

# Focused Attention Improves Document-Grounded Generation

**Shrimai Prabhume**, Kazuma Hashimoto,  
Yingbo Zhou, Alan W Black, Ruslan Salakhutdinov



**Carnegie Mellon University**

Language Technologies Institute



# Document Grounded Generation

# Document Grounded Generation



Context

# Document Grounded Generation



Context

+



Document

# Document Grounded Generation



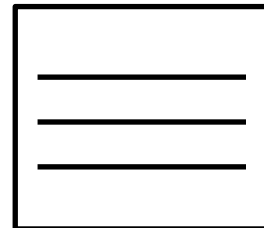
Context

+



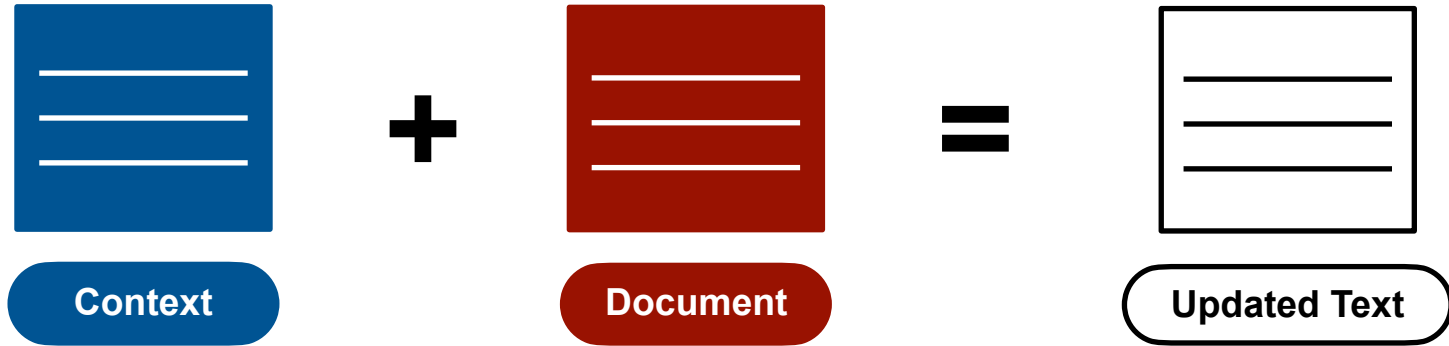
Document

=



Updated Text

# Document Grounded Generation



- generate text given a context and a document
- generated text should *coherently fit* the context and contain new information from the document

# Wikipedia Update Generation

# Wikipedia Update Generation



The screenshot shows the Wikipedia article page for "Monkey selfie copyright dispute". At the top, it indicates the user is "Not logged in" and provides links for "Talk", "Contributions", "Create account", and "Log in". Below this is a navigation bar with "Article" and "Talk" tabs, and a search box labeled "Search Wikipedia". The article title "Monkey selfie copyright dispute" is prominently displayed, followed by the subtext "From Wikipedia, the free encyclopedia". The main text begins with "The **monkey selfie copyright dispute** is a series of disputes about the **copyright** status of **selfies** taken by **Celebes crested macaques** using equipment belonging to the British nature photographer David Slater. The disputes involve **Wikimedia Commons** and the blog *Techdirt*, which have hosted the images following their publication in newspapers in July 2011 over Slater's objections that he holds the copyright, and **People for the Ethical Treatment of Animals** (PETA), who have argued that the macaque should be assigned the copyright." To the right of the text is a photograph of a macaque with its mouth open, showing its teeth. Below the image is a caption: "One of the monkey selfies at issue". On the left side of the page, there is a sidebar with various navigation links such as "Main page", "Contents", "Current events", "Random article", "About Wikipedia", "Contact us", "Donate", "Contribute", "Help", "Community portal", "Recent changes", "Upload file", "Tools", "What links here", "Related changes", and "Special pages".



# Wikipedia Update Generation

WIKIPEDIA  
The Free Encyclopedia

Main page  
Contents  
Current events  
Random article  
About Wikipedia  
Contact us  
Donate

Contribute  
Help  
Community portal  
Recent changes  
Upload file

Tools  
What links here  
Related changes  
Special pages

Not logged in | Talk | Contributions | Create account | Log in

Article [Talk](#) [Read](#) [Edit](#) [View history](#)

## Monkey selfie copyright dispute

From Wikipedia, the free encyclopedia

The **monkey selfie copyright dispute** is a series of disputes about the [copyright](#) status of [selfies](#) taken by [Celebes crested macaques](#) using equipment belonging to the British nature photographer David Slater. The disputes involve [Wikimedia Commons](#) and the blog *Techdir*, which have hosted the images following their publication in newspapers in July 2011 over Slater's objections that he holds the copyright, and [People for the Ethical Treatment of Animals](#) (PETA), who have argued that the macaque should be assigned the copyright.

One of the monkey selfies at issue

Slater has argued that he has a valid copyright claim, as he engineered the situation that resulted in the pictures by travelling to Indonesia, befriending a group of wild macaques, and setting up his camera equipment in such a way that a "selfie" picture might come about. The [Wikimedia Foundation's](#) 2014 refusal to remove the pictures from its Wikimedia Commons image library was based on the understanding that copyright is held by the creator, that a non-human creator (not being a [legal person](#)) cannot hold copyright, and that the images are thus in the [public domain](#).

Amateur Photographer

Subscribe now Digital

Facebook Twitter Instagram YouTube Google+ RSS

Home Latest Opinion Reviews How To Inspiration AP Awards The Video Mode

Home / Latest / Latest news

## Ape-rture priority photographer plays down monkey reports

Chris Cheesman  
July 5, 2011

Facebook Twitter Pinterest 0 shares

**A photographer who says he witnessed monkeys taking pictures of themselves, tells Amateur Photographer (AP) that much of the media coverage has been exaggerated.**

Wildlife photographer David Slater today played down newspaper reports that suggest a bunch of Indonesian monkeys grabbed his camera and began taking self-portraits.

And he revealed that the shots, featuring the endangered crested black macaque monkey, were taken three years ago.

Speaking to AP, David explained that his camera had been mounted on a tripod when the primates

### Latest reviews

#### Leica M10-R review

Andy Westlake takes a detailed look at Leica's latest high-resolution...

#### Canon EOS R6 review: hands-on first look

Look out Sony - Canon may just have got full-

# Wikipedia Update Generation



WIKIPEDIA  
The Free Encyclopedia

Article [Talk](#)

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Read [Edit](#) [View history](#)

## Monkey selfie copyright dispute

From Wikipedia, the free encyclopedia

The **monkey selfie copyright dispute** is a series of disputes about the [copyright](#) status of [selfies](#) taken by [Celebes crested macaques](#) using equipment belonging to the British nature photographer David Slater. The disputes involve [Wikimedia Commons](#) and the blog *Techdir*, which have hosted the images following their publication in newspapers in July 2011 over Slater's objections that he holds the copyright, and [People for the Ethical Treatment of Animals](#) (PETA), who have argued that the macaque should be assigned the copyright.



Slater has argued that he has a valid copyright claim, as he engineered the situation that resulted in the pictures by travelling to Indonesia, befriending a group of wild macaques, and setting up his camera equipment in such a way that a "selfie" picture might come about. The [Wikimedia Foundation's](#) 2014 refusal to remove the pictures from its [Wikimedia Commons](#) image library was based on the understanding that copyright is held by the creator, that a non-human creator (not being a [legal person](#)) cannot hold copyright, and that the images are thus in the [public domain](#).

One of the monkey selfies at issue

The following day, *Amateur Photographer* reported that Slater gave them further explanation as to how the photographs were created, downplaying the way newspaper articles had described them; Slater said reports that a monkey ran off with his camera and "began taking self-portraits" were incorrect and that the portrait was shot when his camera had been mounted on a tripod, with the primates playing around with a remote cable release as he fended off other monkeys.<sup>[13]</sup>



Amateur Photographer

Subscribe now Digital

Home Latest Opinion Reviews How To Inspiration AP Awards The Video Mode

Home / Latest / Latest news

## Ape-rture priority photographer plays down monkey reports

Chris Cheesman  
July 5, 2011

0 shares

**A photographer who says he witnessed monkeys taking pictures of themselves, tells Amateur Photographer (AP) that much of the media coverage has been exaggerated.**

Wildlife photographer David Slater today played down newspaper reports that suggest a bunch of Indonesian monkeys grabbed his camera and began taking self-portraits.

