

# Can machines assist humans in discovering new knowledge?

Fabio Petroni

## Bio

- Researcher at Thomson Reuters until 2018
- Researcher in FAIR until September 2022
- Co-Founder of



# Samaya

<https://samaya.ai>

**Knowledge-intensive NLP:** tasks that requires – even for humans – access to a large body of information.

# Collab orators



Ledell Wu



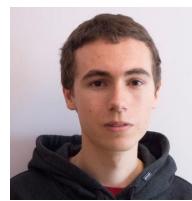
Kashyap Popat



Naman Goyal



Nicola Cancedda



Mikel Artetxe



Luke Zettlemoyer



Mikhail Plekhanov



Michele Bevilacqua



Nicola De Cao



Sebastian Riedel



Yacine Jernite



Vladimir Karpukhin



Jean Maillard



Vassilis Plachouras



Tim Rocktäschel



Scott Wen-tau Yih



Angela Fan



Samuel Broscheit



Edouard Grave



Lucas Hosseini



Gautier Izacard



Patrick Lewis



Pierre-Emmanuel  
Mazare



Maria Lomeli



Majid Yazdani



Giuseppe  
Ottaviano



James Thorne



Ola Piktus



Armand Joulin



Timo Schick



Jane Yu

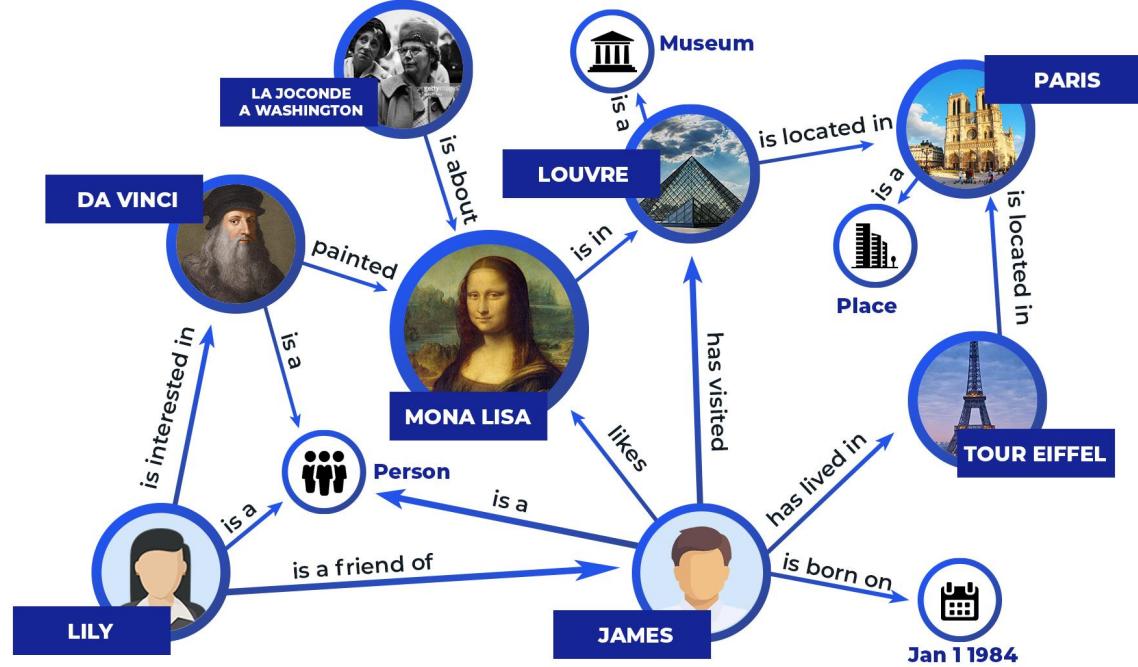
For decades, AI researchers have searched for a representation of knowledge that is most useful for machines

