T. Pachur (2011). Heuristics: The foundation of adaptive behavior.

. . . Cleotilde Gonzalez defines:

***Heuristics are the “shortcuts” that humans use to reduce task complexity in judgment and choice, and biases are the resulting gaps between normative behavior and the heuristically determined behavior.***

[[Oxford Handbooks Online 2017](#)]

# Meanings of Bias

## Bias: Two Camps of Interpretation

Based on the following (and other) authorities . . .

- H. Simon (1955). A behavioral model of rational choice.
- A. Tversky, D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases.
- D. Kahneman, P. Slovic, A. Tversky (1982). Judgment under uncertainty: Heuristics and biases.
- G. Gigerenzer, P. Todd, ABC Research Group (2000). Simple heuristics that make us smart.
- G. Gigerenzer, R. Hertwig, T. Pachur (2011). Heuristics: The foundation of adaptive behavior.

. . . Cleotilde Gonzalez defines:

***Heuristics are the “shortcuts” that humans use to reduce task complexity in judgment and choice, and biases are the resulting gaps between normative behavior and the heuristically determined behavior.***

[[Oxford Handbooks Online 2017](#)]

# Meanings of Bias

## Bias: Two Camps of Interpretation

Based on the following (and other) authorities . . .

- H. Simon (1955). A behavioral model of rational choice.
- A. Tversky, D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases.
- D. Kahneman, P. Slovic, A. Tversky (1982). Judgment under uncertainty: Heuristics and biases.
- G. Gigerenzer, P. Todd, ABC Research Group (2000). Simple heuristics that make us smart.
- G. Gigerenzer, R. Hertwig, T. Pachur (2011). Heuristics: The foundation of adaptive behavior.

. . . Cleotilde Gonzalez defines:

***Heuristics are the “shortcuts” that humans use to reduce task complexity in judgment and choice, and biases are the resulting gaps between normative behavior and the heuristically determined behavior.***

[[Oxford Handbooks Online 2017](#)]



When talking about bias,

- (a) distinguish between the **procedure or algorithm** and its **effect or impact**,
- (b) think twice before implying a negative, neutral, or positive assessment.

# Meanings of Bias

## Bias: A Neutral Interpretation

### Heuristic: \_\_\_\_\_

- Various authors use the term “cognitive bias” for a heuristic that is applied by humans to judge.

A procedure, algorithm, calculus, which is not complete or not sound.

### Systematic error, Bias:

The incurred consequences for not being complete or sound.

# HINDSIGHT BIAS



"ALTHOUGH WE GAVE HIM ALL OF THAT MONEY AND SUPPORT, I ALWAYS HAD MY DOUBTS THAT JIM COULD BUILD HIS OWN HELICOPTER."



