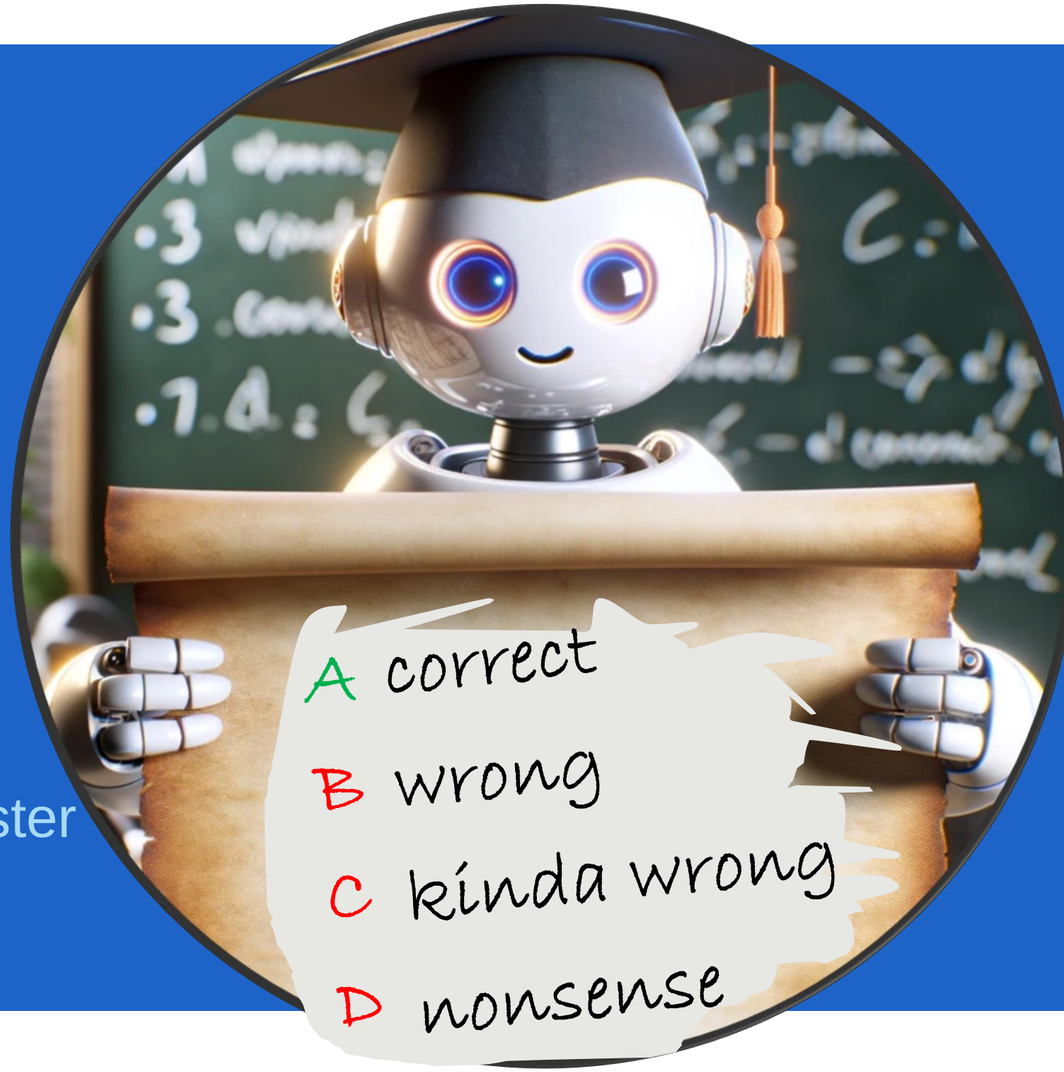
FACULTY OF ENGINEERING
AND ARCHITECTURE

Department of Information Technology
Text-to-Knowledge team, IDLab research group

Language Model Adaptation with Applications in AI for Education



Ph.D. candidate: Semere Kiros Bitew
Promotors: prof. Chris Develder, prof. Thomas Demeester
March 08th, 2024

GHENT UNIVERSITY · imec

Outline

- AI in Education
- Research Motivation
- Language Modeling
- Research Contributions
  - Distractor generation task
  - Gap-filling grammar exercise generation
- Conclusion

Let’s do a quick poll: Personal Trivia

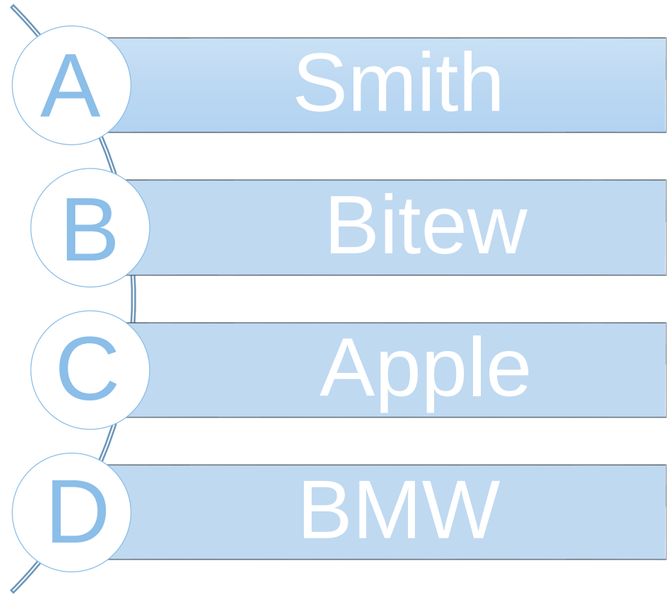
Let’s do a quick poll: Personal Trivia

What is my last name?

Let’s do a quick poll: Personal Trivia

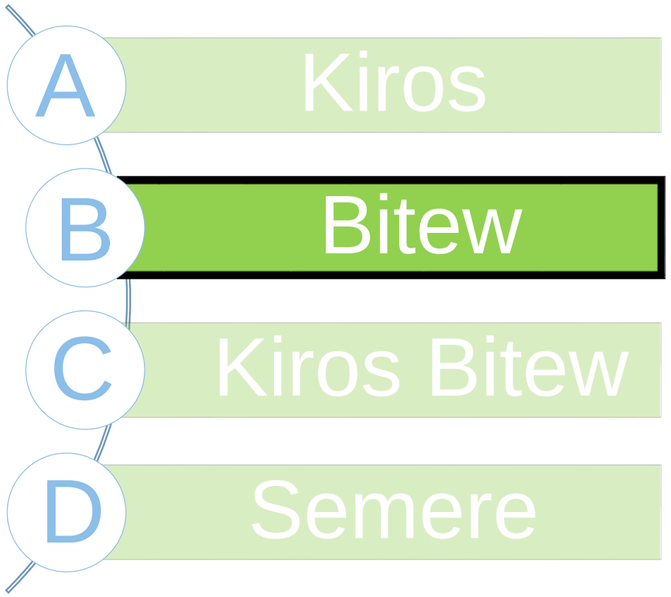
What is my last name?

What is my last name?

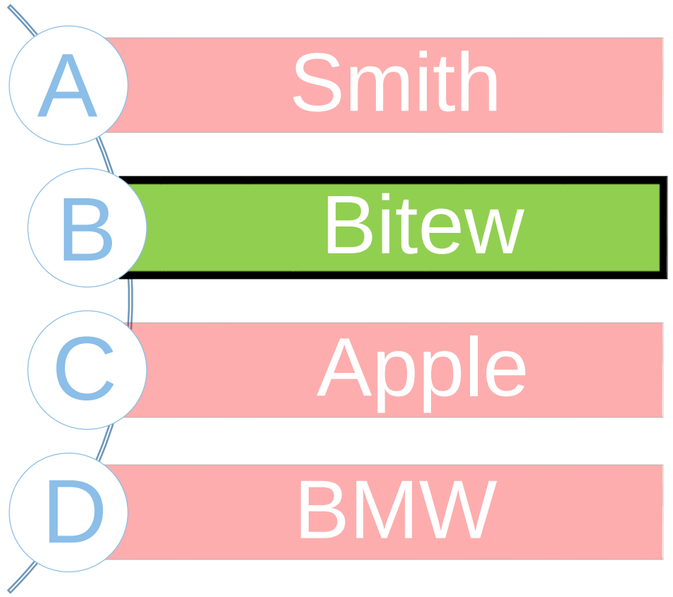


Let’s do a quick poll: Personal Trivia

What is my last name?

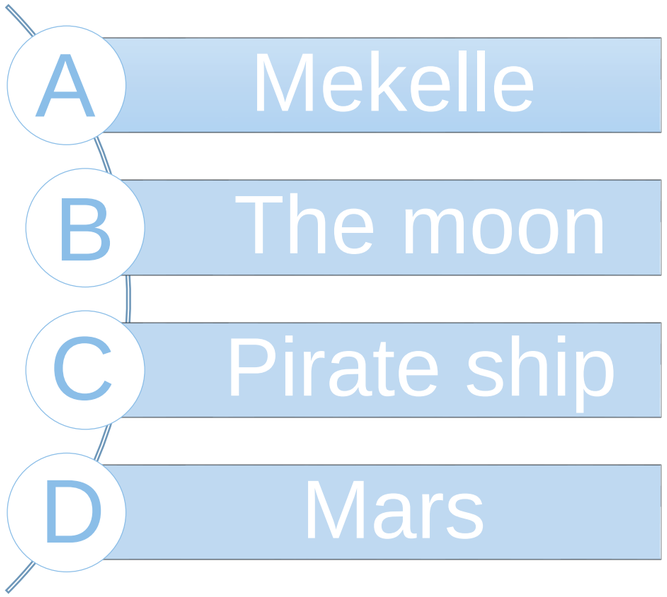


What is my last name?

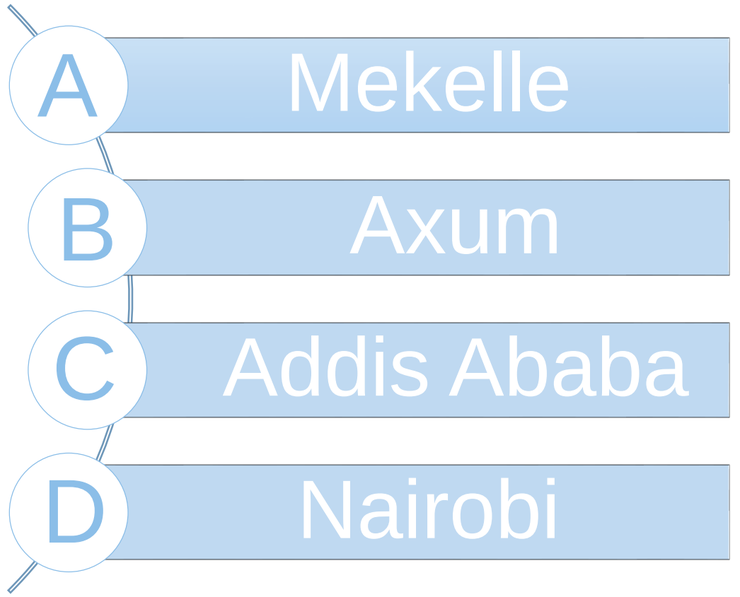


Let’s do a quick poll: Personal Trivia

Where am I originally from?

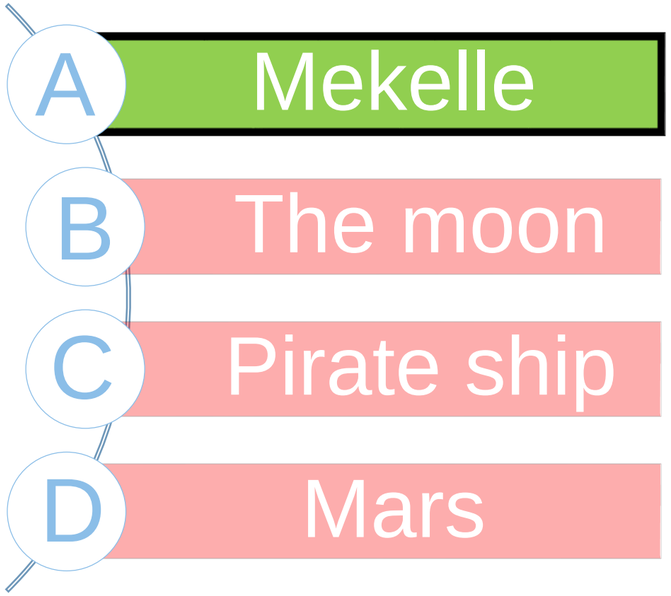


Where am I originally from?

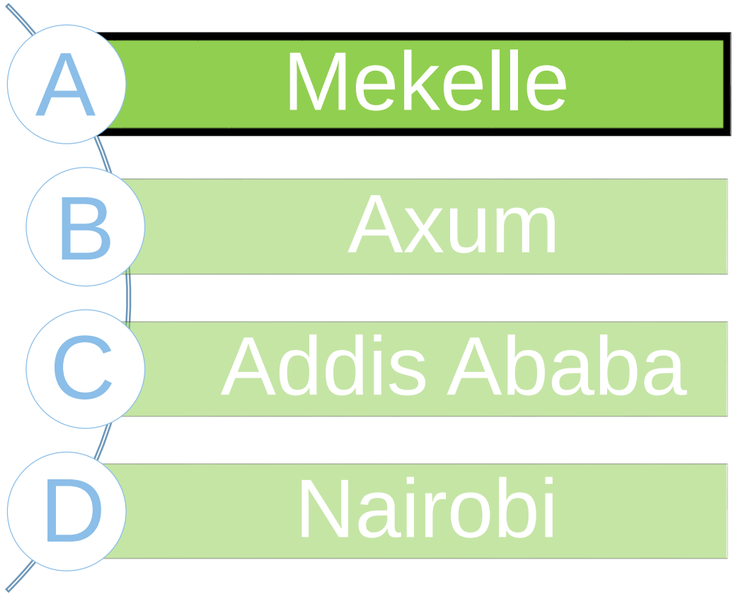


Let’s do a quick poll: Personal Trivia

Where am I originally from?



Where am I originally from?



Let’s do a quick poll: Personal Trivia

What year was I born?

A 1776

B 1892

C 2050

D 1992

What year was I born?

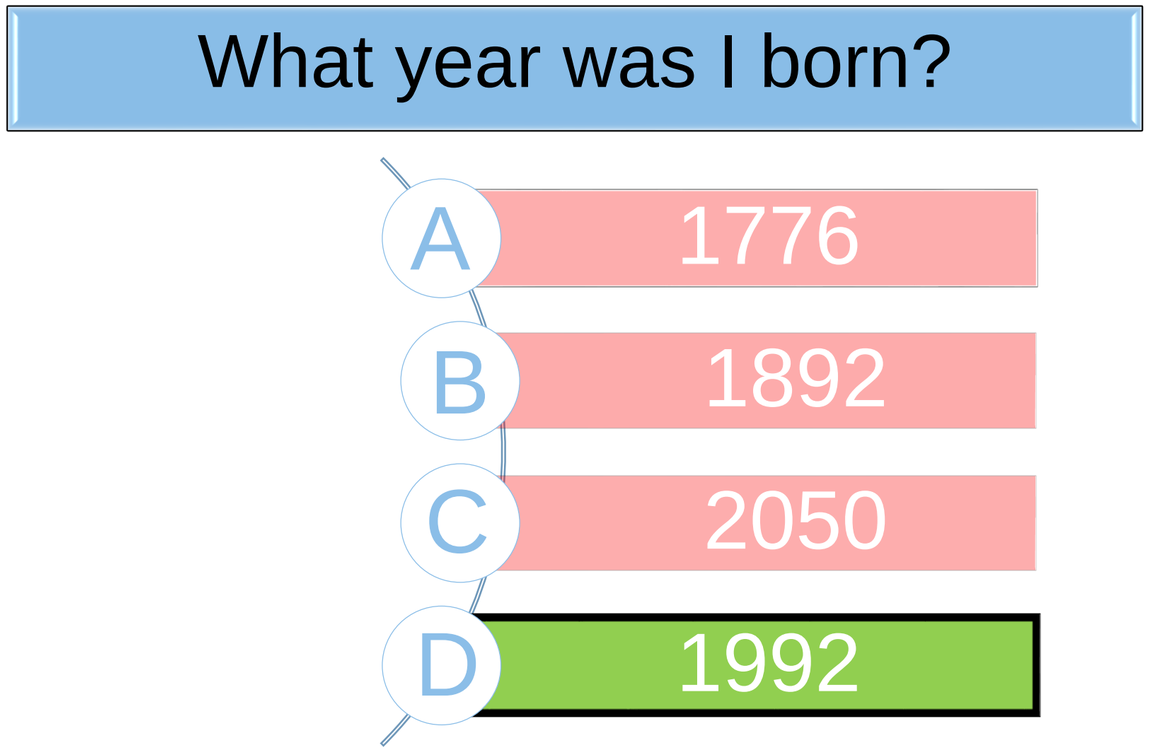
A 1984

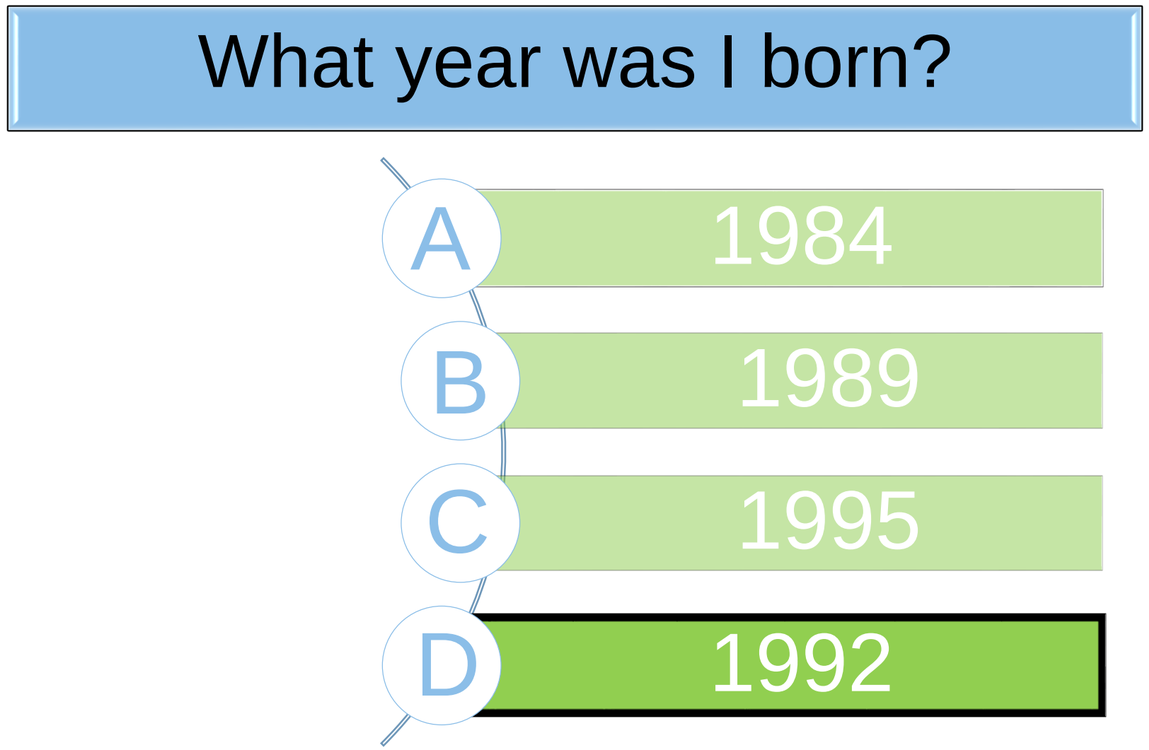
B 1989

C 1995

D 1992

Let’s do a quick poll: Personal Trivia





Let’s do a quick poll: Personal Trivia

My last name is **Bitew**, born in **1992**, and I am from **Mekelle**!

Introduction to

AI in Education

AI in Education

Support for teaching, learning, and school administration.

AI in Education

AI in Education

AI in Education

AI in Education

AI in Education

Examples

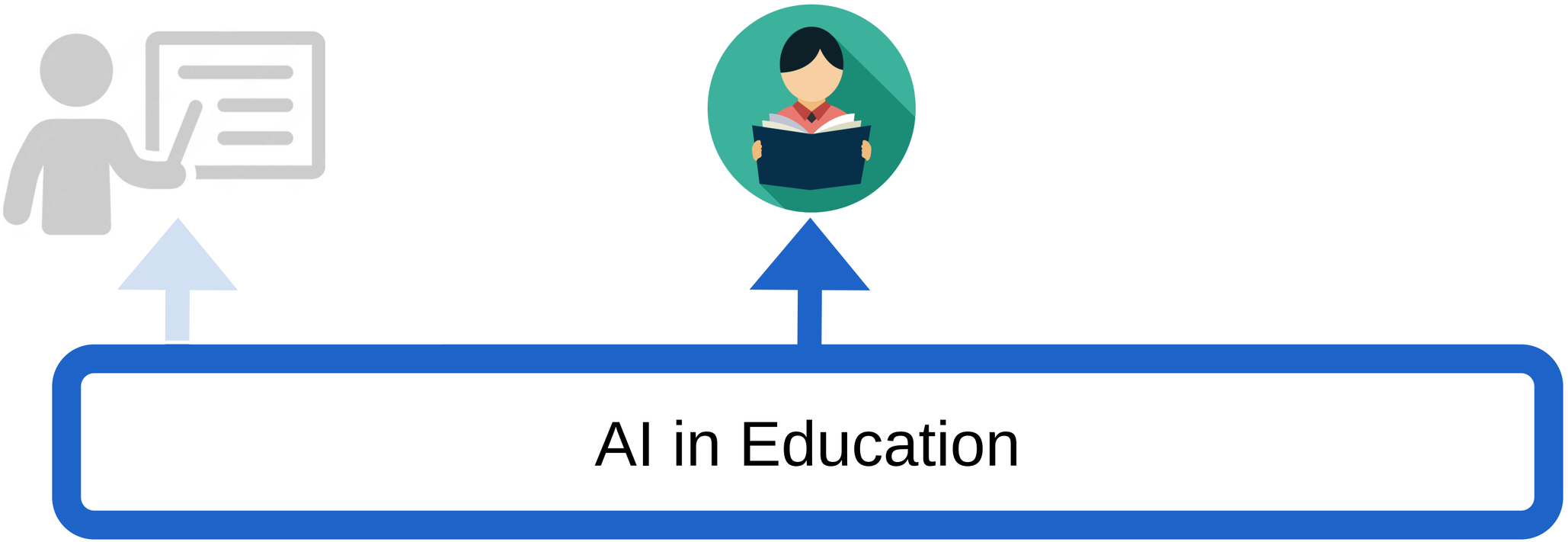
- FAQ Chatbot
- Essay Scoring

AI in Education

AI in Education

Examples

- FAQ Chatbot
- Essay Scoring

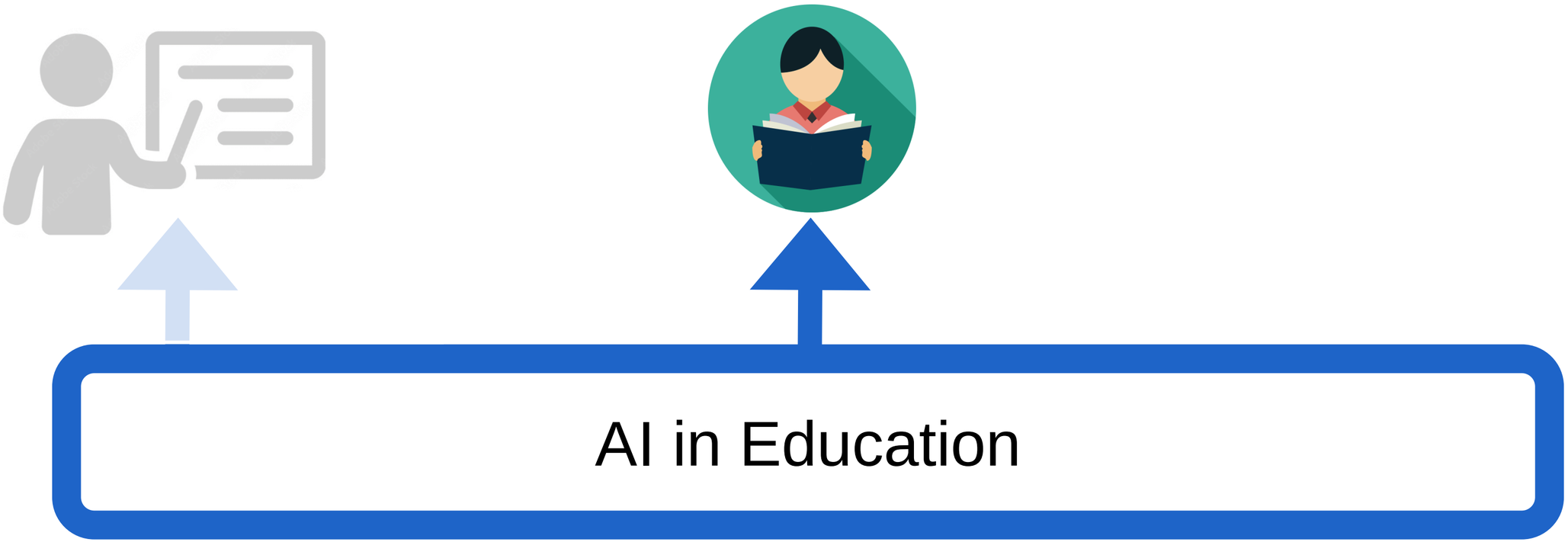


AI in Education

Examples

- FAQ Chatbot
- Essay Scoring

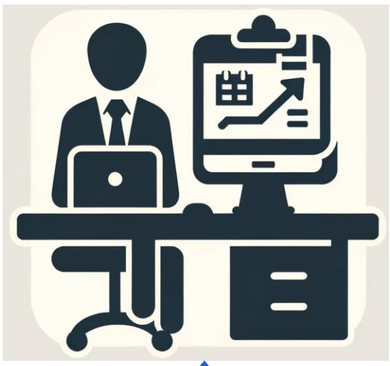
- duolingo
- Formative assessment



AI in Education

Examples

- FAQ Chatbot
- Essay Scoring
- duolingo
- Formative assessment

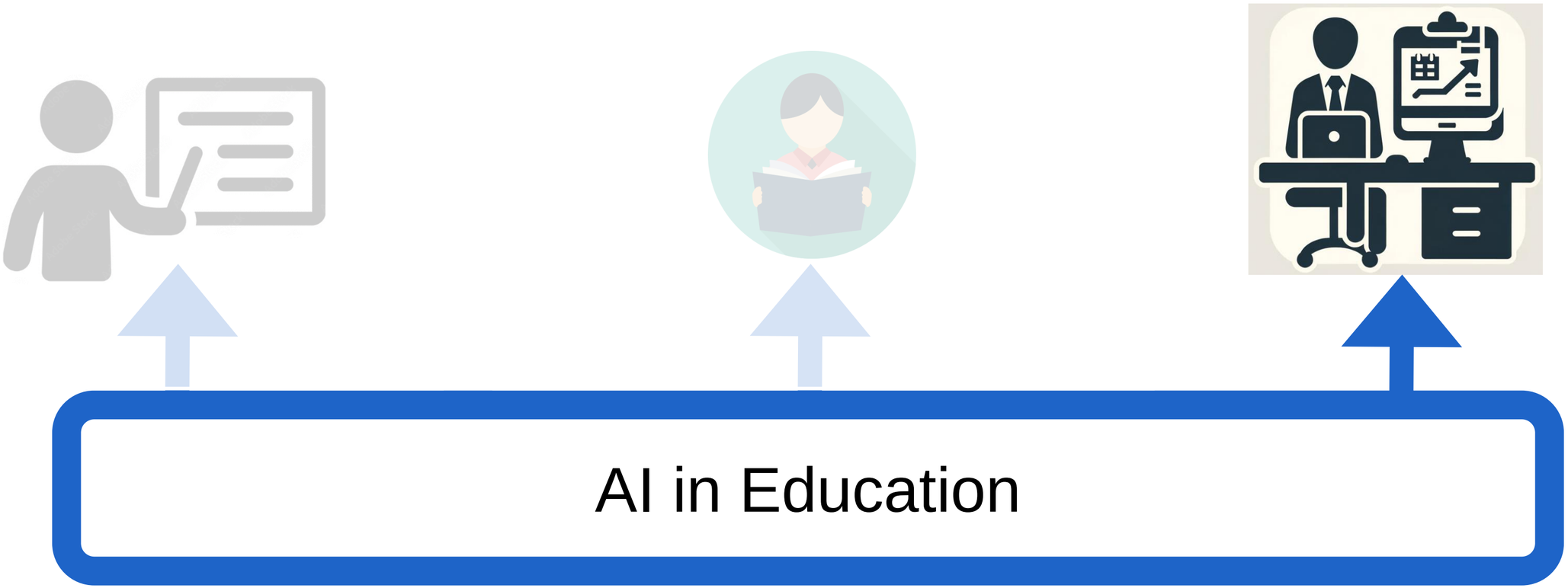


AI in Education

AI in Education

Examples

- FAQ Chatbot
- Essay Scoring
- duolingo
- Formative assessment
- Resource allocation



