



NER Caption Accuracy

NER: Under the Hood

What is it? How does it work? And how can you use it?

AI-Media's Commitment to Quality Captions

AI-Media is unwavering in its commitment to providing high-quality and accurate captions through the integration of cutting-edge technologies and methodologies, particularly the Number Edition Recognition (NER) system. Recognizing the significance of accessibility and inclusivity in the media landscape, AI-Media has embraced a mission to ensure that their captioning services not only meet but exceed industry standards.

The NER system employed by AI-Media plays a pivotal role in achieving this commitment. NER is a sophisticated methodology that excels in identifying and categorizing errors within a captioned piece/transcript. By leveraging the capabilities of NER, AI-Media enhances the accuracy and contextuality of captions, ensuring that viewers receive a comprehensive and nuanced understanding of the content. This system's ability to discern specific entities and their relationships contributes to more precise and contextually relevant captions, addressing the diverse needs of viewers, including those with hearing impairments or language barriers.

AI-Media's choice to utilize the NER system of measurement reflects a strategic commitment to staying at the forefront of technological advancements. This dedication not only enhances the overall quality of our captioning services but also underscores our commitment to promote inclusive and accessible media experiences for all audiences. By embracing such advanced tools, AI-Media remains steadfast in its mission to make information universally accessible and comprehensible, fostering a more inclusive and equitable media landscape.

The Evolution of Caption Accuracy Measurement

Traditionally, the Word Error Rate (WER) model dominated caption accuracy assessments. However, with a growing need for more precise and viewer-centric evaluation, the NER (Number Edition Recognition) model emerges as the new standard, offering a more comprehensive and user-focused approach. Rather than solely taking words spoken vs. words captioned into account, the NER model focuses on overall meaning and idea units. The WER model takes less effort and time to complete, however, it does not provide an unbiased and accurate score of the content. The WER model also provides no details on the nature of errors and further work is therefore required to identify the main source(s) of errors and meaning, which NER accomplishes.

NER: Under the Hood

Introducing NER

Developed by Dr. Pablo Romero-Fresco at the University of Roehampton in the UK, the NER model serves as an objective tool for evaluating live caption quality. Adopted globally, NER is highly repeatable, making it applicable across various industries and services. The NER Model quantifies caption quality and helps AI-Media develop our ASR (Automatic Speech Recognition) solution and products. In addition to sharing NER scores with customers, results and data from hundreds of NERs allows AI-Media to develop LEXI's software, LEXI's Core Dictionary, and Topic Models. We also partner with the founder of the NER model, Pablo Romero-Fresco, to ensure alignment. This allows us to be confident in our evaluation and development of our products, compared to other competitors.

Benefits of the NER System

- **Viewer-Centric Focus:** NER addresses quality from the viewer's perspective, aligning with exhaustive surveys correlating NER scores with subjective viewer reports.
- **Repeatability:** Different assessments by various entities yield similar results, enabling the establishment of standardized benchmarks across industries.
- **Efficiency:** A ten-minute sample analysis can reliably indicate the NER score for the entire program.
- **Constructive Feedback:** The breakdown of Edition and Recognition errors offers valuable insights for captioners to improve their craft.



NER Measurement in Brief

The NER model calculates the accuracy of live captions through a unique equation, providing a quality score ranging from 0 to 100. A perfect score of 100 indicates flawless captioning.

> 99.5%	Exceptional quality, a seamless viewing experience.
99.0% – 99.5%	Excellent viewing experience, a few glitches but otherwise very clear.
98.5% – 99.0%	Great. Viewer can understand most of the content with some issues.
98.0% – 98.5%	Good. The viewer will concentrate more on the captions than the program.
< 98.0%	Needs Work. Caption issues that severely hinder the viewing experience.

What Constitutes a Good NER Score?

The benchmark for quality live captioning is set at 98%, a threshold acknowledged by regulatory authorities worldwide, including the UK's Office of Communications and the Canadian Radio-television, Telecommunications Commission. The United States FCC (Federal Communications Commission), while not stating a specific minimum accuracy % requirement, does have rules that must be adhered to as shown below... AI-Media also meets all these requirements.