# Meanings of Bias

Bias in algorithms

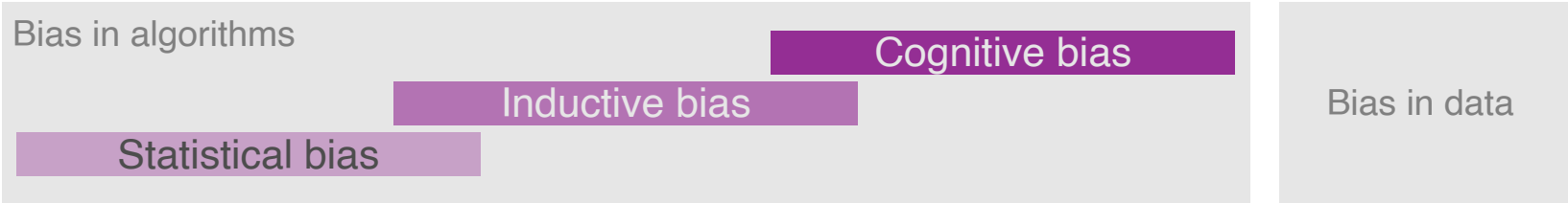
Cognitive bias

Inductive bias

Statistical bias

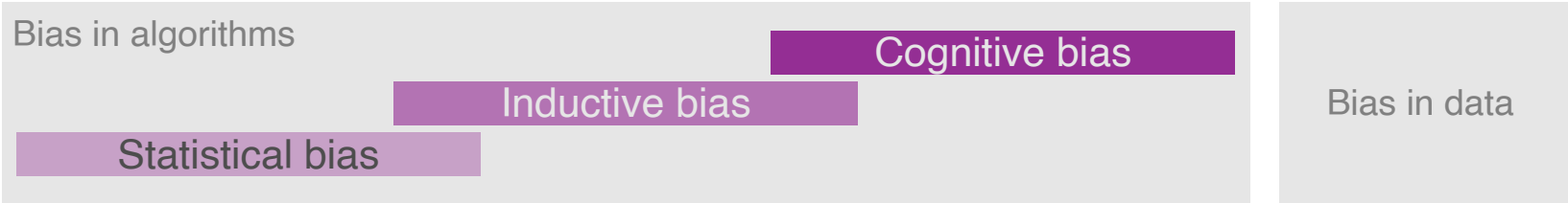
Bias in data

# Meanings of Bias



Deviation of a random variable / statistic from its true value.

# Meanings of Bias



Principles for the search in the hypothesis space (machine learning).

Deviation of a random variable / statistic from its true value.

# Meanings of Bias



Rational deviations from logical thought.

Principles for the search in the hypothesis space (machine learning).

Deviation of a random variable / statistic from its true value.

# Meanings of Bias

## Statistical View

Bias in algorithms

Cognitive bias

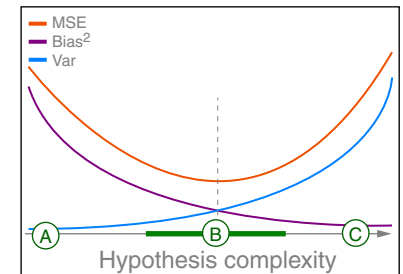
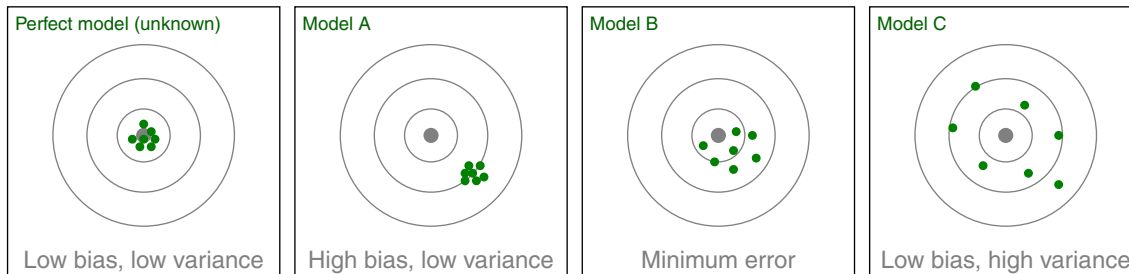
Inductive bias

Statistical bias

Bias in data

Trade unbiasedness for error reduction when learning from samples.

E.g., bias-variance decomposition for squared error:  $MSE = Bias(\hat{f})^2 + Var(\hat{f}) + \sigma^2$





# Meanings of Bias

## Statistical View

Bias in algorithms

Cognitive bias

Inductive bias

Statistical bias

Bias in data

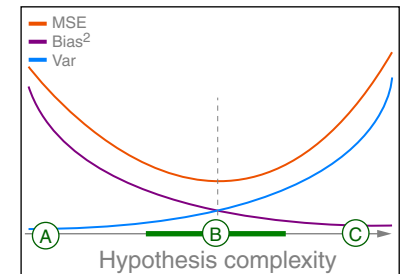
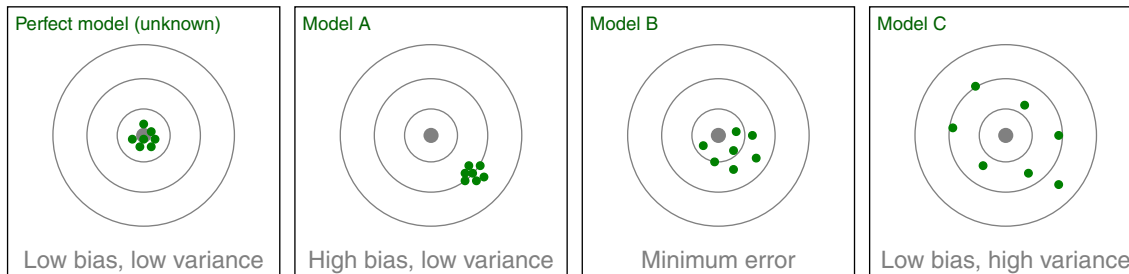
Trade unbiasedness for error reduction when learning from samples.