AI in Education

Focus of my research

AI in Education

Research Motivation

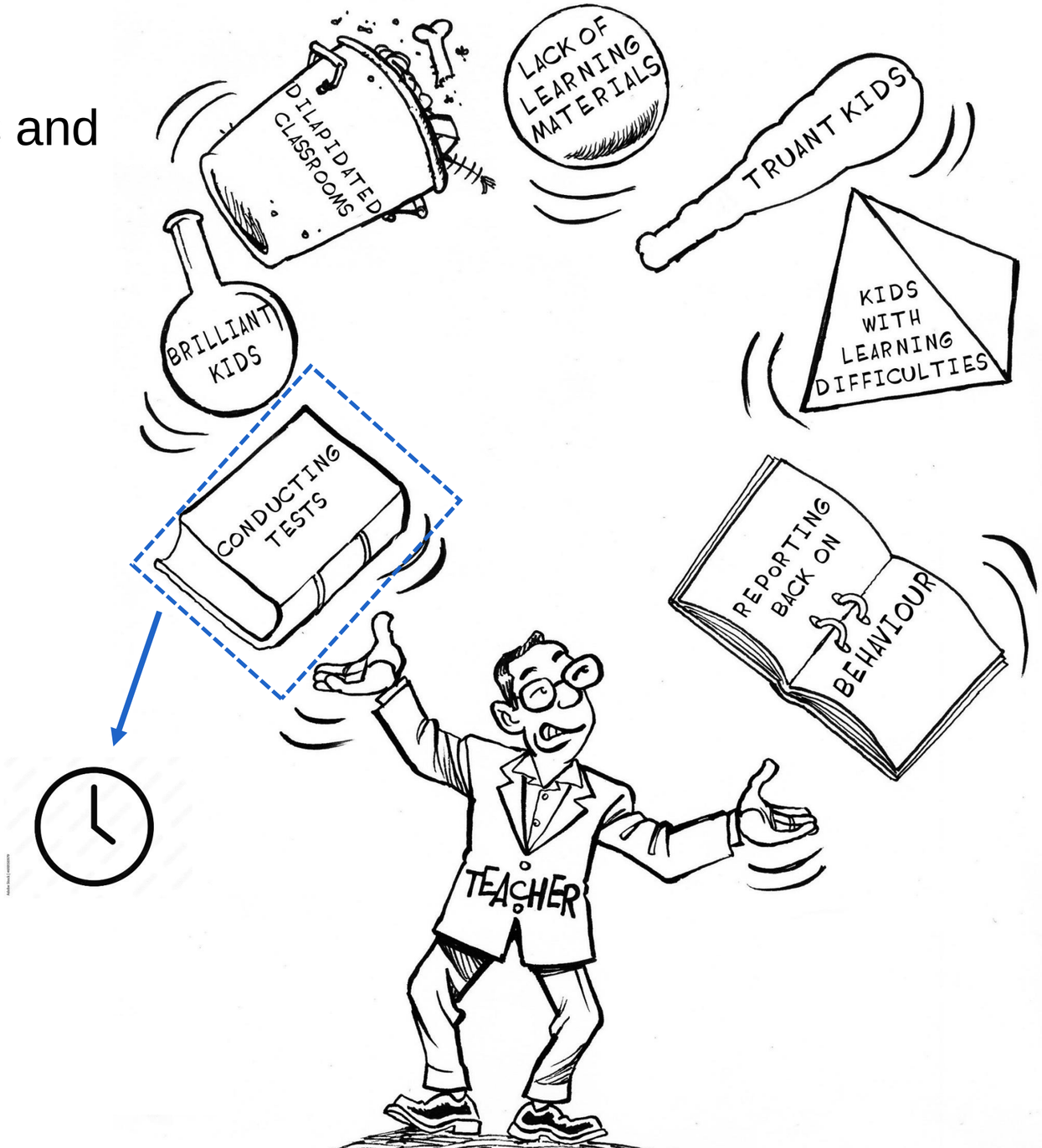
Burden on educators

- Heavy workload from creating exercises and tests

Burden on educators

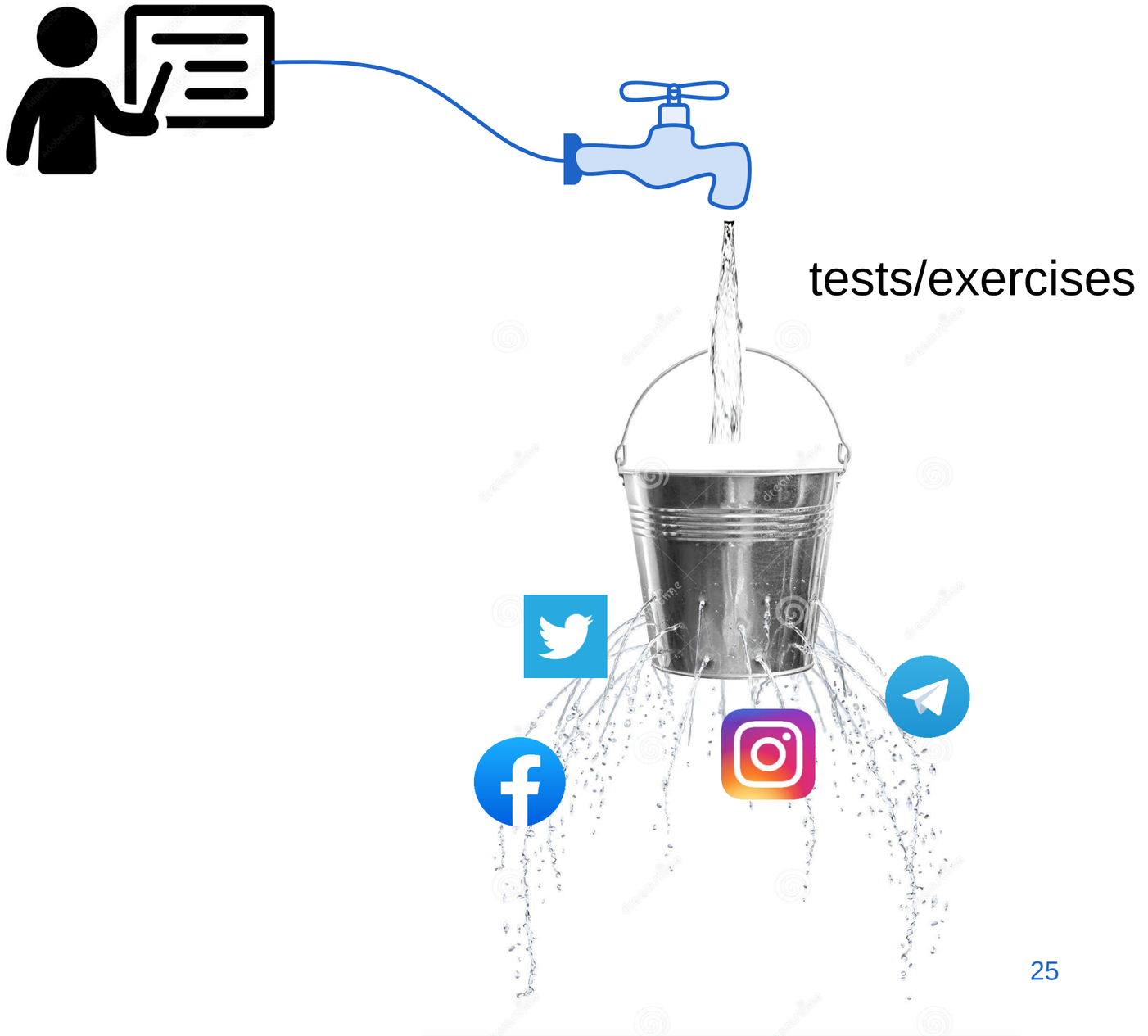
- Heavy workload from creating exercises and tests

- Formulating question requires
  - Skill
  - Experience
  - Time



Single-use dilemma

- Question sharing and leakage



Single-use dilemma

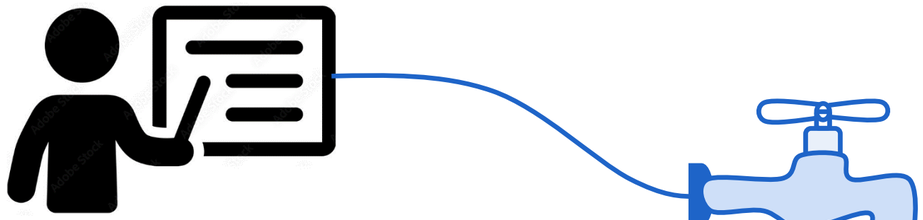
- Question sharing and leakage

- Problem in high-stakes:
  - Certification, Job tests

Single-use dilemma

- Question sharing and leakage

- Problem in high-stakes:
  - Certification, Job tests



tests/exercises

ETS TOEFL

USMLE SM
United States Medical Licensing Exam

Shift to digital education

Rapid digitization of educational materials, especially after COVID19

Shift to digital education

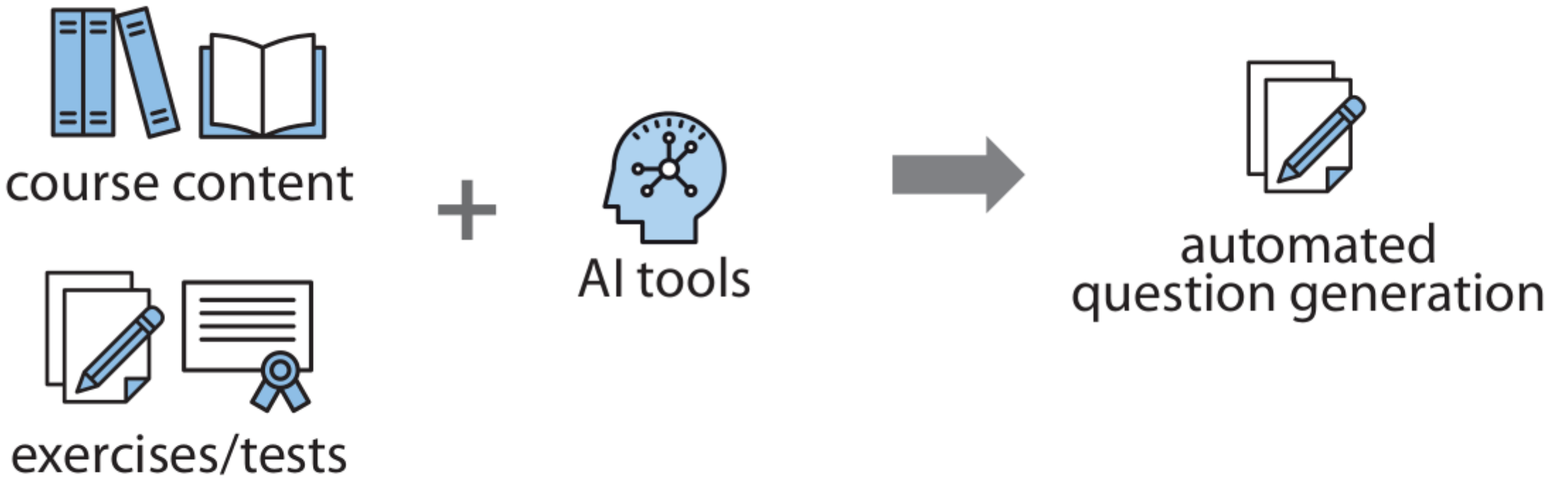
Rapid digitization of educational materials, especially after COVID19

course content

exercises/tests

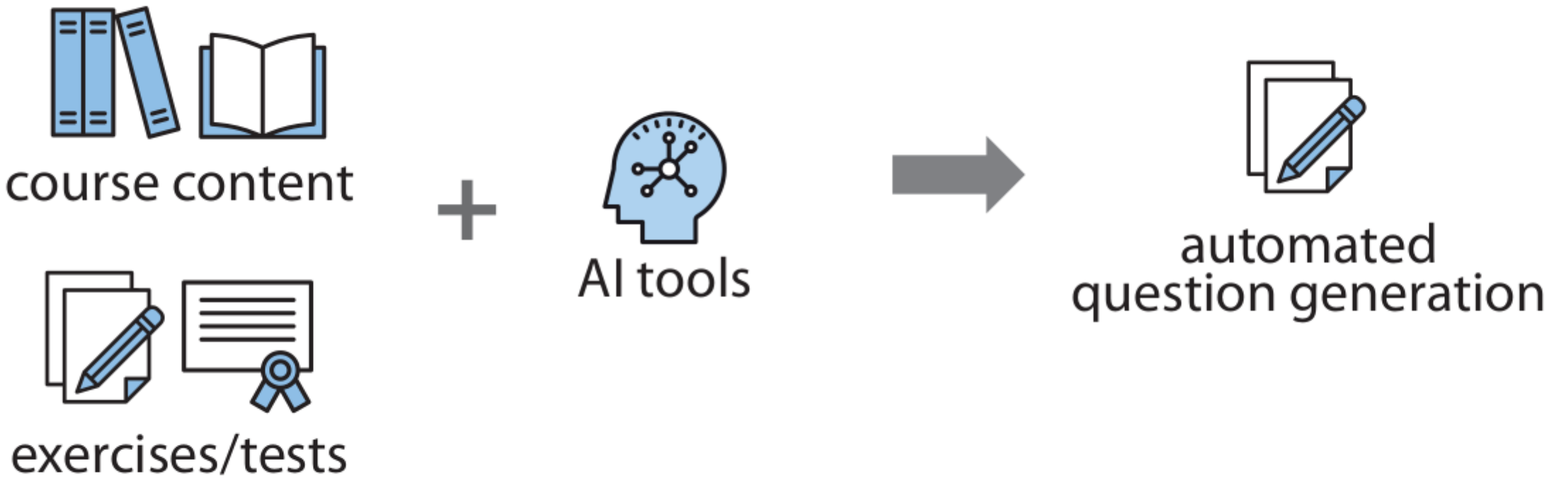
Shift to digital education

Rapid digitization of educational materials, especially after COVID19



Shift to digital education

Rapid digitization of educational materials, especially after COVID19

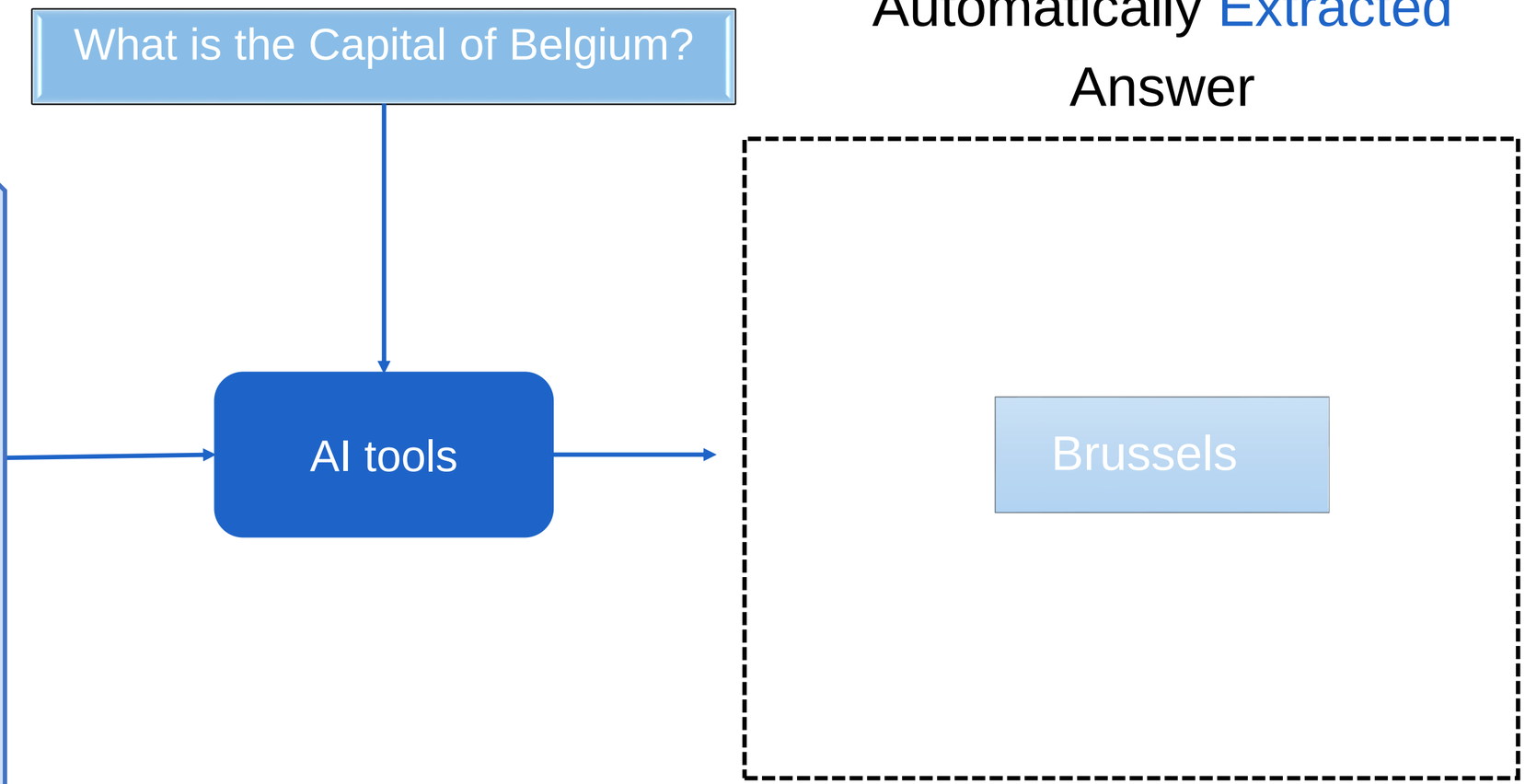


Goal: automatic educational question generation to support teachers & students

Concrete Example 1 (Question Answering)

Paragraph about Belgium

Belgium is a small country in Europe distinguished by its remarkable combination of historical significance and contemporary governance. Brussels is the capital city of Belgium. Other major cities include Antwerp, which is known for its diamond trade; Ghent, which has intact medieval buildings; and Liège, a city with a rich industrial past and culinary fame for its delicious waffles.



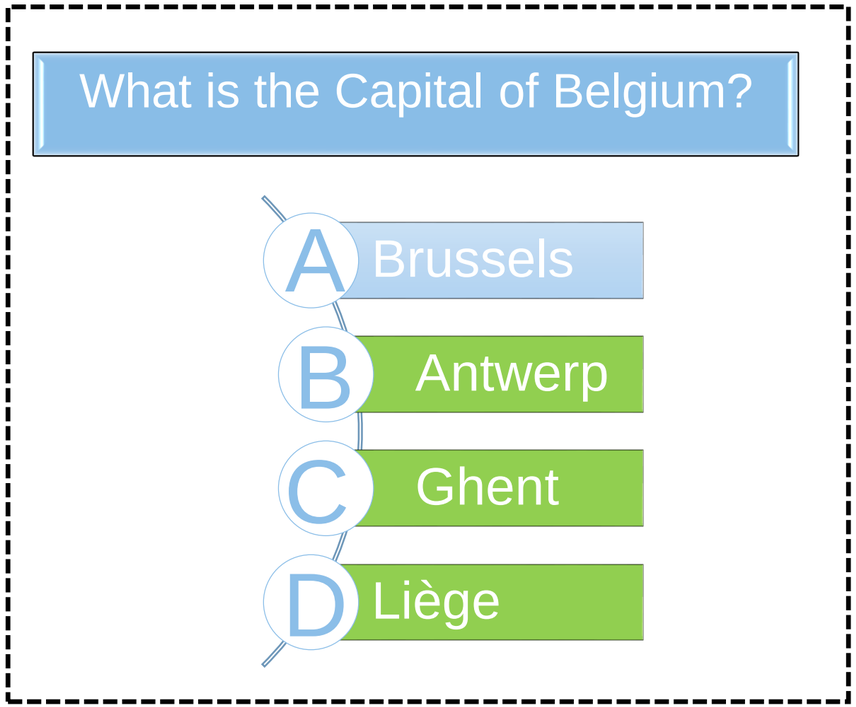
Concrete Example 2 (Question Generation)

Paragraph about Belgium

Belgium is a small country in Europe distinguished by its remarkable combination of historical significance and contemporary governance. Brussels is the capital city of Belgium. Other major cities include Antwerp, which is known for its diamond trade; Ghent, which has intact medieval buildings; and Liège, a city with a rich industrial past and culinary fame for its delicious waffles.

AI tools

Automatically Generated Multiple-Choice Question



Input and output in human language

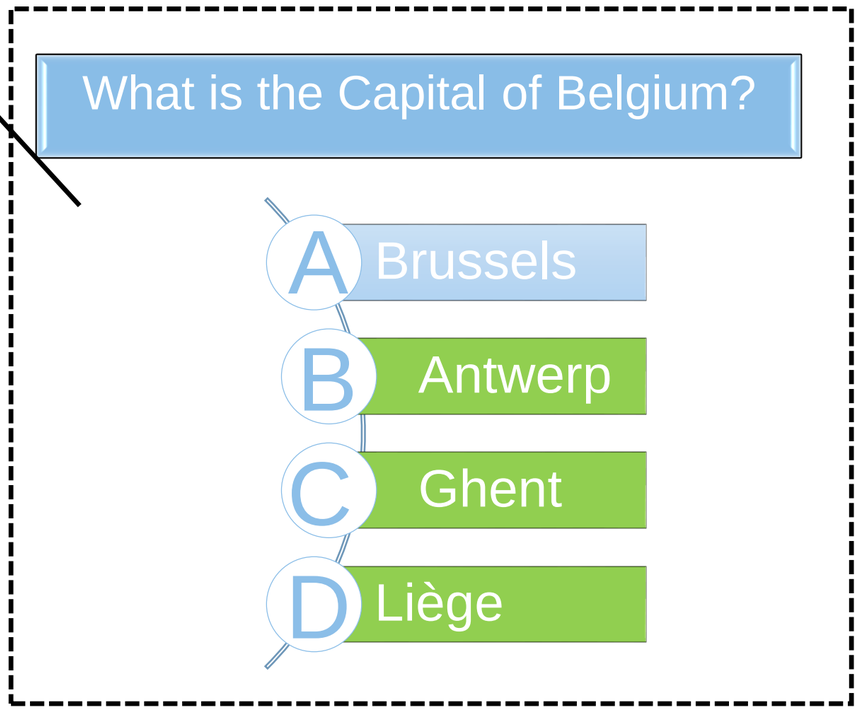
Paragraph about Belgium

Belgium is a small country in Europe distinguished by its remarkable combination of historical significance and contemporary governance. Brussels is the capital city of Belgium. Other major cities include Antwerp, which is known for its diamond trade; Ghent, which has intact medieval buildings; and Liège, a city with a rich industrial past and culinary fame for its delicious waffles.

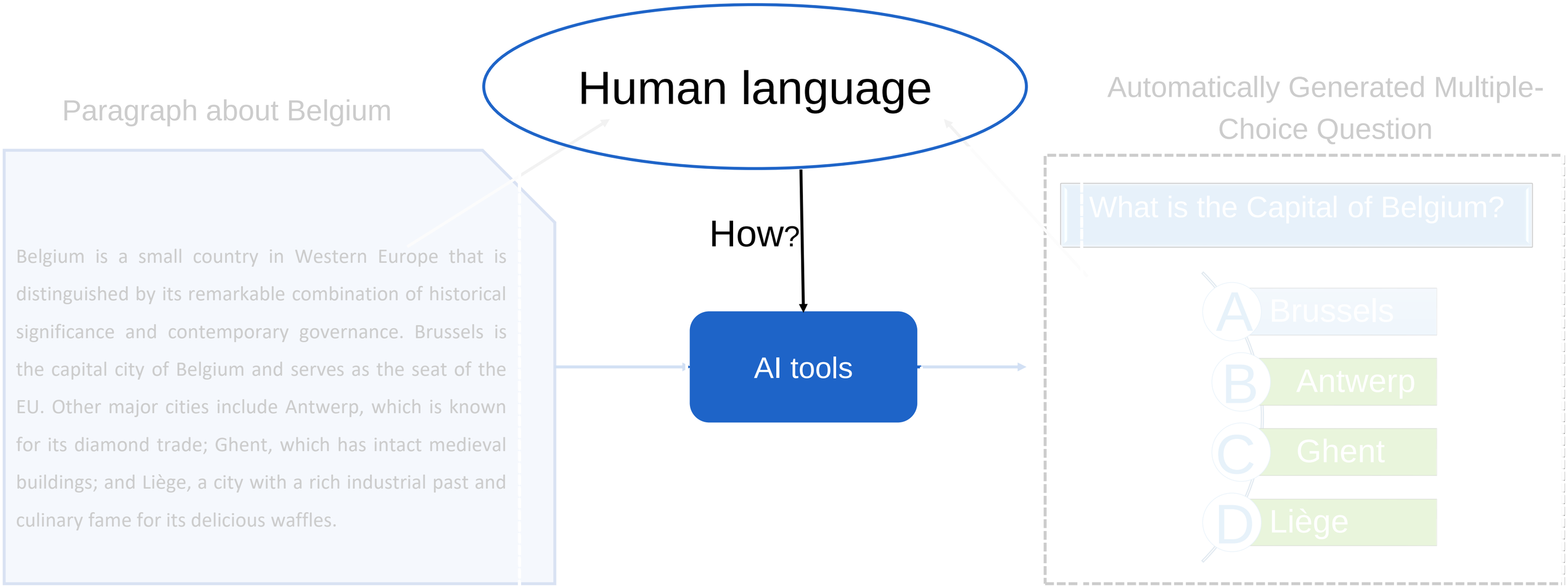
Human language

AI tools

Automatically Generated Multiple-Choice Question



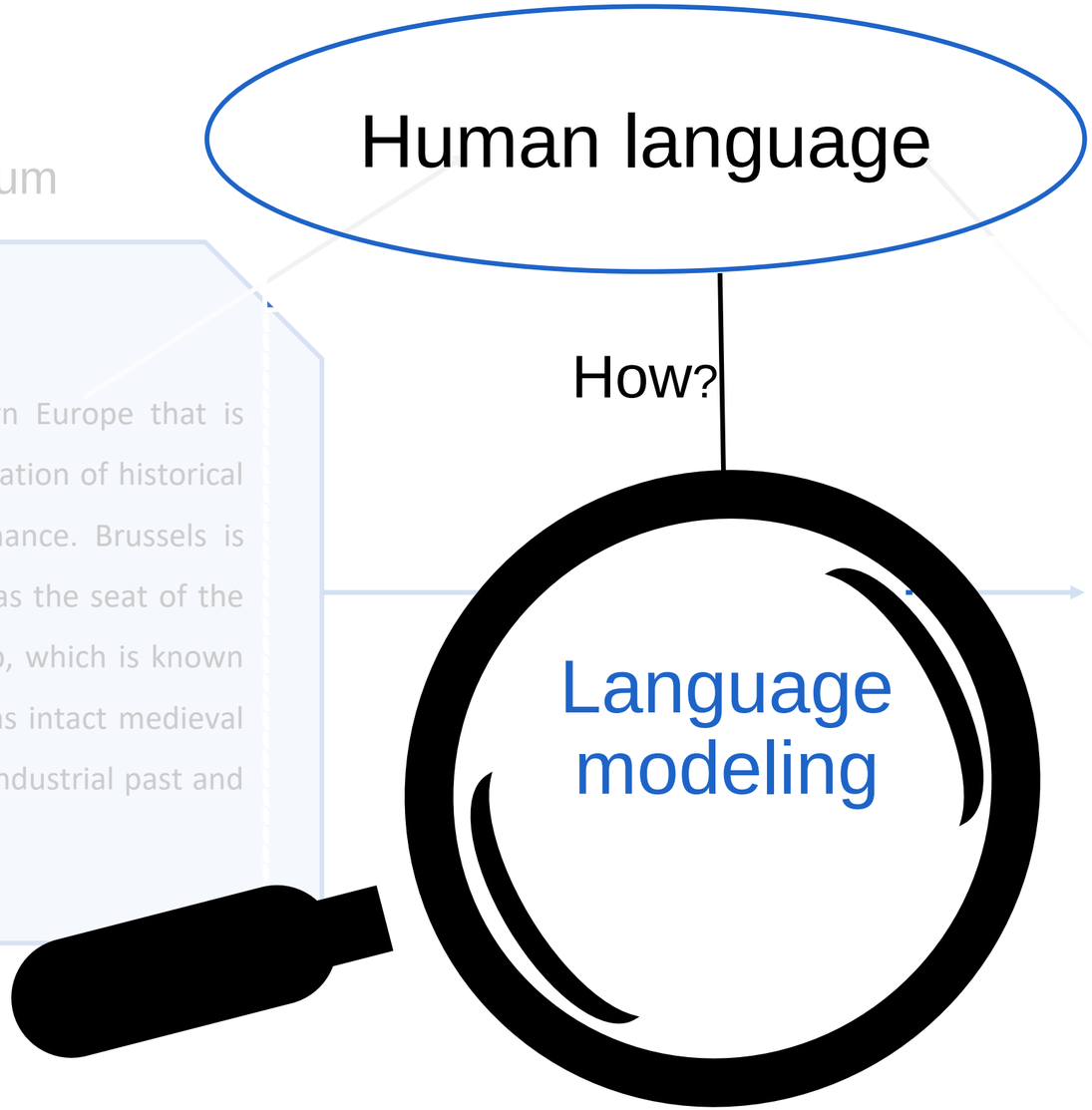
How do computers understand language?



Language models

Paragraph about Belgium

Belgium is a small country in Western Europe that is distinguished by its remarkable combination of historical significance and contemporary governance. Brussels is the capital city of Belgium and serves as the seat of the EU. Other major cities include Antwerp, which is known for its diamond trade; Ghent, which has intact medieval buildings; and Liège, a city with a rich industrial past and culinary fame for its delicious waffles.



Automatically Generated Multiple-Choice Question

What is the Capital of Belgium?

A Brussels

B Antwerp

C Ghent

D Liège

Language Modeling

Language models (LMs)

Language model is a *probabilistic model* of the human language.

Language models (LMs)

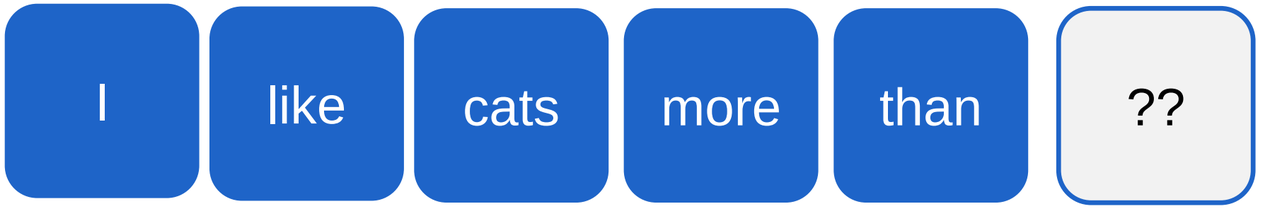
Language model is a *probabilistic model* of the human language.

1. Predicting the next word

Language models (LMs)

Language model is a *probabilistic model* of the human language.

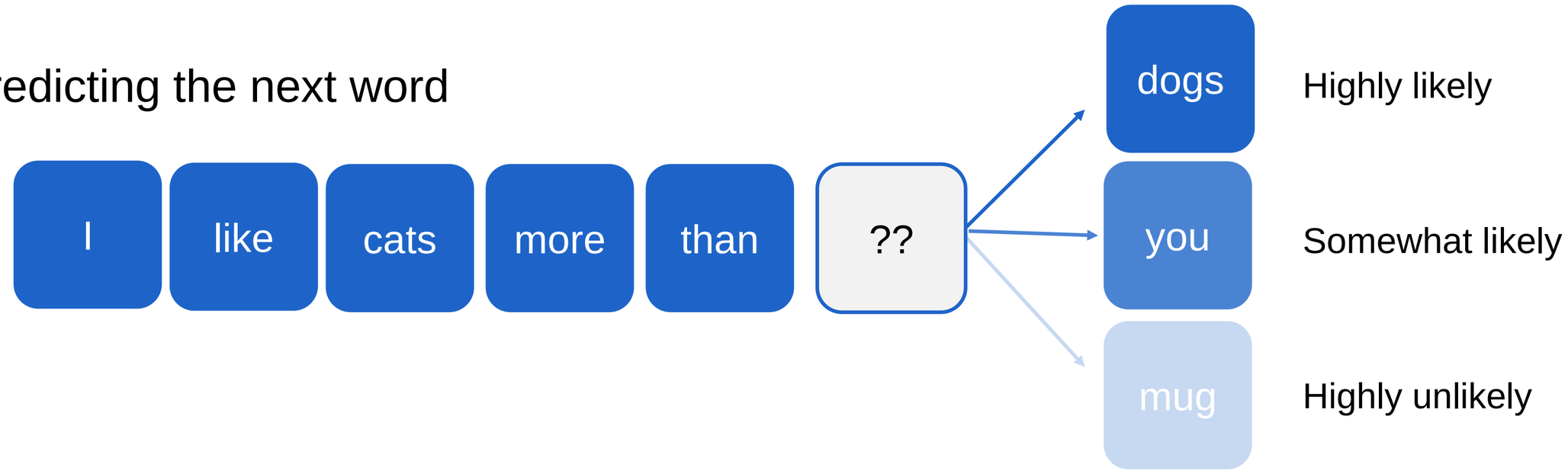
1. Predicting the next word



Language models (LMs)

Language model is a *probabilistic model* of the human language.