And he revealed that the shots, featuring the endangered crested black macaque monkey, were taken three years ago.

Speaking to AP, David explained that his camera had been mounted on a tripod when the primates

### Latest reviews

**Leica M10-R review**

Andy Westlake takes a detailed look at Leica's latest high-resolution...

**Canon EOS R6 review: hands-on first look**

Look out Sony - Canon may just have got full-

# Wikipedia Update Generation



WIKIPEDIA  
The Free Encyclopedia

Article [Talk](#)

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Read [Edit](#) [View history](#)

## Monkey selfie copyright dispute

From Wikipedia, the free encyclopedia

The **monkey selfie copyright dispute** is a series of disputes about the [copyright](#) status of [selfies](#) taken by [Celebes crested macaques](#) using equipment belonging to the British nature photographer David Slater. The disputes involve [Wikimedia Commons](#) and the blog *Techdirt*, which have hosted the images following their publication in newspapers in July 2011 over Slater's objections that he holds the copyright, and [People for the Ethical Treatment of Animals](#) (PETA), who have argued that the macaque should be assigned the copyright.



Slater has argued that he has a valid copyright claim, as he engineered the situation that resulted in the pictures by travelling to Indonesia, befriending a group of wild macaques, and setting up his camera equipment in such a way that a "selfie" picture might come about. The [Wikimedia Foundation's](#) 2014 refusal to remove the pictures from its Wikimedia Commons image library was based on the understanding that copyright is held by the creator, that a non-human creator (not being a [legal person](#)) cannot hold copyright, and that the images are thus in the [public domain](#).

One of the monkey selfies at issue

The following day, *Amateur Photographer* reported that Slater gave them further explanation as to how the photographs were created, downplaying the way newspaper articles had described them; Slater said reports that a monkey ran off with his camera and "began taking self-portraits" were incorrect and that the portrait was shot when his camera had been mounted on a tripod, with the primates playing around with a remote cable release as he fended off other monkeys.<sup>[13]</sup>



Amateur Photographer

Subscribe now Digital

Home Latest Opinion Reviews How To Inspiration AP Awards The Video Mode

Home / Latest / Latest news

## Ape-rture priority photographer plays down monkey reports

Chris Cheesman  
July 5, 2011

0 shares

**A photographer who says he witnessed monkeys taking pictures of themselves, tells Amateur Photographer (AP) that much of the media coverage has been exaggerated.**

Wildlife photographer David Slater today played down newspaper reports that suggest a bunch of Indonesian monkeys grabbed his camera and began taking self-portraits.

And he revealed that the shots, featuring the endangered crested black macaque monkey, were taken three years ago.

Speaking to AP, David explained that his camera had been mounted on a tripod when the primates

### Latest reviews

**Leica M10-R review**

Andy Westlake takes a detailed look at Leica's latest high-resolution...

**Canon EOS R6 review: hands-on first look**

Look out Sony - Canon may just have got full-

Generate text that is:

1. **coherent** to the Wikipedia context
  2. contains **new events** from the article
- Dataset Size of 636k!**

# Document Grounded Dialogue

# Document Grounded Dialogue

## Chat History

**Alice:** Yes this is a Christmas classic ...

**Bob:** Very well said! Some of the critics were ...

**Alice:** Yes I found the first half very entertaining.

**Bob:** I feel you! So hard to believe it was  
released in 1990!

# Document Grounded Dialogue

## Chat History

**Alice:** Yes this is a Christmas classic ...

**Bob:** Very well said! Some of the critics were ...

**Alice:** Yes I found the first half very entertaining.

**Bob:** I feel you! So hard to believe it was released in 1990!



WIKIPEDIA  
The Free Encyclopedia

Article **Talk**

## *Home Alone*

From Wikipedia, the free encyclopedia

*Home Alone* is a 1990 American [comedy film](#) written and produced by [John Hughes](#) and directed by [Chris Columbus](#). It stars [Macaulay Culkin](#), [Joe Pesci](#), [Daniel Stern](#), [John Heard](#), and [Catherine O'Hara](#). Culkin stars as eight-year-old Kevin, who defends his home from burglars Harry and Marv (Pesci and Stern) after his family accidentally leaves him behind on their vacation.

*Home Alone* was filmed between February and May 1990 in locations throughout [Illinois](#). It was originally set to be distributed by [Warner Bros.](#), but transferred to [20th Century Fox](#) when Hughes exceeded Warner's \$10 million budget.

## Document



# Document Grounded Dialogue

## Chat History

**Alice:** Yes this is a Christmas classic ...

**Bob:** Very well said! Some of the critics were ...

**Alice:** Yes I found the first half very entertaining.

**Bob:** I feel you! So hard to believe it was released in 1990!

## Response

**Alice:** I like the quote "**sadistic festival of adult-bashing**". Between's **Kevin's parents forgetting him and the burglars** this movie doesn't make adult looks too bright.



WIKIPEDIA  
The Free Encyclopedia

Article **Talk**

## *Home Alone*

From Wikipedia, the free encyclopedia

*Home Alone* is a 1990 American [comedy film](#) written and produced by [John Hughes](#) and directed by [Chris Columbus](#). It stars [Macaulay Culkin](#), [Joe Pesci](#), [Daniel Stern](#), [John Heard](#), and [Catherine O'Hara](#). Culkin stars as eight-year-old Kevin, who defends his home from burglars Harry and Marv (Pesci and Stern) after his family accidentally leaves him behind on their vacation.

*Home Alone* was filmed between February and May 1990 in locations throughout [Illinois](#). It was originally set to be distributed by [Warner Bros.](#), but transferred to [20th Century Fox](#) when Hughes exceeded Warner's \$10 million budget.

## Document



# Document Grounded Dialogue

**Alice:** Yes this is a Christmas classic ...

**Bob:** Very well said! Some of the critics were ...

**Alice:** Yes I found the first half very entertaining

Chat History

- Generate response that is:
  1. *coherent* to the dialogue context (chat history)
  2. contains *information* from the document
- CMU Document Grounded Conversations (**CMU\_DoG**) Dataset
- Dataset size of **130k utterances!**

**Alice:** I like the quote "**sadistic festival of adult-bashing**". Between's **Kevin's parents forgetting him and the burglars** this movie doesn't make adult looks too bright.



Document





# Models

- Pre-trained encoder-decoder model
  - provide a stronger baseline for the task

# Pre-trained Models

## DOC

...

Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.

# Pre-trained Models

## DOC

...

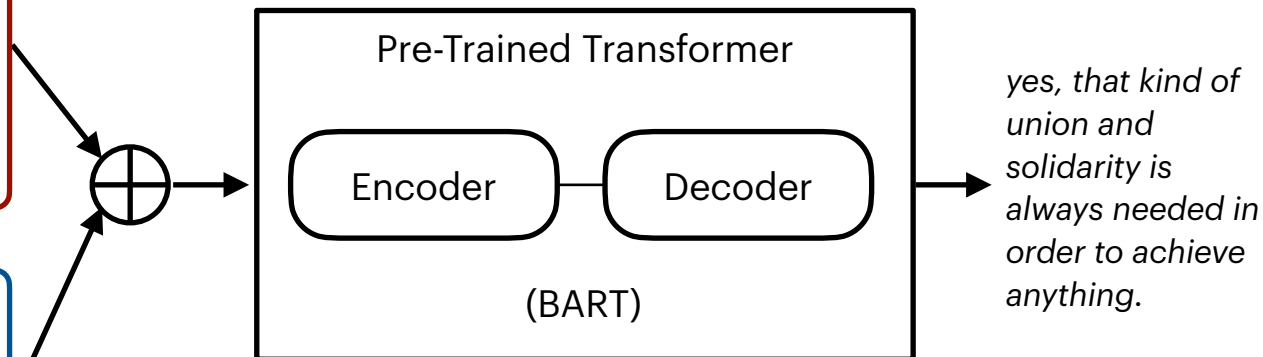
Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house.*

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.