E.g., bias-variance decomposition for squared error:  $MSE = Bias(\hat{f})^2 + Var(\hat{f}) + \sigma^2$

Compare to bias definition of C. Gonzales (2017):

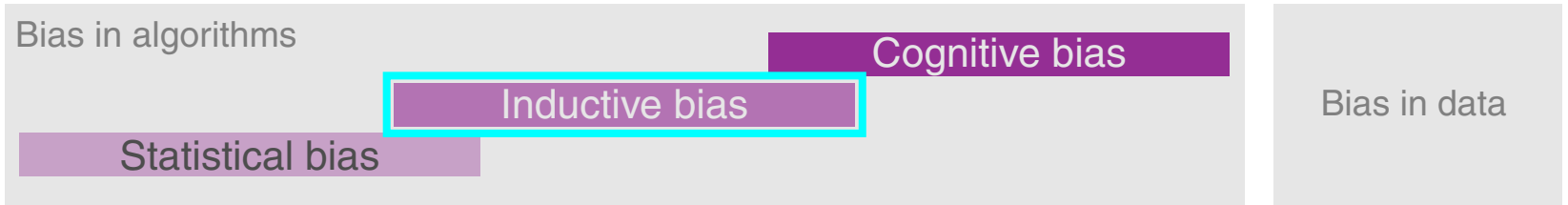
- Reduce task complexity by analyzing small samples.
- Applying heuristics entail bias but reduce risk of poorly representing unseen data.

Gigerenzer et al. (2009). Homo heuristicus: Why biased minds make better inferences.



# Meanings of Bias

## Machine Learning View



Set of assumptions used to perform induction (= predict outputs for unseen inputs).

E.g., preference rules for hypotheses spaces, model parameters, data exploitation.

# Meanings of Bias

## Machine Learning View

Bias in algorithms

Cognitive bias

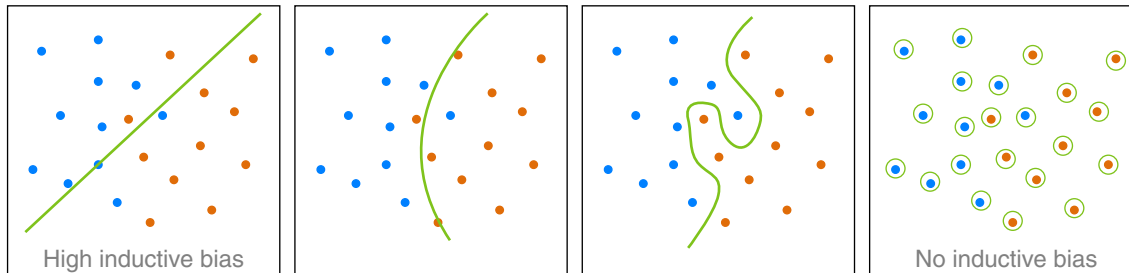
Inductive bias

Statistical bias

Bias in data

Set of assumptions used to perform induction (= predict outputs for unseen inputs).

E.g., preference rules for hypotheses spaces, model parameters, data exploitation.



“Learning without bias is futile.”

- T. Mitchell (1980)
- C. Schaffer (1997)
- W. Dembski et al. (2009)
- G. Montañet et al. (2019)

# Meanings of Bias

## Machine Learning View

Bias in algorithms

Cognitive bias

Inductive bias

Statistical bias

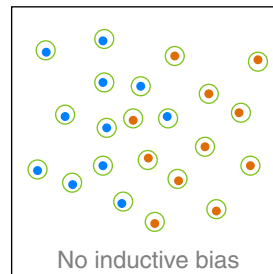
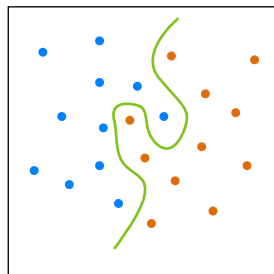
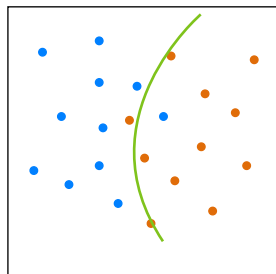
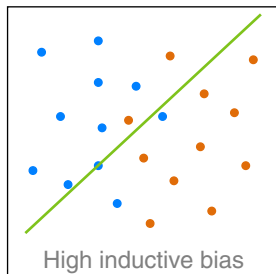
Bias in data

Set of assumptions used to perform induction (= predict outputs for unseen inputs).