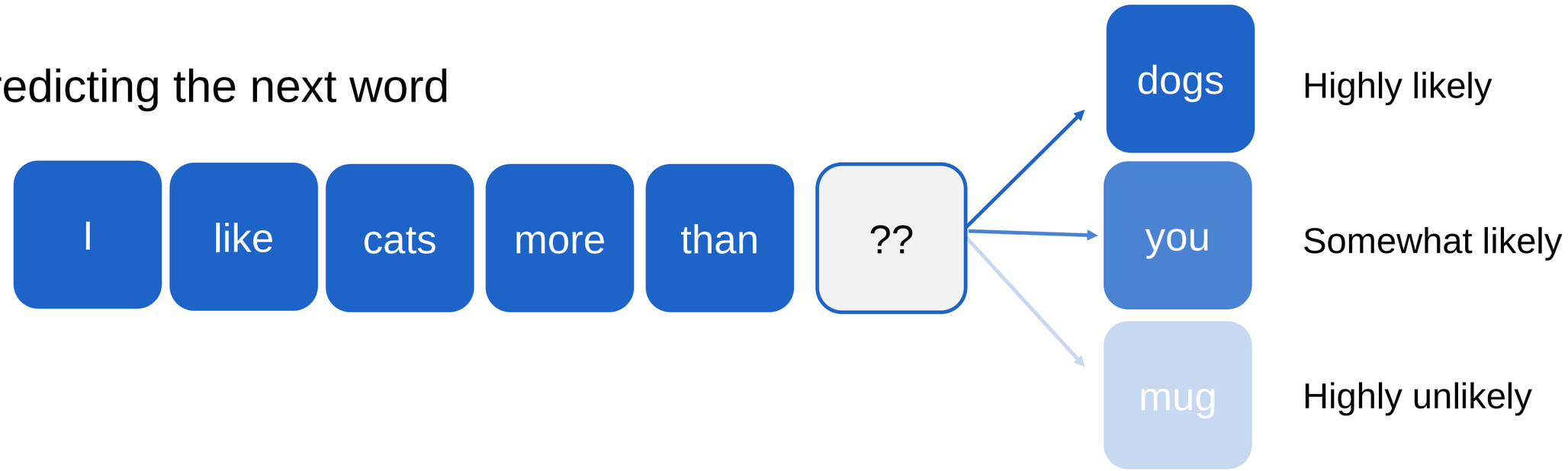
1. Predicting the next word



Language models (LMs)

Language model is a *probabilistic model* of the human language.

1. Predicting the next word

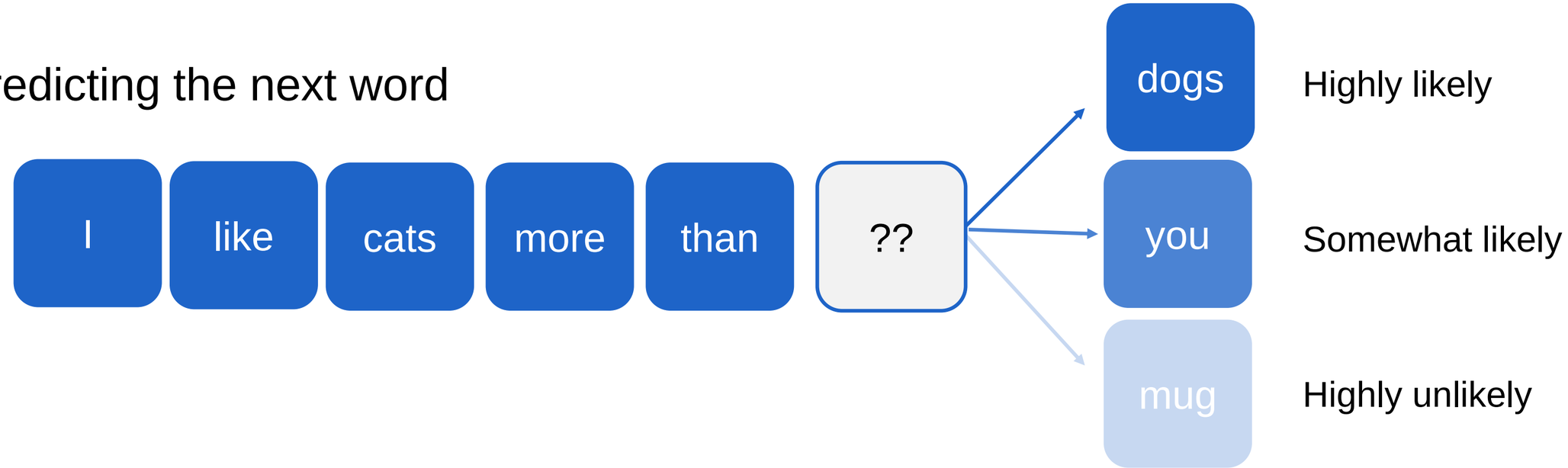


2. Likelihood of the sentence

Language models (LMs)

Language model is a *probabilistic model* of the human language.

1. Predicting the next word



2. Likelihood of the sentence

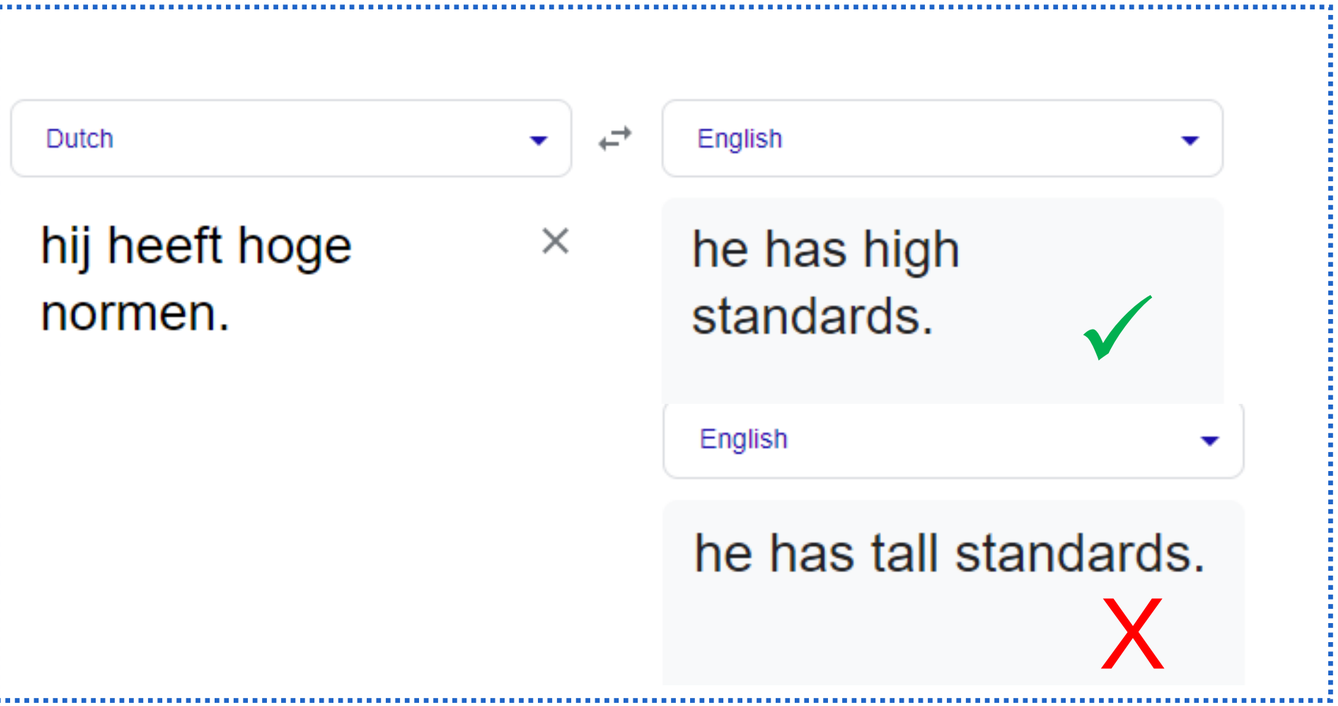
P (The sky is the limit.) >> P (limit is The the sky.)

Why assign probabilities?

Why assign probabilities?

Real-world applications

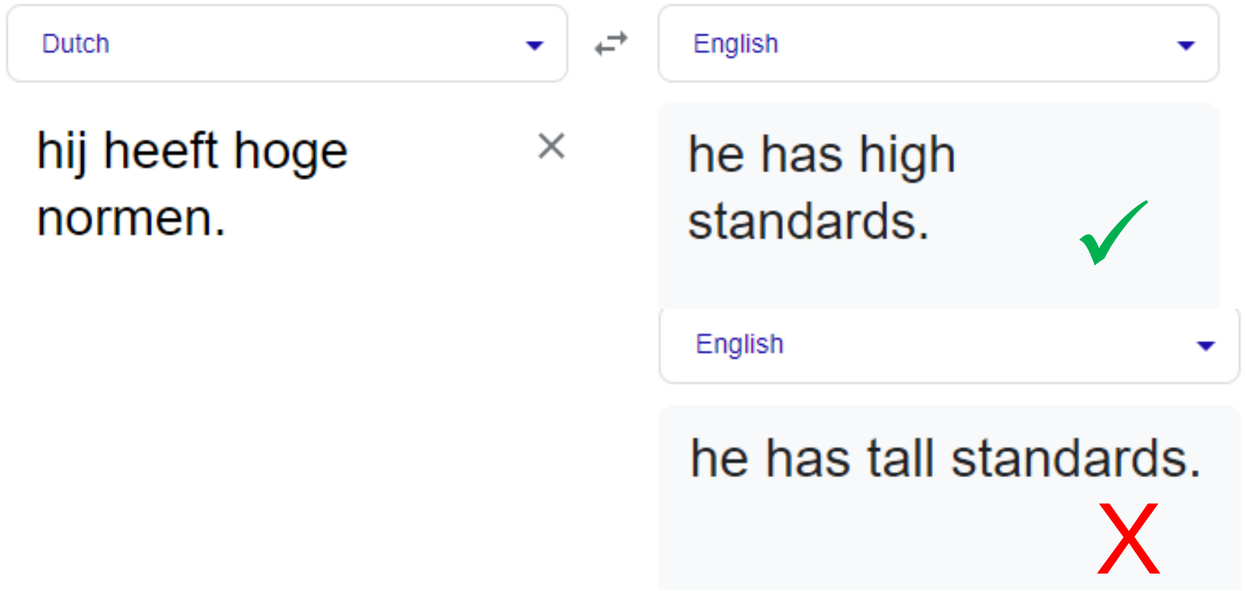
Machine Translation



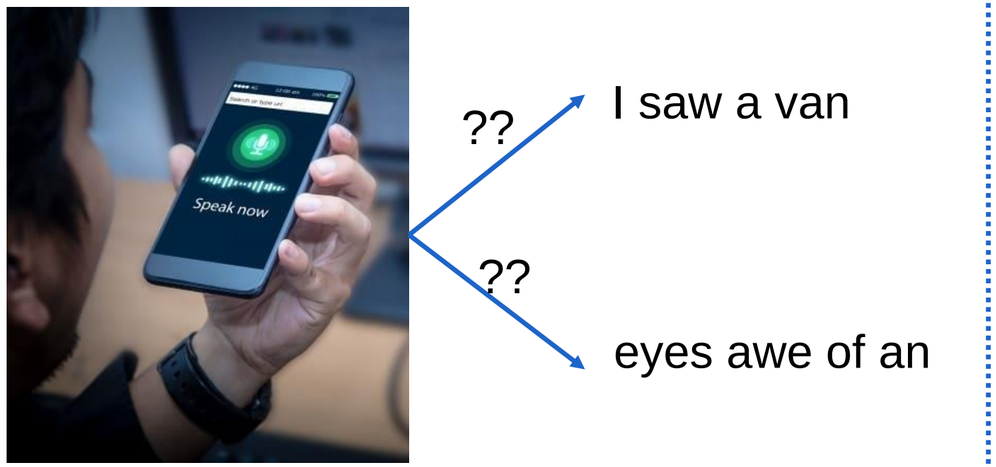
Why assign probabilities?

Real-world applications

Machine Translation



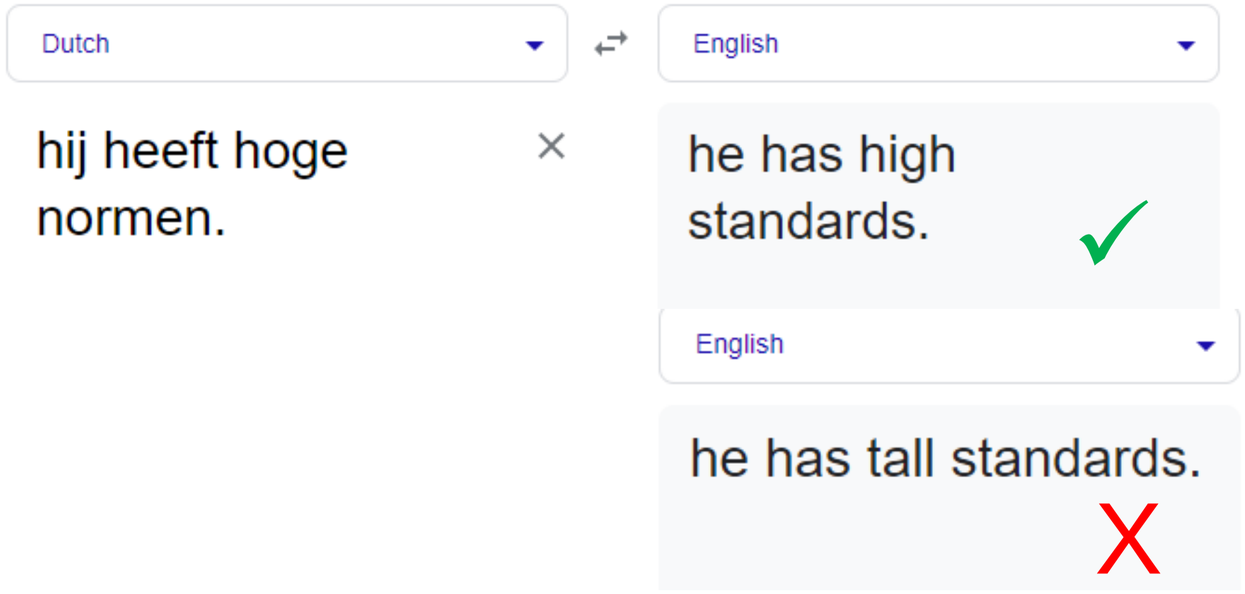
Speech recognition



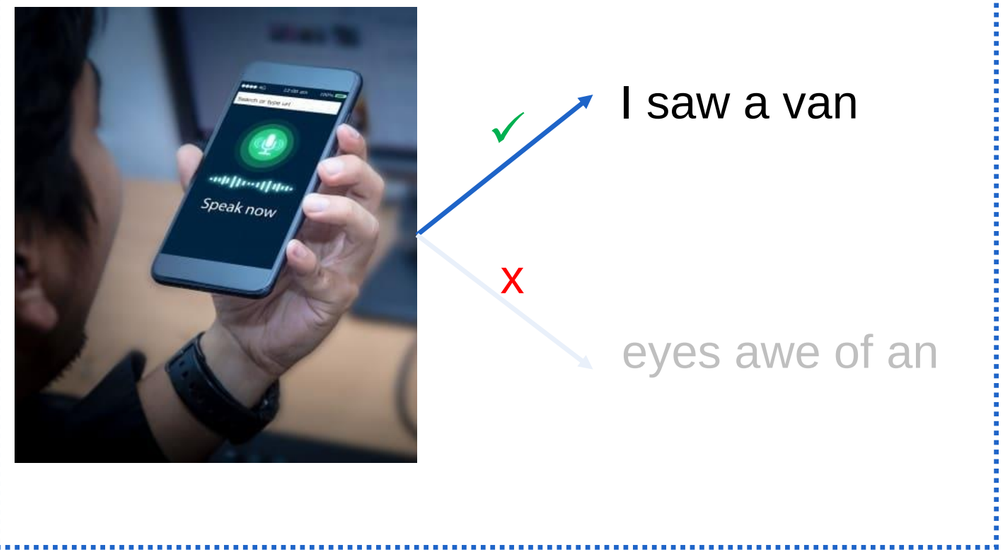
Why assign probabilities?

Real-world applications

Machine Translation



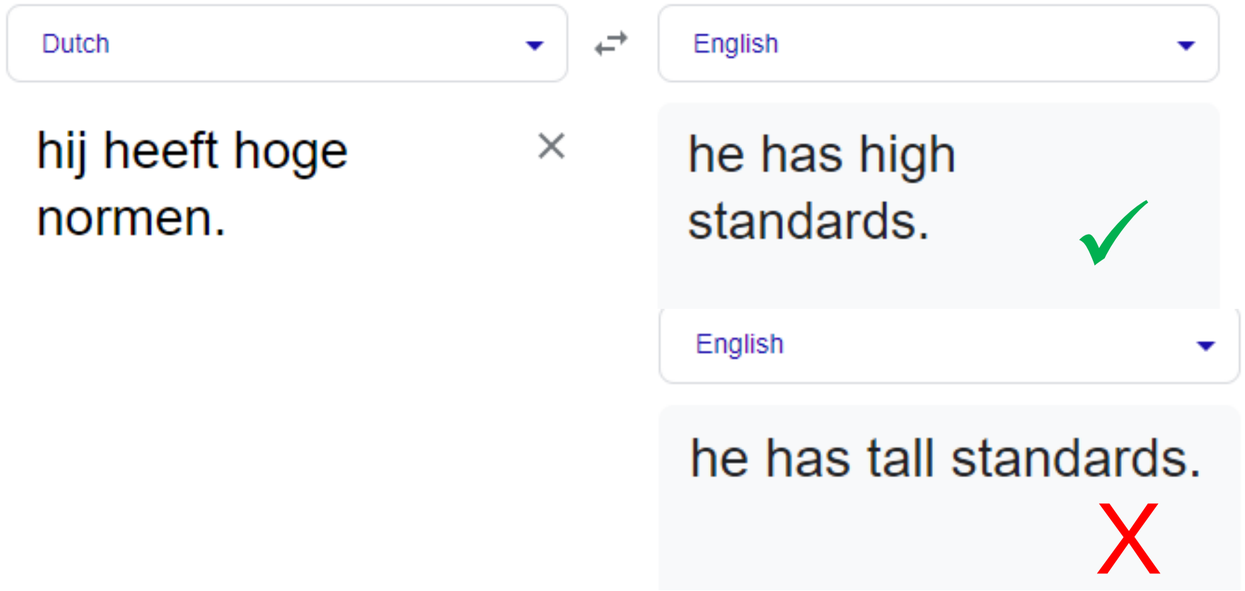
Speech recognition



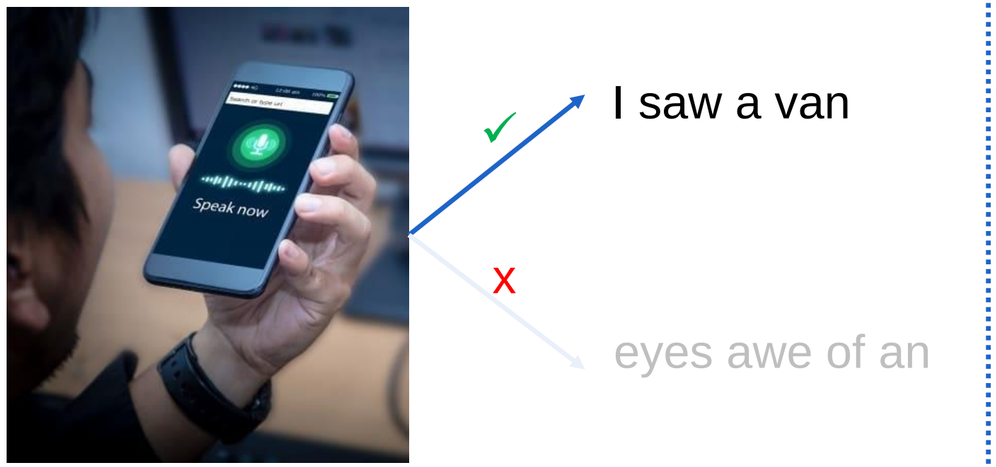
Why assign probabilities?

Real-world applications

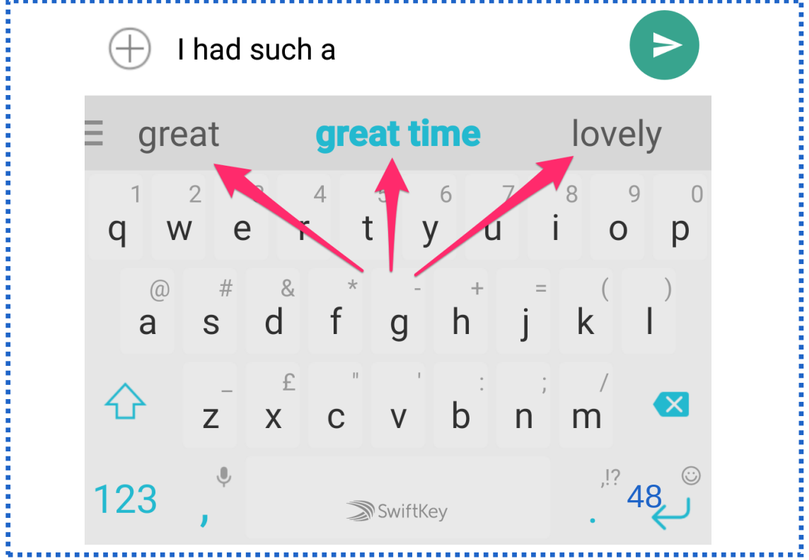
Machine Translation



Speech recognition



Sentence auto-completion



GHENT UNIVERSITY imec

How do LMs assign probabilities?

How do LMs ~~assign probabilities?~~ Learn?

How do LMs Learn?

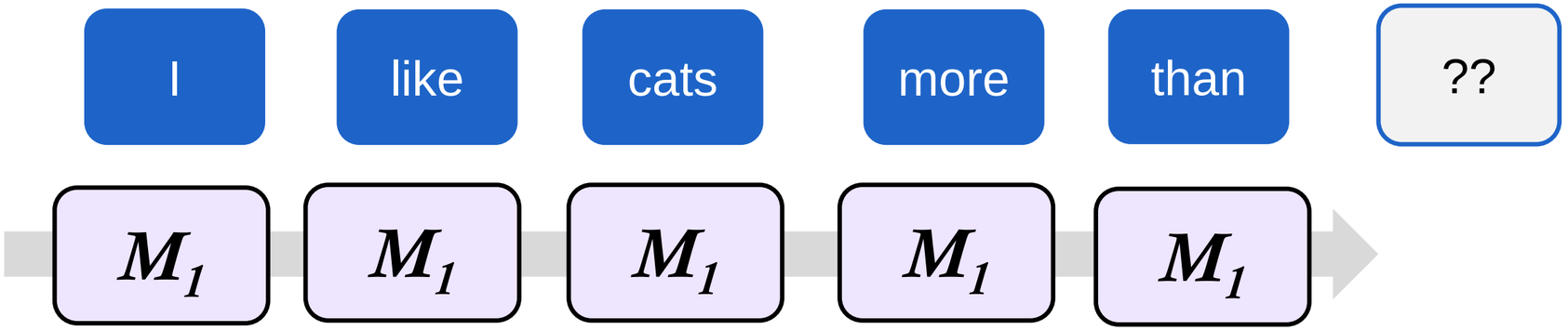
Two common tasks to teach LMs

1. *Causal Language modeling*

2. *Masked language modeling*

1 Causal Language Modeling

Predict the next word from a vocabulary using an untrained model M_1

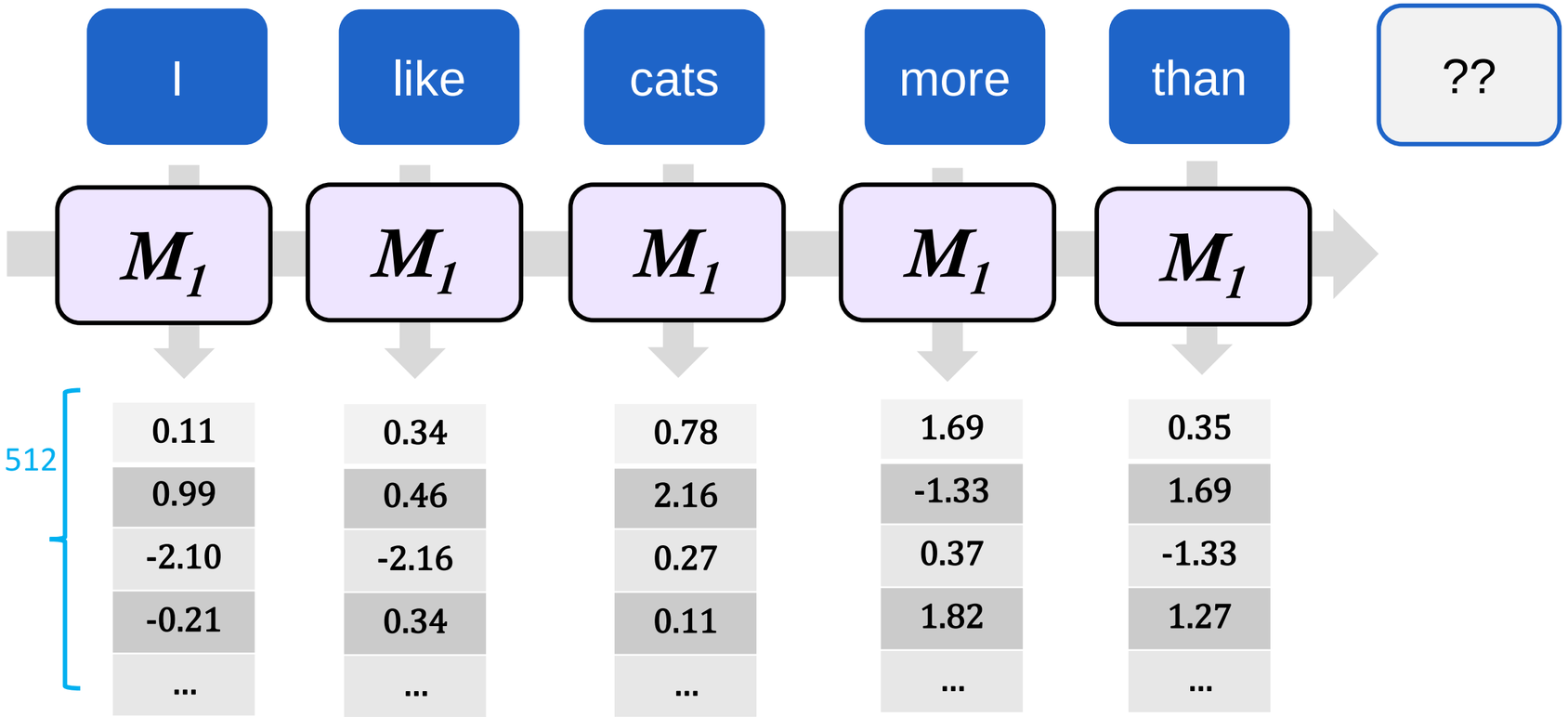


Vocabulary

dogs
potatoes
mug
you
anything
-
-
-

1 Causal Language Modeling

First the model M_0 converts each word to a numerical vector (e.g., 512)