How Is NER Calculated?

NER Formula

The NER score is determined using the formula:

$$\text{NER} = 100 - \frac{E + R}{N \times 100}$$

where:

- N is the total number of words and punctuation in the captioned piece.
- E is the sum of Edition errors*.
- R is the sum of Recognition errors**.

*Edition errors occur when spoken words are omitted or added incorrectly, while

**Recognition errors involve incorrect words in the captions.

Definitions:

Number = Number of words and punctuation in a captioned piece/transcript

Edition = The sum of edition errors, meaning punctuation' errors, formatting errors, or when a word was added or omitted from the captioned transcript period.

Recognition = Meaning a misspelling.

Notes:

The weightings for errors are as follows:

- 0.0 The discrepancy has no impact on comprehension.
- 0.25 The discrepancy has a small impact on comprehension.
- 0.5 The discrepancy impacts on comprehension.
- 1.0 The discrepancy misleads the viewer with false information.

A **0.0 weighting** is assigned when spoken filler words such as *you know*, *at the end of the day*, *really*, *I think* are edited from the captions on purpose, or when a phrase is reworded with fewer words but conveys the same meaning.

A 0.25 weighting is assigned to errors such as clear misspellings e.g. *there, their, they're*, or to capitalization or punctuation errors which lead to confusion.

A 0.5 weighting is assigned to errors such as missing information which would otherwise help the viewer's understanding ("There were road closures in Dunoon today," spoken but missing from the captions) or errors where a word is mistranslated ("...this famous gardener..." represented in the captions as "...this famous guard dog...").

A 1.0 weighting is assigned to errors where the captions may appear to be correct to the viewer, but are actually providing misinformation, such as an incorrect number (1,000 vs. 1,000,000) or the mistranslation of *now* ("Malcolm Turnbull is not the new leader..." instead of "Malcolm Turnbull is now the new leader...")

What is a minor (0.25) error?

Minor errors can be misrecognitions where the intended word can be inferred by the viewer (i.e. 'there' vs 'their'). Minor errors can also be small omissions or misused punctuation. In any case, a minor error may cause the viewer momentary confusion, but it does not obscure the intended meaning.

What is a standard (0.5) error?

Standard errors result in the client missing out on a unit of information. For example, a misrecognition that is so confusing it obscures the intended meaning, or a piece of information that is missing from the captions altogether (including unsuccessful paraphrasing).

What is a serious (1.0) error?

A serious error is one which misleads the client. For example, a misrecognition, paraphrase, or unit of missing information that completely changes the original meaning but appears feasible in the context.

How is the score calculated?

Once all errors have been weighted, NER calculates the accuracy score as follows:

Accuracy = Number of words minus total weight of errors divided by number of words multiplied by 100.

In other words, the number of edition and recognition errors is deducted from the total number of words in the live captions. This number is then divided by the total number of words in the verbatim transcript, and finally multiplied by one hundred.

99.5%	Exceptional quality, a seamless viewing experience.
99.0% – 99.5%	Excellent viewing experience, a few glitches but otherwise very clear.
98.5% – 99.0%	Great. Viewer can understand most of the content with some issues.
98.0% – 98.5%	Good. The viewer will concentrate more on the captions than the program.
< 98.0%	Needs work. Caption issues can severely hinder the viewing experience.

Viewer-Centric Impact

The NER model prioritizes the viewer's experience by assigning weightings to Edition and Recognition errors based on their impact on comprehension. That is, the NER model is a more holistic approach, taking into account Edition and Recognition errors by evaluating punctuation, text, content, meaning, and circumstance, to ensure that the score is an accurate representation of the quality of captions. Weightings range from 0.0 (no impact) to 1.0 (misleading information), ensuring a nuanced assessment of caption quality.

By focusing on idea units rather than individual words, the NER system aligns with a viewer-centric approach, providing a fair evaluation that benefits both viewers and captioners.

