



34th Annual **INCOSE**
international symposium

hybrid event

Dublin, Ireland
July 2 - 6, 2024



Journalistic best practices for Integrating IV&V into a Generative AI Enterprise
Work Culture

IV&V and GenAI

Jason Blevins
CSEP#00161

2-6 July 2024

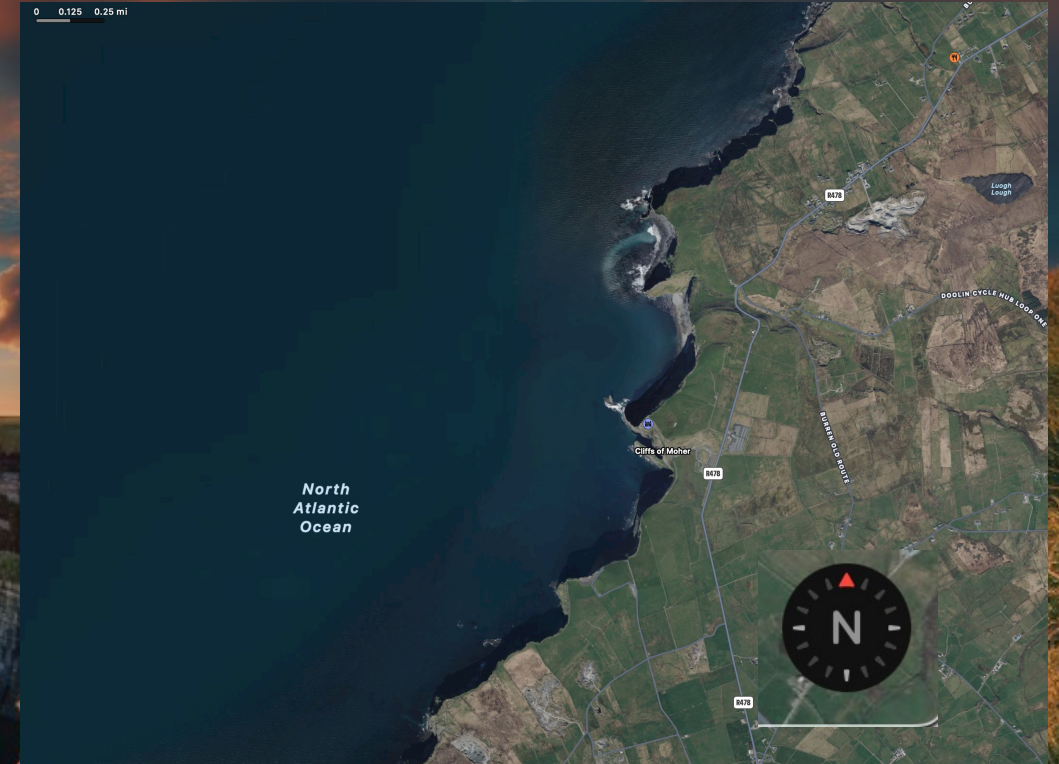
www.incose.org/symp2024 #INCOSEIS



/Imagine Prompt: Cliffs of Moher in Ireland, at sunset, golden hour, soft lighting, hyper photo realistic widescreen format 4k resolution --ar 4:3