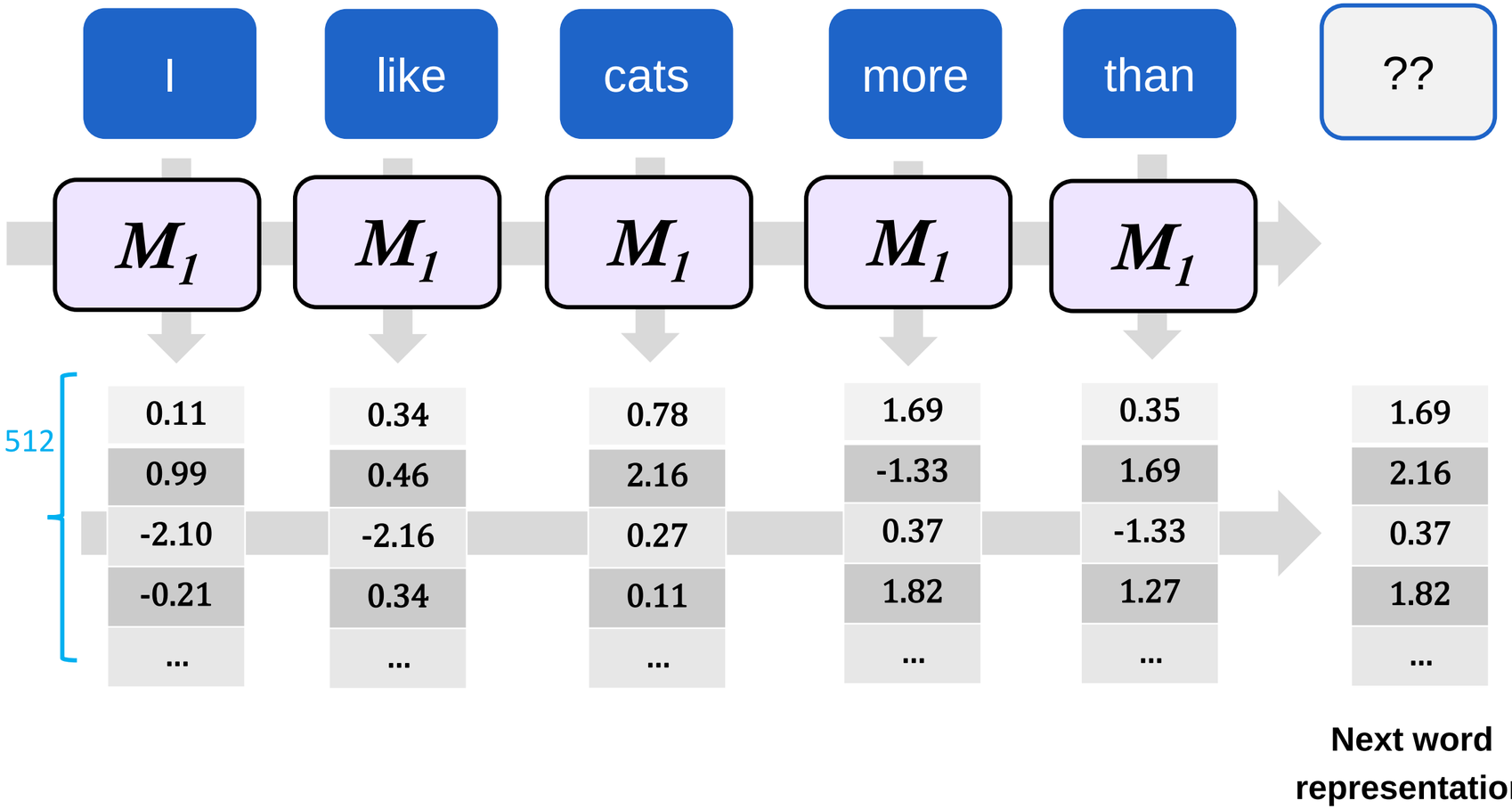


Vocabulary

dogs
potatoes
mug
you
anything
-
-
-

1 Causal Language Modeling

Combine all the word vectors into one **next word** representation vector

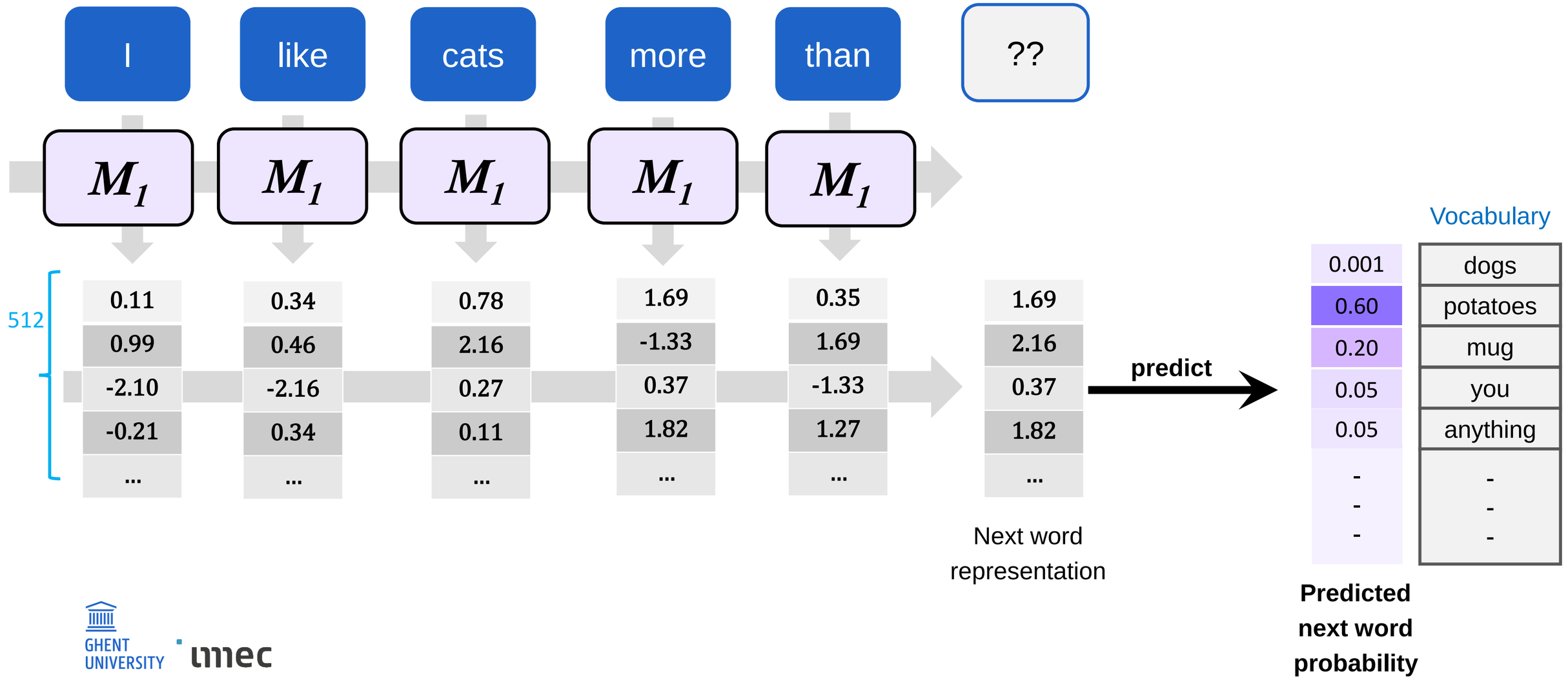


Vocabulary

dogs
potatoes
mug
you
anything
-
-
-

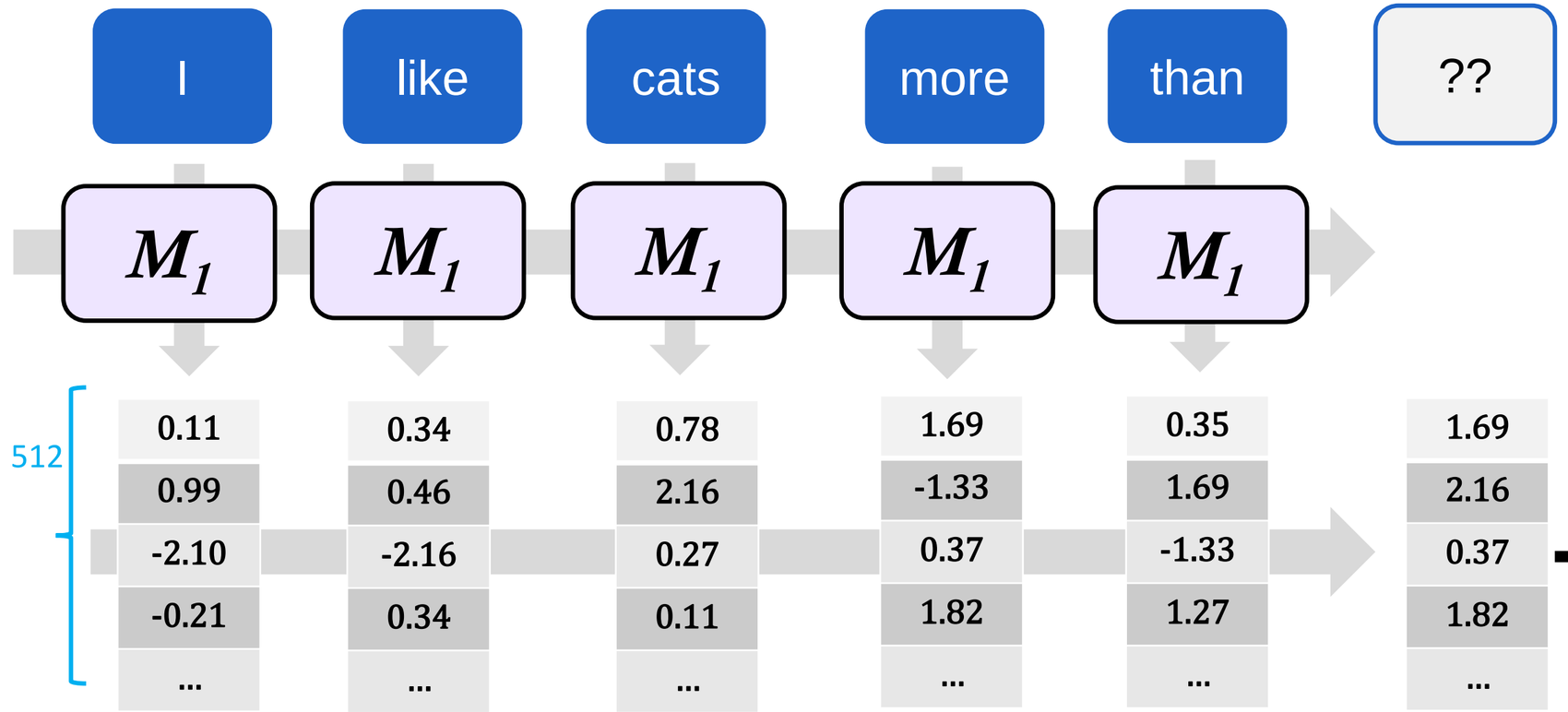
1 Causal Language Modeling

Using the next word representation, **predict** the probability of next word



1 Causal Language Modeling

Mismatch between prediction and truth. How to **solve** this?



Prediction != truth

predict

Vocabulary

0.001	dogs	1
0.60	potatoes	0
0.20	mug	0
0.05	you	0
0.05	anything	0
- - -	- - -	- - -

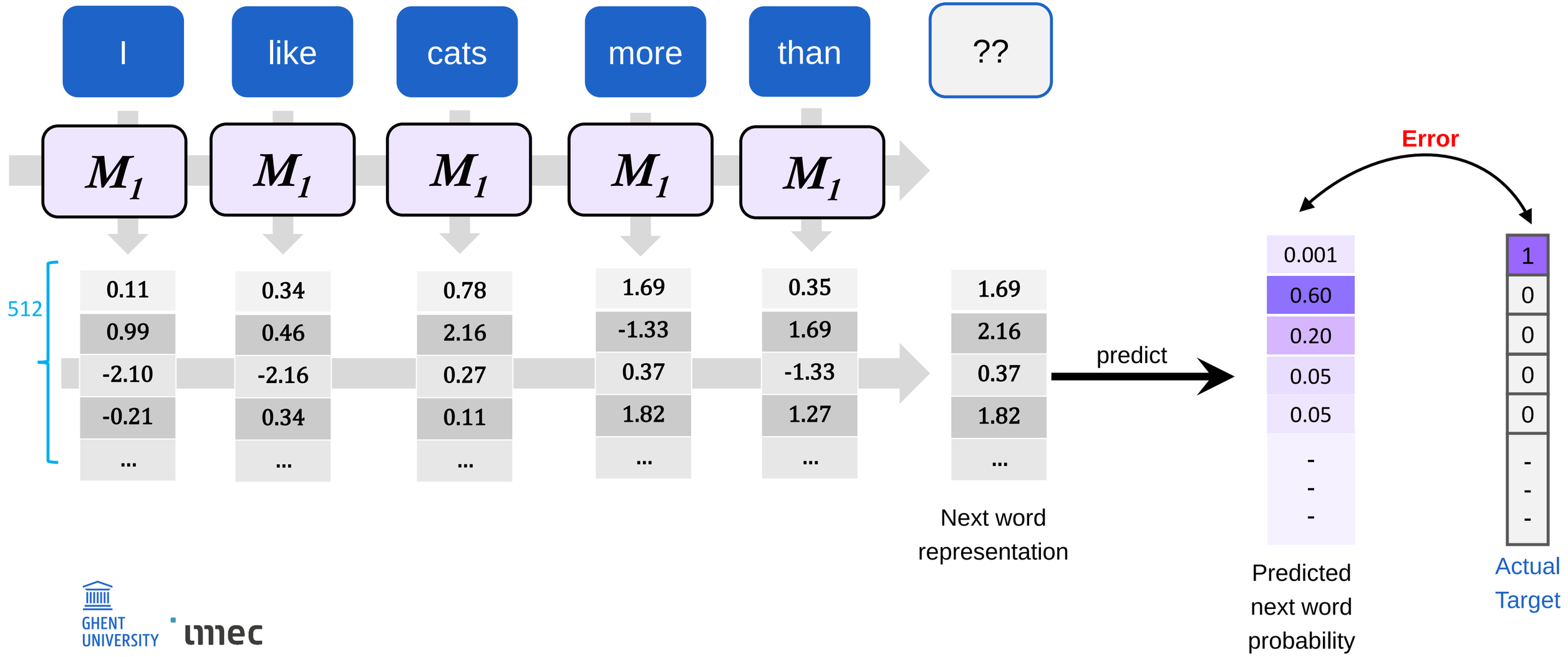
Predicted next word probability

Actual Target

GHENT UNIVERSITY imec

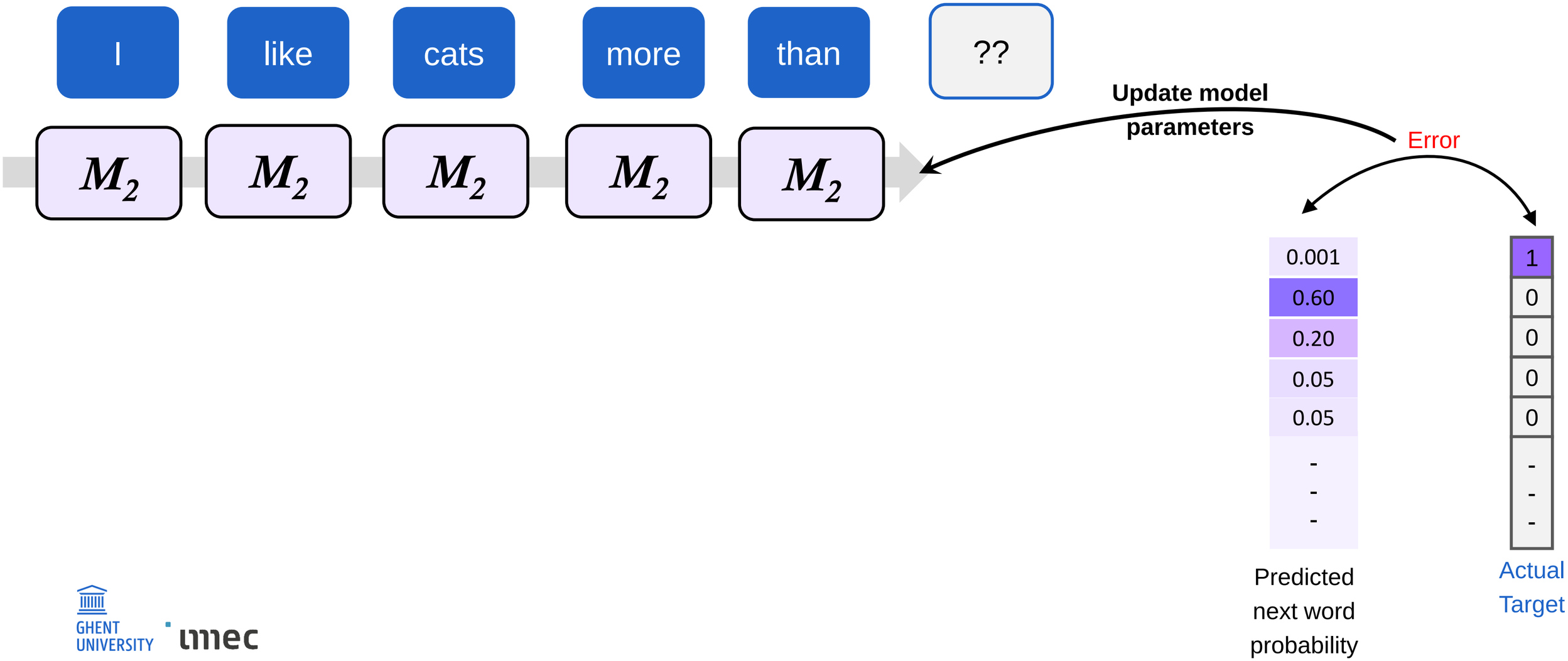
1 Causal Language Modeling

Calculate the **error (loss)** between the prediction and the actual target



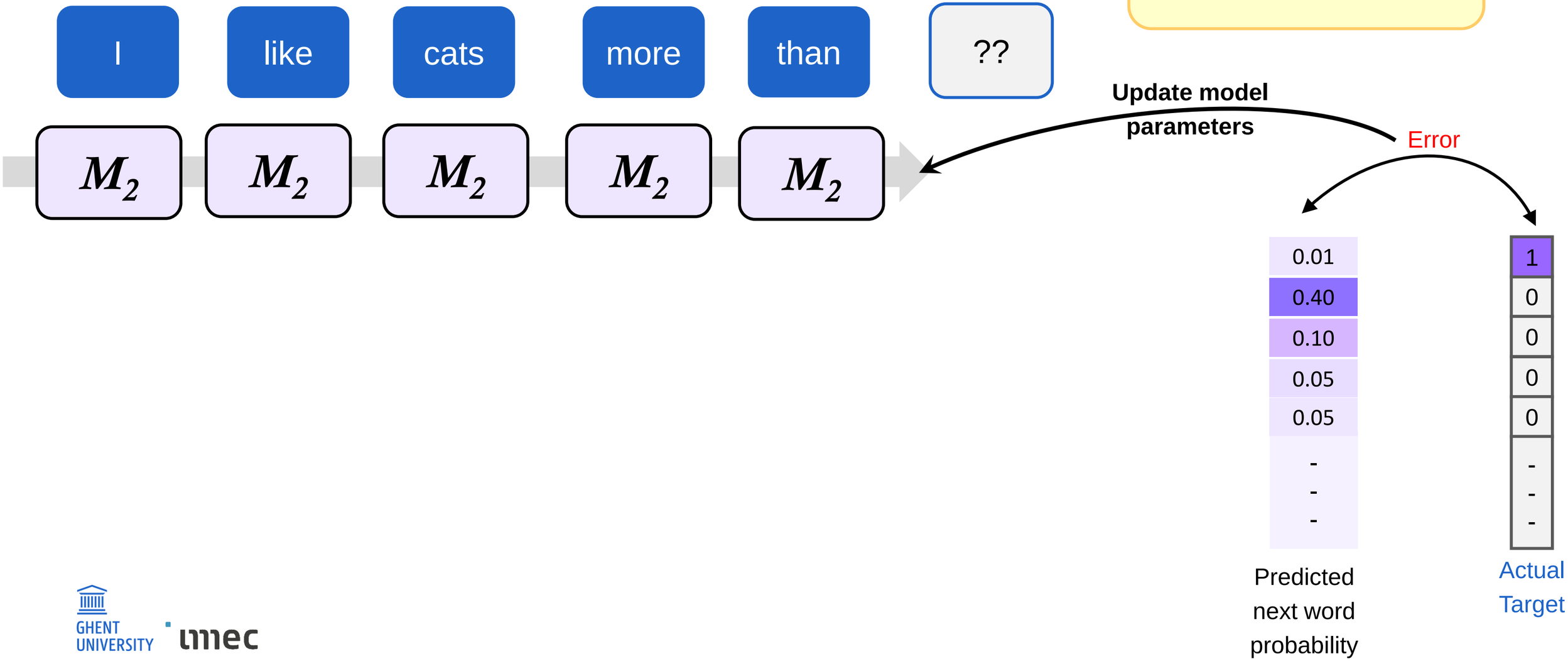
1 Causal Language Modeling

Update the model's parameters M: $M_1 \rightarrow M_2$



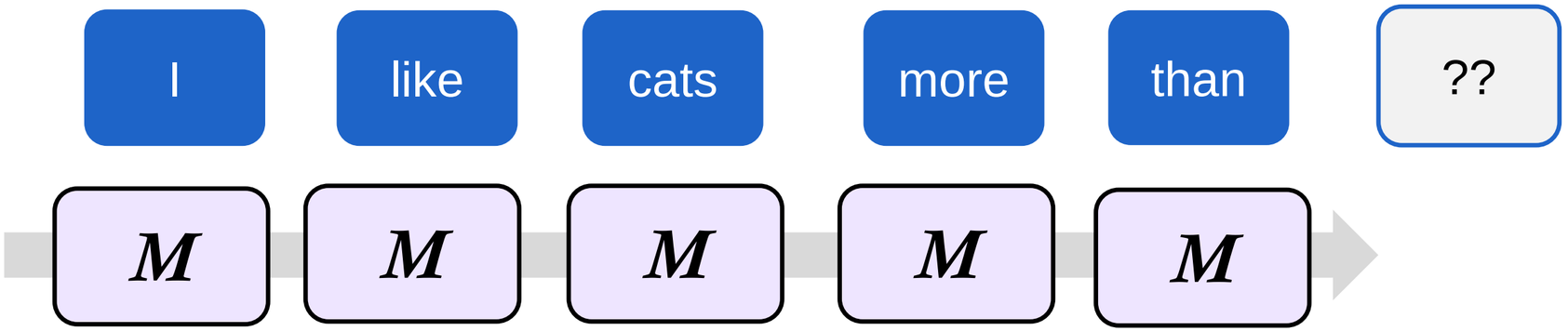
1 Causal Language Modeling

Repeat this process to minimize the error



1 Causal Language Modeling

The model is trained after several iterations (e.g., 100 iterations)



LM prediction similar to truth

predict

Vocabulary

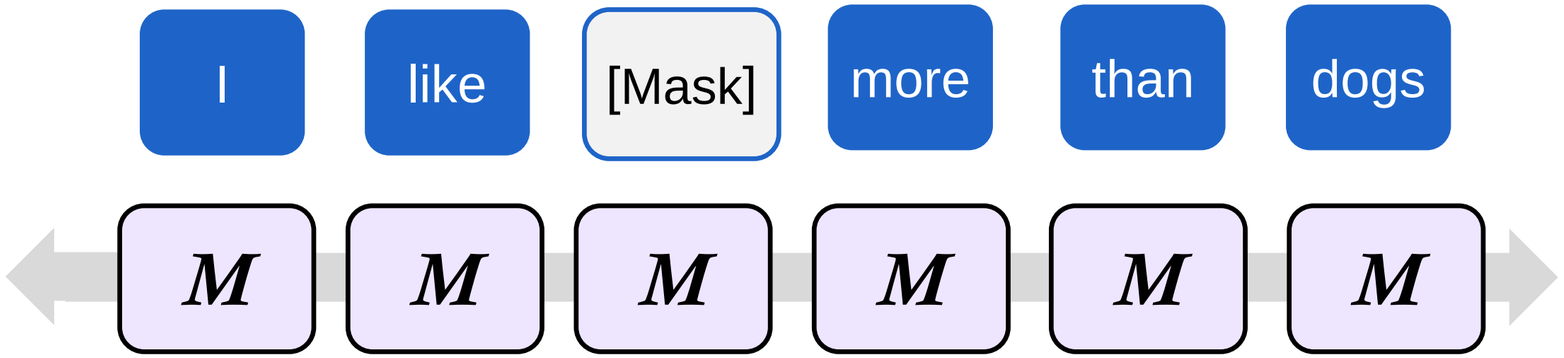
Predicted next word probability	Actual Target	
0.93	dogs	1
0.0001	potatoes	0
0.0002	mug	0
0.03	you	0
0.01	anything	0
-	-	-
-	-	-
-	-	-

1 Causal Language Modeling

- Processes text from left to right
- Natural for generative tasks (e.g., Question Generation)
- Examples: GPT-3/4, ChatGPT

2 Masked Language Modeling

Predict the masked word

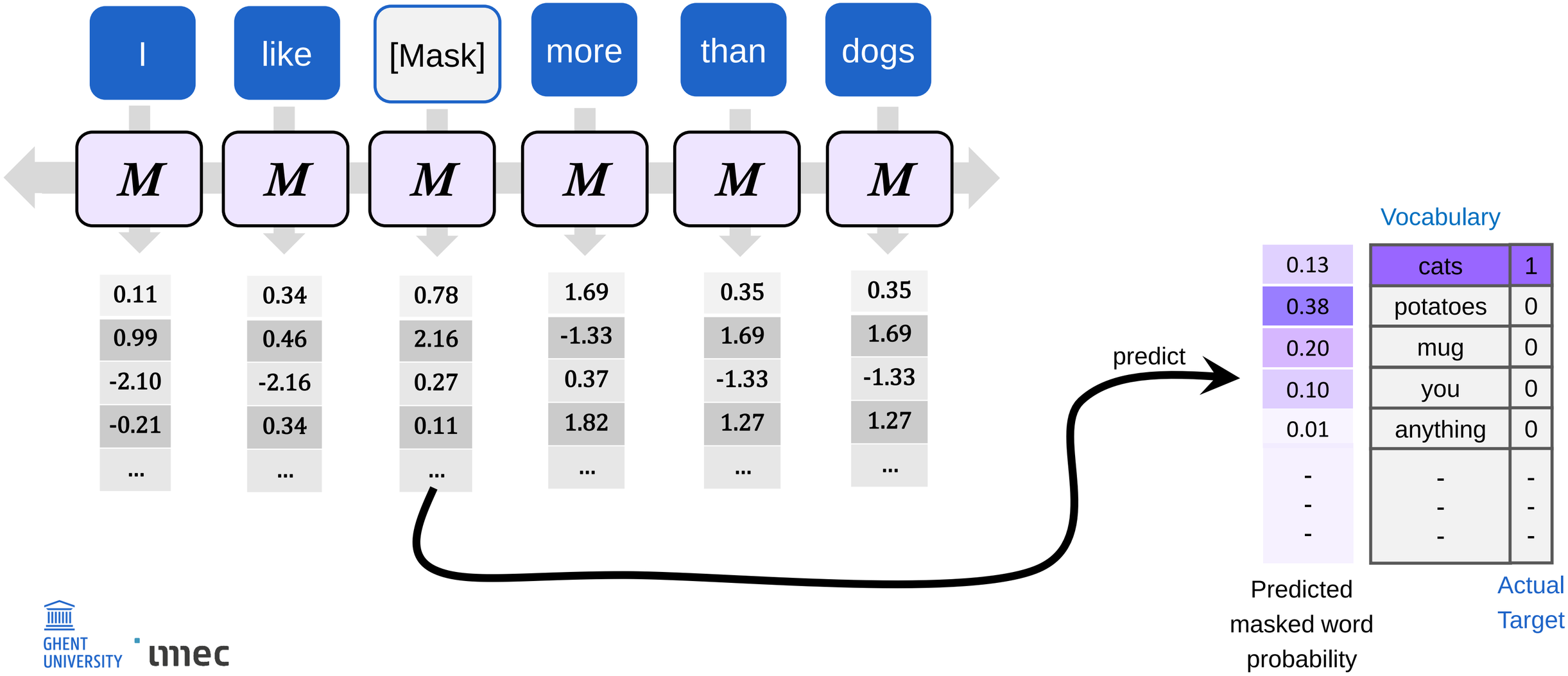


Vocabulary

cats
potatoes
mug
you
anything
-
-
-

2 Masked Language Modeling

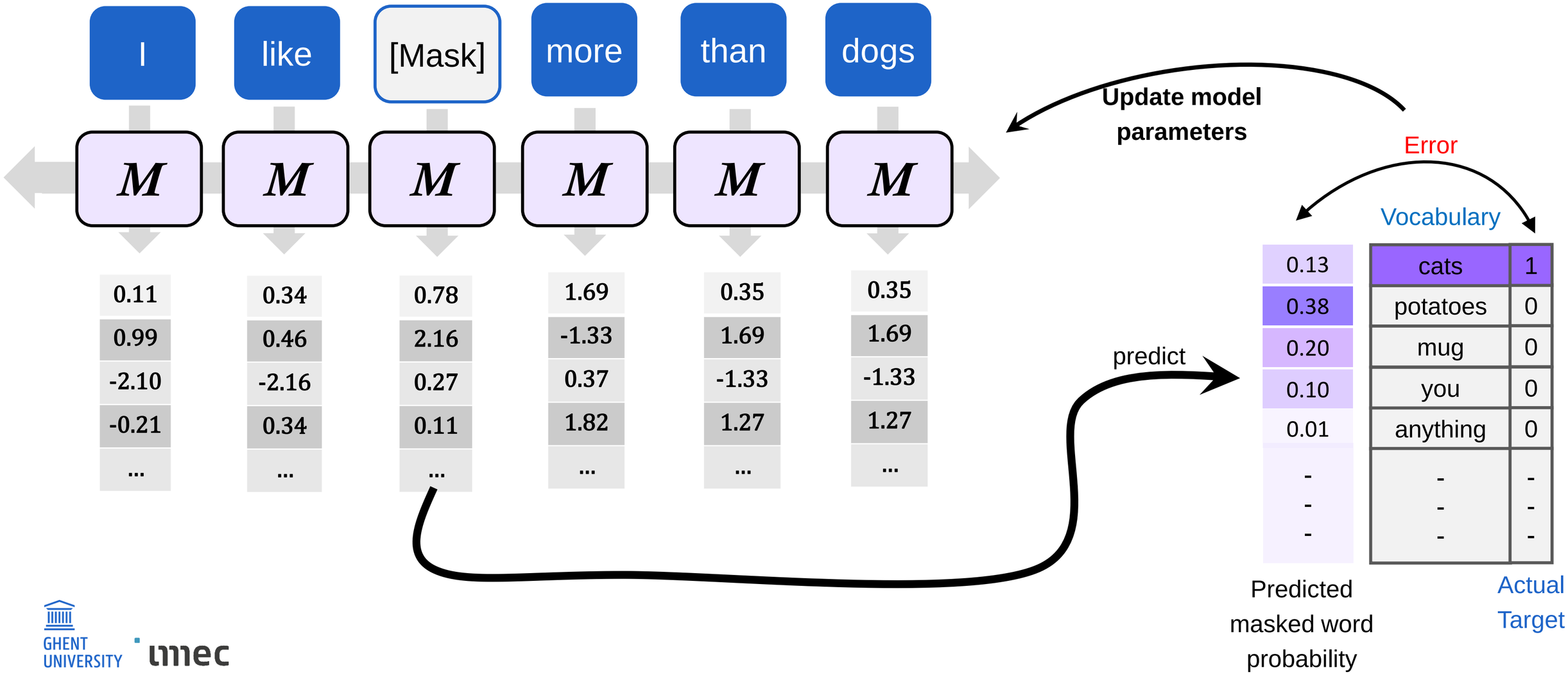
Predict the masked word



GHENT UNIVERSITY imec

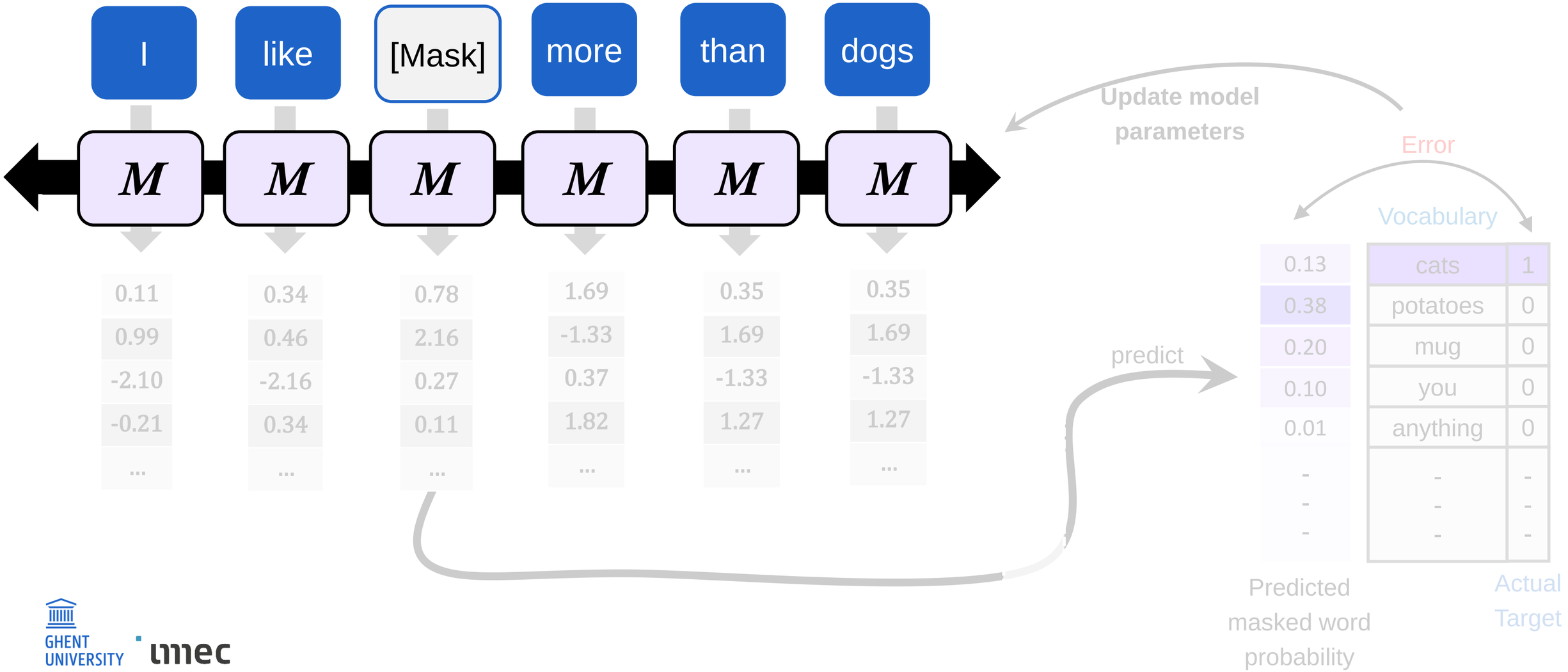
2 Masked Language Modeling

Predict the masked word



2 Masked Language Modeling

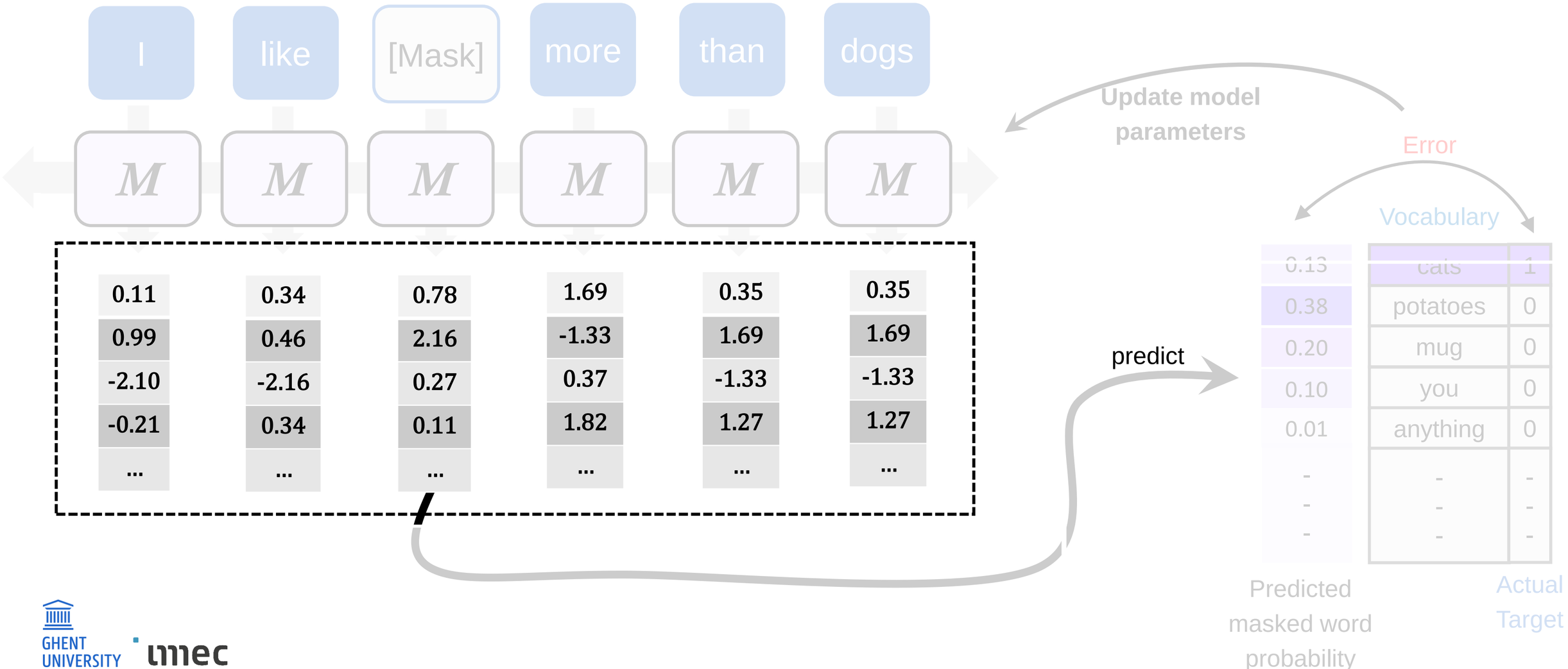
Bidirectional processing of text



GHENT UNIVERSITY imec

2 Masked Language Modeling

- Suitable for text understanding tasks (e.g., question answering, gap-fill prediction)