Machine Reading



Auto-Text to Knowledge

DARPA Machine  
Reading Program



# Limitations of knowledge graphs

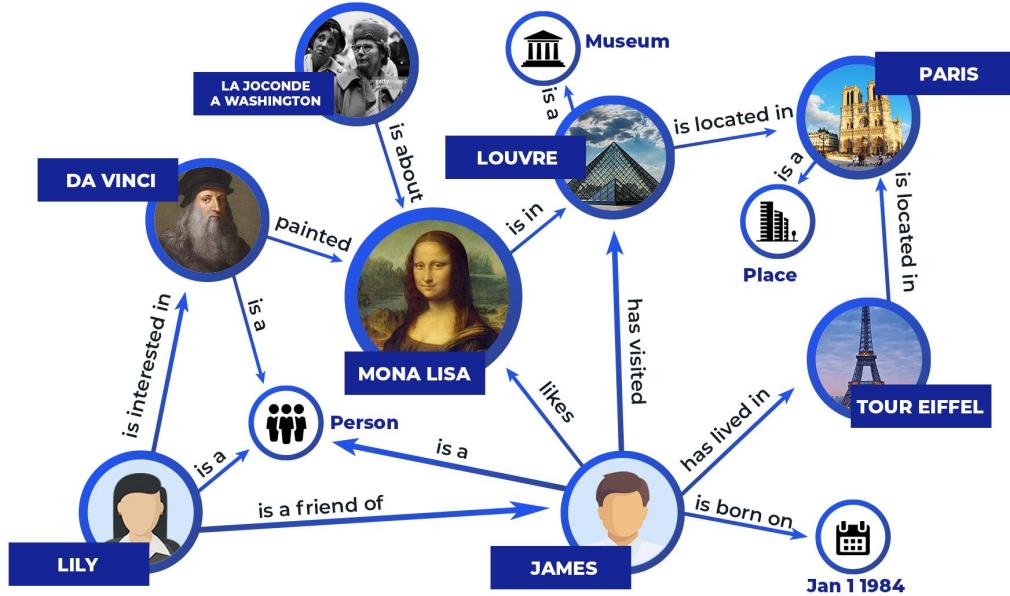
Human supervision

Schema engineering

Predefined class o relations

Difficult to extend to more data

Incomplete



# Language Models as Knowledge Bases?

Fabio Petroni<sup>1</sup> Tim Rocktäschel<sup>1,2</sup> Patrick Lewis<sup>1,2</sup> Anton Bakhtin<sup>1</sup>  
Yuxiang Wu<sup>1,2</sup> Alexander H. Miller<sup>1</sup> Sebastian Riedel<sup>1,2</sup>

<sup>1</sup>Facebook AI Research

<sup>2</sup>University College London

{fabio petroni, rockt, plewis, yolo, yuxiangwu, ahm, sriedel}@fb.com

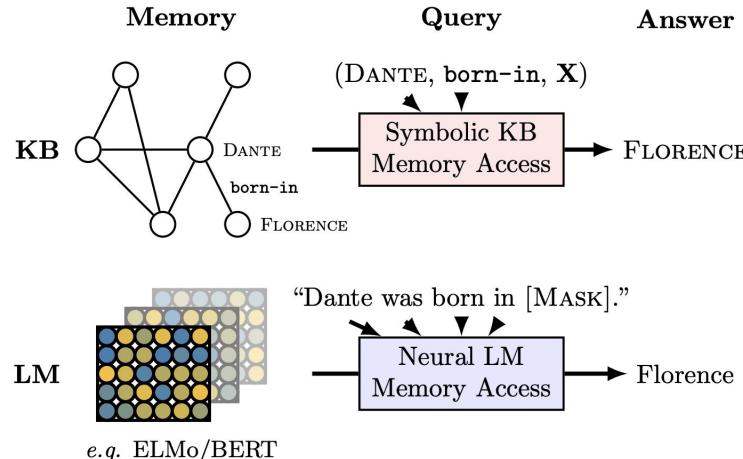


Figure 1: Querying knowledge bases (KB) and language models (LM) for factual knowledge.

# Texts as Knowledge Bases



Christopher Manning

Joint work with Gabor Angeli and Danqi Chen

Stanford NLP Group

[@chrmanning](https://twitter.com/chrmanning) · [@stanfordnlp](https://twitter.com/stanfordnlp)

AKBC 2016

# The KILT benchmark



[kiltbenchmark.com](https://kiltbenchmark.com)

5 task families

11 datasets

1 knowledge source

3.5M datapoints

Fabio Petroni, Aleksandra Piktus, Angela Fan, Patrick Lewis, Majid Yazdani, Nicola De Cao, James Thorne, Yacine Jernite, Vladimir Karpukhin, Jean Maillard, Vassilis Plachouras, Tim Rocktäschel, Sebastian Riedel: **Kilt: a benchmark for knowledge intensive language tasks.** NAACL-HLT 2021

## Slot Filling

INPUT:  
Star Trek [SEP] creator

OUTPUT:  
Gene Roddenberry

PROVENANCE:  
17157886-1

## Open Domain QA

INPUT:  
When did Star Trek go off the air

OUTPUT:  
June 3, 1969

PROVENANCE:  
17157886-5

INPUT:  
Which Star Trek star directed Three Men and a Baby?

OUTPUT:  
Leonard Nimoy

PROVENANCE:  
17157886-4, 596639-7

INPUT:  
Treklanta (formerly "TrekTrax Atlanta") is an annual convention for what American science fiction media franchise?