# Models

- Pre-trained encoder-decoder model
  - provide a stronger baseline for the task
- Context Driven Representation (CoDR)
  - build a representation of the content in document that is ***not present*** in the context
- Document Headed Attention (DoHA)
  - specific ***focus on document*** information

# Context Driven Representation (CoDR)

## DOC

...

Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.

# Context Driven Representation (CoDR)

## DOC

...

Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house.*

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.

Pre-Trained Transformer  
Encoder  
(BART Encoder)



# Context Driven Representation (CoDR)

## DOC

...

Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.

Pre-Trained Transformer  
Encoder  
(BART Encoder)



$h_d$



$h_c$

# Context Driven Representation (CoDR)

## DOC

...

Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.

$h_d$



$h_c$





# Context Driven Representation (CoDR)

## DOC

...

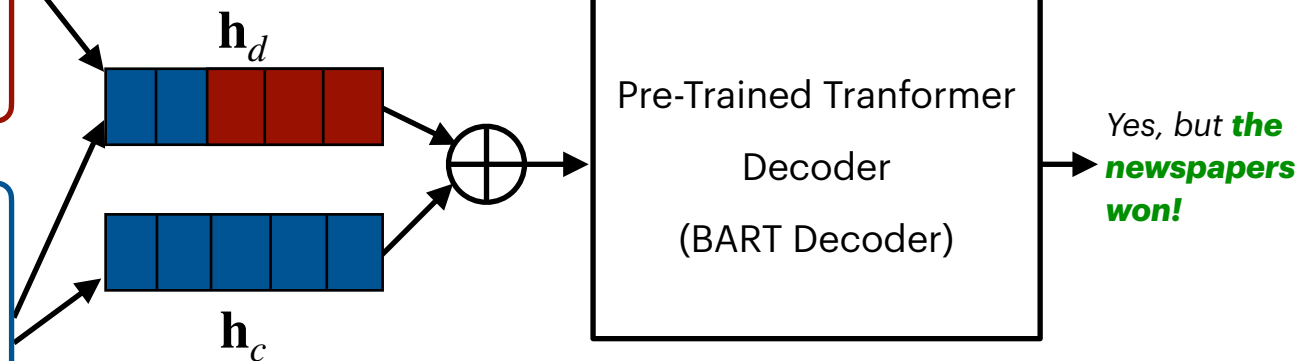
Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.



# Document Headed Attention (DoHA)

## DOC

...

Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.

$h_d$



$h_c$



# Document Headed Attention (DoHA)

## DOC

...

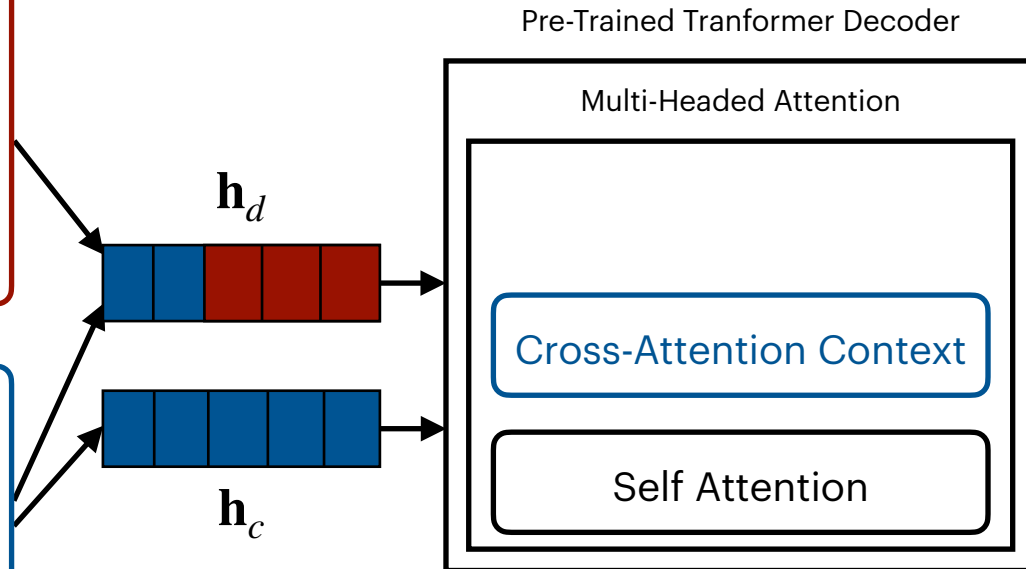
Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.



# Document Headed Attention (DoHA)

## DOC

...

Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.

 $h_d$  $h_c$ 

Pre-Trained Transformer Decoder

Multi-Headed Attention

Cross-Attention Doc

Cross-Attention Context

Self Attention

# Document Headed Attention (DoHA)

## DOC

...

Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house*.

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.

$$\text{CrossAttention\_Doc}(Q, K, V) = [H_1, \dots, H_m] \mathbf{W}^{do}$$

$$\mathbf{H}_i = \text{Attention}(Q \mathbf{W}_i^{dQ}, K \mathbf{W}_i^{dK}, V \mathbf{W}_i^{dV})$$

Pre-Trained Transformer Decoder

Multi-Headed Attention

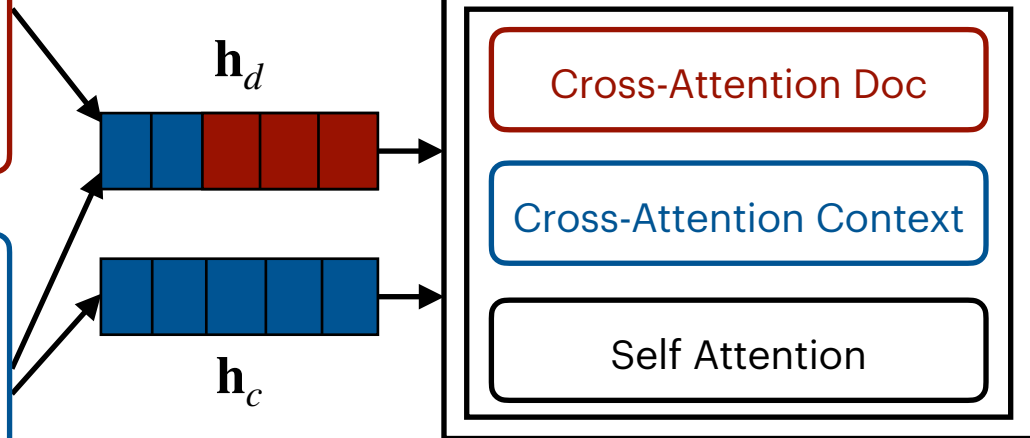
Cross-Attention Doc

Cross-Attention Context

Self Attention

$\mathbf{h}_d$

$\mathbf{h}_c$



# Document Headed Attention (DoHA)

## DOC

...

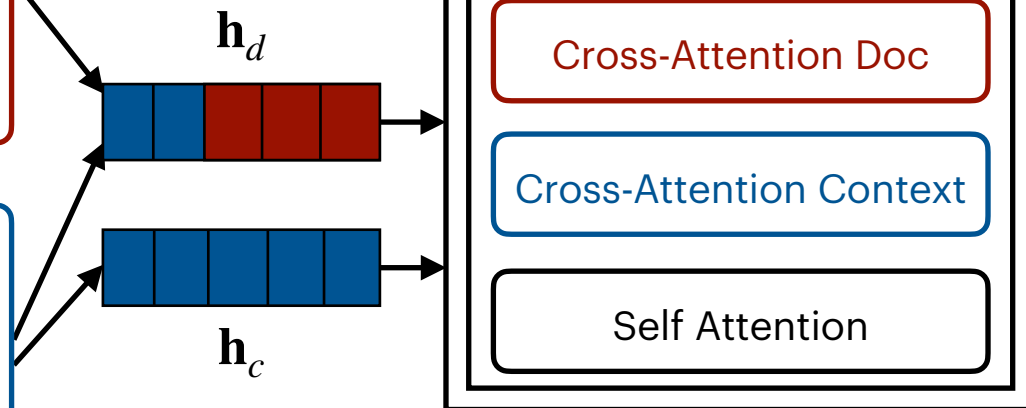
Meanwhile, *newspapers across the country pick up the story in solidarity with the post and times*. The court rules 6 3 in the newspapers' favor, vindicating graham's decision. *Nixon demands that the post should be barred from the white house.*

...