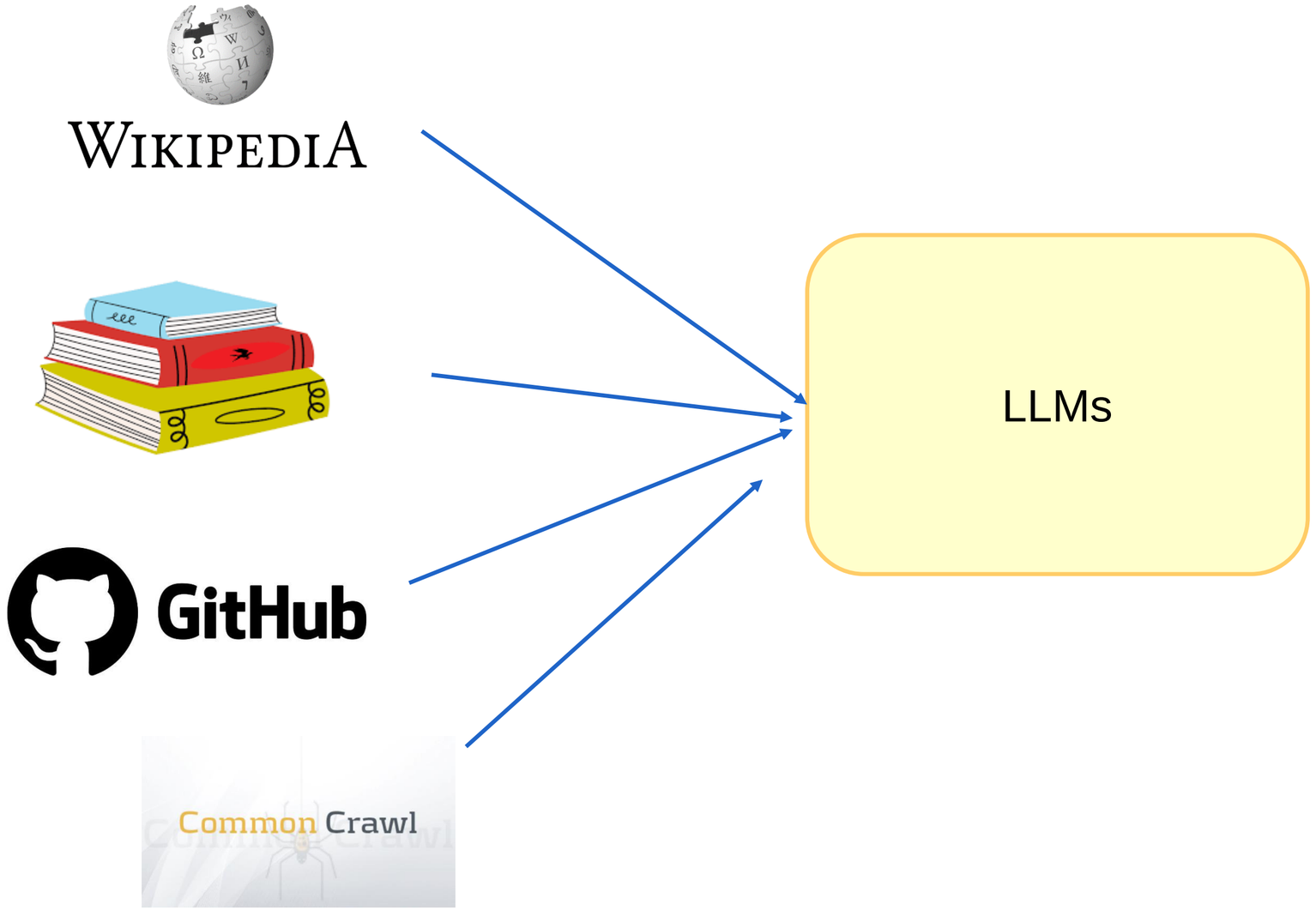
GHENT UNIVERSITY imec

2 Masked Language Modeling

- Suitable for text understanding tasks (e.g., question answering, gap-fill prediction)
- Examples: BERT, RoBERTa

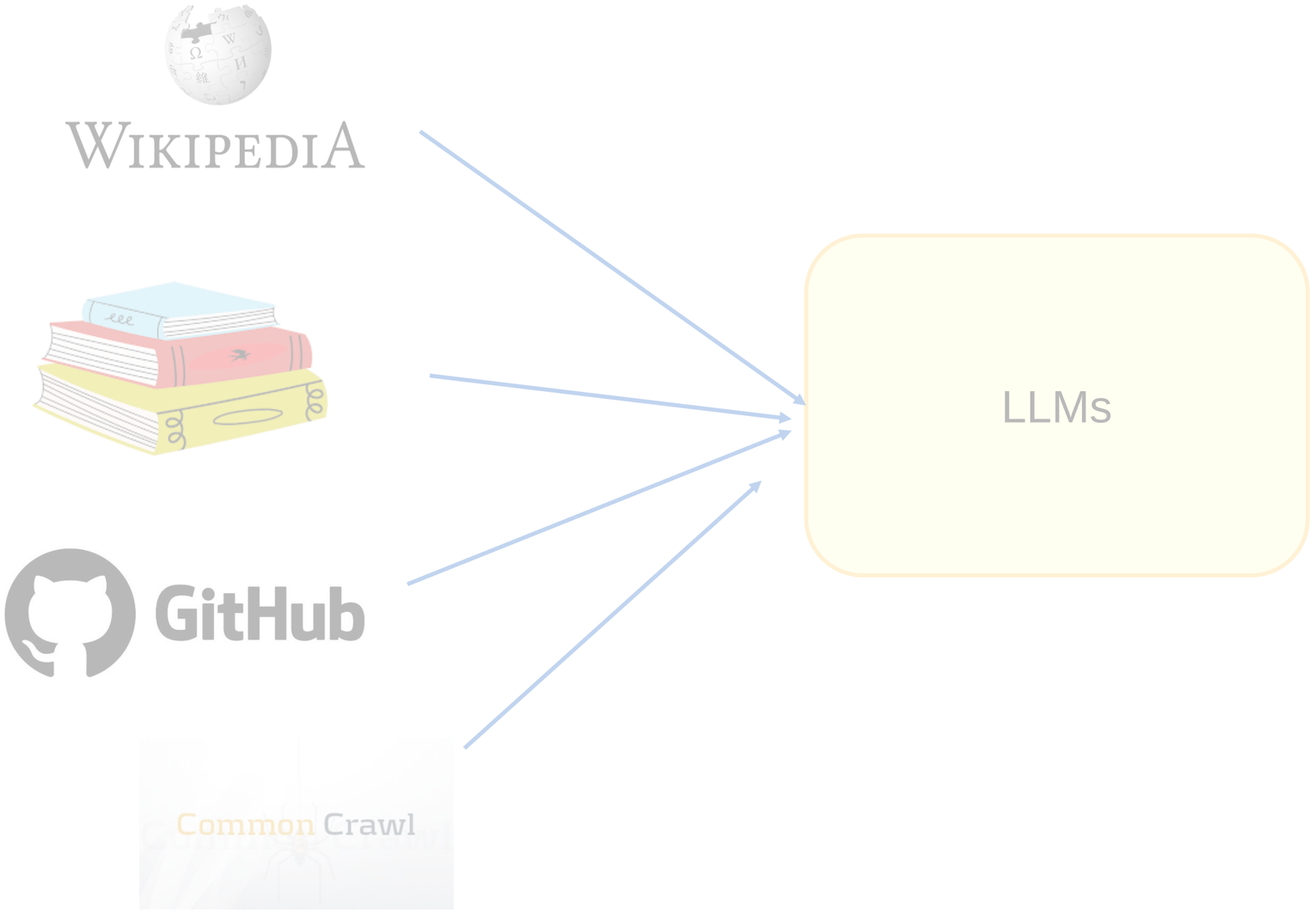
Large Language Models (LLMs)

LLMs are pretrained on tons of textual data



Large Language Models (LLMs)

LLMs are pretrained on tons of textual data



Examples of LLMs

GPT-4

Gemini

LLaMA by Meta

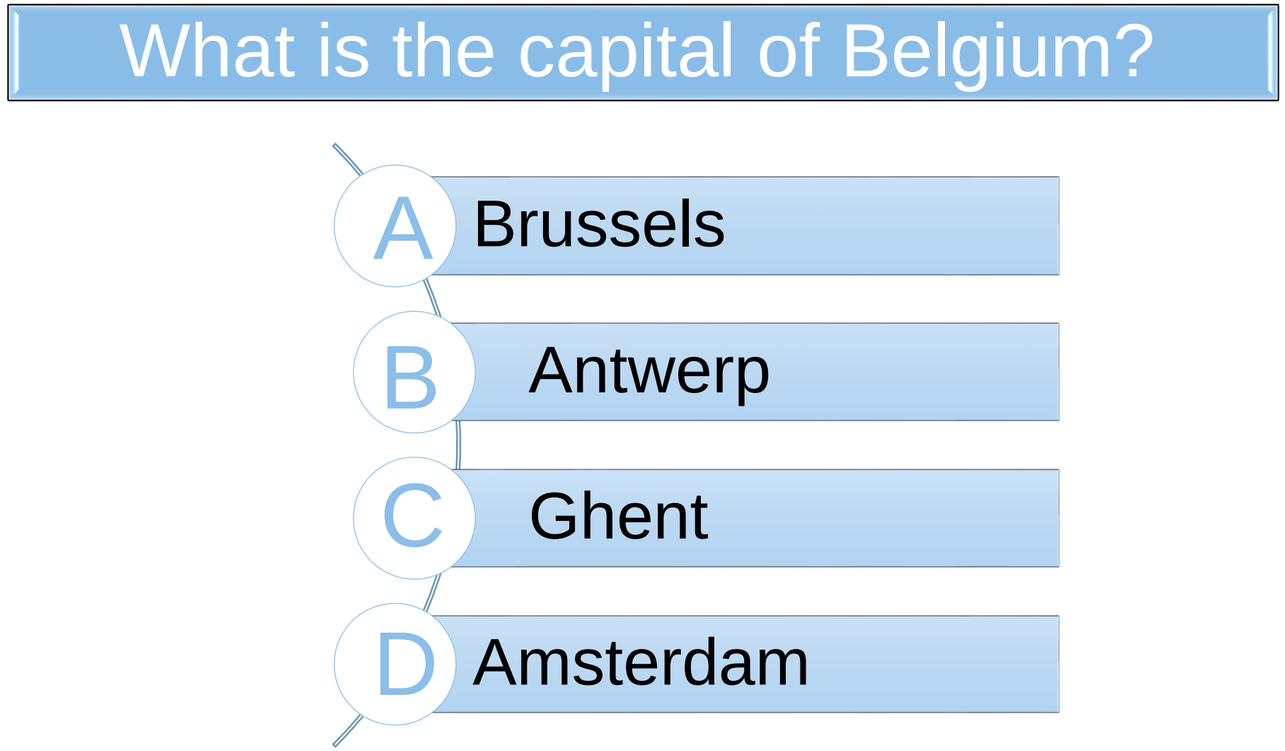
Distractor Generation Task

Distractor Generation

Multiple-choice question (**MCQ**) has

Distractor Generation

Multiple-choice question (**MCQ**) has

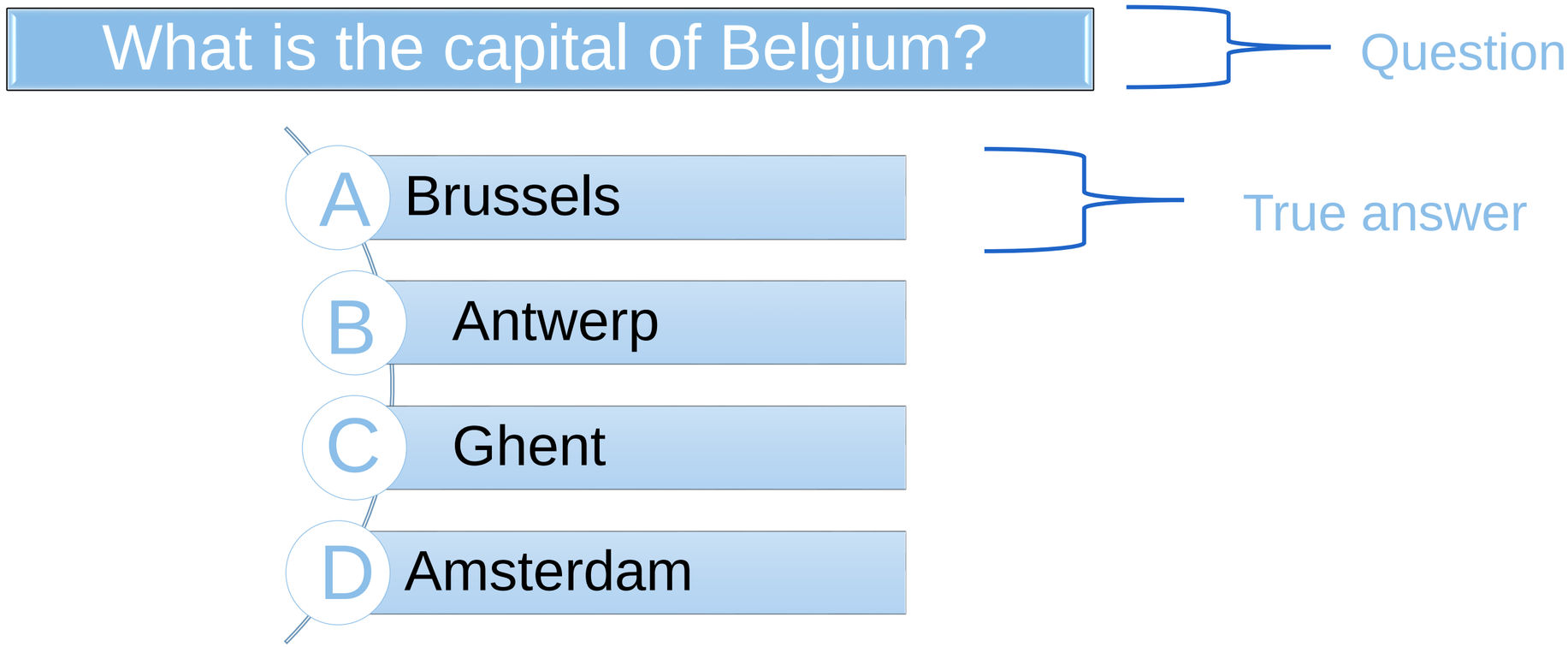


Distractor Generation

Multiple-choice question (**MCQ**) has

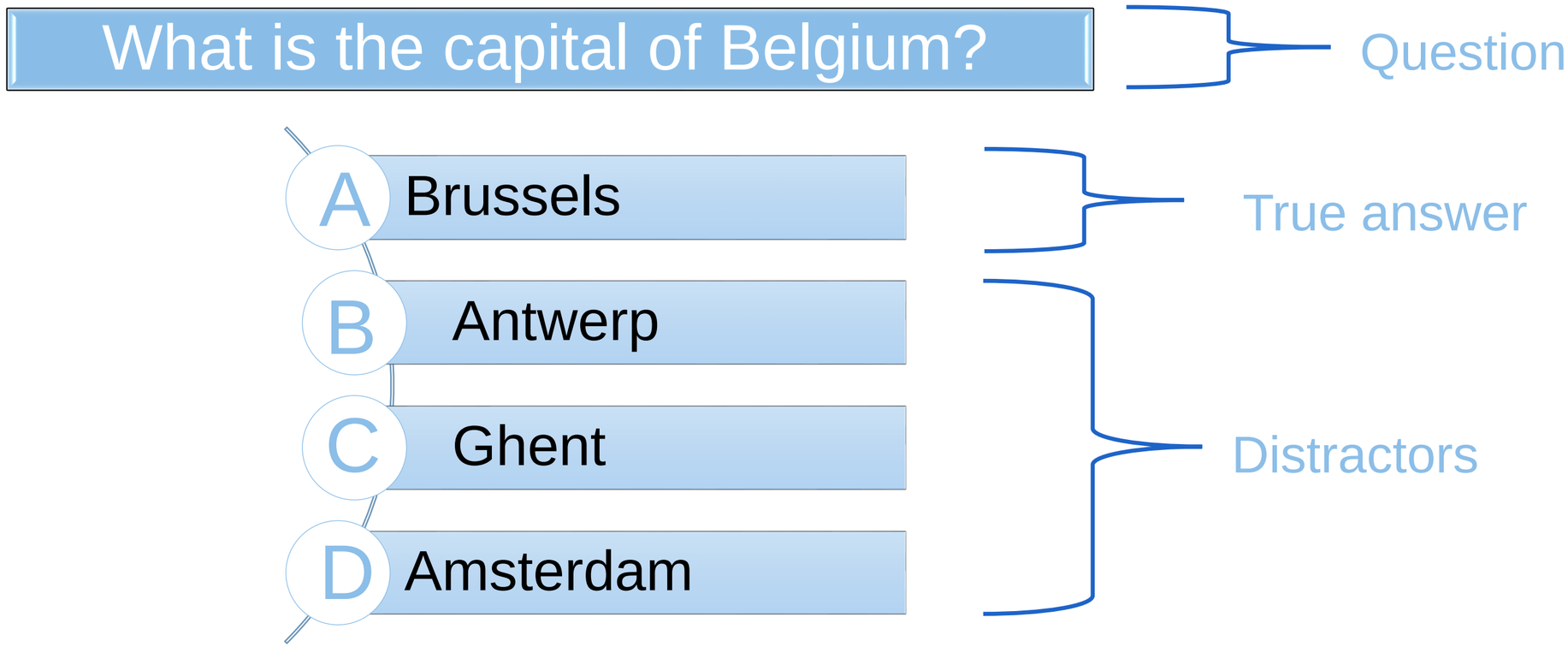
Distractor Generation

Multiple-choice question (**MCQ**) has



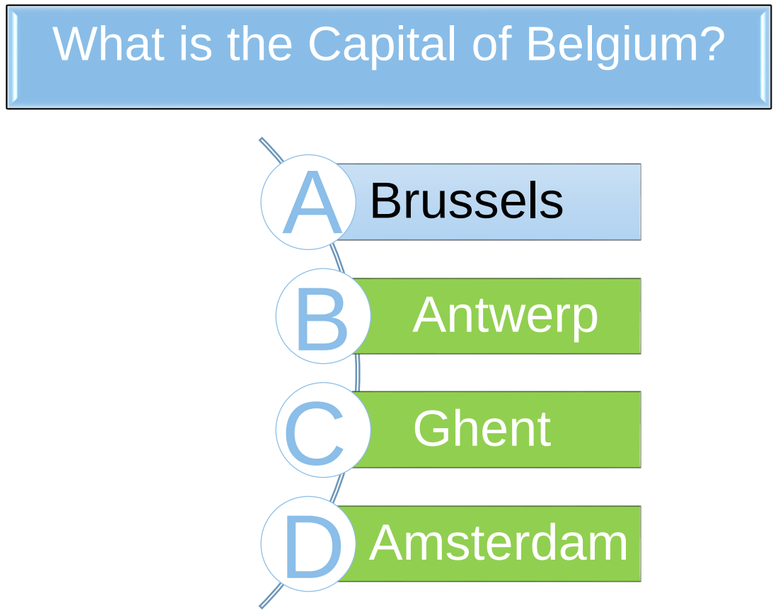
Distractor Generation

Multiple-choice question (**MCQ**) has

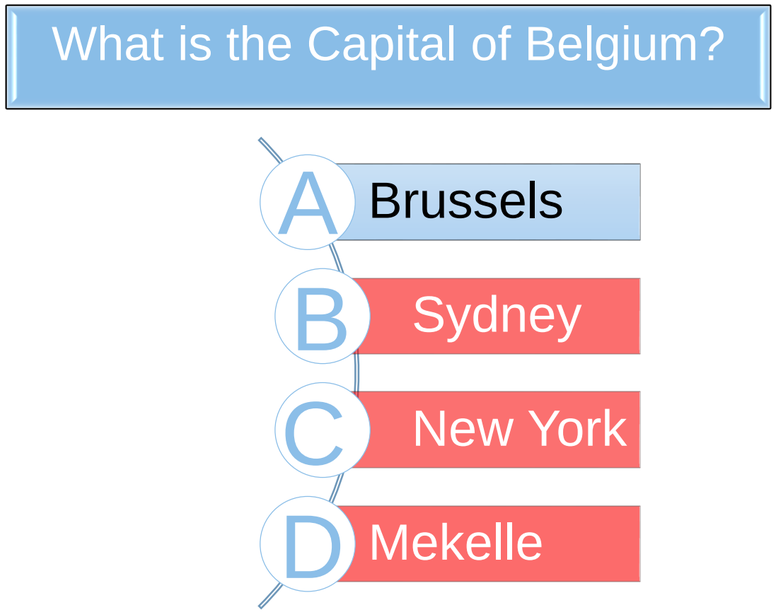


Quality of Distractors

Effective Distractors



Poor Distractors



Quality of Distractors

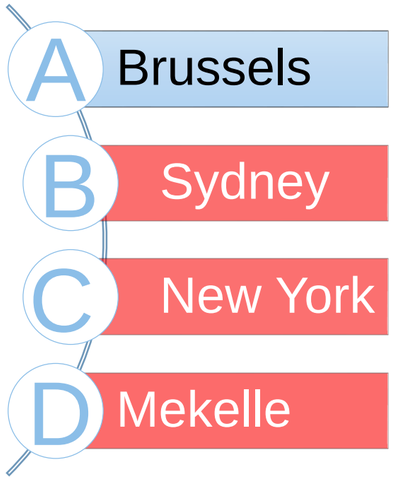
Effective Distractors

What is the Capital of Belgium?

- ✓ Plausible
- ✓ Tests knowledge
- ✓ Good MCQ

Poor Distractors

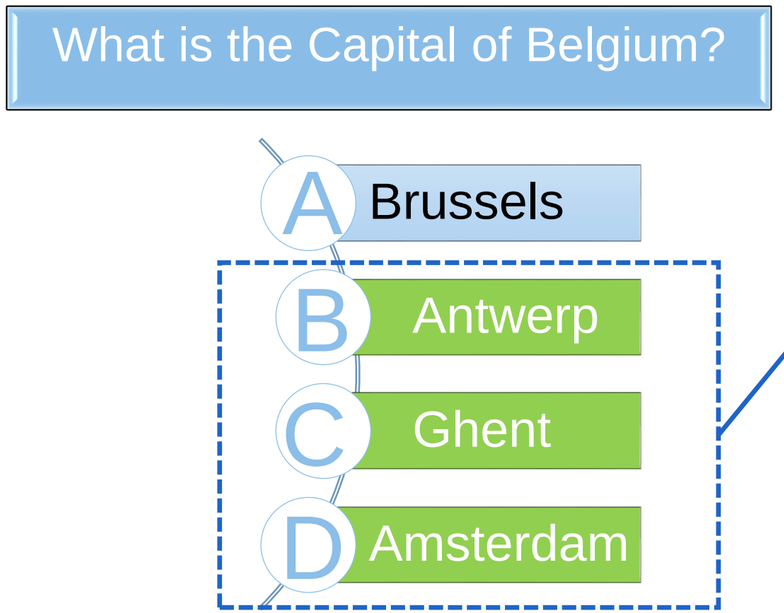
What is the Capital of Belgium?



- × Implausible
- × Easily dismissed
- × Bad MCQ

Quality of Distractors

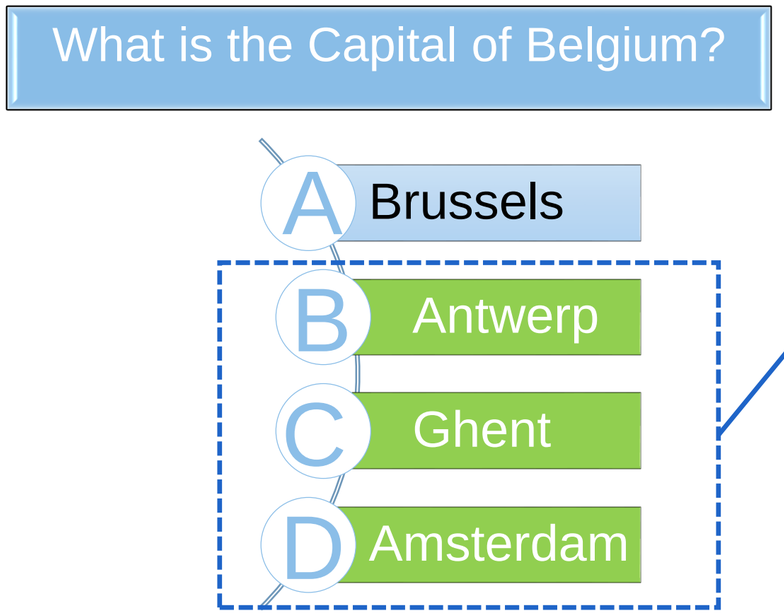
Effective Distractors



Generating distractors → time-consuming task!

Quality of Distractors

Effective Distractors



Generating distractors → time-consuming task!

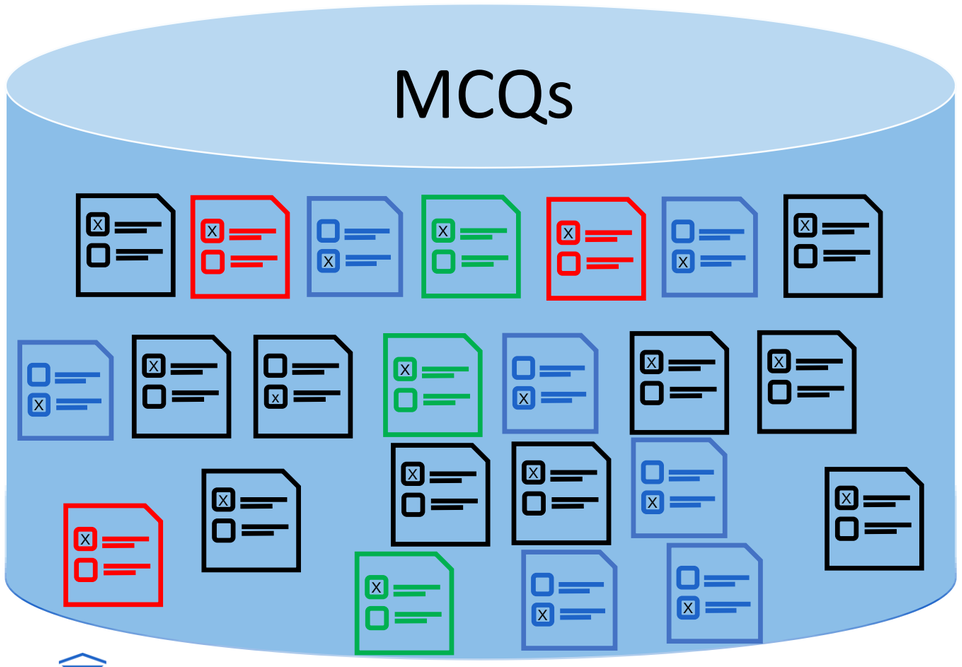
Goal: Can we automate this process to **increase** teachers' efficiency?