For the most accurate and viewer-friendly captions, trust the NER model – setting a new standard in caption accuracy measurement.

AI-Media Regular NER Reporting Via LEXI Lab

AI-Media proactively provides its customers with regular quality reporting, including internal reports for product enhancement and development, as well as external reports for our customers that subscribe to our Caption Quality Services. The benefit of regular reporting is that those evaluations can be used to track caption quality over time, develop further Topic Model curation, and for compliance and regulatory obligations. Those recurring evaluations are sent out to the requesting parties, whether internal departments or to our customers. We can evaluate in multiple languages, and a variety of captioning types, including ASR and human-generated content.

Sample NER results for LEXI Automatic Captioning

Live News Broadcast Applications					
NER Score	Region	Topic Model	Total Word Count	Edition Errors	Recognition Errors
99.73	APAC	Yes	2,050	3	2.5
99.67	EMEA	Yes	1,762	4	1.75
99.65	EMEA	Yes	1,762	4	2.25
99.62	EMEA	Yes	1,992	6	1.5
99.62	APAC	Yes	1,906	4.25	3
99.62	APAC	Yes	1,932	4	3.25
99.59	APAC	Yes	1,750	4	3.25
99.59	EMEA	Yes	1,808	3.5	4
99.57	EMEA	Yes	1,805	6.5	1.25
99.52	APAC	Yes	1,851	5.5	3.25
99.51	EMEA	Yes	1,727	4	4.5
99.51	APAC	Yes	1,833	4.75	4.25
99.50	EMEA	Yes	1,901	6.5	3
99.44	APAC	Yes	2,012	4.75	5.5

Note: Sample transcripts are provided below for NER results shown above in **bold**

Sports Broadcast Applications

NER Score	Region	Topic Model	Total Word Count	Edition Errors	Recognition Errors
99.42	FRENCH	Yes	1,863	3.75	8.5
99.35	USA	Yes	1,876	7.75	4.5
99.17	USA	Yes	1,599	7.75	5.5
99.08	USA	Yes	1,680	10.5	5
99.06	USA	Yes	1,785	12.25	4.5
99.05	USA	Yes	1,943	9	9.5
99.01	USA	Yes	1,547	10.75	4.5
99.00	USA	Yes	1,855	12.75	5.75
99.00	USA	Yes	1,623	6	10.25
98.99	USA	Yes	2,122	15	6.5
98.94	USA	Yes	2,155	13.25	9.5
98.75	CANADA	No	1,582	14.5	5.25
98.73	USA	Yes	2,073	8.75	17.5

Note: Sample transcripts are provided below for NER results shown above in **bold**



VERBATIM TRANSCRIPT

ANOTHER ASSIST FOR TATUM.
THAT'S HIS SIXTH TO GO WITH 21 POINTS, 10 REBOUNDS.
BARNES CONNECTS ON A THREE.
>> HE'S LOOKED COMFORTABLE ALL NIGHT,
HARRISON BARNES.
>> BARNES, 14 POINTS, SIX REBOUNDS.
HAUSER IN AND OUT.
PORZINGIS.
HE JUST OUTHUSTLED SAN ANTONIO THERE.
HAUSER UNABLE TO CASH IN.
WHITE TRACKS THAT ONE.
TRY TO SAVE IT.
COULD NOT.
BUT AGAIN, YOU HAVE SPURS PLAYERS STANDING AROUND
WATCHING, AND CELTICS HAWKING THAT BASKETBALL.
>> YOU KNOW THIS IS WHAT THE CELTICS DO.
THE CORNER CRASHES.
IT IS INCREDIBLY DIFFICULT TO DEAL WITH SOMEBODY
IN A DEAD SPRINT FROM 15 AND BEYOND AS AN OFFENSIVE REBOUNDER.
BUT YOU CANNOT DO WHAT YOU JUST SAID.
STAND AROUND, NO QUESTION.
>> ESPECIALLY , IT'S BEEN AN AREA OF STRUGGLE FOR SAN ANTONIO
THIS SEASON.