OUTPUT:  
Star Trek

PROVENANCE:  
17157886-1, 28789994-6

zsRE

NQ

TQA

HoPo



Knowledge source:  
5.9 Million Wikipedia pages

## Star Trek 17157886

Star Trek is an American media franchise based on the science fiction television series created by Gene Roddenberry.<sup>1</sup> [...] It followed the interstellar adventures of Captain James T. Kirk (William Shatner) and his crew aboard the starship USS "Enterprise", a space exploration vessel built by the United Federation of Planets in the 23rd century.<sup>2</sup> The "Star Trek" canon includes "The Original Series", an animated series, five spin-off television series, the film franchise, and further adaptations in several media.<sup>3</sup> [...] The original 1966–69 series featured William Shatner as Captain James T. Kirk, Leonard Nimoy<sup>4</sup> as Spock, DeForest Kelley as Dr. Leonard "Bones" McCoy, James Doohan as Montgomery "Scotty" Scott, Nichelle Nichols as Uhura, George Takei as Hikaru Sulu, and Walter Koenig as Pavel Chekov. During the series' first run, it earned several nominations for the Hugo Award for Best Dramatic Presentation, and won twice. [...] NBC canceled the show after three seasons; the last original episode aired on June 3, 1969.<sup>5</sup> [...]

## Three Men and a Baby 596639

Three Men and a Baby is a 1987 American comedy film directed by Leonard Nimoy<sup>7</sup> and starring Tom Selleck, Steve Guttenberg, Ted Danson and Nancy Travis. [...]

## Treklanta 28789994

Treklanta is an annual "Star Trek" convention based in Atlanta, Georgia that places special emphasis on fan-based events, activities, programming and productions.<sup>6</sup> [...]

## Dialogue

INPUT:  
I am a big fan of Star Trek, the American franchise created by Gene Roddenberry. I don't know much about it. When did the first episode air?

It debuted in 1966 and aired for 3 seasons on NBC.

What is the plot of the show?

OUTPUT:  
William Shatner plays the role of Captain Kirk. He did a great job.

PROVENANCE:  
17157886-2

WoW

## Fact Checking

INPUT:  
Star Trek had spin-off television series.

OUTPUT:  
Supports

PROVENANCE:  
17157886-3

FEV

## Entity Linking

INPUT:  
[...] Currently the site offers five movie collections ranging from \$149 for 10 [START\_ENT]Star Trek [END\_ENT] films to \$1,125 for the eclectic Movie Lovers' Collection of 75 movies. [...]

OUTPUT:  
Star Trek

PROVENANCE:  
17157886

CnWn

# Leaderboard

The KILT leaderboard.

Phase: Slot Filling - Zero Shot RE, Split: test

Order by metric

B	- Baseline	*	- Private	V	- Verified	Visible Metrics		
Rank	Participant team	R-Prec	Recall@5 (↑)	Accuracy (↑)	F1 (↑)	KILT-AC	KILT-F1	Last submission at
▼	▼	(↑) ▼	▼	▼	▼	(↑) ▼	(↑) ▼	▼
1	FiD-Light	96.13	96.68	85.30	88.85	83.99	87.21	1 year ago
2	gripRank	99.28	99.66	74.65	80.27	74.29	79.89	4 months ago
3	SEAL (single ngram)	97.99	99.34	74.63	79.66	73.20	78.12	2 years ago
4	IBM Research AI (1) (KGL_1)	98.49	99.23	72.55	77.05	72.31	76.69	2 years ago
5	anon2054 (MetaRAG)	95.81	96.64	71.61	76.60	71.10	75.86	1 year ago
6	IBM Research AI (0) (KGL_0 (reupload))	94.18	95.19	68.97	74.47	68.32	73.45	2 years ago
7	KILT-WEB 2 (Wikipedia)	89.63	97.87	73.96	78.43	67.20	70.99	2 years ago

# Knowledge Discovery

- Making sense of world knowledge remain complex and demanding tasks.
- A researcher working on a new hypothesis still needs to read through literature, write notes, mentally connect the dots
- A knowledge worker needs years of experience, research and reading to spot a novel correlation
- Abundant data often includes conflicting points, necessitating careful analysis and verification.
- **Can machines to assist humans in navigating, interpreting, and verifying world knowledge?**

# YES, they can!

## nature machine intelligence

[Explore content](#) ▾   [About the journal](#) ▾   [Publish with us](#) ▾

---

[nature](#) > [nature machine intelligence](#) > [articles](#) > [article](#)

Article | [Open access](#) | [Published: 19 October 2023](#)

### Improving Wikipedia verifiability with AI

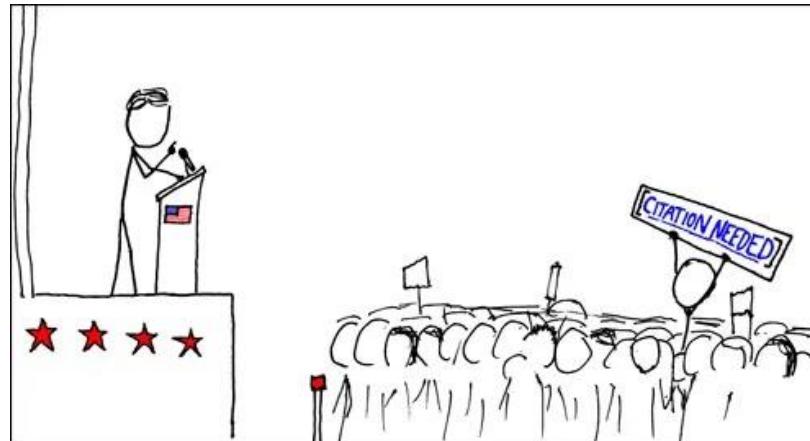
[Fabio Petroni](#)✉, [Samuel Broscheit](#), [Aleksandra Piktus](#), [Patrick Lewis](#), [Gautier Izacard](#), [Lucas Hosseini](#),  
[Jane Dwivedi-Yu](#), [Maria Lomeli](#), [Timo Schick](#), [Michele Bevilacqua](#), [Pierre-Emmanuel Mazaré](#), [Armand Joulin](#), [Edouard Grave](#) & [Sebastian Riedel](#)