Research question

1. How to automatically generate distractors?

Televic Dataset

Real-world dataset provided by industry partner



Televic Dataset

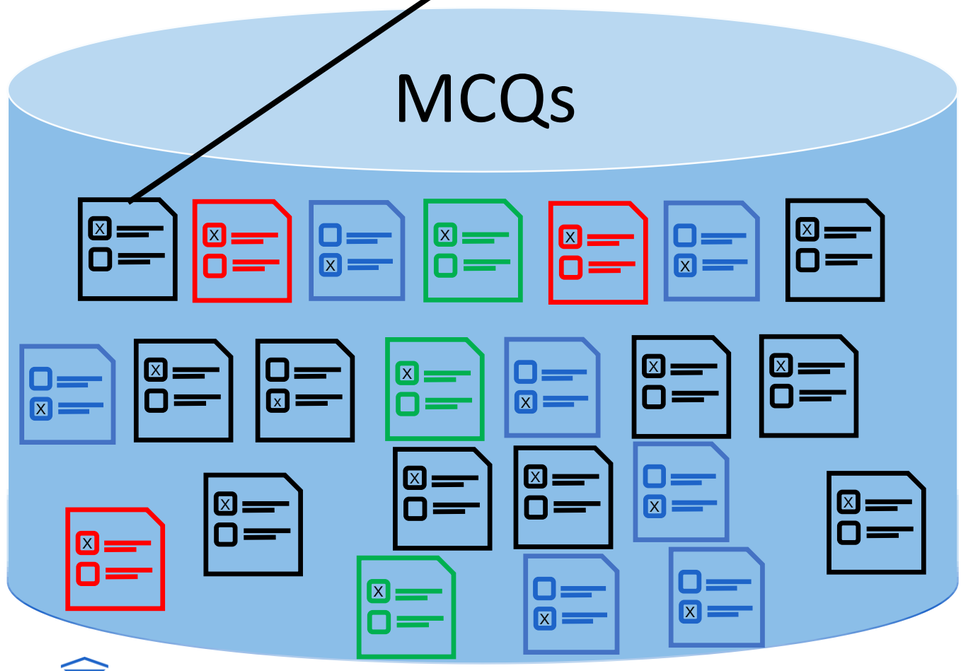
The highest mountain in the world is ___

A. Mont Blanc

B. Kilimanjaro

C. Mount Everest

D. Amba Alaje



GHENT UNIVERSITY imec

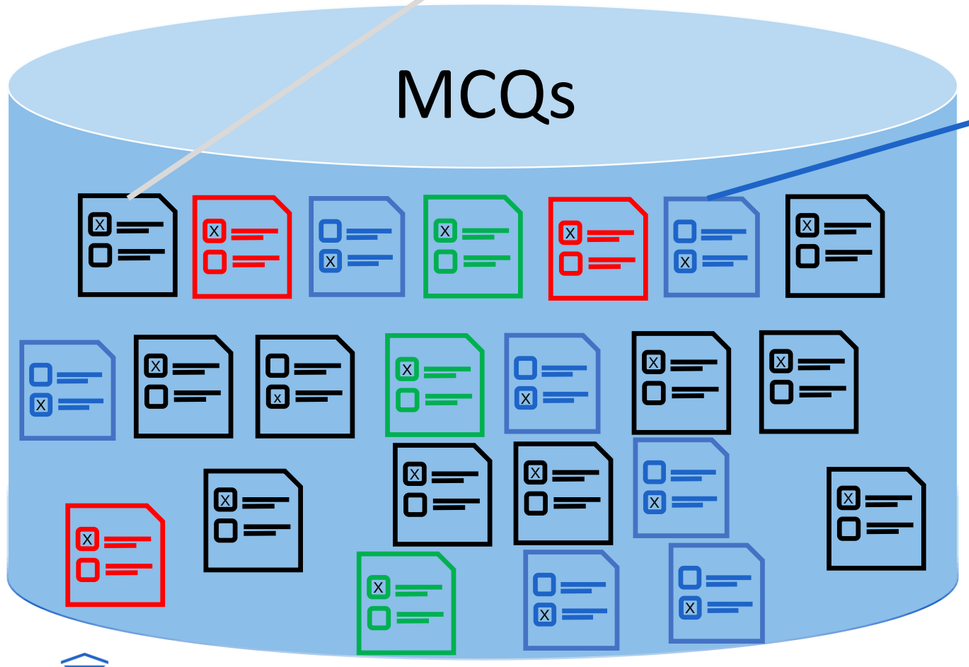
Televic Dataset

The highest mountain in the world is ____

A. Mont Blanc
C. Mount Everest
B. Kilimanjaro
D. Amba Alaje

In welk werelddeel light Noord-Korea?

A. Antarctica
C. Afrika
B. Europa
D. Azie



GHENT UNIVERSITY imec

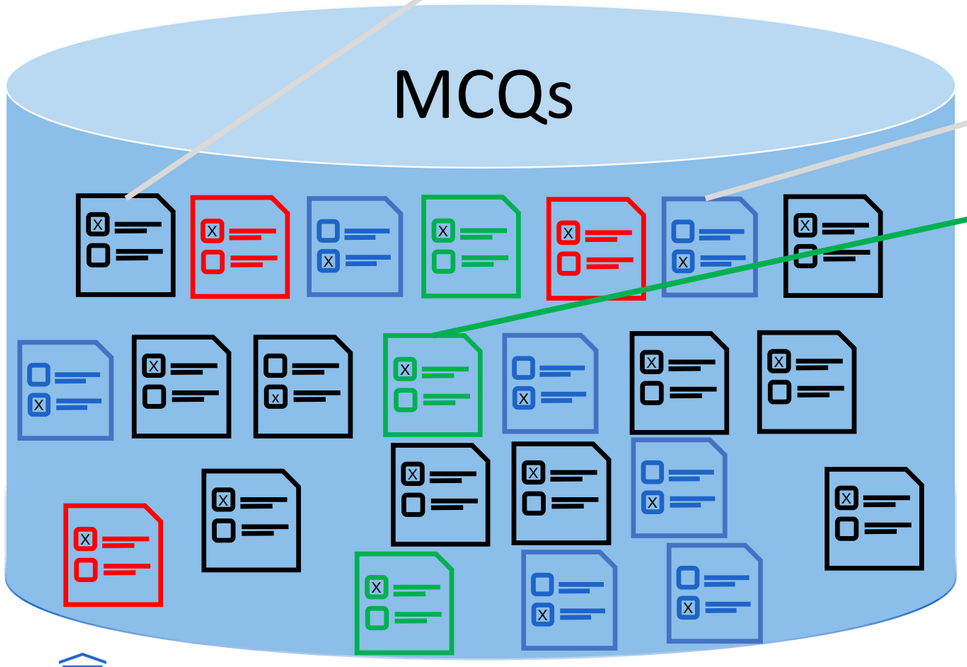
Televic Dataset

The highest mountain in the world is ___

A. Mont Blanc
C. Mount Everest
B. Kilimanjaro
D. Amba Alaje

In welk werelddeel light Noord-Korea?

A. Antarctica
C. Afrika
B. Europa
D. Azie



By this time next month we ___ the designs.

A. will have finished
C. will finish
B. are finishing
D. go to finish

GHENT UNIVERSITY imec

Televic Dataset

The highest mountain in the world is ____

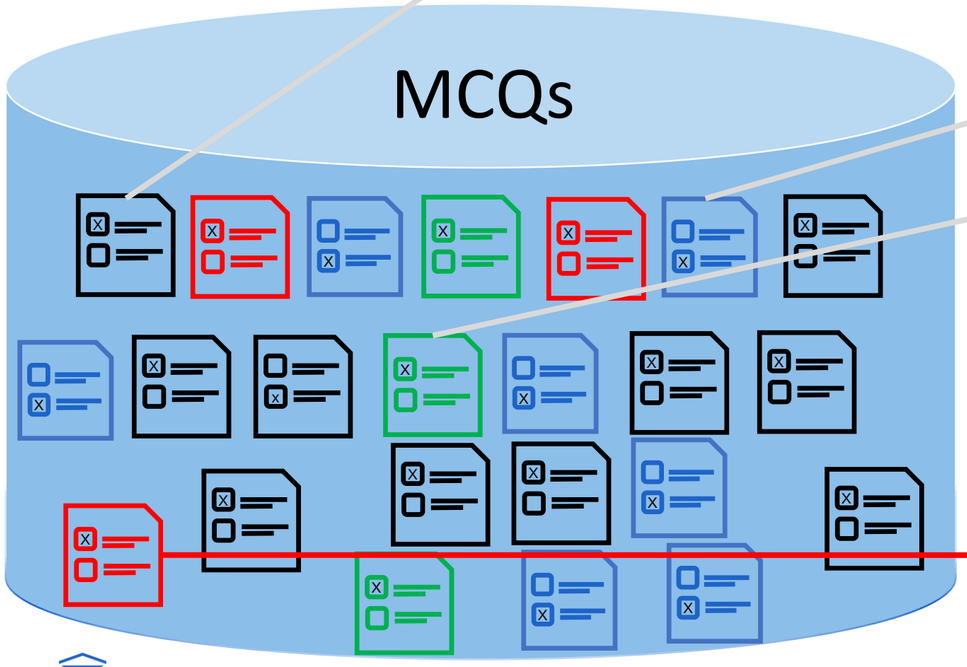
A. Mont Blanc C. Mount Everest

B. Kilimanjaro D. Amba Alaje

In welk werelddeel light Noord-Korea?

A. Antarctica C. Afrika

B. Europa D. Azie



By this time next month we ___ the designs.

A. will have finished C. will finish

B. are finishing D. go to finish

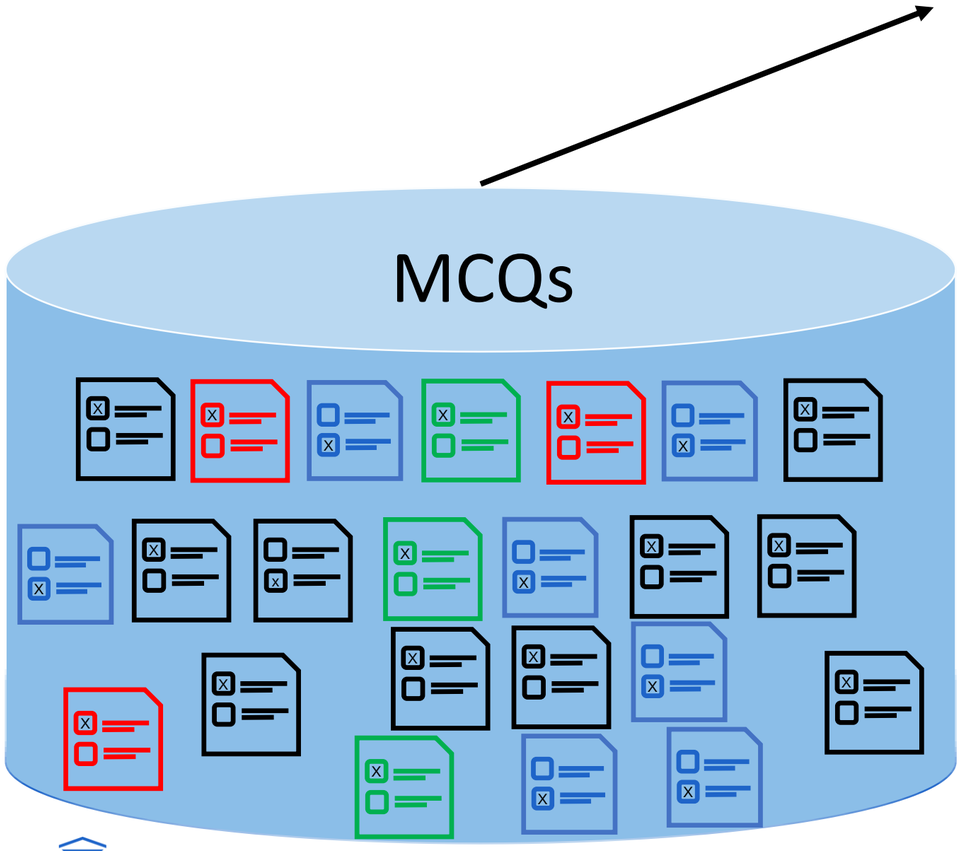
Weet je nog wat de eerste mensensoort was?

A. Homo erectus C. Homo habilis

B. Homo sapiens D. Homo neanderthalensis

GHENT UNIVERSITY imec

Televic Dataset



Total of

~ 62,000 MCQs

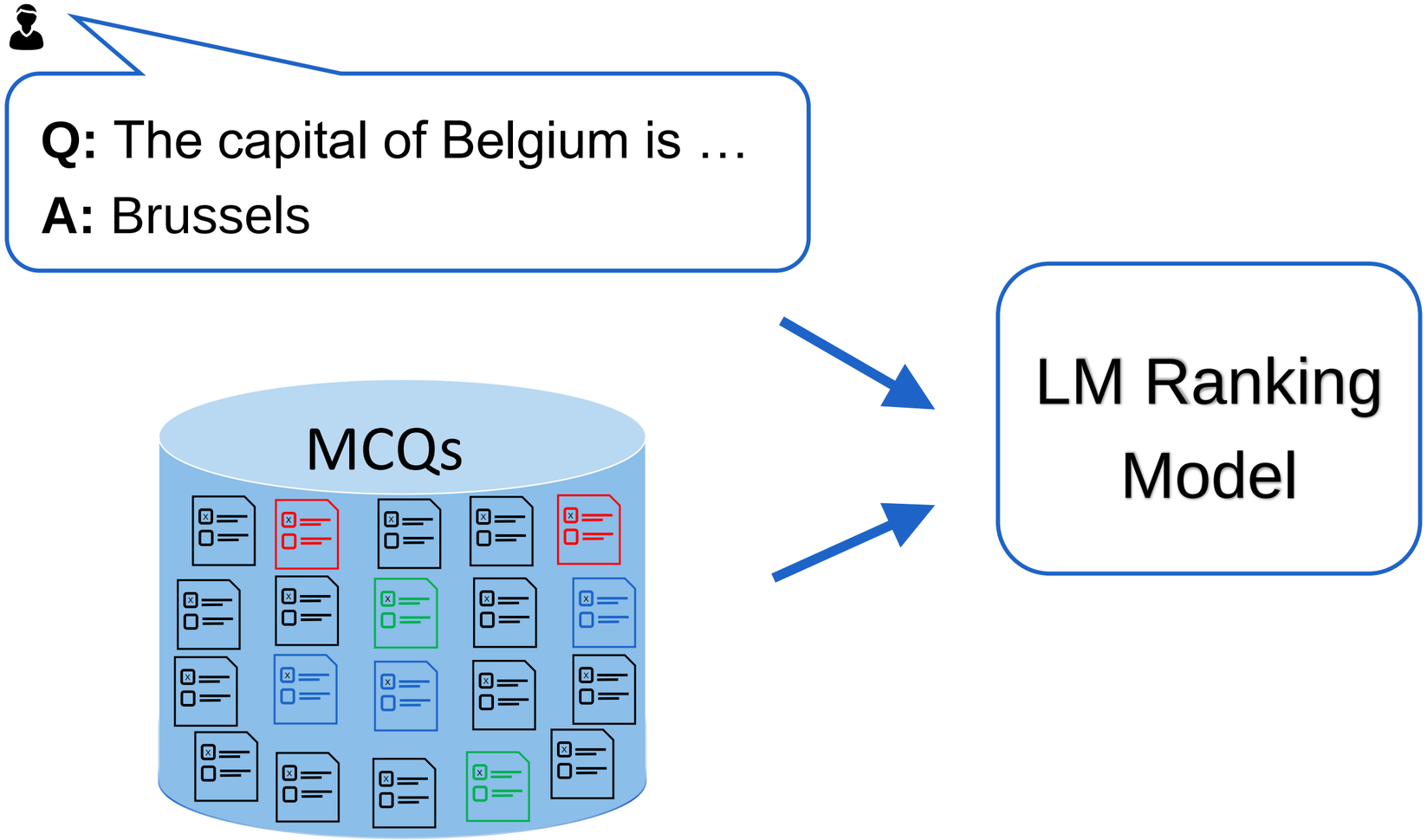
Proposed Solution

Learn to rank the distractors according to their relevance to a question

GHENT UNIVERSITY imec

Learning to rank distractors

Language Model (**LM**) based model ranks existing distractors

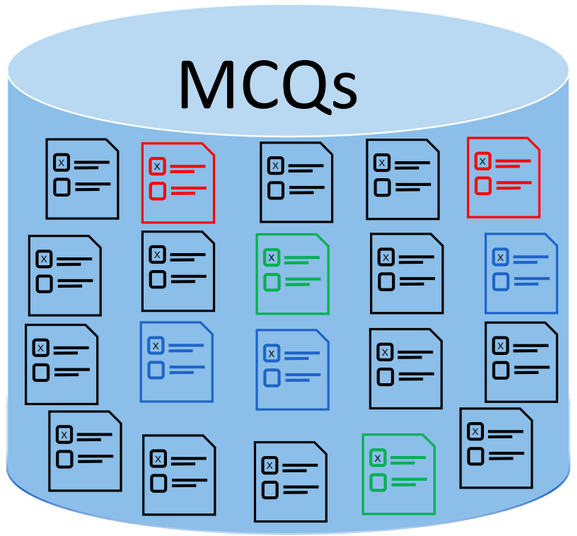


Learning to rank distractors

Language Model (**LM**) based model ranks existing distractors

Q: The capital of Belgium is …

A: Brussels



LM Ranking Model

Ranked Distractors

1. Antwerp
2. Ghent
3. Amsterdam
4.

⋮

77008. will finish

77009. kern

77010. Ribosomen

Our approach

We propose three models

Our approach (Q-SIM)

1. Question Similarity (Q-SIM) model
   - Questions are similar if they share a distractor or an answer

Our approach (Q-SIM)

1 Question Similarity (Q-SIM) model

- Questions are similar if they share a distractor or an answer

What is the capital of the Netherlands?

A. The Hague C. Amsterdam

B. Rotterdam D. Brussels

The capital of Belgium is …

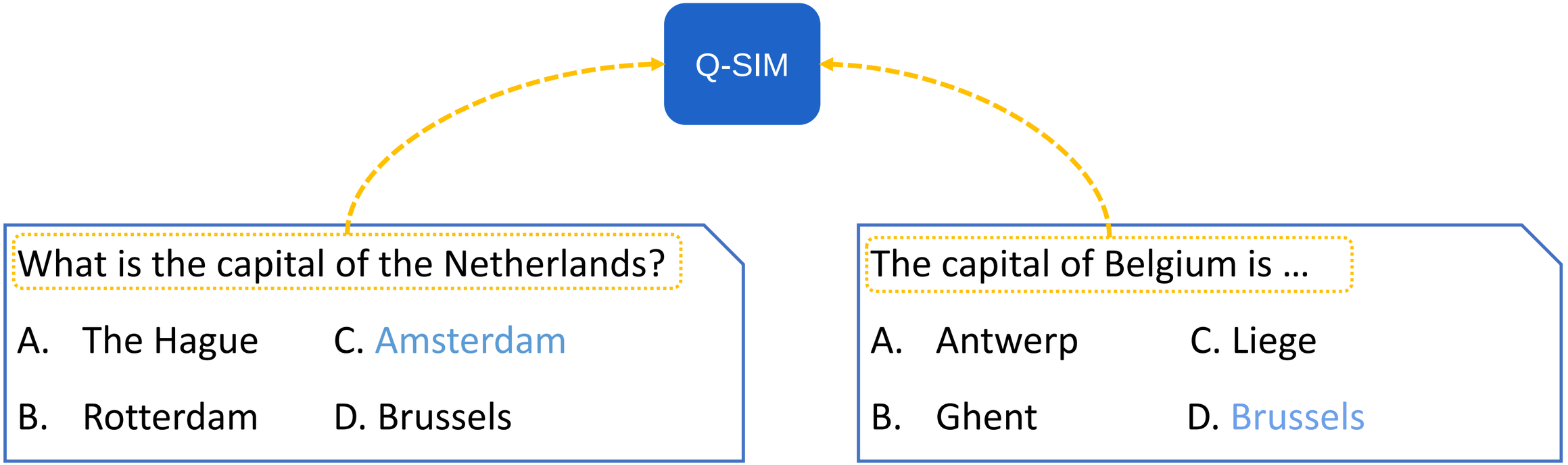
A. Antwerp C. Liege

B. Ghent D. Brussels

Our approach (Q-SIM)

1 Question Similarity (Q-SIM) model

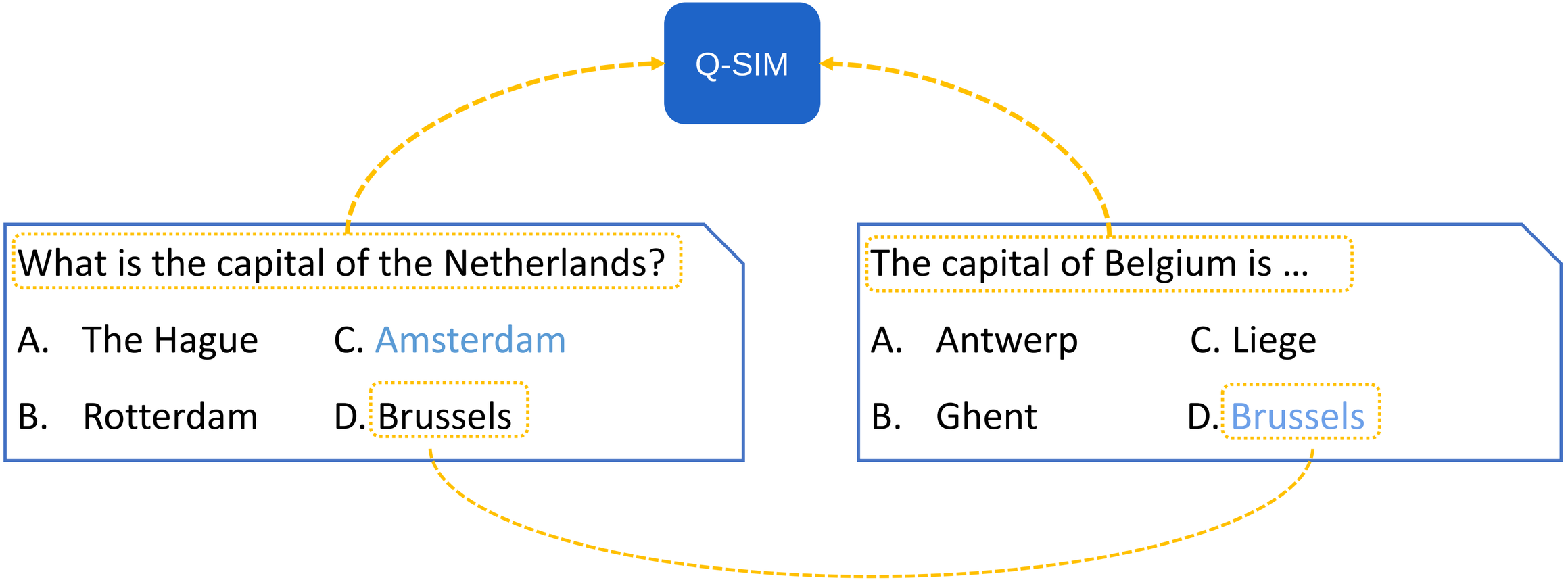
- Questions are similar if they share a distractor or an answer



Our approach (Q-SIM)

1 Question Similarity (Q-SIM) model

- Questions are similar if they share a distractor or an answer



Our approach (Q-SIM)

1 Question Similarity (Q-SIM) model

- Q-SIM clusters similar questions together

Q-SIM

Our approach (D-SIM)

2 Distractor Similarity (D-SIM) model

- Distractors are similar if they co-occur within the same MCQ

Our approach (D-SIM)

2 Distractor Similarity (D-SIM) model

- Distractors are similar if they co-occur within the same MCQ

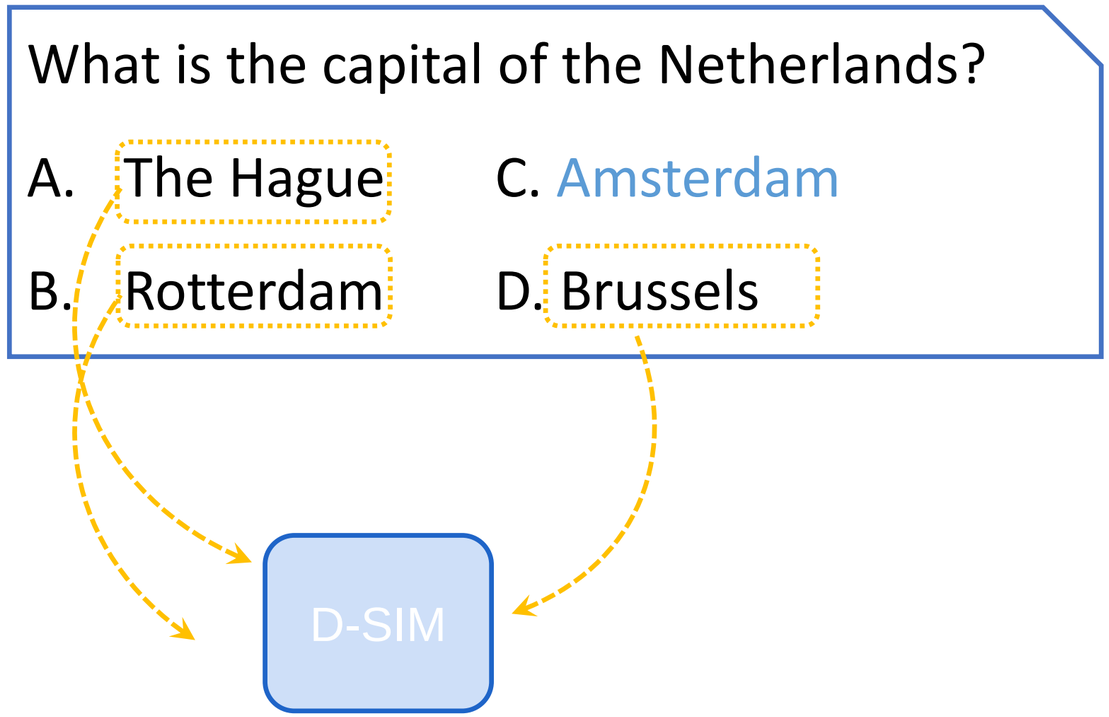
What is the capital of the Netherlands?

A. The Hague
C. Amsterdam
B. Rotterdam
D. Brussels

Our approach (D-SIM)

2 Distractor Similarity (D-SIM) model

- Distractors are similar if they co-occur within the same MCQ

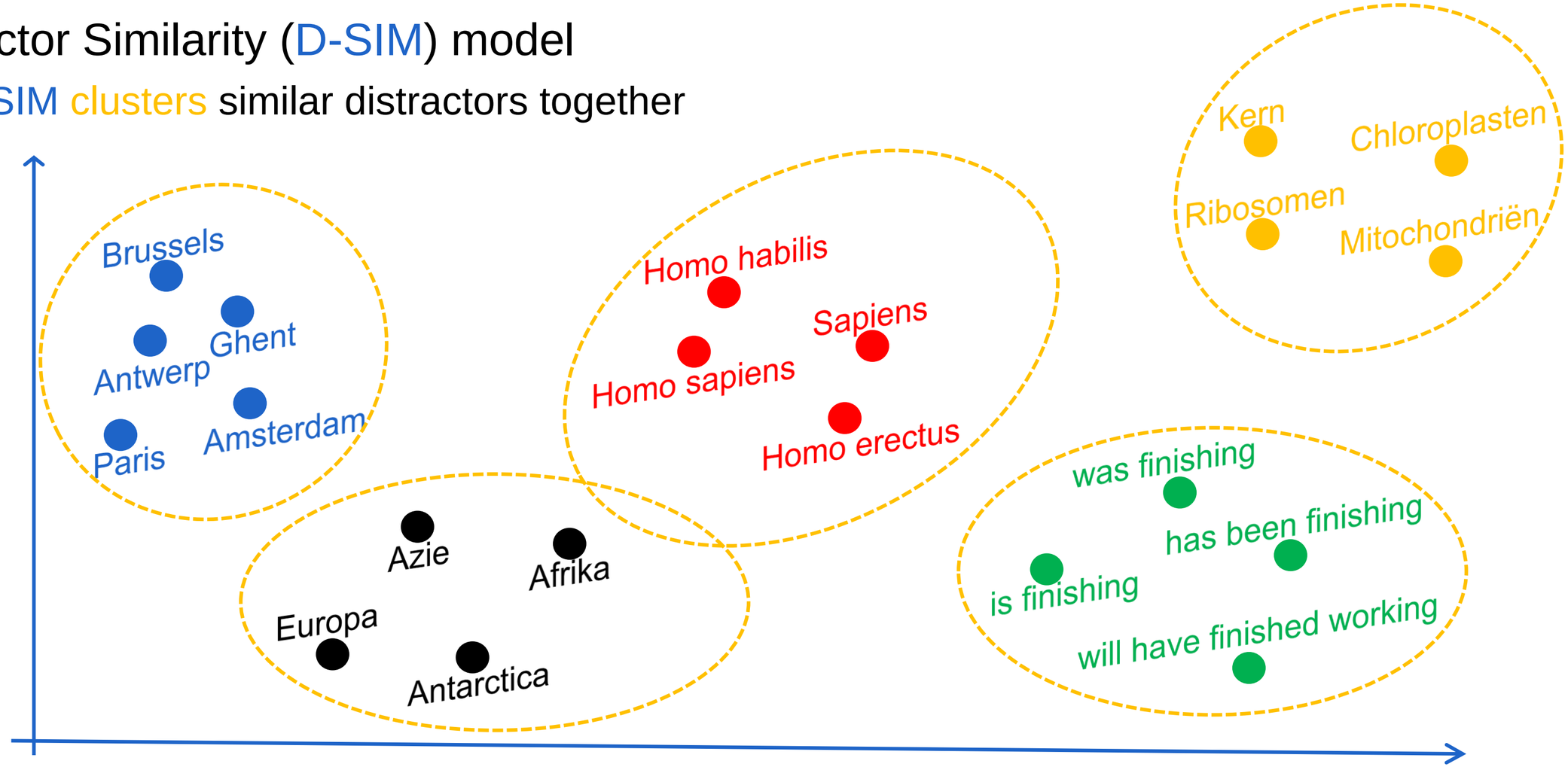


Our approach (D-SIM)