LEXI 3.0 CAPTIONS

ANOTHER ASSIST FOR TATUM.
THAT'S HIS SIXTH TO GO WITH 21 POINTS, **TEN** REBOUNDS.
BARNES CONNECTS ON A THREE.
>> HE'S LOOKED COMFORTABLE ALL NIGHT,
HARRISON BARNES.
>> BARNES, 14 POINTS, **SIX** REBOUNDS.
HAUSER IN AND OUT.
PORZINGIS.
HE JUST OUTHUSTLED SAN ANTONIO THERE.
HAUSER UNABLE TO CASH IN.
WHITE TRACKS THAT ONE.
TRY TO SAVE IT.
COULD NOT.
BUT AGAIN, YOU HAVE SPURS PLAYERS STANDING AROUND
WATCHING AND CELTICS HAWKING THAT BASKETBALL.
>> YOU KNOW THIS IS WHAT THE CELTICS DO.
THE CORNER CRASHES.
IT IS INCREDIBLY DIFFICULT TO DEAL WITH SOMEBODY
IN A DEAD SPRINT FROM 15 AND BEYOND AS AN OFFENSIVE REBOUNDER.
BUT YOU CANNOT DO WHAT YOU JUST SAID.
STAND AROUND, NO QUESTION.
>> ESPECIALLY , IT'S BEEN AN AREA OF STRUGGLE FOR SAN ANTONIO
THIS SEASON.



VERBATIM TRANSCRIPT

AND HE HAD A 12 POINT LEAD.
AGAIN, THIS REMINDS ME OF MIDWAY IN THE THIRD QUARTER
AGAINST CHICAGO.
GET A NICE LEAD AND JUST TAKE YOUR FOOT OFF THE GAS.
>> A 9-0 GRIZZLY RUN.
9:50 TO GO IN THE THIRD.
TIME OUT, SPURS.
>> THE SPURS AND OUR OFFICIAL MILITARY APPRECIATION PARTNER,
USAA, WOULD LIKE TO TAKE A MOMENT TO RECOGNIZE AND THANK
ALL CURRENTLY SERVING MILITARY PERSONNEL, VETERANS, AND
THEIR FAMILIES FOR THEIR SERVICE AND FOR THEIR SACRIFICE.
AND IT'S SALUTE NIGHT ON FRIDAY THE 17TH, THE SPURS TAKE ON
THE GRIZZLIES AGAIN.
8:30 START TIME PRESENTED BY USAA.
MILITARY AND VETS CAN SAVE UP TO 40% OFF TICKETS.
VISIT SPURS.COM/SALUTE TO GET YOUR TICKETS TODAY.
MILITARY MEMBERS IN ATTENDANCE, MAKE SURE TO STOP BY SECTION
105.
SECTION 105 FOR THE NEW BURGERS FOR THE BRAVE VALUE MEAL
PRESENTED BY USAA.

LEXI 3.0 CAPTIONS

AND HE HAD A 12 POINT LEAD **AGAIN.**
THIS REMINDS ME OF MIDWAY IN THE THIRD QUARTER
AGAINST CHICAGO.
GET A NICE LEAD AND JUST TAKE YOUR FOOT OFF THE GAS.
>> **A90** GRIZZLY RUN. **950** TO GO IN THE THIRD.
TIME OUT, SPURS.
>> THE SPURS **IN** OUR OFFICIAL MILITARY APPRECIATION PARTNER,
USAA, WOULD LIKE TO TAKE A MOMENT TO RECOGNIZE AND THANK
ALL CURRENTLY SERVING MILITARY PERSONNEL, VETERANS, AND
THEIR FAMILIES FOR THEIR SERVICE AND FOR THEIR SACRIFICE.
AND IT'S SALUTE NIGHT ON FRIDAY THE 17TH, THE SPURS TAKE ON
THE GRIZZLIES AGAIN.
830 START TIME PRESENTED BY USAA.
MILITARY _ VETS CAN SAVE UP TO 40% OFF TICKETS **IF IT'S FIRST.COM/SALUTE**
TO GET YOUR TICKETS TODAY.
MILITARY MEMBERS IN ATTENDANCE, MAKE SURE TO STOP BY SECTION
105.
SECTION 105 FOR THE NEW BURGERS FOR THE BRAVE VALUE MEAL
PRESENTED BY USAA.