One year later, a security guard discovers a break-in progress at watergate complex.

## CONTEXT

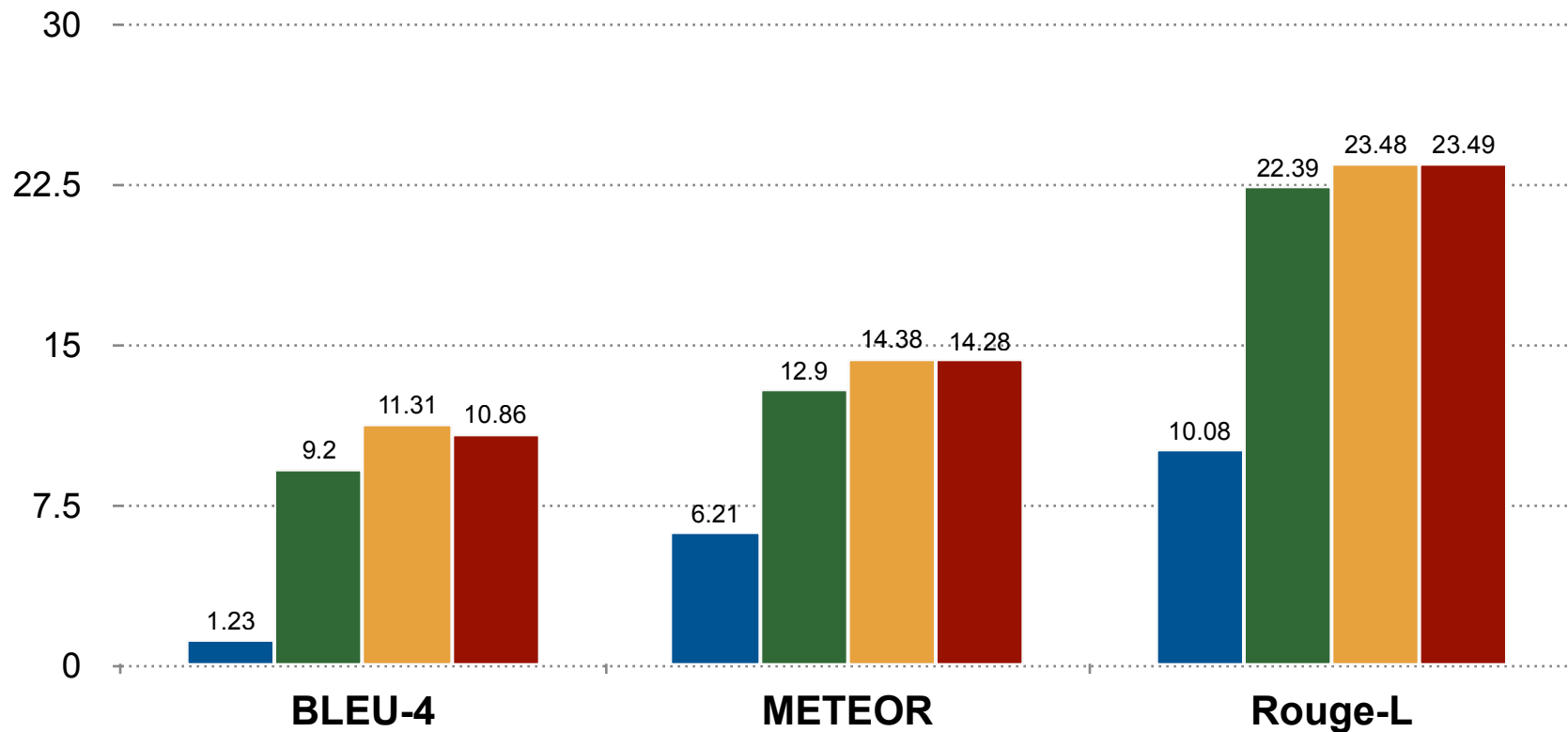
It was honorable for the *other papers to publish in solidarity with the post and times*. That had to prove that the people want to know.



Yes, but it was dangerous for *the white house to ban the post from the whitehouse.*