2 Distractor Similarity (D-SIM) model

- D-SIM clusters similar distractors together

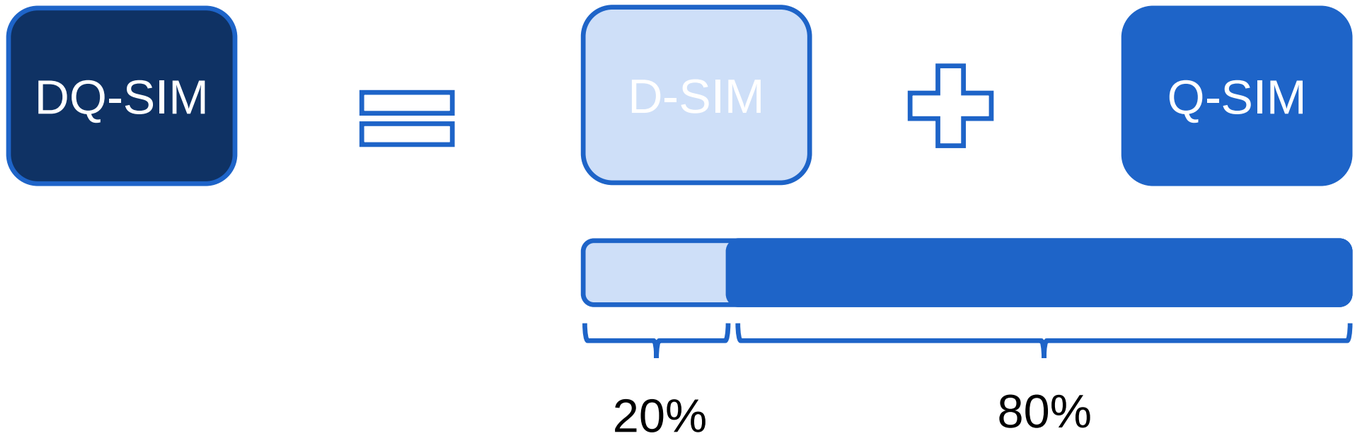
D-SIM



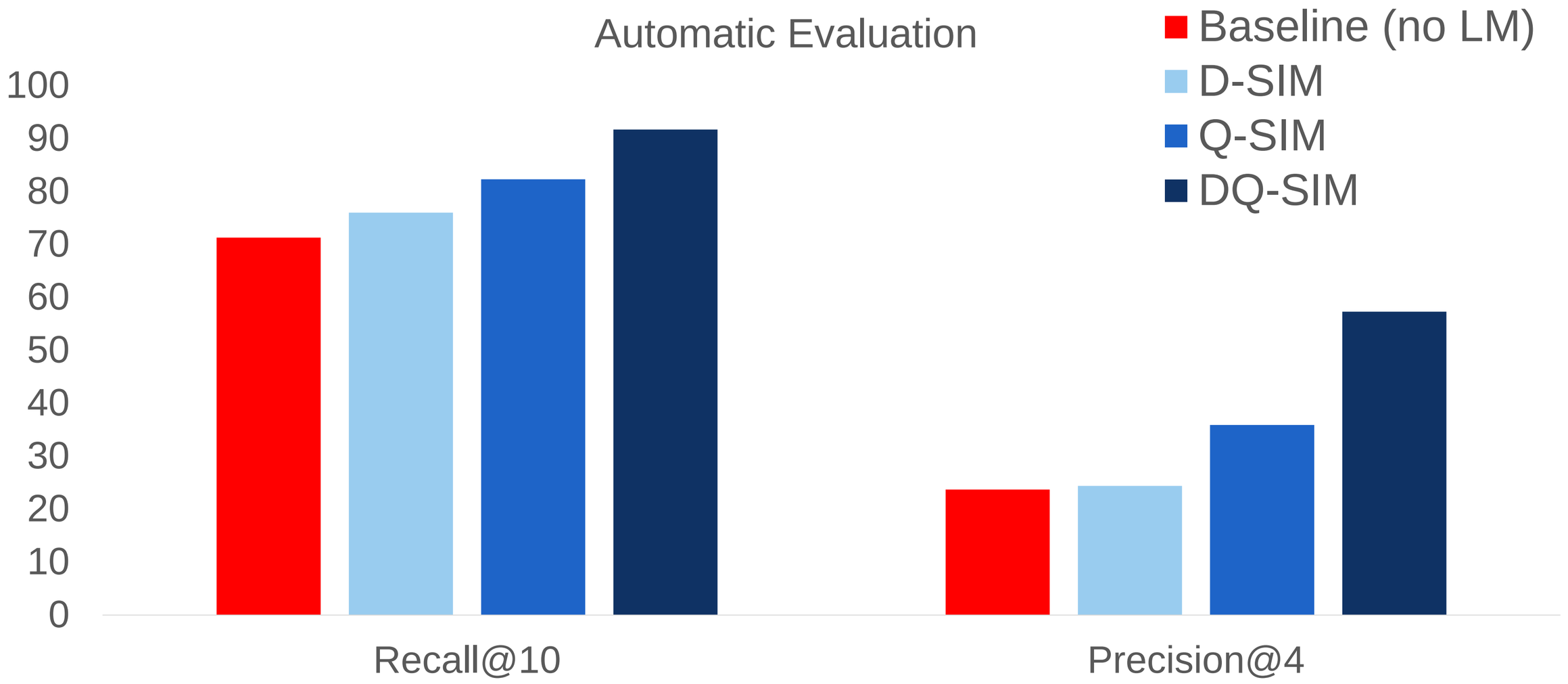
Our approach (DQ-SIM)

3 Distractor-Question Similarity (DQ-SIM) model

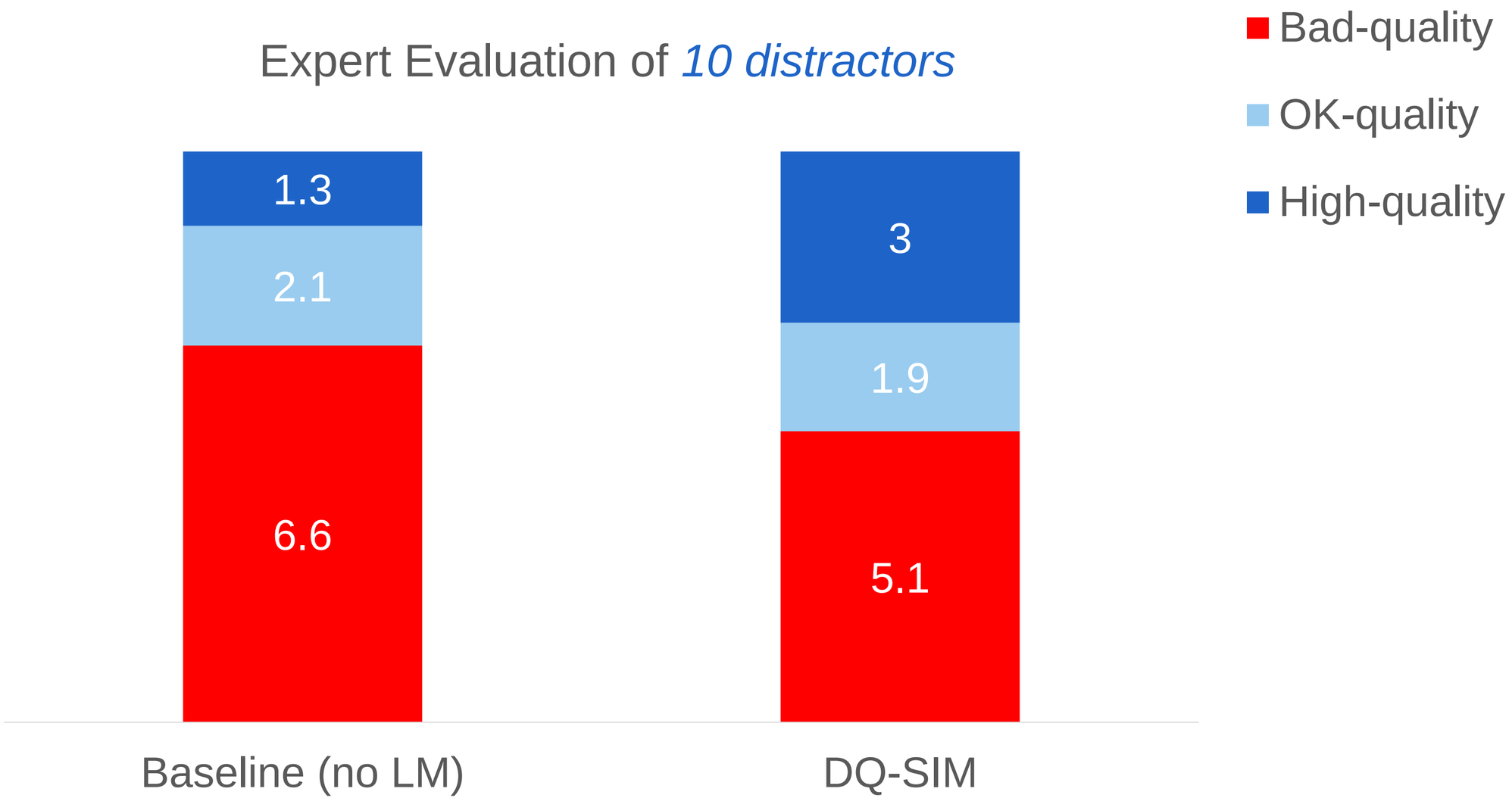
- Combines Q-SIM and D-SIM



Gain in performance using language models



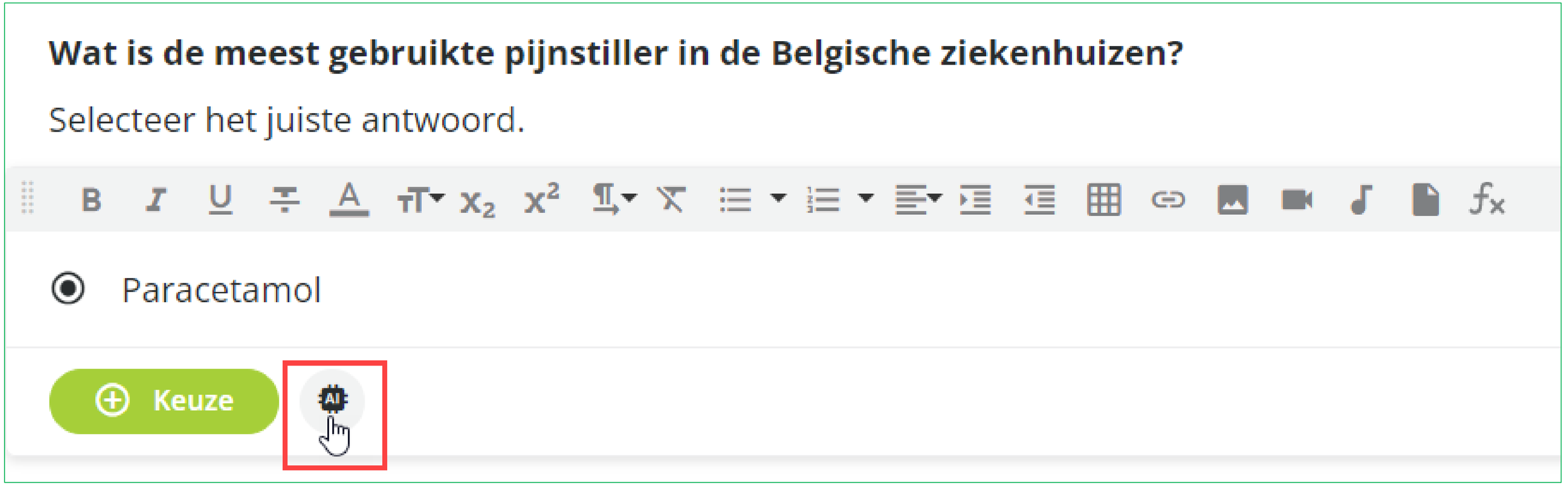
What do teachers say about our models?



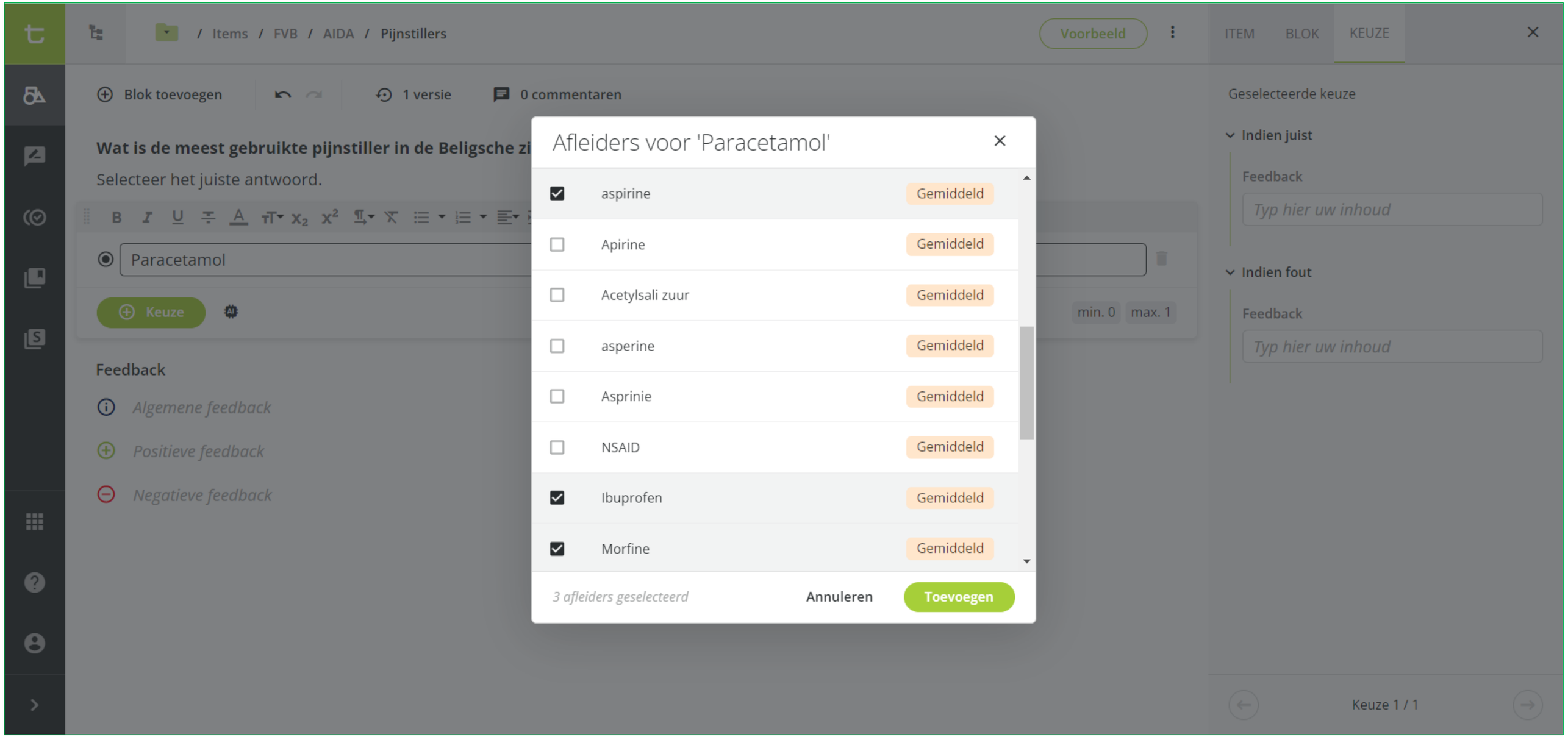
What do teachers say about our models?

3 distractors high-quality out of 10

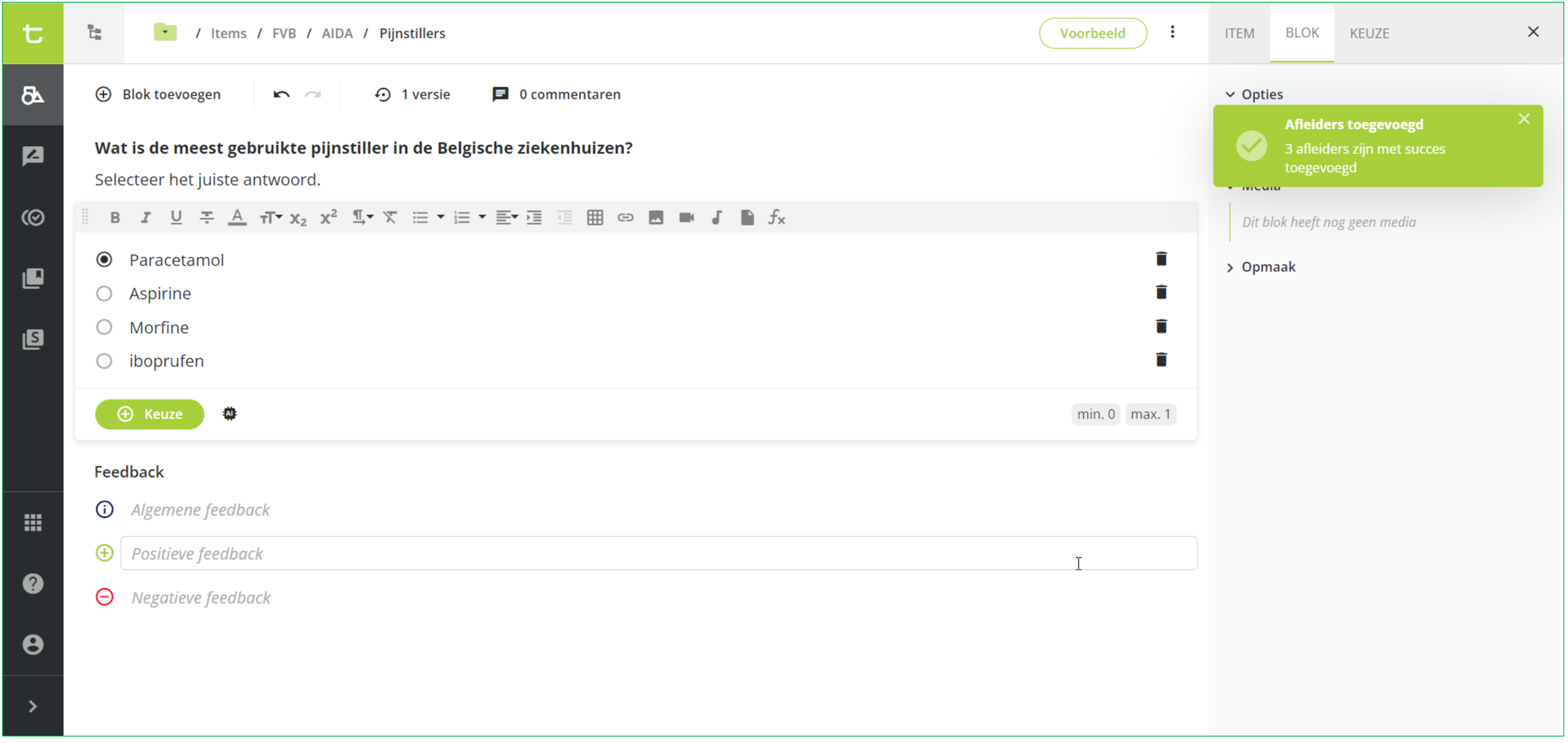
Our model in Televic's platform



Our model in Televic's platform



Our model in Televic's platform



Two months later

released ChatGPT

Distractor Generation

Using Generative language models

Research questions

1. How to automatically generate distractors?

2. How to design a system that generates free-form distractors?

Free-form Distractor Generation

- Recap: the previous methods CANNOT generate brand new distractors

Free-form Distractor Generation

- Recap: the previous methods CANNOT generate brand new distractors
- Solution: Ask LLMs such as ChatGPT to generate distractors (***Zero-ChatGPT***)

Free-form Distractor Generation

- Recap: the previous methods CANNOT generate brand new distractors
- Solution: Ask LLMs such as ChatGPT to generate distractors (***Zero-ChatGPT***)

Original Question and Correct Answer:

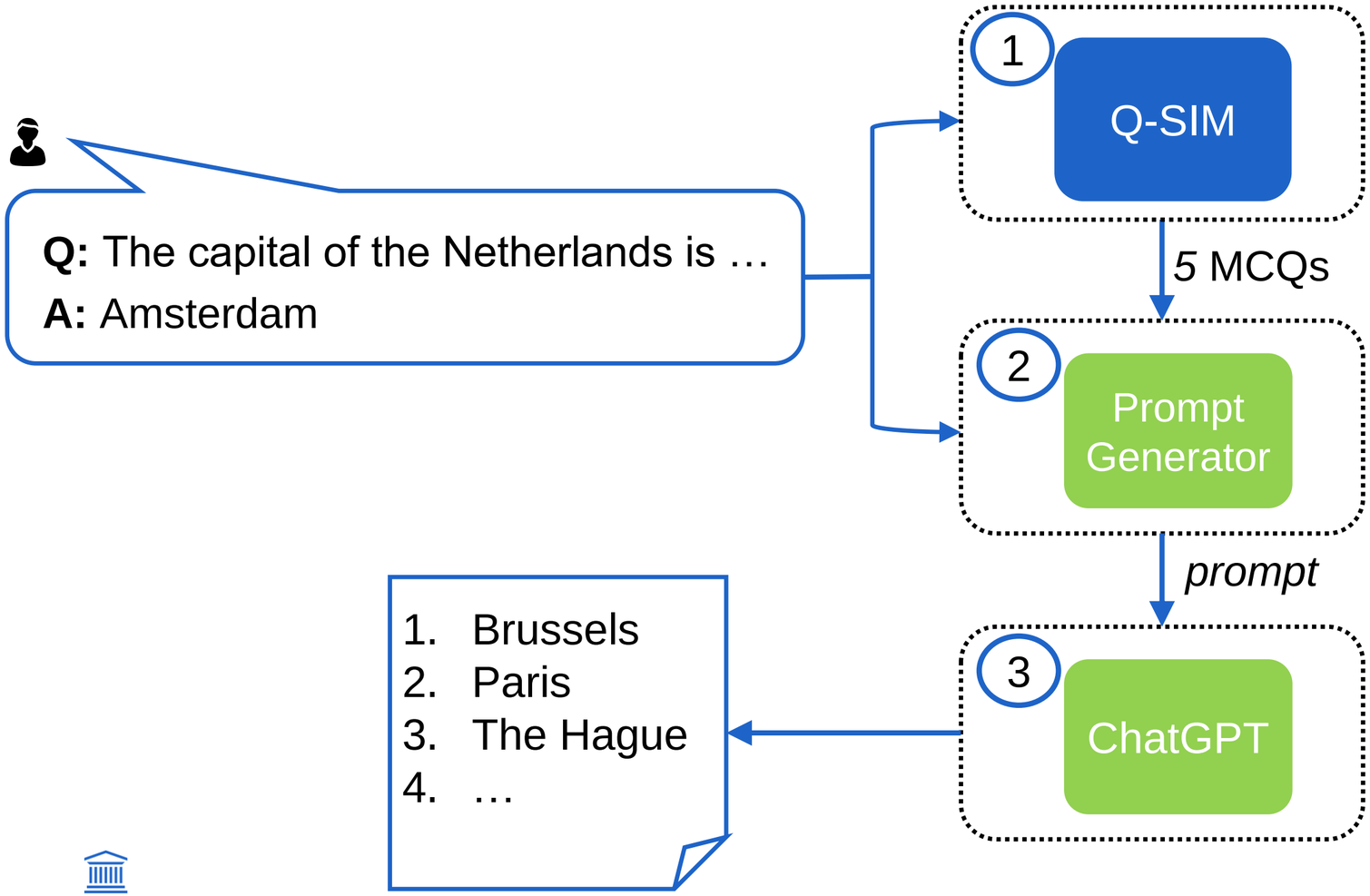
The capital of the Netherlands is …
Amsterdam

Zero-shot prompt input:

Generate 10 plausible but incorrect answers for the following question.
Question: The capital of the Netherlands is …
Answer: Amsterdam

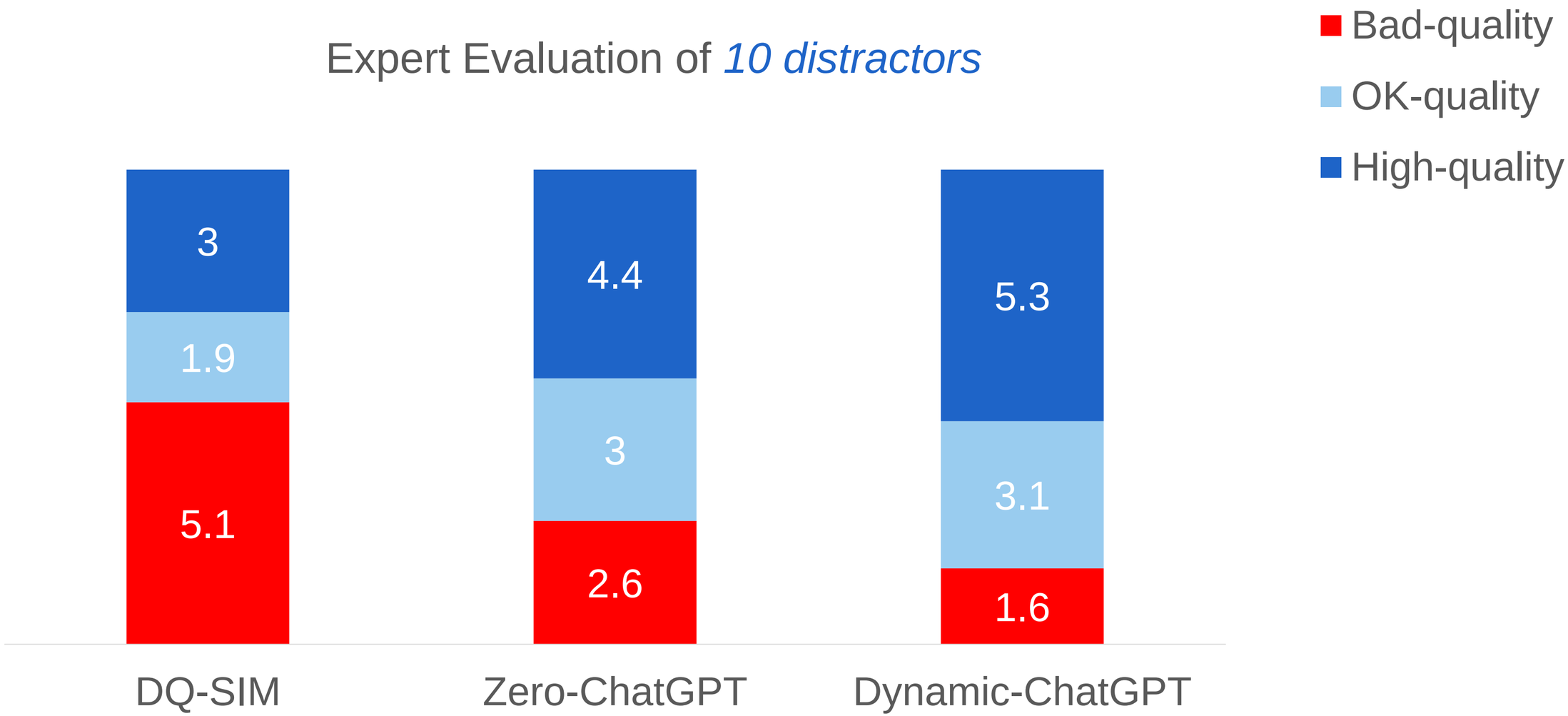
Our approach (Dynamic-ChatGPT)

- We proposed a pipeline approach with three components



GHENT UNIVERSITY imec

Performance gain on quality of generated distractors



Gap-filling Grammar

Exercise Generation

Research questions

1. How to automatically generate distractors?

2. How to design a system that generates free-form distractors?

3. How to adapt a language model to generate gap-fill grammar exercises?

Gap-filling Grammar Exercise Generation

Gap-filling grammar exercise important in formative assessment

Gap-filling Grammar Exercise Generation

new tense

Gap-filling Grammar Exercise Generation

new tense

John arrived from the airport yesterday at 8:00. He checked into the hotel at 9:00, and met his colleagues at 10:00. He will present his work tomorrow.

Gap-filling Grammar Exercise Generation

new tense

Past tense

John *arrived* from the airport yesterday at 8:00. He *checked* into the hotel at 9:00, and *met* his colleagues at 10:00. He will present his work tomorrow.

Gap-filling Grammar Exercise Generation

new tense

Past tense

John *arrived* from the airport yesterday at 8:00. He *checked* into the hotel at 9:00, and *met* his colleagues at 10:00. He will present his work tomorrow.

Input

Guardiola admitted there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It worked to devastating effect.

City will play Chelsea in next week.

Gap-filling Grammar Exercise Generation

new tense

Past tense

John *arrived* from the airport yesterday at 8:00. He *checked* into the hotel at 9:00, and *met* his colleagues at 10:00. He will present his work tomorrow.

Input

Guardiola admitted there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It worked to devastating effect.

City will play Chelsea in next week.

LM-based Exercise Generator

Output

Guardiola *admitted* there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It *worked* to devastating effect.

City will play Chelsea in next week.

Gap-filling Grammar Exercise Generation

new tense

Future tense

John arrived from the airport yesterday at 8:00. He checked into the hotel at 9:00, and met his colleagues at 10:00. He *will present* his work tomorrow.

Input

Guardiola admitted there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It worked to devastating effect.

City will play Chelsea in next week.

LM-based Exercise Generator

Gap-filling Grammar Exercise Generation

new tense

Future tense

John arrived from the airport yesterday at 8:00. He checked into the hotel at 9:00, and met his colleagues at 10:00. He *will present* his work tomorrow.

Input

Guardiola admitted there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It worked to devastating effect.

City will play Chelsea in next week.

LM-based Exercise Generator

Output

Guardiola admitted there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It worked to devastating effect.

City *will play* Chelsea in next week.

Gap-filling Grammar Exercise Generation

combine tenses

Past + Future

John *arrived* from the airport yesterday at 8:00. He *checked* into the hotel at 9:00, and *met* his colleagues at 10:00. He *will present* his work tomorrow.

Input

Guardiola admitted there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It worked to devastating effect.

City will play Chelsea in next week.

LM-based Exercise Generator

Gap-filling Grammar Exercise Generation

combine tenses

Past + Future

John *arrived* from the airport yesterday at 8:00. He *checked* into the hotel at 9:00, and *met* his colleagues at 10:00. He *will present* his work tomorrow.

Input

Guardiola admitted there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It worked to devastating effect.

City will play Chelsea in next week.

LM-based Exercise Generator

Output

Guardiola *admitted* there was a deliberate ploy in this game to use Debruyne to unlock Arsenal. It *worked* to devastating effect.

City *will play* Chelsea in next week.

Our approach

1. Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

City

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

City will

City

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

City will

City

City will play

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

City will

City

City will play Chelsea

City will play

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

will

City will

City will play Chelsea

City

City will play

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

will

City will

City will play Chelsea

City

will play

City will play

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

will

City will

will play Chelsea

City will play Chelsea

City

will play

City will play

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

will play Chelsea in

will

City will

will play Chelsea

City will play Chelsea

City

will play

City will play

Our approach

1 Given input text, extract spans (e.g. size of 4)

Input

City will play Chelsea in next week.

will play Chelsea in

will

City will

will play Chelsea

City will play Chelsea

City

will play

City will play

2 Decide if each span is a suitable gap based on example

Exemplar

He *will present* his work tomorrow.

Our approach

1 Given input text, extract spans (e.g. size of 4)

2 Decide if each span is a suitable gap based on example

Input

City will play Chelsea in next week.

Exemplar

He *will present* his work tomorrow.

Gap?

will play Chelsea in

will

City will

will play Chelsea

City will play Chelsea

City

will play

City will play

Our approach

1 Given input text, extract spans (e.g. size of 4)

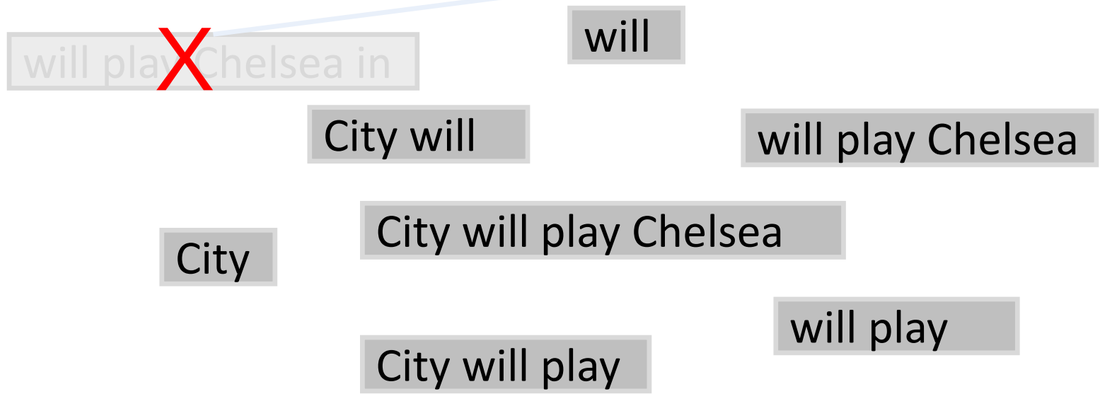
2 Decide if each span is a suitable gap based on example

Input

City will play Chelsea in next week.