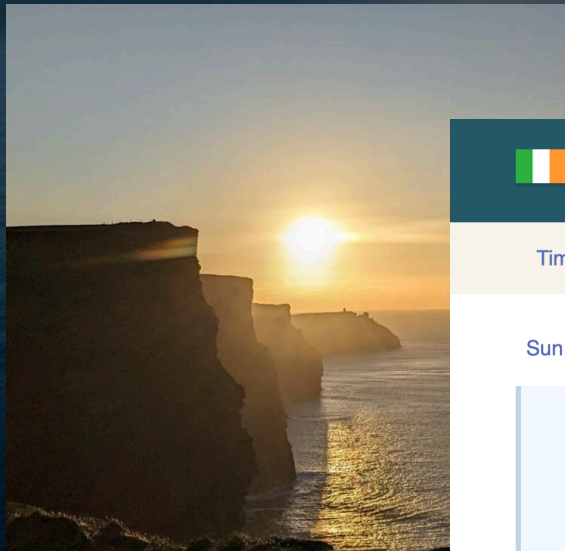


Field Expertise & Secondary Sources

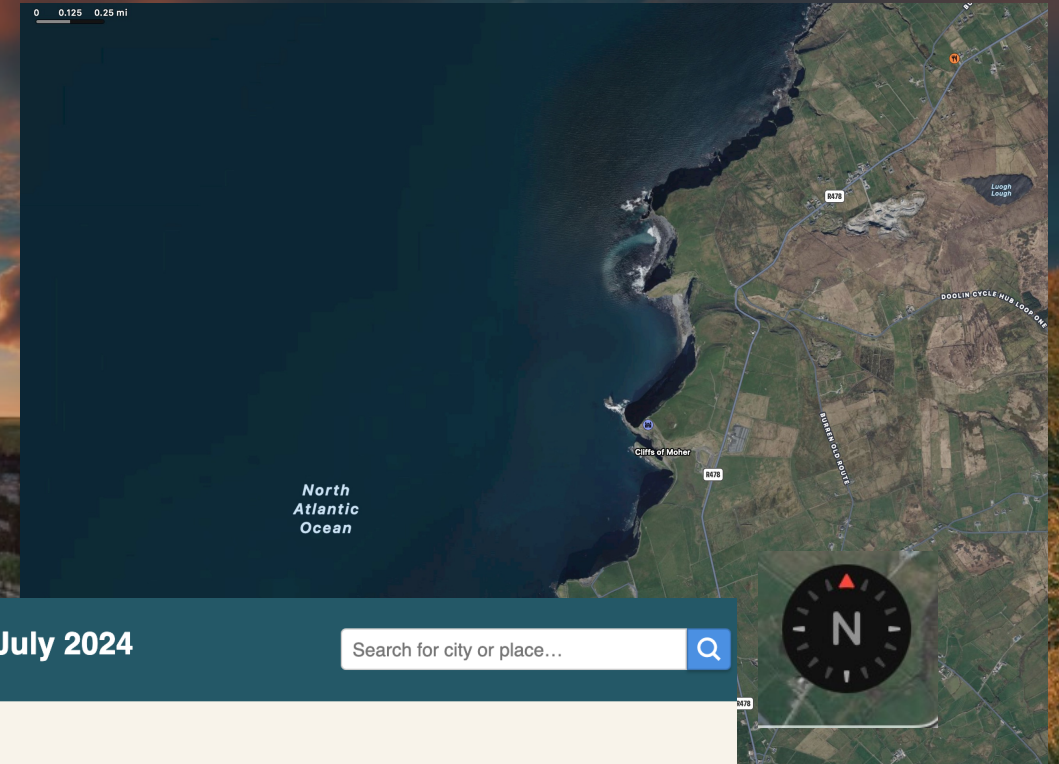


/Imagine Prompt: Cliffs of Moher in Ireland, at sunset, golden hour, soft lighting, hyper photo realistic widescreen format 4k resolution --ar 4:3

Field Expertise & Secondary Sources



/Imagine Prompt: Cliffs of Moher at sunset
photo realistic widescreen for



Ireland, Ireland — Sunrise, Sunset, and Daylength, July 2024

Search for city or place... 

[Time/General](#) [Weather](#) [Time Zone](#) [DST Changes](#) [Sun & Moon](#)

[Sun & Moon Today](#) [Sunrise & Sunset](#) [Moonrise & Moonset](#) [Moon Phases](#) [Eclipses](#) [Night Sky](#)



Daylight

05:12 – 21:59
16 hours, 47 minutes

| | |
|-----------------------|------------------------------|
| Current Time: | 3 Jul 2024, 20:24:48 |
| Sun Direction: | 293.30° WNW ↙ |
| Sun Altitude: | 11.27° |
| Sun Distance: | 152.099 million km |
| Next Equinox: | 22 Sep 2024 13:43 (Autumnal) |
| Sunrise Today: | 05:12 ↗ 48° Northeast |
| Sunset Today: | 21:59 ↙ 312° Northwest |



2024 Sun Graph for Ireland

[Rise/Set Times](#) [Day/Night Length](#)

Overview

- GenAI Opportunities & Challenges
- Models & Hallucinations
- Translation Example
- Inaccuracy Tolerance
- Calculation Example
- Increasing Reliability with Diversity
- Lessons from Journalism
- Journalistic Best Practices

Is Our Culture Ready for Generative AI?