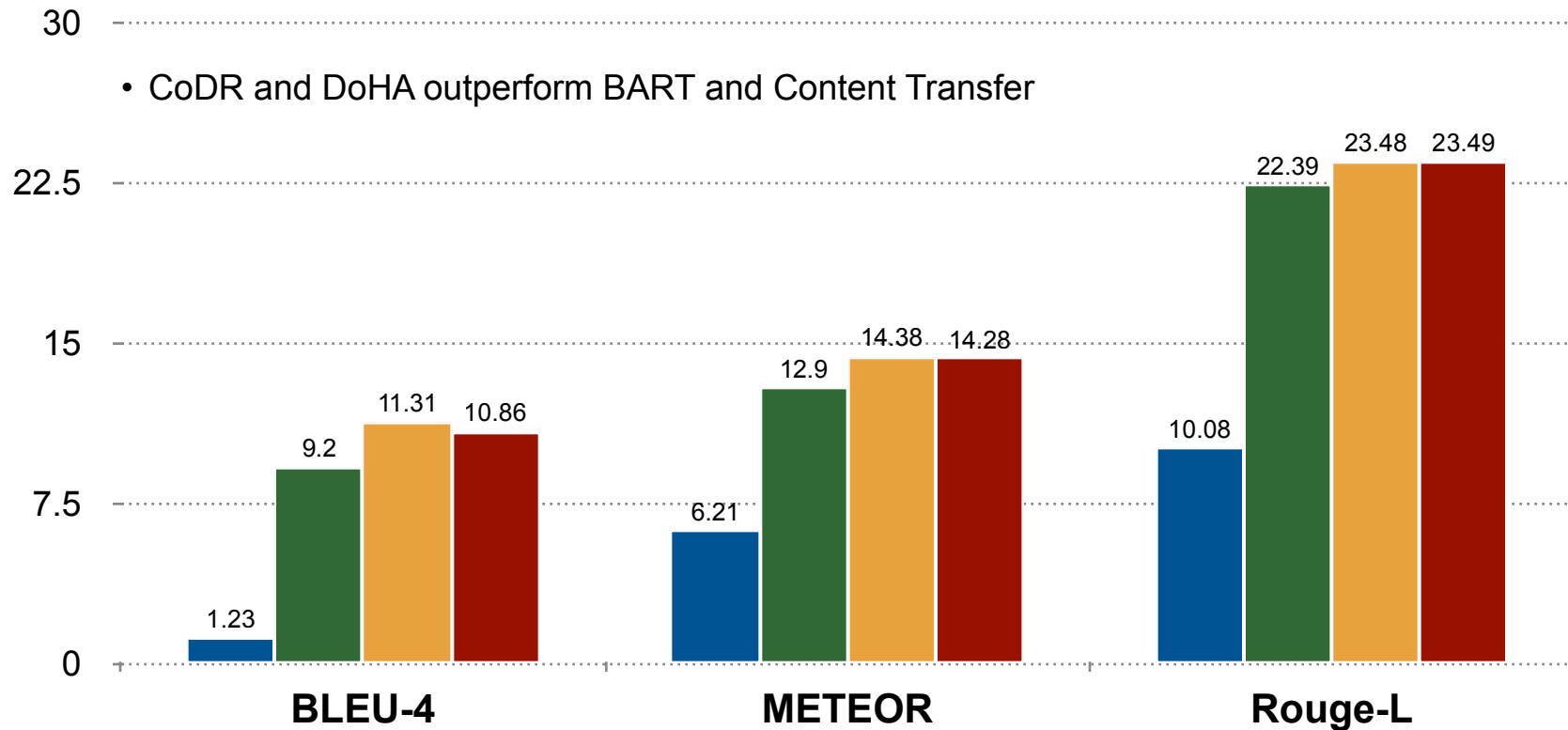
# Wikipedia Update Generation

■ Content Transfer (Prabhumoye et al, NAACL 2019) ■ BART ■ CoDR ■ DoHA



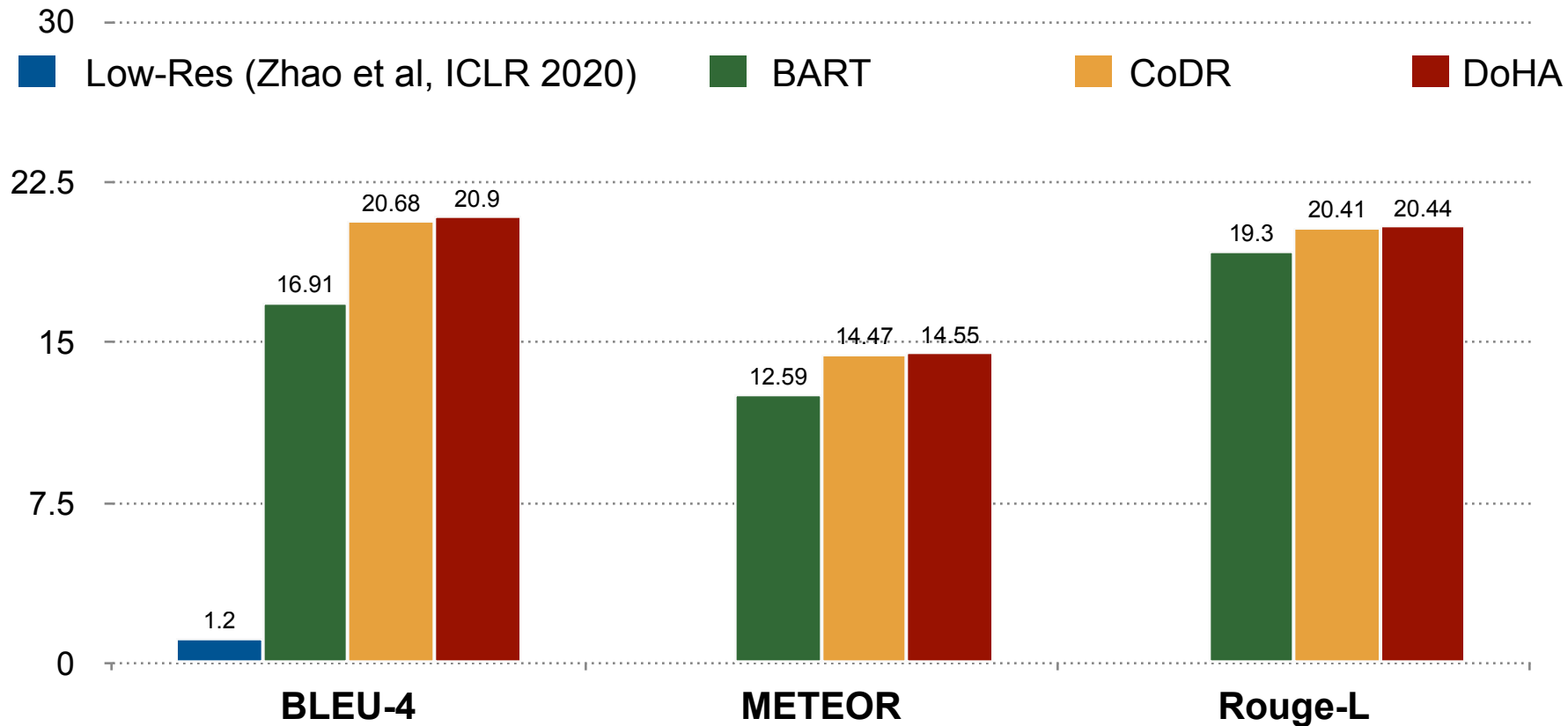
# Wikipedia Update Generation

■ Content Transfer (Prabhumoye et al, NAACL 2019) ■ BART ■ CoDR ■ DoHA

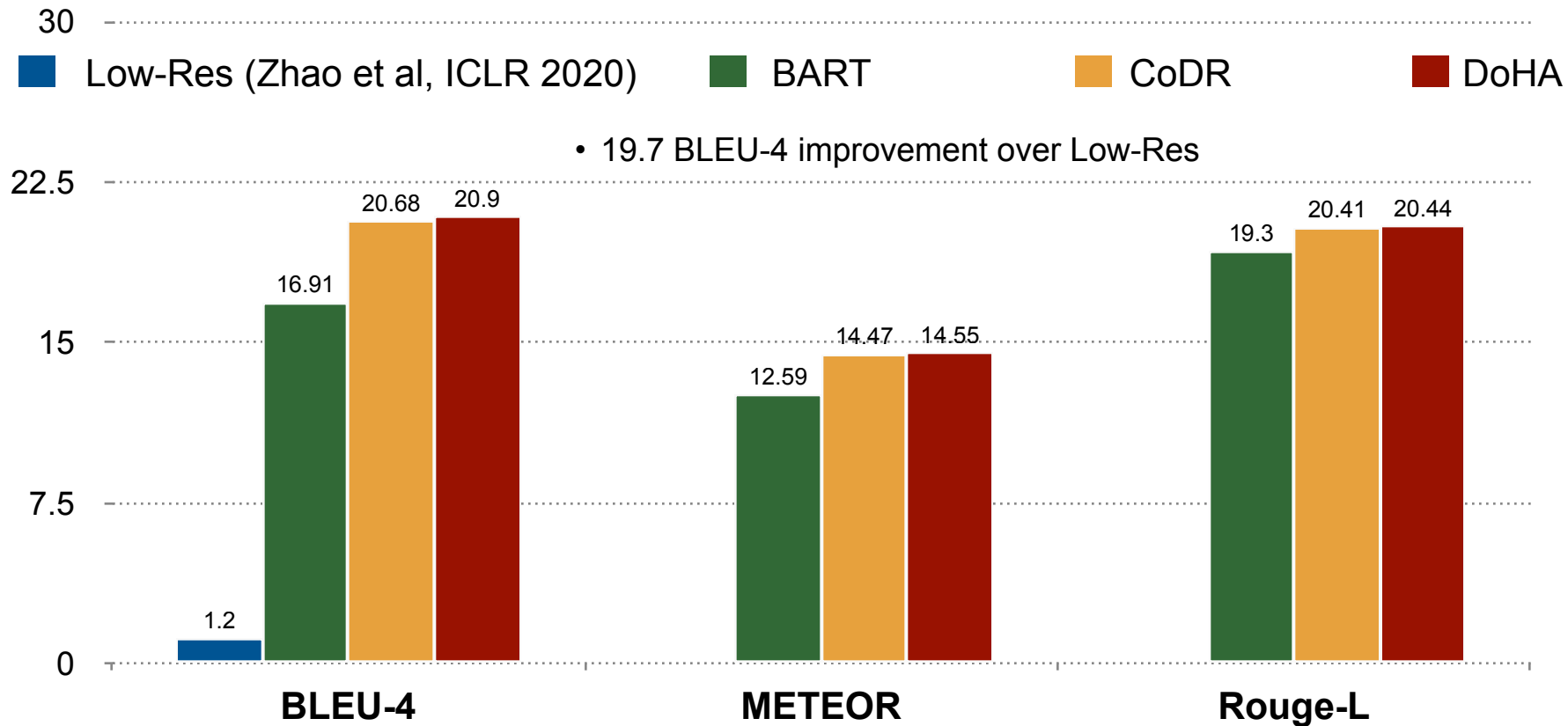




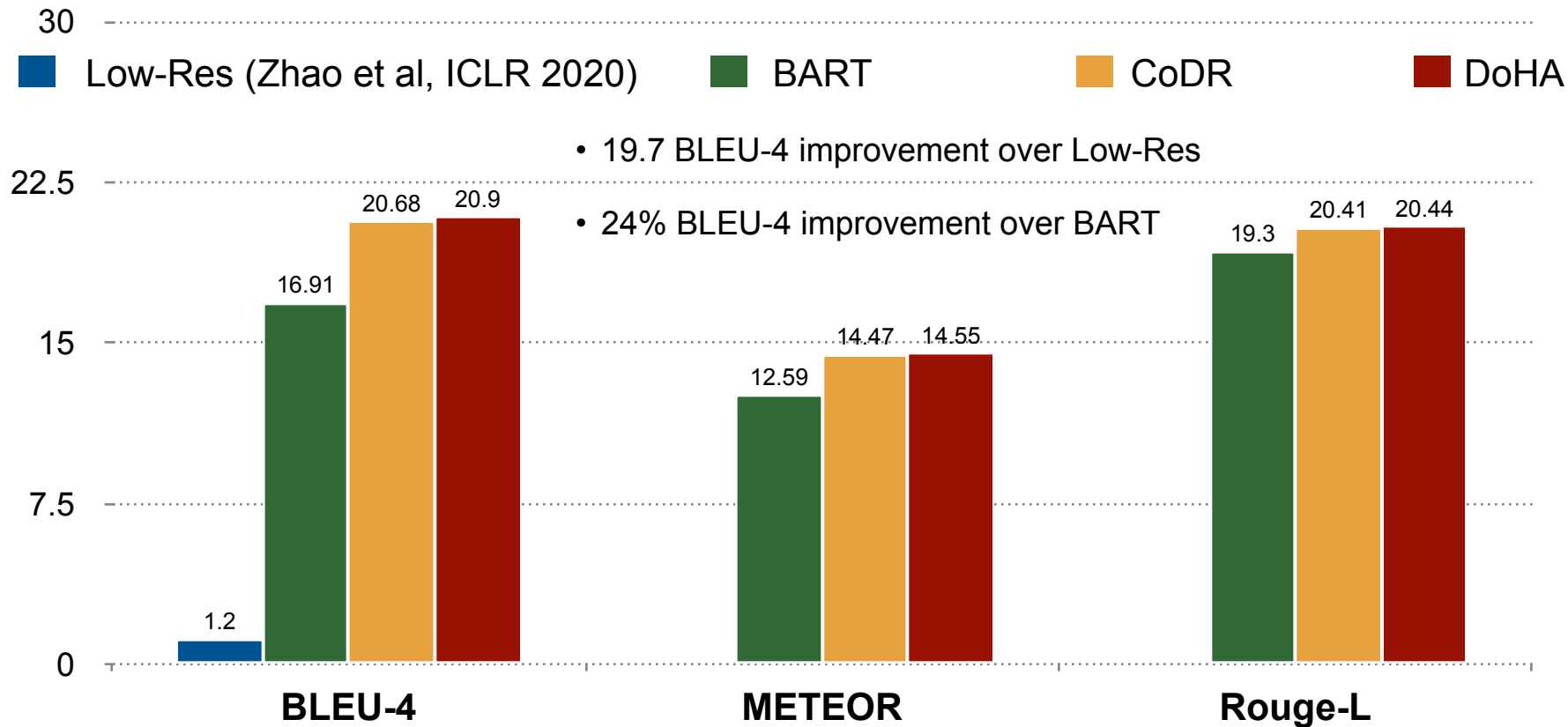
# CMU\_DoG



# CMU\_DoG

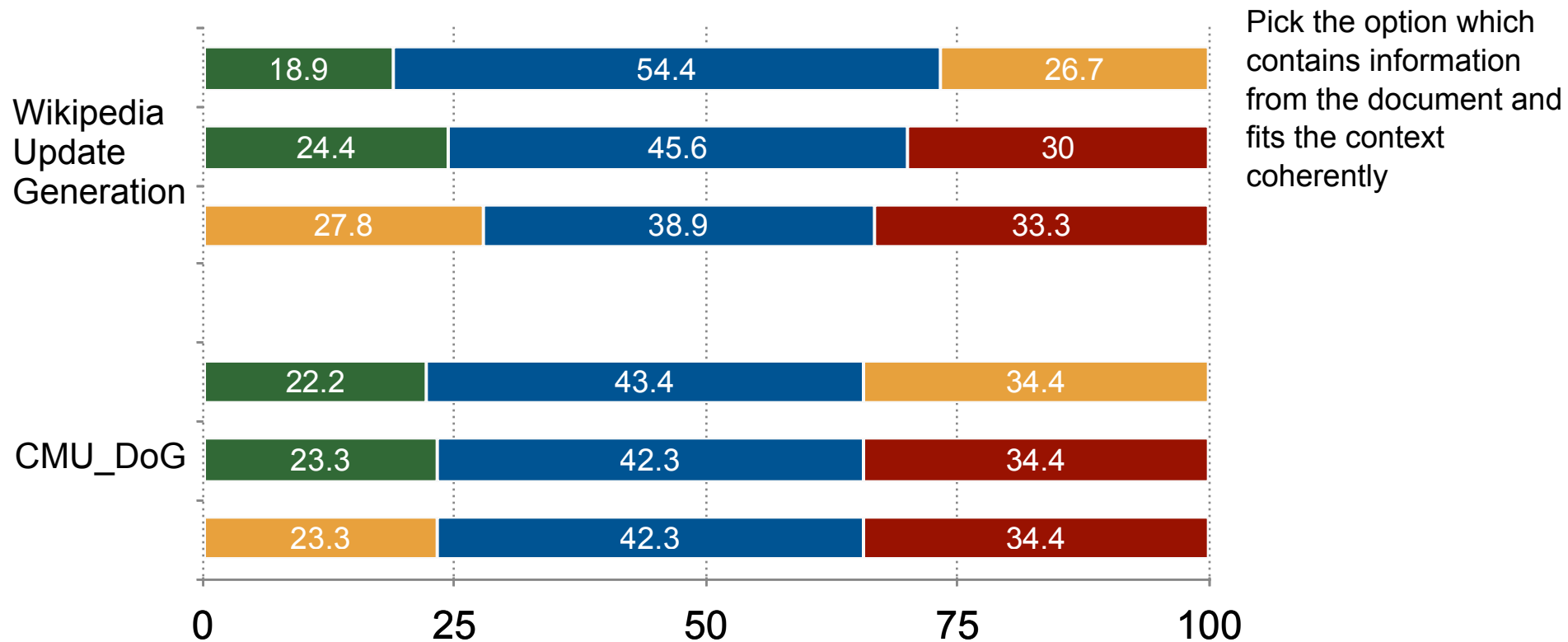


# CMU\_DoG



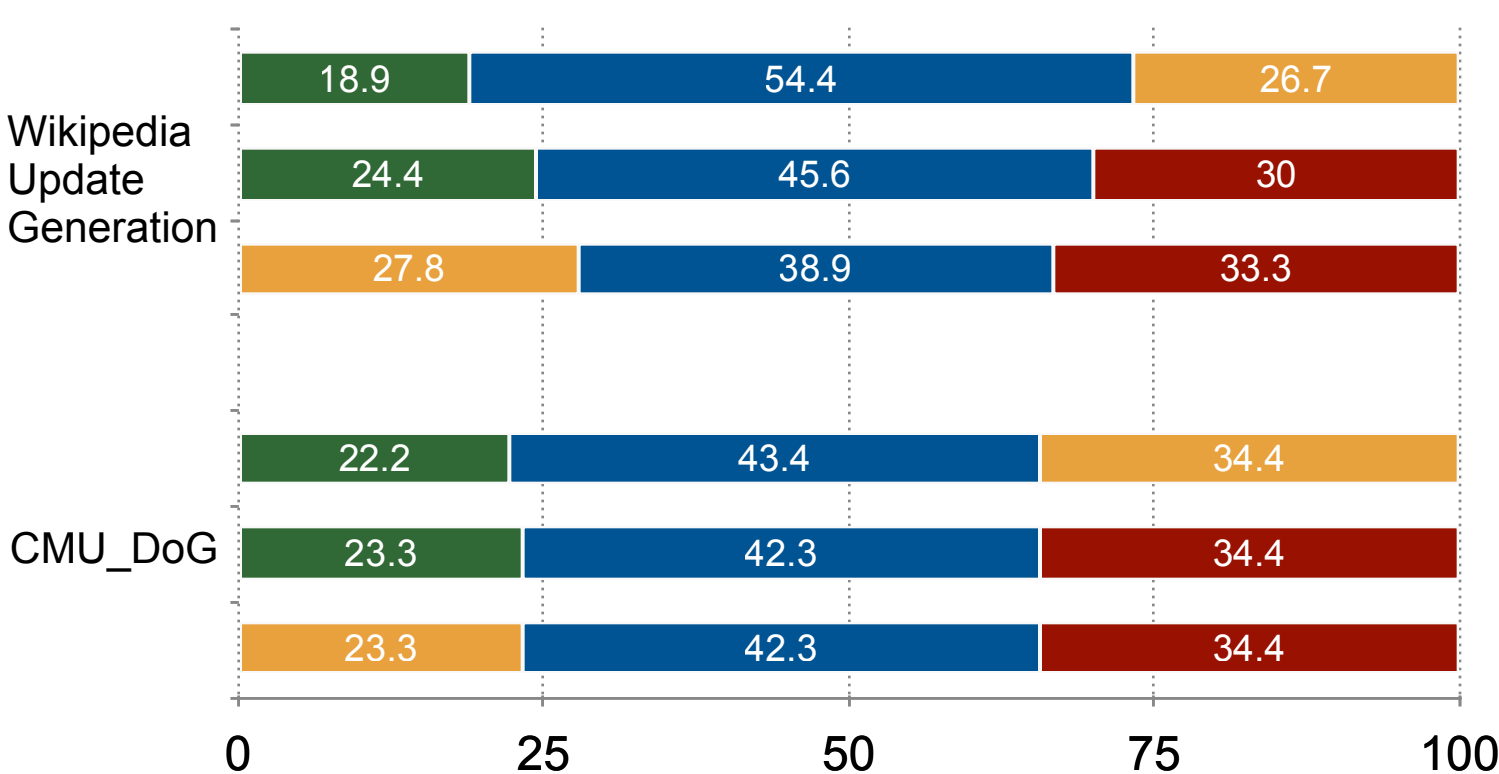
# Relevance Human Evaluation

■ BART      ■ No Pref      ■ CoDR      ■ DoHA



# Relevance Human Evaluation

■ BART      ■ No Pref      ■ CoDR      ■ DoHA

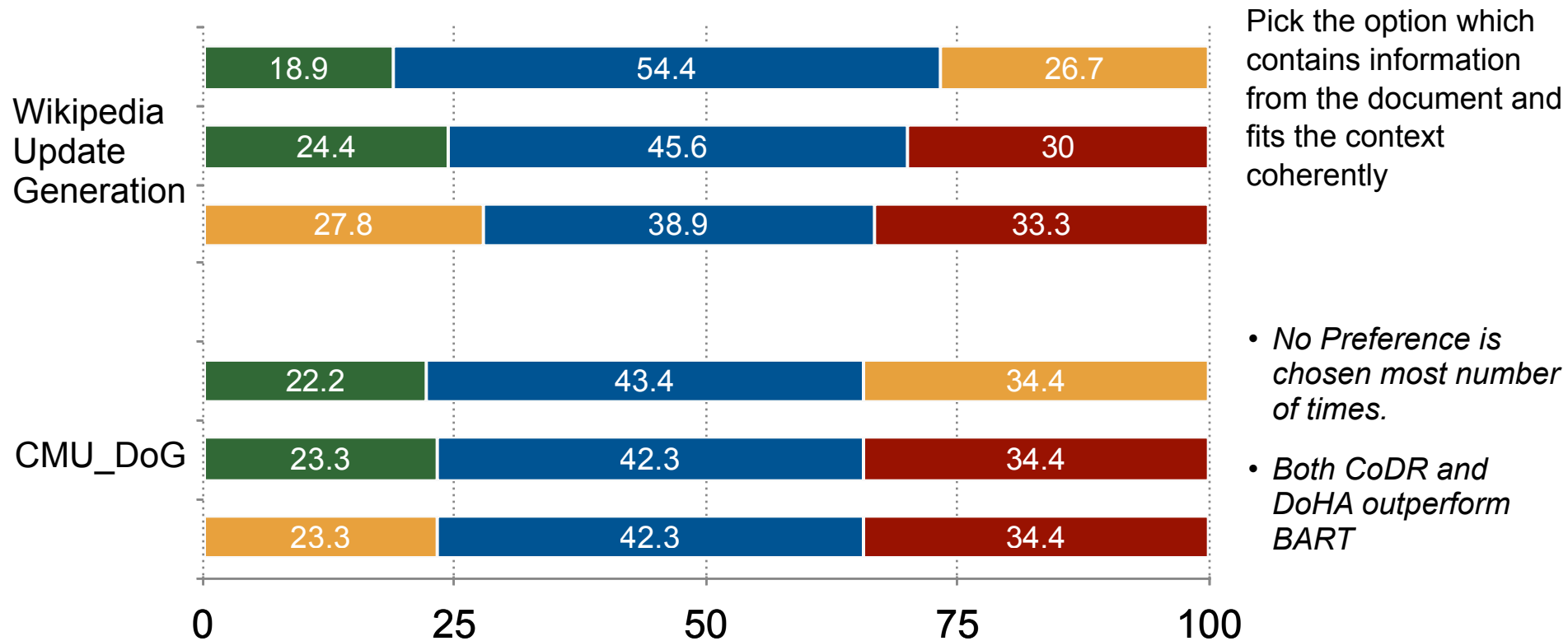


Pick the option which contains information from the document and fits the context coherently

• *No Preference is chosen most number of times.*

# Relevance Human Evaluation

■ BART
 ■ No Pref
 ■ CoDR
 ■ DoHA



# Wikipedia Update Generation

# Wikipedia Update Generation

Error Class	%	Reference	Generation	R