[Nature Machine Intelligence](#) 5, 1142–1148 (2023) | [Cite this article](#)

[Metrics](#)

# Wikipedia Verifiability

- Verifiability is a core content policy of Wikipedia!
- Claims that are likely to be challenged need to be backed by citations.
- Finding relevant sources is a difficult task.
- Many Wikipedia claims do not have any references that support them.
- Even existing citations might not support a given claim or become obsolete.



# Ensuring Verifiability

The task of ensuring verifiability of Wikipedia is titanic!

Expert editors currently rely on the work of volunteers:

- (1) identify and tag claims likely to fail verification by performing fact and reference checks between the wikipedia claim and the cited source;
- (2) provide assistance by suggesting replacement for a failed verification citation with a reliable source that corroborates the claim

At the time of writing, over 500,000 statements on Wikipedia are marked as “Citation needed”.

# Verify Wikipedia

## Mirella Lapata

---

From Wikipedia, the free encyclopedia

**Mirella Lapata FRSE** is a computer scientist and Professor in the [School of Informatics](#) at the [University of Edinburgh](#).<sup>[3]</sup> Working on the general problem of extracting semantic information from large bodies of text, Lapata develops computer algorithms and models in the field of [natural language processing](#) (NLP).<sup>[1]</sup>

### Awards and honors

- In 2009 Lapata became the first recipient of the [Microsoft British Computer Society](#) (BCS)/BCS IRSG Karen Spärck Jones Award for information retrieval and natural language processing; the award commemorates the life and work of [Karen Spärck Jones](#).
- In 2012 Lapata won an [Empirical Methods in Natural Language Processing](#) (EMNLP)-CoNLL 2012 Best Reviewer Award.<sup>[11]</sup>
- In 2016 Lapata, with Eneko Agirre and Sebastian Riedel, won the EMNLP Best Data Set Paper Award.<sup>[12]</sup>
- In 2018 Lapata was awarded, alongside Li Dong, an [Association for Computational Linguistics](#) (ACL) Best Paper Honorable Mention.
- In 2019 Lapata was elected a Fellow of the [Royal Society of Edinburgh](#)<sup>[14]</sup>
- In 2020 Lapata was elected to the [Academia Europaea](#).<sup>[15]</sup>

# Verify Wikipedia



emnlp<sub>2016</sub>

SIGDAT, the Association for Computational Linguistics special interest group on linguistic data and corpus-based approaches to NLP, invites you to participate in EMNLP 2016.

The conference will be held on **November 1–5, 2016** (Tue–Sat) in Austin, Texas, USA.

## Best Paper Committee

**Best Paper & Honorable Mention:** Stephen Clark, Hal Daumé III, Chris Dyer, and Julia Hockenmaier

**Best Short Paper:** Stefan Riezler, Anoop Sarkar, and Noah Smith

**Best Data Set Paper:** Eneko Agirre, Mirella Lapata and Sebastian Riedel

**Chair:** Xavier Carreras and Kevin Duh

# We can build machines to help knowledge-workers

In the paper, we show that a machine can assist Wikipedia editors with both tasks:

- **discovering problematic citations** and
- **improving their verifiability.**



# SIDE

Claim to verify



## Blackfoot Confederacy

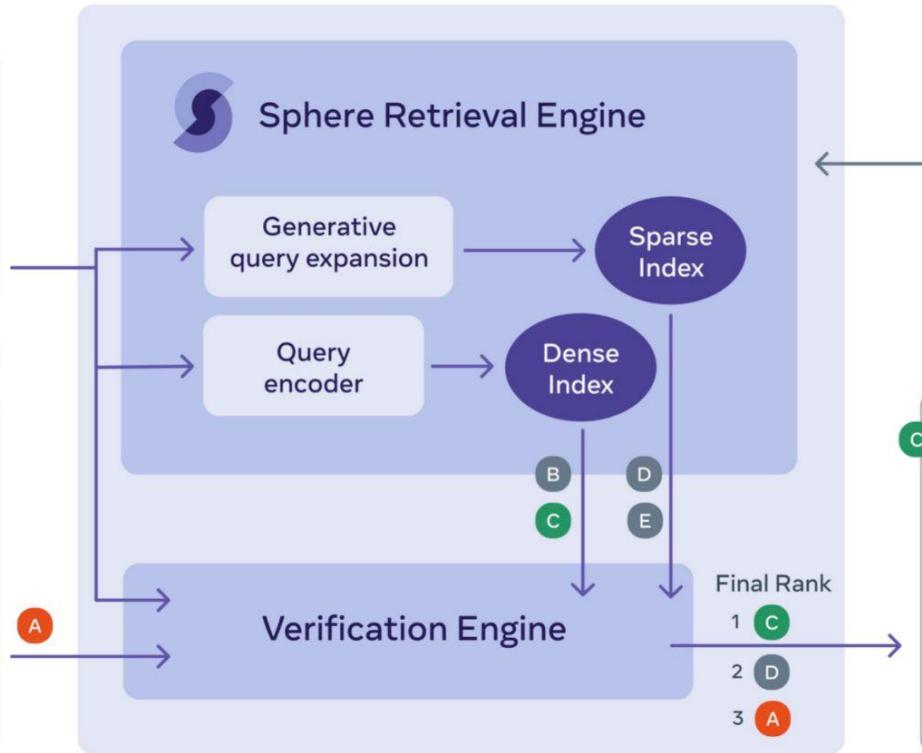
Joe Hipp, Heavyweight boxer, the first Native American to compete for the WBA World Heavyweight Title.[73]

73. "Blackfoot Culture and History".

Existing citation Failed verification

## A Blackfoot Culture and History

Ancient history is interesting and important, but the Blackfoot Indians are still here today, too, and we try to feature modern writers as well as traditional folklore, contemporary art as well as museum pieces, and the life and struggles of today as well as the tragedies of yesterday.



## Sphere corpus

CommonCrawl snapshot  
134M web articles  
906.3M passages

Suggested citation

## C Warrior in the ring: Boxing with Marvin Camel

In 1989 at the twilight of his career, Camel fought Joe Hipp of the Blackfeet Nation. Hipp, who became the first Native American to challenge for the world heavyweight championship, said the fight was one of the weirdest of his career.

# Side demo

# verifier.sideeditor.com

Side ANNOTATION INTERFACE

Detailed Instructions

Select the source on the right that you think best matches with the citation:

**Hillary Clinton**

**Clinton Foundation, "Hard Choices," and speeches**

2015 report said that while "There has never been a better time in history to be born a woman ... this data shows just how far we still have to go." The foundation began accepting new donations from foreign governments, which it had stopped doing while she was secretary. However, even though the Clinton Foundation had stopped taking donations from foreign governments, they continued to take large donations from foreign citizens who were sometimes linked to their governments.

She began work on another volume of memoirs and made appearances on the paid speaking circuit. There she received \$200,000–\$225,000 per engagement, often appearing before Wall Street firms or at business conventions. She also made some unpaid speeches on behalf of the foundation. For the fifteen months ending in March 2015, Clinton earned over \$11 million from her speeches. For the overall period 2007–14, the Clintons earned almost \$141 million, paid some \$56 million in federal and state taxes and donated about \$15 million to charity. [, she was estimated to be worth over \\$30 million on her own, or \\$45–53 million with her husband.](#) [citation needed]

Clinton resigned from the foundation's board in April 2015, when she began her presidential campaign and the foundation said it would accept new foreign governmental donations from six Western nations only.

2016 presidential campaign

On April 12, 2015, Clinton formally announced her candidacy for the presidency in the 2016 election. She had a campaign-in-waiting already in place, including a large donor network, experienced operatives and the Ready for Hillary and Priorities USA Action political action committees and other

**Sources**

**How Hillary Clinton went from 'dead broke' to millionaire status | Fox Business**  
<https://www.foxbusiness.com/features/how-hillary-clinton-went-from-dead-broke-to-millionaire-status>

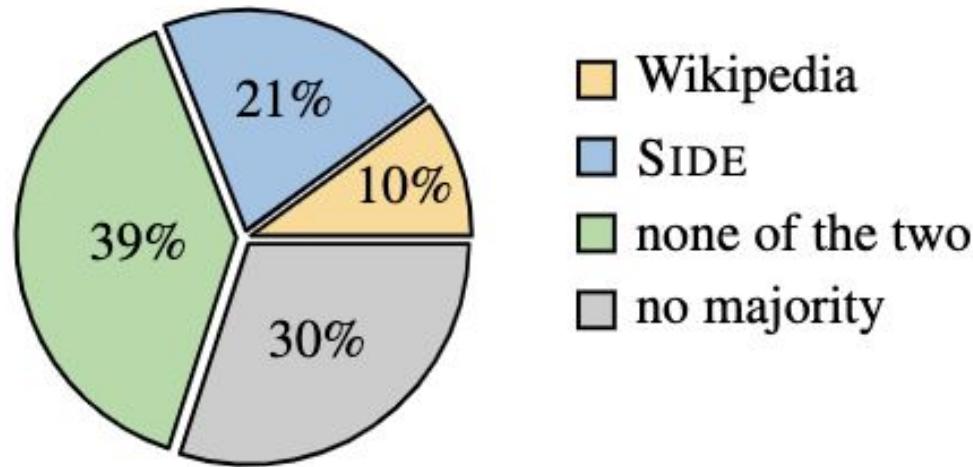
How Hillary Clinton went from 'dead broke' to millionaire status By Mitch Strohm Published August 29, 2016 FeaturesBankrate.com Net worth: \$32 million (as of August 2016) Education: Wellesley College, Yale University George Lucas' net worth How To Invest Extra Savings Money Market Accounts 101 How Hillary Clinton's net worth was built Hillary Clinton's net worth is in the range of \$11 million to \$53 million, giving her an average net worth of \$32 million with no apparent liabilities, according to federal records. Hillary says she and Bill Clinton were "dead broke" after they left the White House in 2001. But...