Gen AI Opportunities

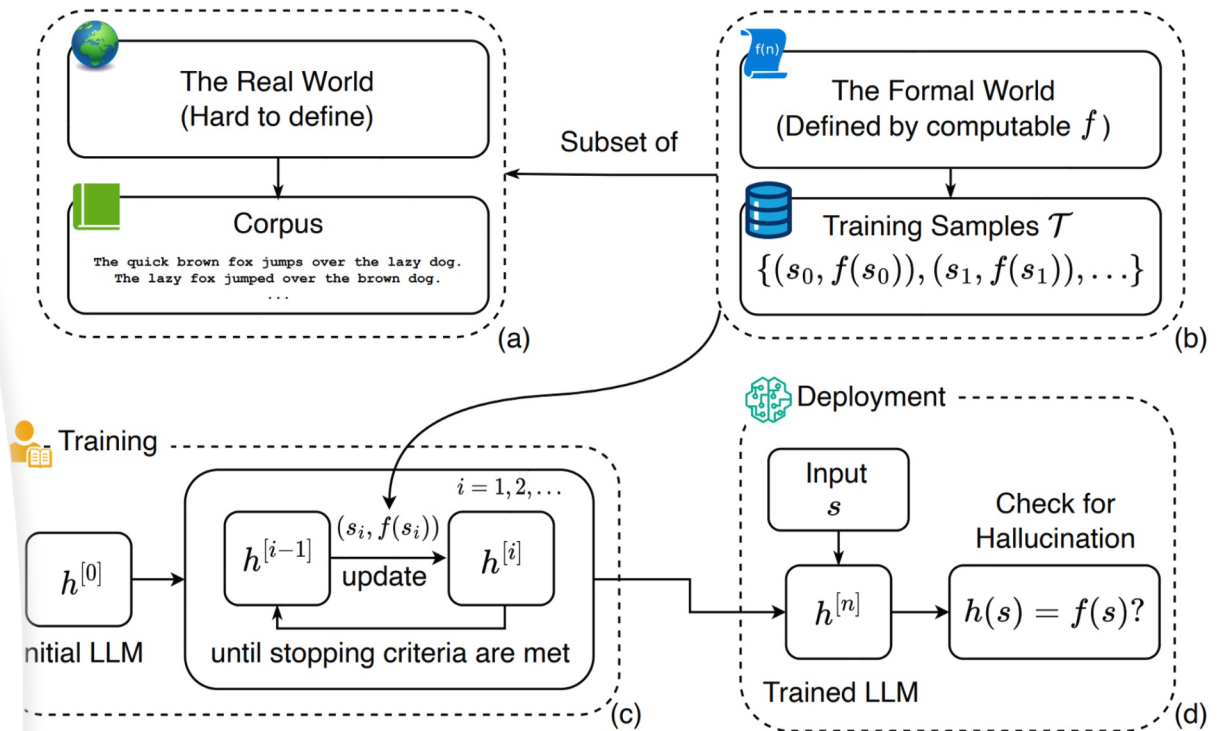
- Rapid Probabilistic Information Access
- Automatic “Accurate” Summarization
- Good at Rough Draft Exposition
- Excellent Translation of Language and Code
- Amplify the Experts with Human AI Alignment

Challenges

- Potential for Atrophy
- Probably Right, Possibly Very Wrong
- Guesses Mistaken for Insight
- Training Data May Become Poisoned
- Requires Subject Matter Expertise to Validate and Verify

“All models are wrong, but some are useful”.
George E. P. Box

Hallucination is Inevitable: An Innate Limitation of Large Language Models



Hallucination is Inevitable: An Innate Limitation of Large Language Models

Ziwei Xu, Sanjay Jain, Mohan Kankanhalli

Hallucination has been widely recognized to be a significant drawback for large language models (LLMs). There have been many works that attempt to reduce the extent of hallucination. These efforts have mostly been empirical so far, which cannot answer the fundamental question whether it can be completely eliminated. In this paper, we formalize the problem and show that it is impossible to eliminate hallucination in LLMs. Specifically, we define a formal world where hallucination is defined as inconsistencies between a computable LLM and a computable ground truth function. By employing results from learning theory, we show that LLMs cannot learn all of the computable functions and will therefore always hallucinate. Since the formal world is a part of the real world which is much more complicated, hallucinations are also inevitable for real world LLMs. Furthermore, for real world LLMs constrained by provable time complexity, we describe the hallucination-prone tasks and empirically validate our claims. Finally, using the formal world framework, we discuss the possible mechanisms and efficacies of existing hallucination mitigators as well as the practical implications on the safe deployment of LLMs.