Exemplar

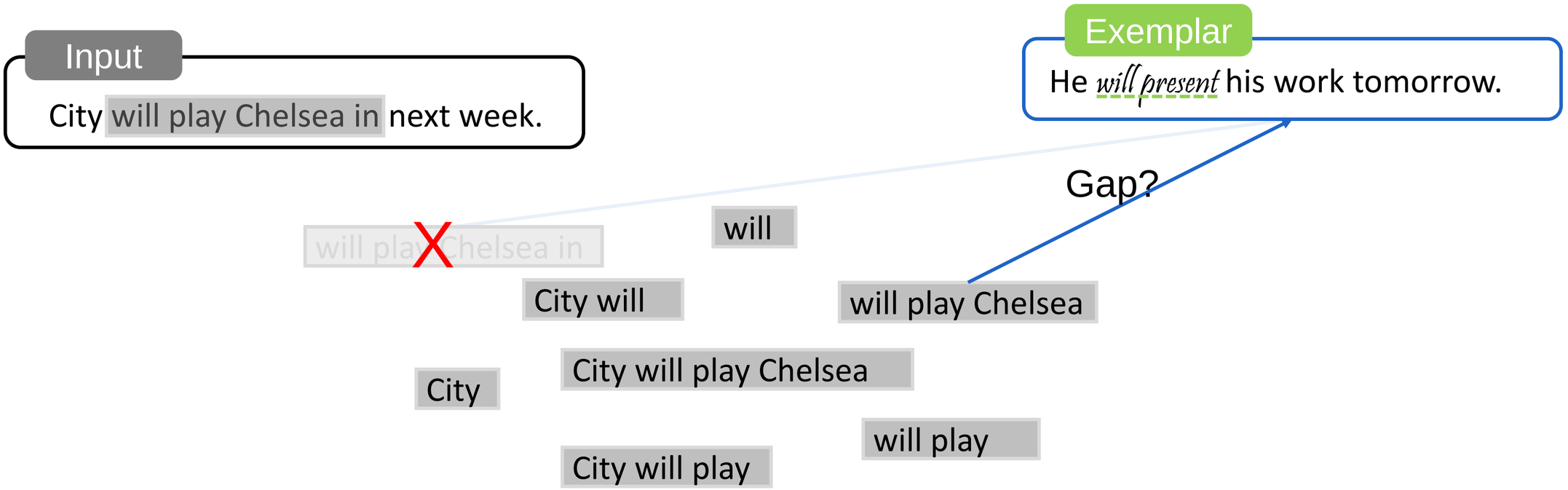
He *will present* his work tomorrow.



Our approach

1 Given input text, extract spans (e.g. size of 4)

2 Decide if each span is a suitable gap based on example



Our approach

1 Given input text, extract spans (e.g. size of 4)

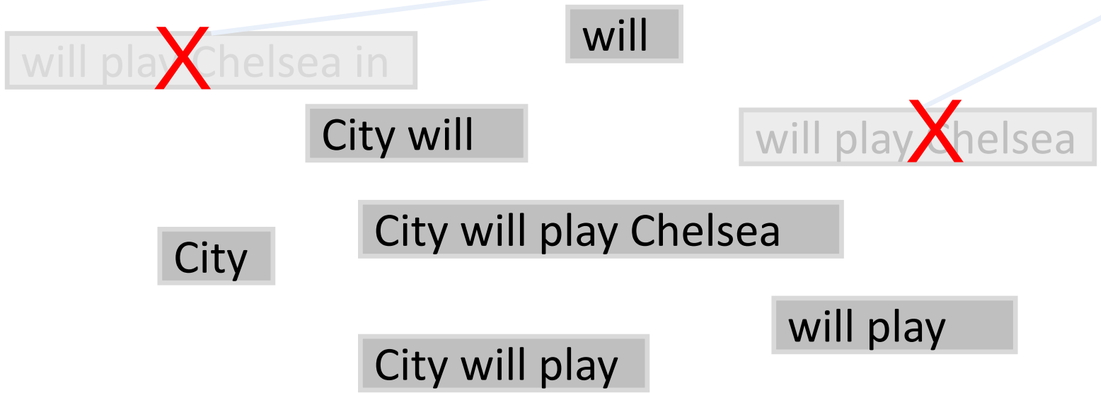
2 Decide if each span is a suitable gap based on example

Input

City will play Chelsea in next week.

Exemplar

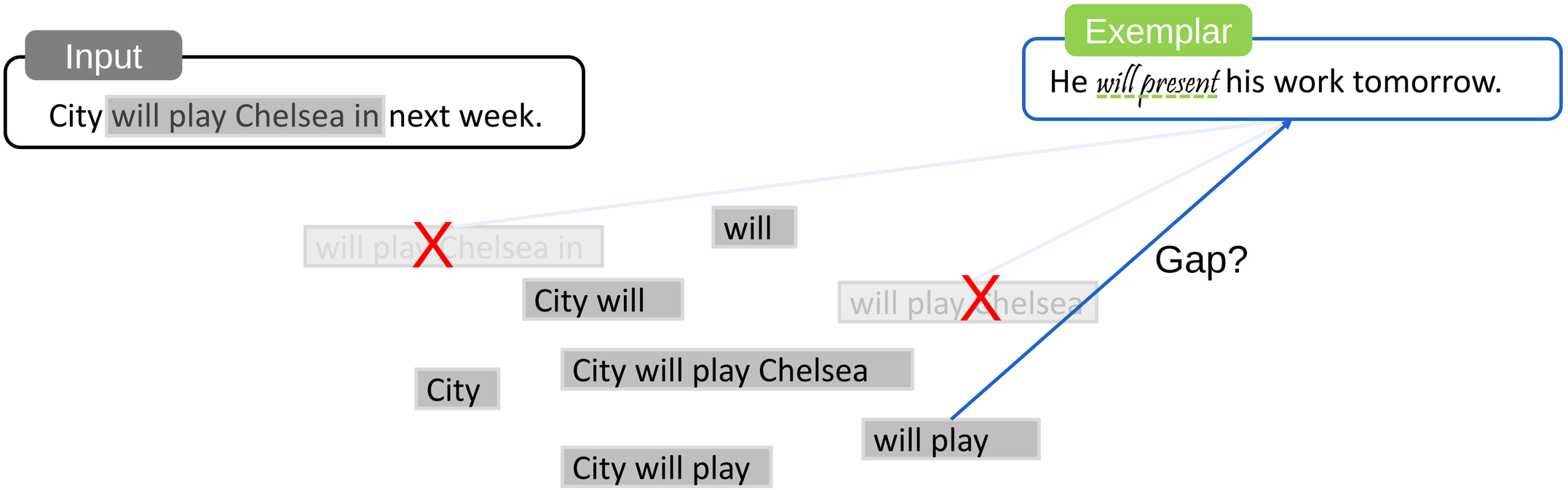
He *will present* his work tomorrow.



Our approach

1 Given input text, extract spans (e.g. size of 4)

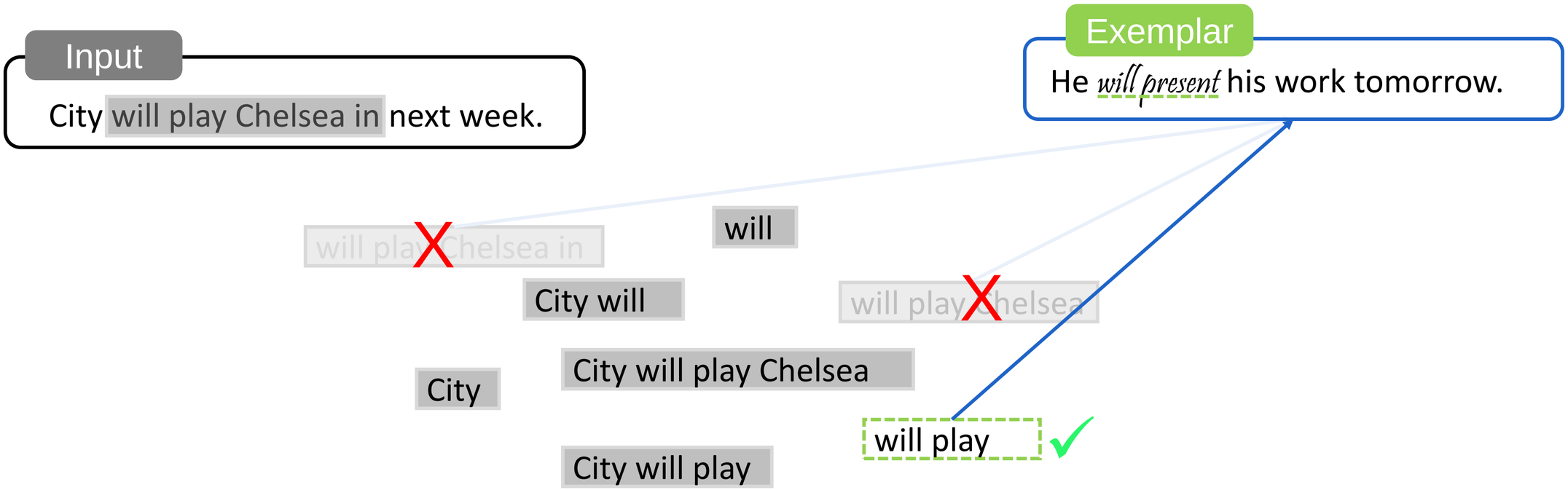
2 Decide if each span is a suitable gap based on example



Our approach

1 Given input text, extract spans (e.g. size of 4)

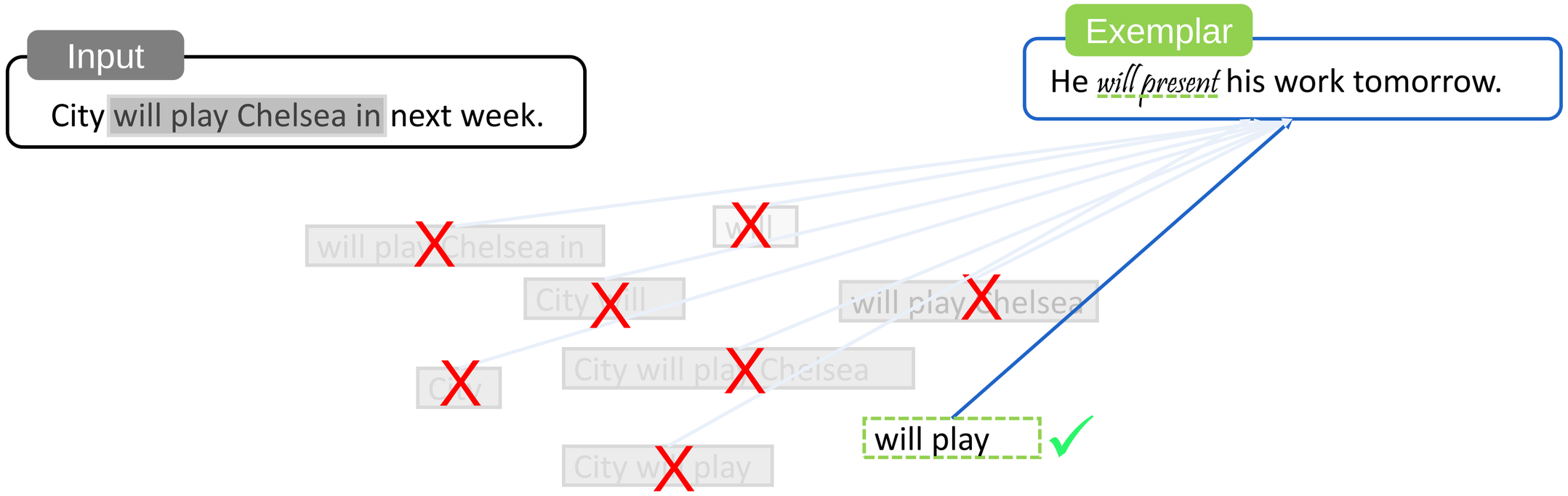
2 Decide if each span is a suitable gap based on example



Our approach

1 Given input text, extract spans (e.g. size of 4)

2 Decide if each span is a suitable gap based on example



Our approach

1 Given input text, extract spans (e.g. size of 4)

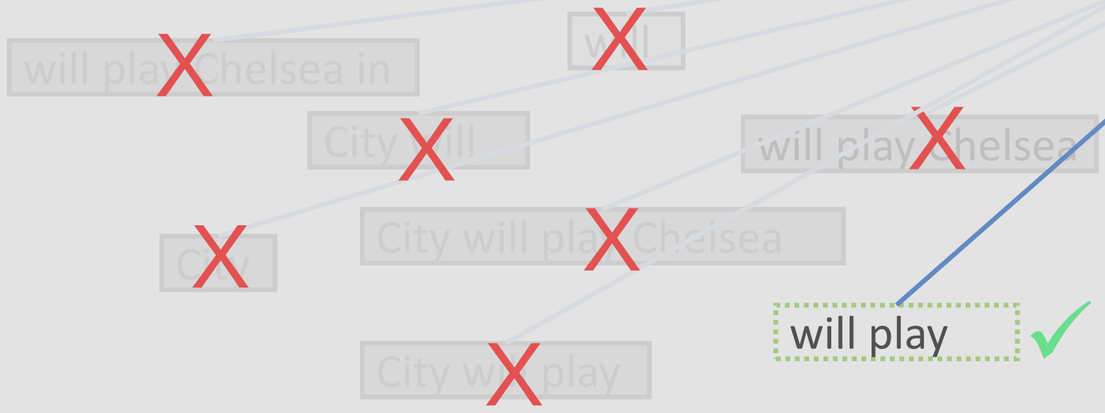
2 Decide if each span is a suitable gap based on example

Input

City will play Chelsea in next week.

Exemplar

He *will present* his work tomorrow.



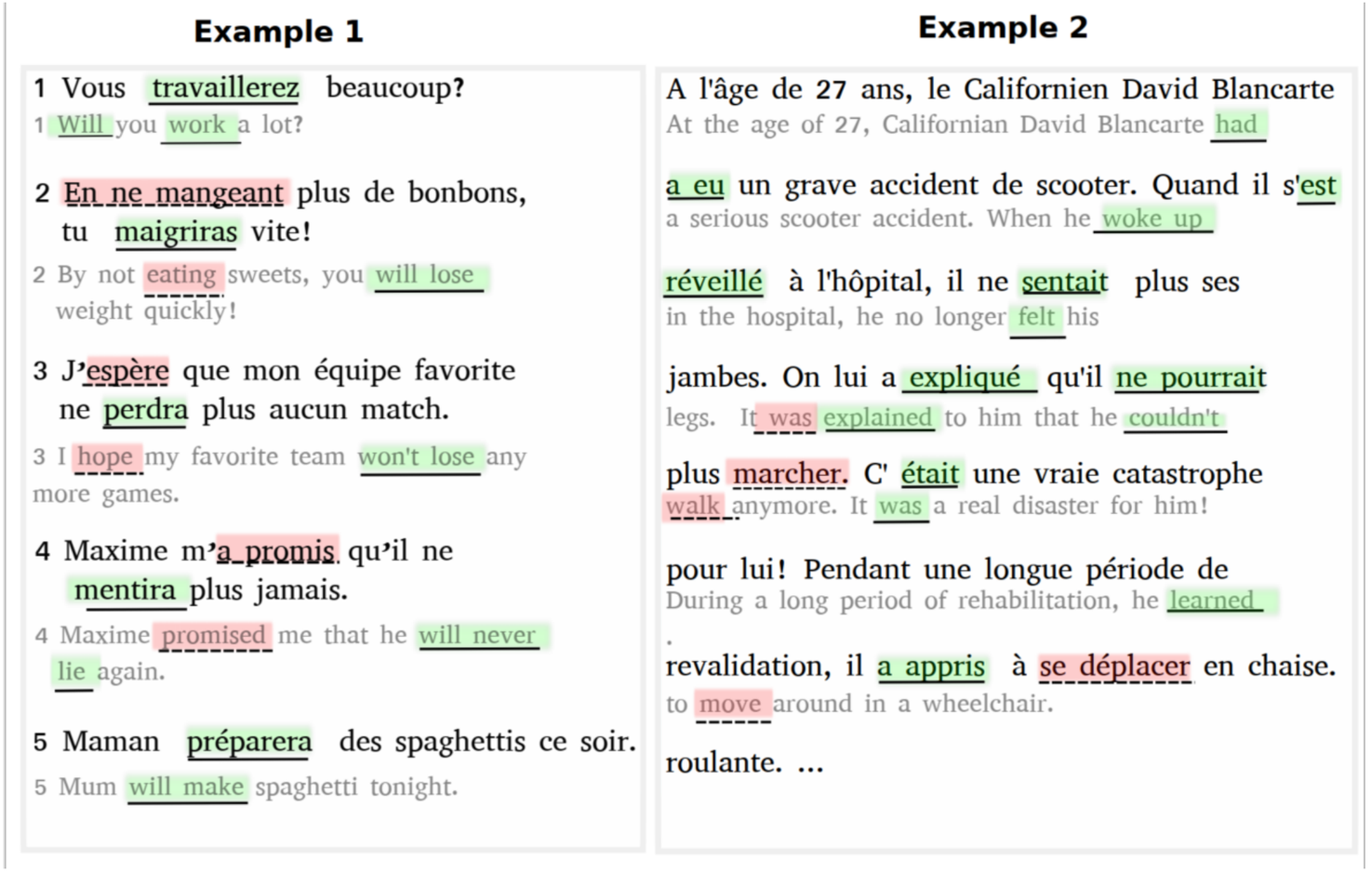
New Exercise

City *will play* Chelsea in next week.

GF2 Grammar Dataset

A real-world corpus of existing exercises on mixed grammar topics

- ~ 800 exercise documents
- French grammar exercises



Example 1

1 Vous travaillerez beaucoup?
1 Will you work a lot?

2 En ne mangeant plus de bonbons, tu maigriras vite!
2 By not eating sweets, you will lose weight quickly!

3 J'espère que mon équipe favorite ne perdra plus aucun match.
3 I hope my favorite team won't lose any more games.

4 Maxime m'a promis qu'il ne mentira plus jamais.
4 Maxime promised me that he will never lie again.

5 Maman préparera des spaghettis ce soir.
5 Mum will make spaghetti tonight.

Example 2

A l'âge de 27 ans, le Californien David Blancarte
At the age of 27, Californian David Blancarte had

a eu un grave accident de scooter. Quand il s'est
a serious scooter accident. When he woke up

réveillé à l'hôpital, il ne sentait plus ses
in the hospital, he no longer felt his

jambes. On lui a expliqué qu'il ne pourrait
legs. It was explained to him that he couldn't

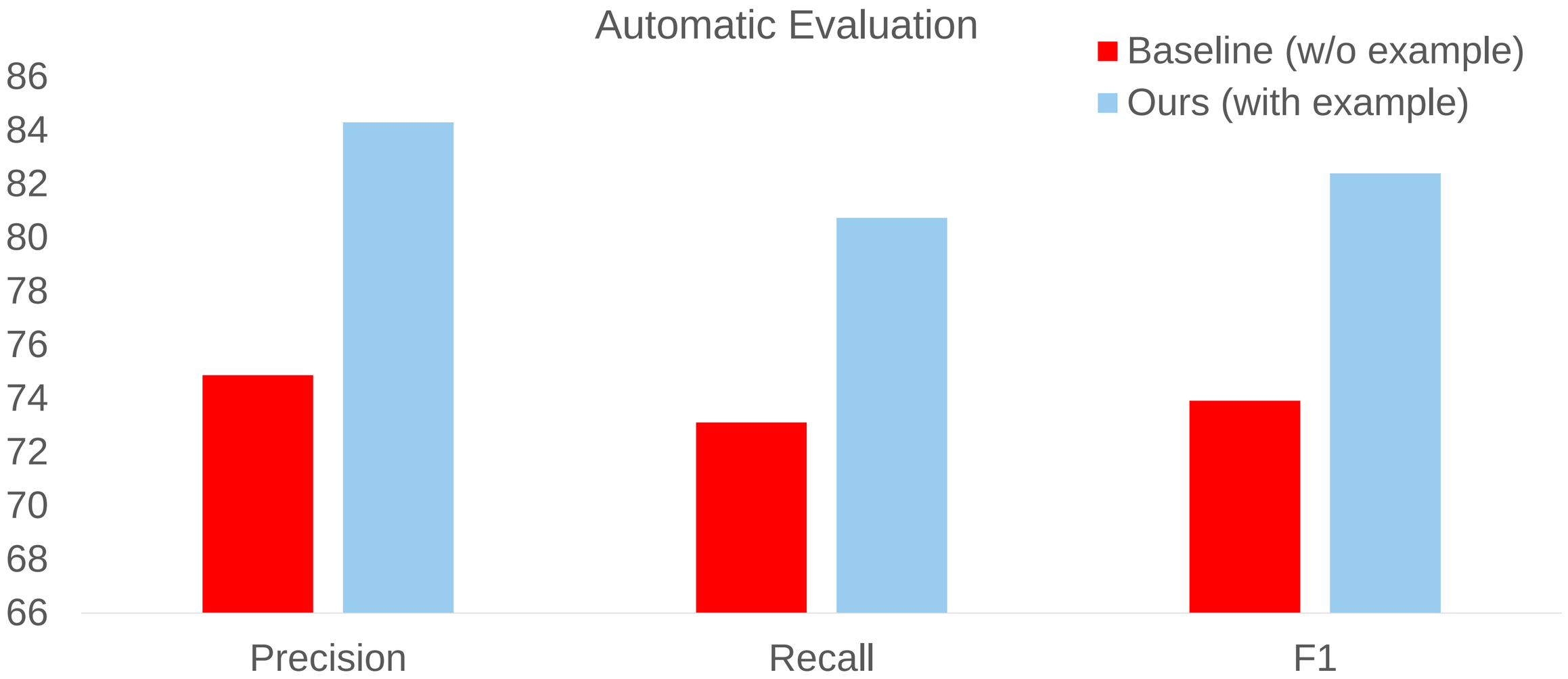
plus marcher. C' était une vraie catastrophe
walk anymore. It was a real disaster for him!

pour lui! Pendant une longue période de
During a long period of rehabilitation, he learned

.
revalidation, il a appris à se déplacer en chaise.
to move around in a wheelchair.

roulante. ...

Gain in performance using our approach



Take-away Messages

Burden on educators

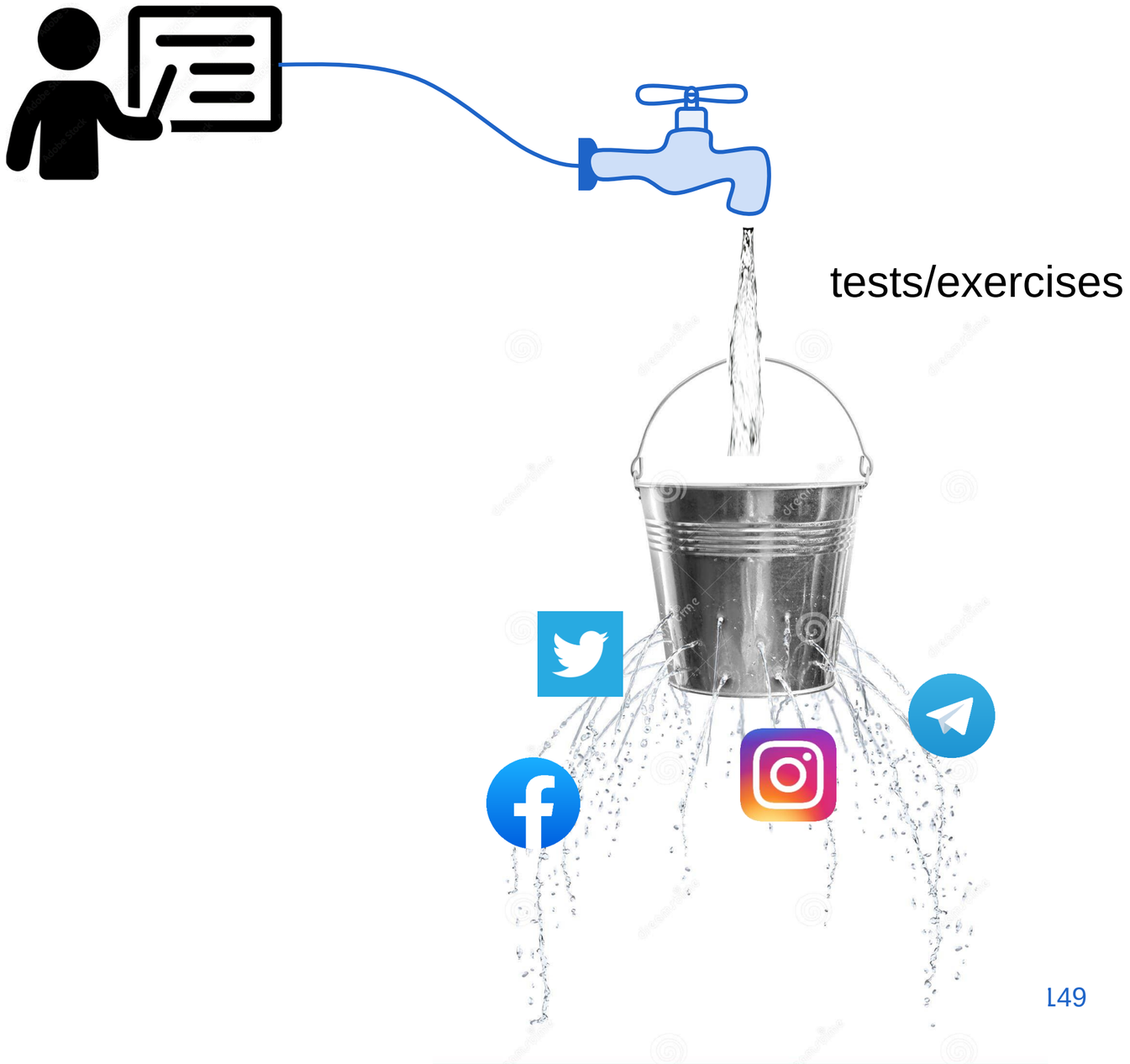
GHENT UNIVERSITY imec

Burden on educators

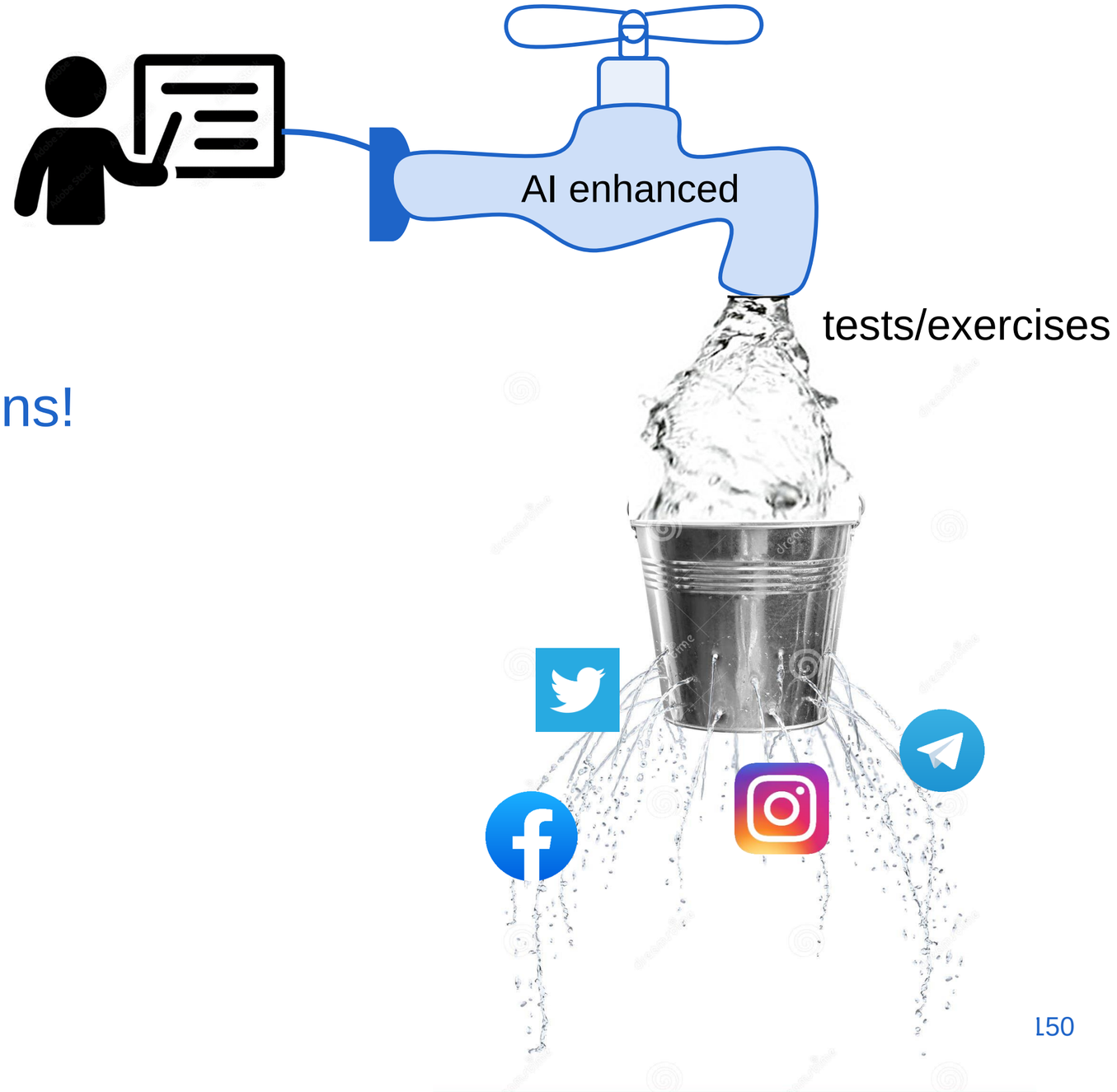
Tiny bit less burden!

GHENT UNIVERSITY imec

Single-use dilemma



Single-use dilemma



Ability to generate many questions!

Distractor Generation

- Smartly combining locally fine-tuned models with LLMs such as ChatGPT improves distractor generation

Distractor Generation

- Smartly combining locally fine-tuned models with LLMs such as ChatGPT improves distractor generation
- Teacher evaluations: critical for assessing quality of AI educational tools

Gap-filling exercise generation

- Our example-aware gap-filling exercise generator better than baseline that doesn’t integrate examples

Gap-filling exercise generation

- Our example-aware gap-filling exercise generator better than baseline that doesn’t integrate examples
- Facilitates targeted **formative** assessment
  - Teachers can use it for introducing new grammar exercises
  - Students can practice on grammar exercises they want to master

Ethical and practical consideration

- Deployment of AI tools in education:
  - Rigorous critical evaluation of AI models and human supervision
  - Emphasis on privacy of students and educators!

Ethical and practical consideration

- Deployment of AI tools in education:
  - Rigorous critical evaluation of AI models and human supervision
  - Emphasis on privacy of students and educators!

- Address data bias in LLMs

Ethical and practical consideration

- Deployment of AI tools in education:
  - Rigorous critical evaluation of AI models and human supervision
  - Emphasis on privacy of students and educators!

- Address data bias in LLMs

- Consider environmental impact during the development of these tools!

Thanks!

Let’s do a quick poll: Personal Trivia

What is my last name?

A Kiros

C Semere

B Bitew

D Kiros Bitew

Let’s do a quick poll: Personal Trivia

What is my last name?

A BMW

C Gent

B Bitew

D Kiros Bitew