[SHOW MORE ↑](#)

**The Richest And Poorest Presidential Candidates: From Hillary's Millions To Marco Rubio's Debts**  
<https://www.forbes.com/sites/afontevecchia/2015/09/29/the-richest-and-poorest-presidential-candidates-from-hillarys-millions-to-marco-rubios-debts/>

...show and appearances on Fox News, radio gigs, books and speeches. He's now worth \$9 million. And then there's Hillary Clinton. As my colleague Dan Alexander explains in his well-researched piece, Bill and

## Results (Wikipedia users)

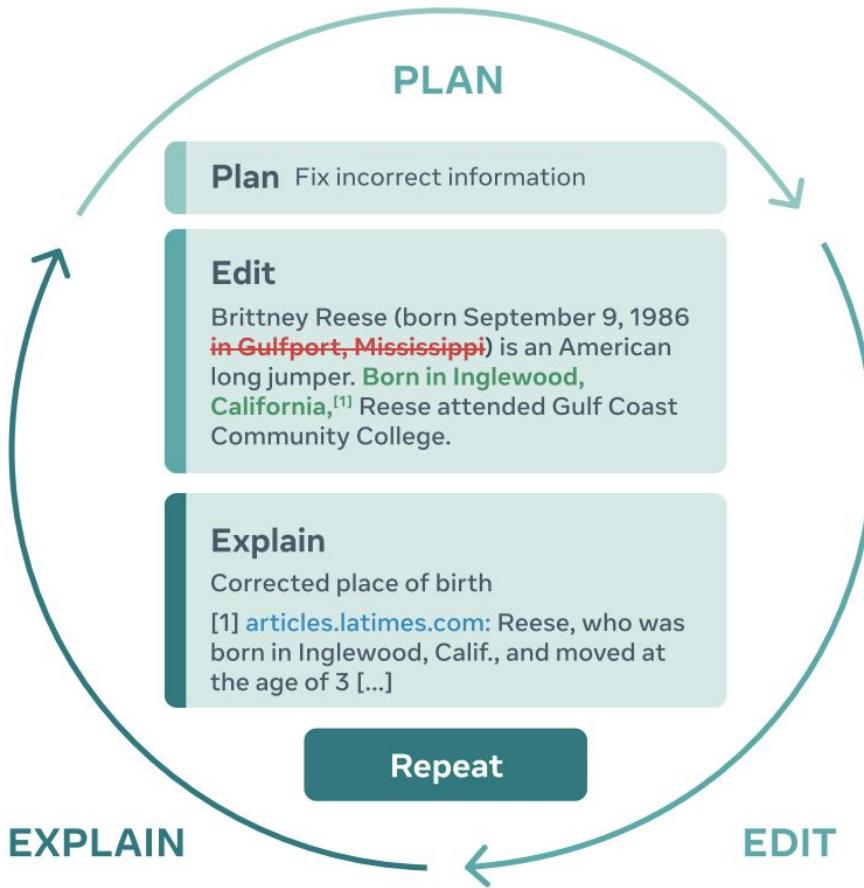


**Figure 4.** Wikipedia users annotations via our demo.



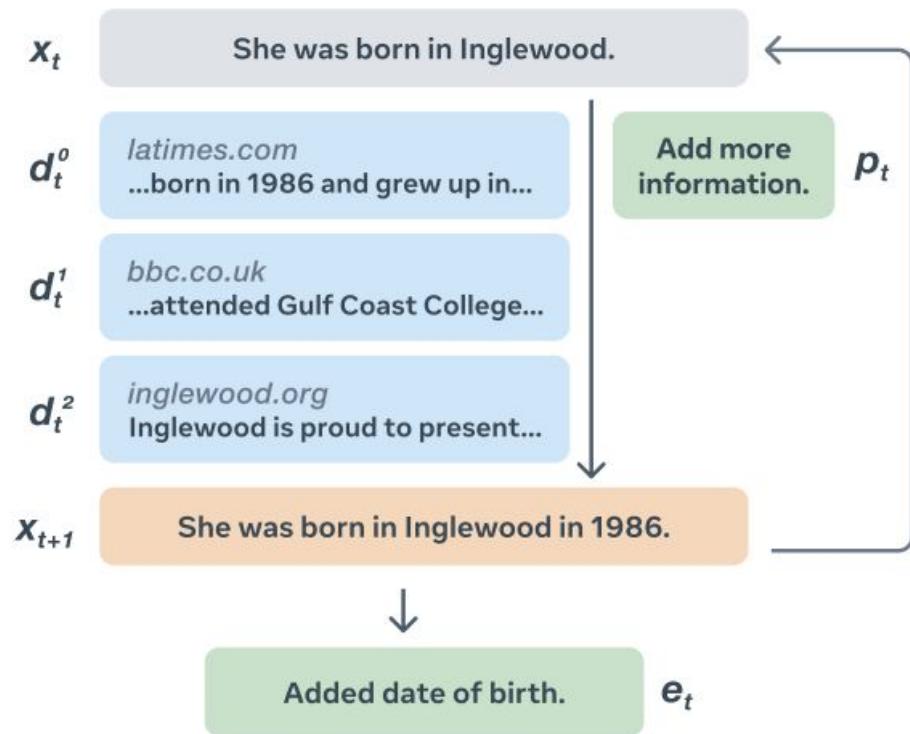
# Propose Wikipedia edits

Timo Schick, Jane Dwivedi-Yu,  
Zhengbao Jiang, Fabio Petroni,  
Patrick Lewis, Gautier Izacard, Qingfei  
You, Christoforos Nalmpantis,  
Edouard Grave, Sebastian Riedel -  
**PEER: A Collaborative Language  
Model**



# Peer process

- Trained on Wikipedia edit history (7M datapoints)
- Retrieved supported documents for each edit using the Verify Wikipedia pipeline