Subjects: Computation and Language (cs.CL); Artificial Intelligence (cs.AI); Machine Learning (cs.LG)

Cite as: arXiv:2401.11817 [cs.CL]
(or arXiv:2401.11817v1 [cs.CL] for this version)
<https://doi.org/10.48550/arXiv.2401.11817>

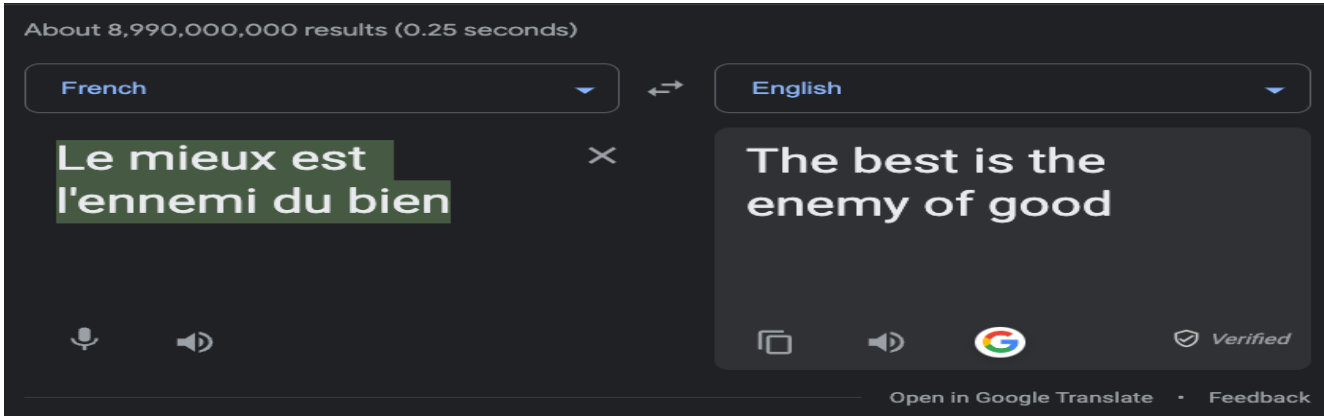
Submission history

From: Ziwei Xu [view email]
[v1] Mon, 22 Jan 2024 10:26:14 UTC (291 KB)

Journalism: A Short History Lesson



“Le mieux est l'ennemi du bien” – Voltaire 1764



ChatGPT 4: The exact quote “The Perfect is the enemy of the good” in the original French is "Le mieux est l'ennemi du bien." This quote is often attributed to Voltaire, coming from his poem, "La Bégueule," which appeared in his "Dictionnaire philosophique" (Philosophical Dictionary). The English translation of this quote is indeed "Perfect is the enemy of the good," conveying the idea that insisting on perfection often prevents the completion or achievement of good things.



Midjourney Bot
/imagine prompt: hyper photo realistic portrait of voltaire in rococo style

Generative Code + Translation

Python

```
•# Function to capture user information
•def capture_user_info():
•print("Please enter the following
information:")

•# Capture name
•name = input("Name: ")

•# Capture address
•address = input("Address: ")

•# Capture email address
•email = input("Email Address: ")

•# Capture time zone
•time_zone = input("Time Zone: ")

•return name, address, email, time_zone

•# Example usage
•if __name__ == "__main__":
•user_name, user_address, user_email,
user_time_zone = capture_user_info()

•# Print captured information
•print("\nUser Information:")
•print(f"Name: {user_name}")
•print(f"Address: {user_address}")
•print(f"Email Address: {user_email}")
•print(f"Time Zone: {user_time_zone}")
```