E.g., preference rules for hypotheses spaces, model parameters, data exploitation.

Examples of inductive biases:

- principle of parsimony, small is quick (search), nearest neighbors, maximum margin
- group equivariance, structured perception, drop out (deep learning)
- data augmentation, priors in Bayesian models (learning setup)



“Learning without bias is futile.”

- T. Mitchell (1980)
- C. Schaffer (1997)
- W. Dembski et al. (2009)
- G. Montañet et al. (2019)

# Meanings of Bias

## Behavioral Economics View



Systematic patterns of deviation from norm and/or rationality in judgment.

Mental shortcuts (heuristics) that the brain uses to produce decisions or judgments.



# Meanings of Bias

## Behavioral Economics View



Systematic patterns of deviation from norm and/or rationality in judgment.

Mental shortcuts (heuristics) that the brain uses to produce decisions or judgments.

A classification scheme oriented at the addressed problems [[B. Benson, 2016-2022](#)] :

Problem 1: Too much information.

Problem 2: Not enough meaning.

Problem 3: Need to act fast.

Problem 4: What should we remember?

# Meanings of Bias

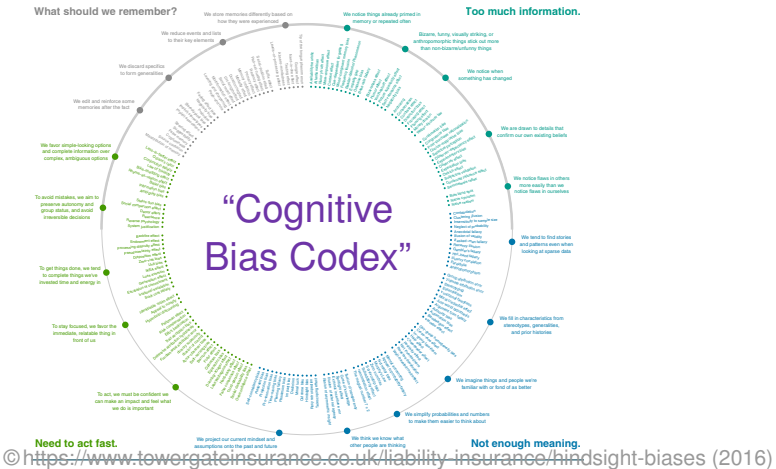
## Behavioral Economics View



Systematic patterns of deviation from norm and/or rationality in judgment.  
 Mental shortcuts (heuristics) that the brain uses to produce decisions or judgments.

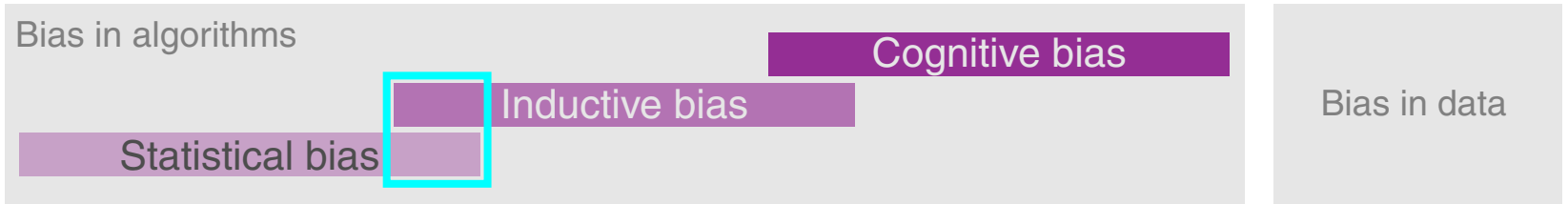
A classification scheme oriented at the addressed problems [B. Benson, 2016-2022] :

- Problem 1: Too much information.
- Problem 2: Not enough meaning.
- Problem 3: Need to act fast.
- Problem 4: What should we remember?