- The plain is given by the comment associated with the edits in Wikipedia



# Edit Eval

*The benchmark for text improvements*

Barack Obama and Hillary Clinton were both ~~running for the Democratic party nomination for~~ candidate nominees of the Democratic party in the 2008 Presidential election.

Rephrase this text

Make the text neutral

On June 3rd, Obama received enough endorsements to ~~rightly~~ clinch the nomination. Obama went on to win the general election against Republican John McCain and became the ~~current~~ 44th President of the United States of America.

Update the article

Simplify the text

Clinton went on to serve as the 67th United States Secretary of State, ~~serving~~ within the Obama administration. In the 2016 Presidential election, she

Fix grammar errors

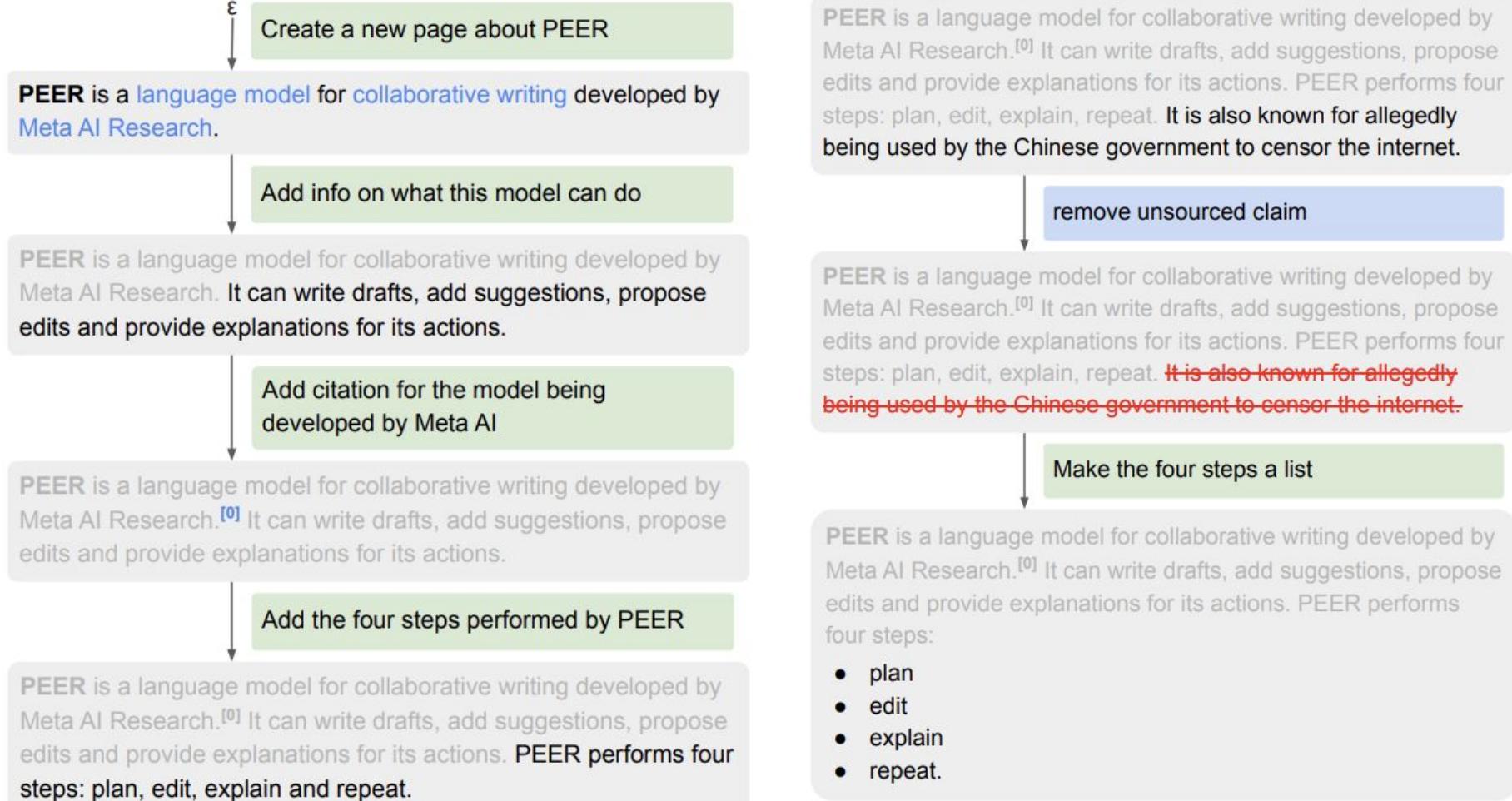
became the nominee of the Democratic party, becoming the first woman to be a nominee of a major U.S. political party. Clinton, however, ~~lose~~ lost the general election to Donald Trump. Clinton's third memoir, *What Happened*, features an account of her loss in the 2016 election.