<b>Linguistic Variation:</b> Generation is appropriate but a linguistic variation of the reference.	43	December 12 - The Smiths play Brixton Academy, their last ever gig before their dissolution.	December 12 - The Smiths perform their final show, at Brixton Academy in London.	41

Generation is a paraphrase of the Reference  
Still gets a Rouge-L score of 41!  
Inadequate automated metrics



# Wikipedia Update Generation

Error Class	%	Reference	Generation	R
<b><i>Partial Hallucination:</i></b> Generation is either missing or hallucinates some information	23	America Online and Prodigy (online service) offered access to the World Wide Web system for the first time this year, releasing browsers that made it easily accessible to the general public.	The World Wide Web was first introduced on January 17, 1995 on Prodigy.	17

The generation has captured some information of the reference but missed some.

# Wikipedia Update Generation

Error Class	%	Reference	Generation	R
<b>Incoherent</b> <b>Reference:</b> The reference does not coherently follow the context	22	“The Naked Ape” by Desmond Morris, is published.	Zoologist Desmond Morris publishes “The Naked Ape”.	26

Wikipedia pages that are in the form of lists  
*1340s, Timeline of DC Comics (1950s)*  
 50% times the generation is grounded in the document and close to reference

# Wikipedia Update Generation

Error Class	%	Reference	Generation	R
<b><i>Incorrect:</i></b> The generation is either not appropriate or is not grounded (completely hallucinates the information).	7	The year 2000 is sometimes abbreviated as “Y2K” (the “Y” stands for “year”, and the “K” stands for “kilo-” which means “thousand”)	The Y2K conspiracy theory claimed that a secret nuclear attack by the United States on 2 January 2000 was planned to begin World War 2.	9

About 85% times, the generation is either completely or partially grounded if the reference is grounded