Powershell

```
•# Function to capture user information
•function Capture-UserInfo {
•Write-Host "Please enter the following information:"

•# Capture name
•$name = Read-Host "Name"

•# Capture address
•$address = Read-Host "Address"

•# Capture email address
•$email = Read-Host "Email Address"

•# Capture time zone
•$timeZone = Read-Host "Time Zone"

•# Create a hashtable to store the captured information
•$userInfo = @{}
•Name = $name
•Address = $address
•Email = $email
•TimeZone = $timeZone
•}

•return $userInfo
•}

•# Example usage
•$userInfo = Capture-UserInfo

•# Print captured information
•Write-Host "`nUser Information:"
•Write-Host "Name: $($userInfo.Name)"
•Write-Host "Address: $($userInfo.Address)"
•Write-Host "Email Address: $($userInfo.Email)"
•Write-Host "Time Zone: $($userInfo.TimeZone)"
```

Java

```
•import java.util.Scanner;

•public class CaptureUserInfo {
•public static void main(String[] args) {
•Scanner scanner = new Scanner(System.in);

•System.out.println("Please enter the following
information:");

•// Capture name
•System.out.print("Name: ");
•String name = scanner.nextLine();

•// Capture address
•System.out.print("Address: ");
•String address = scanner.nextLine();

•// Capture email address
•System.out.print("Email Address: ");
•String email = scanner.nextLine();

•// Capture time zone
•System.out.print("Time Zone: ");
•String timeZone = scanner.nextLine();

•// Close the scanner
•scanner.close();

•// Print captured information
•System.out.println("\nUser Information:");
•System.out.println("Name: " + name);
•System.out.println("Address: " + address);
•System.out.println("Email Address: " + email);
•System.out.println("Time Zone: " + timeZone);
•}
•}
```

Cobol

```
•IDENTIFICATION DIVISION.
•PROGRAM-ID. CaptureUserInfo.

•DATA DIVISION.
•WORKING-STORAGE SECTION.
•01 NAME PIC X(50).
•01 ADDRESS PIC X(100).
•01 EMAIL PIC X(50).
•01 TIME-ZONE PIC X(50).

•PROCEDURE DIVISION.
•DISPLAY "Please enter the following
information:".

•ACCEPT NAME PROMPT "Name: ".
•ACCEPT ADDRESS PROMPT "Address: ".
•ACCEPT EMAIL PROMPT "Email
Address: ".
•ACCEPT TIME-ZONE PROMPT "Time
Zone: ".

•DISPLAY "User Information:".
•DISPLAY "Name: " NAME.
•DISPLAY "Address: " ADDRESS.
•DISPLAY "Email Address: " EMAIL.
•DISPLAY "Time Zone: " TIME-ZONE.

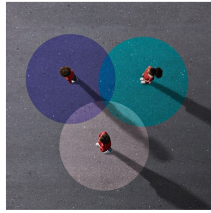
•STOP RUN.
```

Understand Your Inaccuracy Tolerance



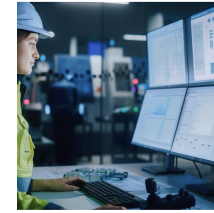
Good/Maith

- Summarization
- Exposition
- Tone/Sentiment
- Graphics
- Internal communications
- Include:
- Human Validation



Verifiable/ Fearr

- Public Chatbots
- Public Facing Models
- M/L Models
- Generative Code
- Include:
- Expert Verification
- Rigorous Testing



Perfect/Foirfe

- Flight controls
- Safety Systems
- Autonomous Systems
- Include:
- Full Testing

Airline chatbot story

<https://www.bbc.com/travel/article/20240222-air-canada-chatbot-misinformation-what-travellers-should-know>

Perfect or Good?

| Availability (%) | Downtime per Year | Downtime per Month | Downtime per Week | Downtime per Day |
|-------------------------|-------------------|---------------------|--------------------|-------------------|
| 99% (Two Nines) | 3.65 days | 7.20 hours | 1.68 hours | 14.4 minutes |
| 99.9% (Three Nines) | 8.76 hours | 43.8 minutes | 10.1 minutes | 1.44 minutes |
| 99.99% (Four Nines) | 52.56 minutes | 4.32 minutes | 1.01 minutes | 8.64 seconds |
| 99.999% (Five Nines) | 5.26 minutes | 25.9 seconds | 6.05 seconds | 864 milliseconds |
| 99.9999% (Six Nines) | 31.5 seconds | 2.59 seconds | 259 milliseconds | 259 milliseconds |
| 99.99999% (Seven Nines) | 3.15 seconds | 262.97 milliseconds | 60.48 milliseconds | 8.64 milliseconds |

98% Accuracy on GPT Text summarization

Perfect or Good?