Add missing information

# Results

Model	Fluency		Clarity ITR-L	Coherence ITR-O	Para. STS	Simplification		Neutral. WNC	Updating	
	JFL	ITR-F				TRK	AST		FRU	WFI
Copy	26.7 / 40.5	32.3 / 86.0	29.5 / 62.9	31.3 / 77.2	21.1	26.3	20.7	31.9 / 0.0	29.8 / 0.0	33.6 / -
Tk	31.8 / 39.0	32.4 / 61.6	<b>38.4 / 58.4</b>	<b>33.8 / 70.4</b>	30.2	32.8	29.9	31.3 / 0.4	12.6 / 3.6	1.3 / 4.5
T0	42.0 / 38.8	24.6 / 34.9	32.6 / 30.2	22.2 / 21.6	34.3	34.4	32.3	22.3 / 0.0	14.2 / 9.6	5.1 / 16.3
T0++	34.7 / 43.2	35.3 / 75.8	37.6 / 56.5	32.7 / 59.9	28.4	32.9	28.2	29.3 / 0.3	12.6 / 3.7	4.4 / 8.1
PEER-3	55.5 / 54.3	51.4 / 84.3	32.1 / 47.1	32.1 / 59.8	28.6	32.5	30.5	53.3 / 21.6	39.1 / 30.9	34.4 / 18.7
PEER-11	55.8 / 54.3	<b>52.1 / 85.2</b>	32.5 / 51.3	32.7 / 62.7	28.2	32.1	29.5	<b>54.5 / 22.8</b>	<b>39.6 / 31.4</b>	<b>34.9 / 20.4</b>
OPT	47.3 / 47.5	34.7 / 70.6	31.5 / 31.5	27.6 / 36.1	29.1	32.6	31.8	31.2 / 0.4	35.9 / 27.3	26.7 / 11.2
GPT-3	50.3 / 51.8	32.1 / 56.7	33.5 / 39.7	26.9 / 36.1	27.2	33.0	30.5	31.7 / 0.6	36.0 / 21.5	27.2 / 10.6
InsGPT	<b>61.8 / 59.3</b>	48.8 / 82.7	35.1 / 48.4	<b>35.9 / 60.2</b>	<b>42.5</b>	<b>38.8</b>	<b>38.0</b>	35.4 / 2.2	36.3 / 24.7	23.6 / 16.1
SotA	- / 62.4	37.2 / -	46.2 / -	38.3 / -	-	34.4	37.2	- / 45.8	- / 47.4	- / -

Table 3: Results for all datasets, averaged across prompts. Tk-Instruct and InstructGPT are shorthanded as Tk and InsGPT, respectively. The first numbers for each task are SARI scores; additional metrics are GLEU for fluency, clarity, and coherence, EM for neutralization, Update-R1 for updating. Supervised scores from left to right are from Ge et al. (2018), Du et al. (2022), Martin et al. (2020), Pryzant et al. (2020) and Logan IV et al. (2021), respectively. The best result for each dataset is shown in bold.



# Summary

- AI can assist humans in knowledge discovery and creation
- **It is time to build those machines!**



@Fabio\_Petroni



We are hiring!

<https://samaya.ai/careers>