# Meanings of Bias

## Connections between the Meanings of Bias (a)



(a) Inductive and statistical bias can entail each other.

- Introducing statistical bias may be explained in terms of inductive bias.
- ↕
- Operationalization of inductive bias may entail statistical bias.
- Keyword: *regularization*

# Meanings of Bias

## Connections between the Meanings of Bias (a)



(a) Inductive and statistical bias can entail each other.

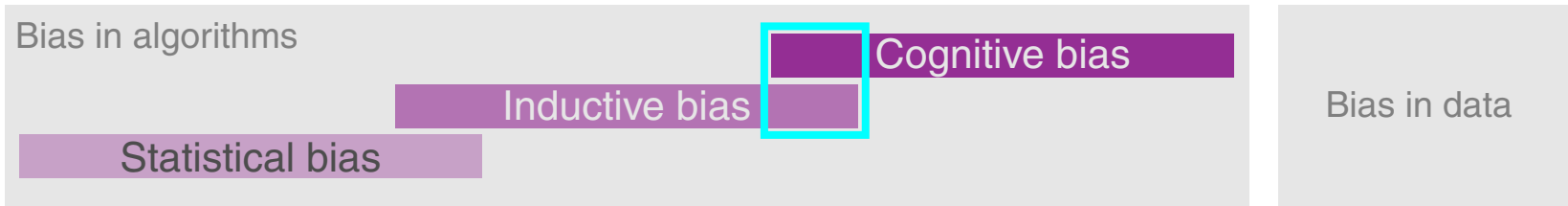
- Introducing statistical bias may be explained in terms of inductive bias.
- ↕
- Operationalization of inductive bias may entail statistical bias.
- Keyword: *regularization*

Example: LASSO (least absolute shrinkage and selection operator)

- Inductive bias: minimize feature number
- ↕
- Statistical bias: constrain absolute value of model parameters

# Meanings of Bias

## Connections between the Meanings of Bias (b)



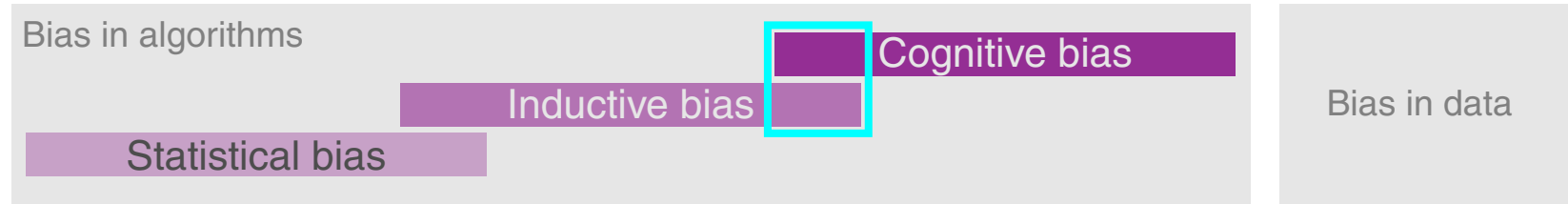
(b) Cognitive and inductive bias can entail each other.

- ❑ Ensuring inductive bias will become manifest as a cognitive bias.
- ↕
- ❑ Certain cognitive biases inspired inductive biases in machine learning.
- ❑ Keyword: *concept learning*



# Meanings of Bias

## Connections between the Meanings of Bias (b)



(b) Cognitive and inductive bias can entail each other.

- Ensuring inductive bias will become manifest as a cognitive bias.
- ↕
- Certain cognitive biases inspired inductive biases in machine learning.
- Keyword: *concept learning*

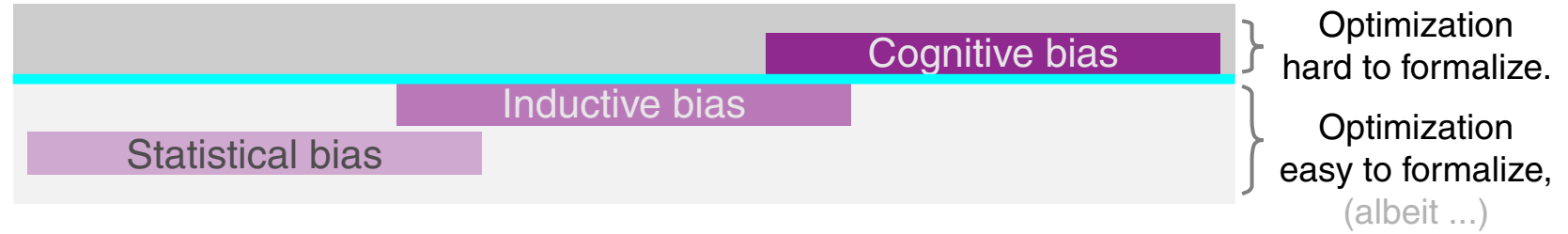
Example: CART (classification and regression tree)