| Availability (%) | Downtime per Year | Downtime per Month | | |
|-------------------------|-------------------|---------------------|--------------------|-------------------|
| 99% (Two Nines) | 3.65 days | 7.20 hours | | |
| 99.9% (Three Nines) | 8.76 hours | 43.8 minutes | | |
| 99.99% (Four Nines) | 52.56 minutes | 4.32 minutes | | |
| 99.999% (Five Nines) | 5.26 minutes | 25.9 seconds | | |
| 99.9999% (Six Nines) | 31.5 seconds | 2.59 seconds | | |
| 99.99999% (Seven Nines) | 3.15 seconds | 262.97 milliseconds | 60.48 milliseconds | 8.64 milliseconds |

99% Availability

- Redundancy, monitoring, backups, failover, SLAs.

99.9% Availability

- Load balancing, geo-redundancy, auto-scaling, disaster recovery, high-quality hardware.

99.99% Availability

- Microservices, immutable infrastructure, hot swapping, enhanced monitoring, zero-downtime deployments.

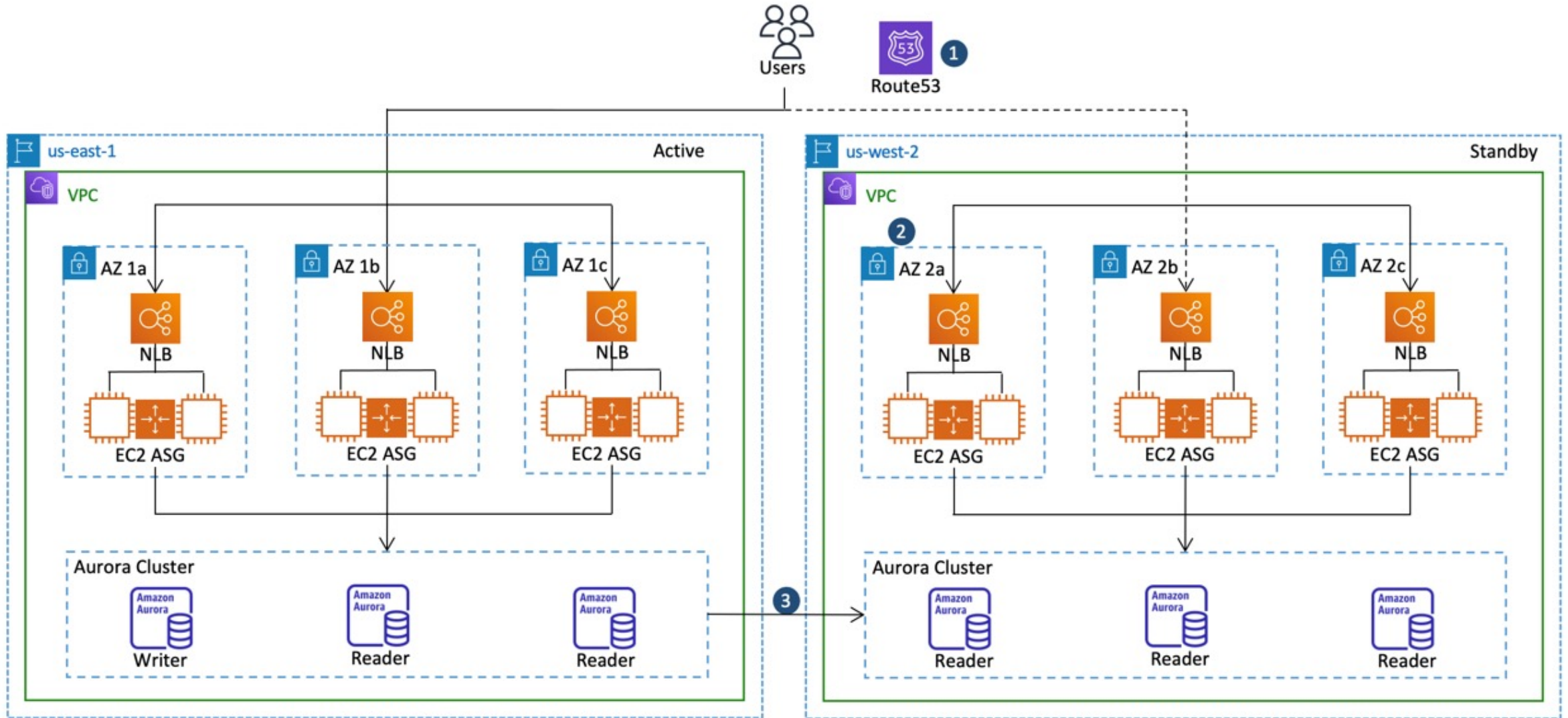
99.999% Availability

- Fault tolerance, predictive maintenance, self-healing, advanced load balancing, synchronous replication.