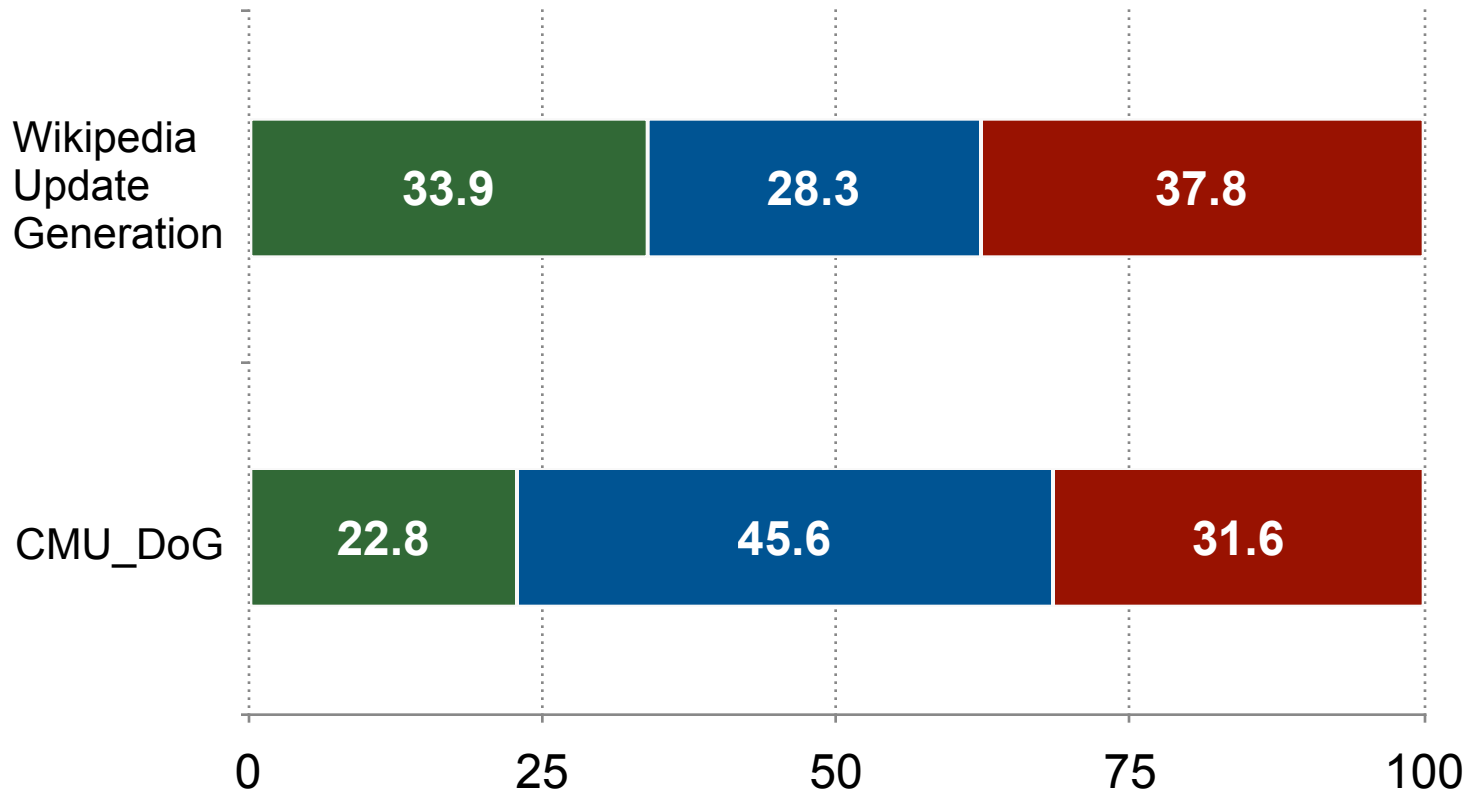
# Wikipedia Update Generation

Error Class	%	Reference	Generation	R
<i>Reference is not grounded</i>	5	This was achieved under dead calm conditions as an additional safety measure, whereas the Wrights flew in a 25 mph+ wind to achieve enough airspeed on their early attempts.	This was verified by a video crew present at the test flight.	14

Dataset Noise!

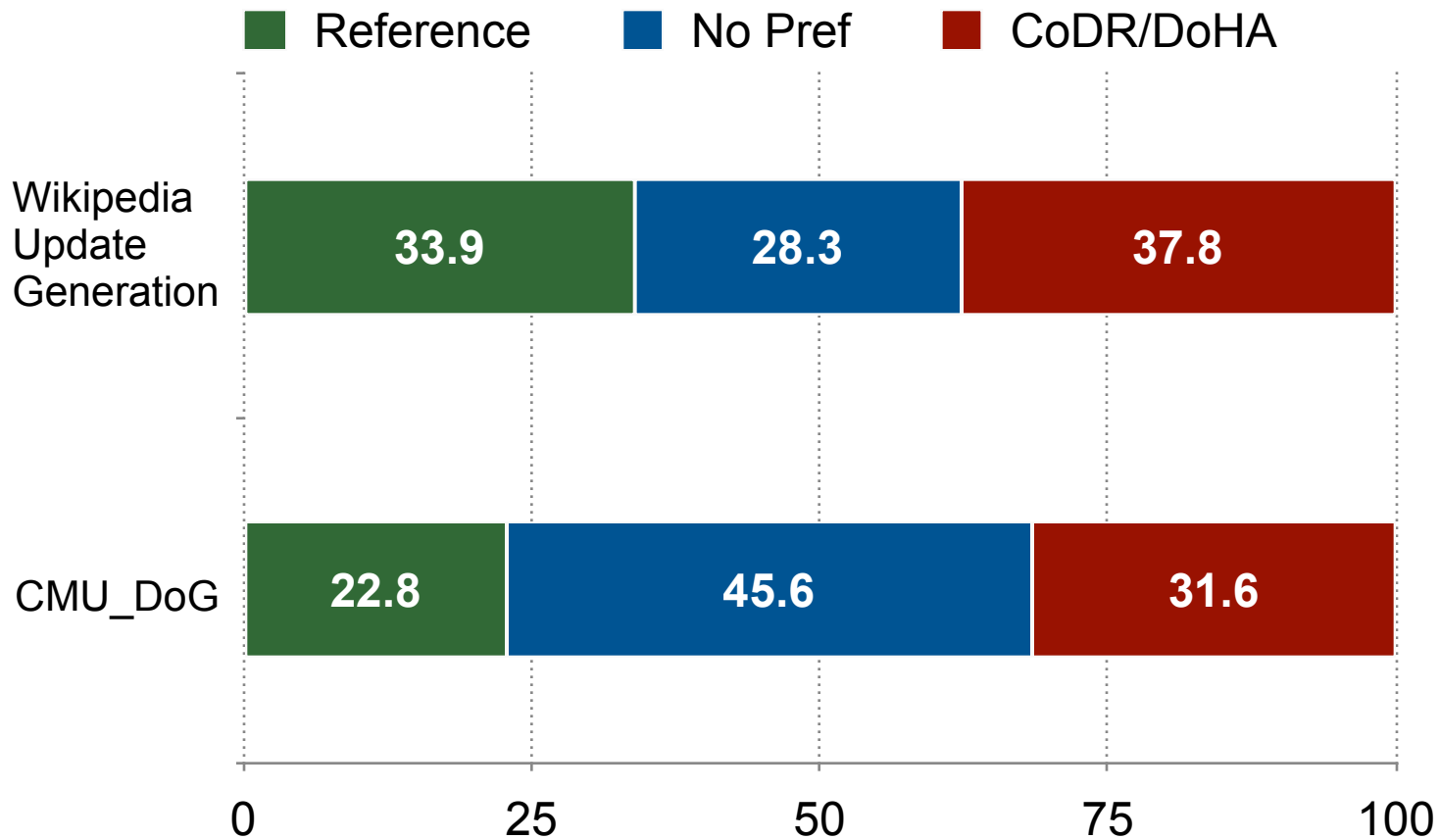
# Reference vs CoDR/DoHA

■ Reference   ■ No Pref   ■ CoDR/DoHA



Pick the option that is most appropriate to the given context.

# Reference vs CoDR/DoHA



Pick the option that is most appropriate to the given context.

1. Automated metrics are inadequate
2. Sole reference should not be considered as the only correct response to the context

# Summary

- New baseline stronger than previous work
- Two new proposed techniques
  - Context Driven Representation
  - Document Headed Attention
- Substantial improvements on automated and human evaluation results on three different datasets
- Comprehensive manual analysis
- Code and trained models are available at <https://github.com/shrimai/Focused-Attention-Improves-Document-Grounded-Generation>