VERBATIM TRANSCRIPT

Health and Social Care Secretary Wes Streeting is promising to finally grasp the nettle on social care reform with an independent commission due to begin work in April. But the commission, chaired by Baroness Louise Casey, is not due to publish its final report until 2028. Social care means help for older or disabled people with day to day tasks like washing, dressing, medication, and eating. Liberal Democrat leader Sir Ed Davey says things need to move at a faster pace.

>> Yes because Liberal Democrats have led the call for a cross-party commission.

But given the crisis in social care and the impact that has on the NHS, we do need to move much faster.

So the Liberal Democrats will be calling on for many more immediate measures to tackle the chronic shortage of staff in our care system and for the pace of this commission to be accelerated significantly.

We have to tackle this.

The conservatives left the care system, the NHS in a total mess. So we know the crisis is deep, but we've had many reports before.

LEXI 3.0 CAPTIONS

Health and Social Care Secretary Wes Streeting is promising to finally grasp the nettle on social care reform with an independent commission due to begin work in April. But the commission, chaired by Baroness Louise Casey, is not due to publish its final report until 2028. Social care means help for older or disabled people with day to day tasks like washing, dressing, medication, and eating. Liberal Democrat leader Sir Ed Davey says things need to move at a faster pace.

>> Yes because Liberal Democrats have led the call for a cross-party commission.

But given the crisis in social care and the impact that has on the NHS, we do need to move much faster.

So the Liberal Democrats will be calling on for many more immediate measures to tackle the chronic shortage of staff in our care system and for the pace of this commission to be accelerated significantly.

We have to tackle this.

The conservatives left **the left**, the care system, the NHS in a total mess. So we know the crisis is deep, but we've had many reports before.



VERBATIM TRANSCRIPT

Indigenous students are lagging far behind in education. While more than 84% of non-Indigenous students in the Northern Territory attend school regularly, the official data shows that numbers drops below 52% for Aboriginal children. The real numbers in this space are even worse. The Northern Territory's Education Minister says children who fail to attend school for 20 days are taken off the official roll and not included in the attendance data. The NT government started fining parents if they don't send their kids to school, and kids need to understand that as well, because we are not going to be able to move forward and we are not going to be able to get better results in education, and for them to be able to get a job, if they do not attend. The funding will have a focus on policing and community safety, something that has been front of mind for Alice Springs residents. We know that community safety is the number one priority of Territorians, no matter where they live. 12 child and family centres will be built in remote communities with a focus on child safety.

LEXI 3.0 CAPTIONS

Indigenous students are lagging far behind in education. While more than 84% of non-Indigenous students in the Northern Territory attend school regularly, the official data shows that numbers drops below 52% for Aboriginal children. The real numbers in this space are even worse. The Northern Territory's education minister says children who fail to attend school for 20 days had taken off the official roll and not included in the attendance data. The NT government started finding parents if they don't send their kids to school, and kids need to understand that as well, because we are not going to be able to move forward and we are not going to be able to get better results in education, and for them to be able to get a job, if they do not attend. The funding will have a focus on policing and community safety, something that has been front of mind for Alice Springs residents. We know that community safety is the number one priority of Territorians, no matter where they live. 12 child and family centres will be built in remote communities with a focus on child safety.



AI MEDIA

For more LEXI NER results and examples of content for each test, scan the QR code below.



AI MEDIA

Find out more about AI-Media's
captioning solutions at [AI-MEDIA.TV](https://ai-media.tv)
or contact SALES@AI-MEDIA.TV