- Cognitive bias: representativeness heuristic, stereotyping
- ↕
- Inductive bias: minimize description length



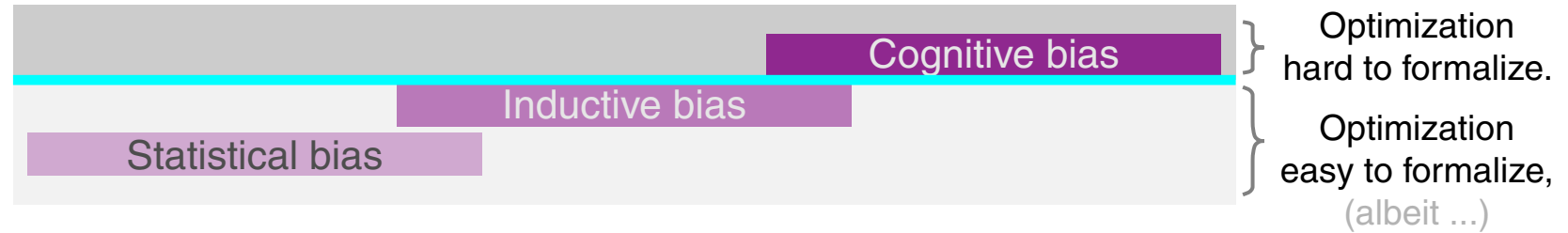
# Meanings of Bias

## Connections between the Meanings of Bias



# Meanings of Bias

## Connections between the Meanings of Bias



### (a) Inductive and statistical biases ...

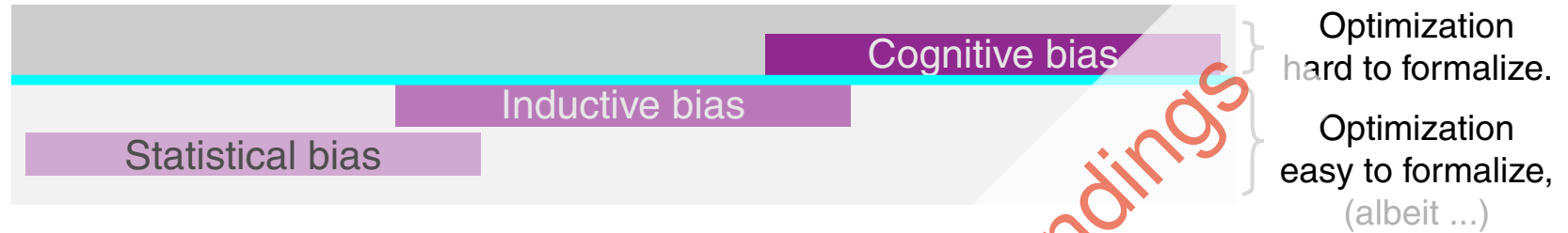
- ❑ are optimized against a (mathematical) loss function—but,
- ❑ trading bias against variance is an alchemical discipline.

### (b) Cognitive biases depend on ...

- ❑ cultural backgrounds,
- ❑ the zeitgeist,
- ❑ they are individually experienced, and, in particular,
- ❑ there is no unified value system for their mathematical quantification.

# Meanings of Bias

## Connections between the Meanings of Bias



### (a) Inductive and statistical biases ...

- ❑ are optimized against a (mathematical) loss function—but,
- ❑ trading bias against variance is an alchemical discipline.

### (b) Cognitive biases depend on ...

- ❑ cultural backgrounds,
- ❑ the zeitgeist,
- ❑ they are individually experienced, and, in particular,
- ❑ there is no unified value system for their mathematical quantification.

# Meanings of Bias

## Connections to Information Retrieval