98%

Accuracy on GPT Text summarization

Diversity = Reliability



Perfect is the Enemy of the Good

To estimate the amount of jet fuel required to fly a Boeing 747 from Newark Liberty International Airport (EWR) to Dublin Airport (DUB), we'll consider the following factors:

1. Flight Distance (D):

3180mi/5115 kilometers

2. Average Ground Speed (V):

560mph/901.2kmh }

3. Flight Time (T): T=5.68 hours

4. Fuel Flow Rate (FFR): 10.5 tons/hour

5. Base Fuel Consumption (BFC): 59.64 tons

6. Adjust for Takeoff, Climb, and Descent:

Additional Fuel=5.96 tons

7. Reserve Fuel: 2.98 tons

8. Total Fuel Burn (TFB): 68.58 tons

- The estimated fuel burn for the flight from Newark to Dublin, is approximately 68.58 tons.



HOW MUCH FUEL DOES A JUMBO JET BURN?

HOW MUCH FUEL DOES A JUMBO JET (BOEING 747-400) BURN?

What fuel does a jumbo jet use between London and New York?

HOW MUCH FUEL DOES A JUMBO JET BURN?

The four engines of the Boeing 747 Jumbo Jet burn approximately 10 to 11 tonnes of fuel on hour when in the cruise. This equates to roughly 1 gallon (approximately 4 litres) of fuel every second. It can carry a maximum of 238,604 liters of fuel and it has a range of about 7,790 nautical miles. A Jumbo Jet (Boeing 747-400) flying from London to New York burns approximately 70,000 kilograms of fuel. Jet fuel has an approximate specific gravity of 0.85, which therefore equates to 82,353 litres. Therefore, the cost of the fuel (based on 1 litre costing 31 pence) required to fly from London to New York is approximately £18,500 (€23,600). The cost of fuel for a a jumbo jet carrying 450 passengers, would work out as about £41 (€52) per person.

The equation used to calculate fuel burn in aviation is often referred to as the **Breguet Range Equation**. This equation is particularly useful for determining the range of an aircraft based on its fuel consumption and efficiency. Here is the Breguet Range Equation in its general form:

Breguet Range Equation

$$R = \frac{V}{C} \times \ln \left(\frac{W_i}{W_f} \right)$$

Where:

- R = Range of the aircraft (distance it can fly)
- V = Velocity or average ground speed of the aircraft
- C = Specific fuel consumption (fuel flow rate per unit thrust, often in units of fuel per unit distance or time)
- W_i = Initial weight of the aircraft (including fuel)
- W_f = Final weight of the aircraft (after fuel has been consumed)
- \ln = Natural logarithm

This equation provides a theoretical basis for understanding how factors like speed, fuel efficiency, and weight affect the distance an aircraft can travel.

Lessons from Journalism



1605 First Newspaper Printing press



1780s First Amendment protects freedom of the press



1830s Penny Press make news affordable to the general public



1923 American Society of Newspaper Editors (ASNE) standards for responsibility in reporting.



1980s 24-hour news cycles and cable news channels, brought new challenges in speed, accuracy, and ethics



2000s Digital/Social media platforms, raising issues related to misinformation, fake news

← Networked Computing



human content creation & algorithmic search



probabilistic generative content →

TCP/IP

WWW

GPT 3.5/4o

Pre-Digital Era Journalism



- Reliance on Primary Sources
- Limited Public Fact-Checking
- Ample Time for Verification
- Public Trust

Digital Transformation

- **Access to Online Databases:** A wealth of digital archives, databases, & resources allowing for quicker verification of facts.
- **Real-Time Verification Tools:** Tools & software have been developed to verify digital content, including reverse image searches, video analysis tools, & social media monitoring platforms.
- **Increased Speed and Pressure:** The 24/7 news cycle & immediacy of online publishing have increased pressure on journalists to report quickly, often at the expense of thorough verification.

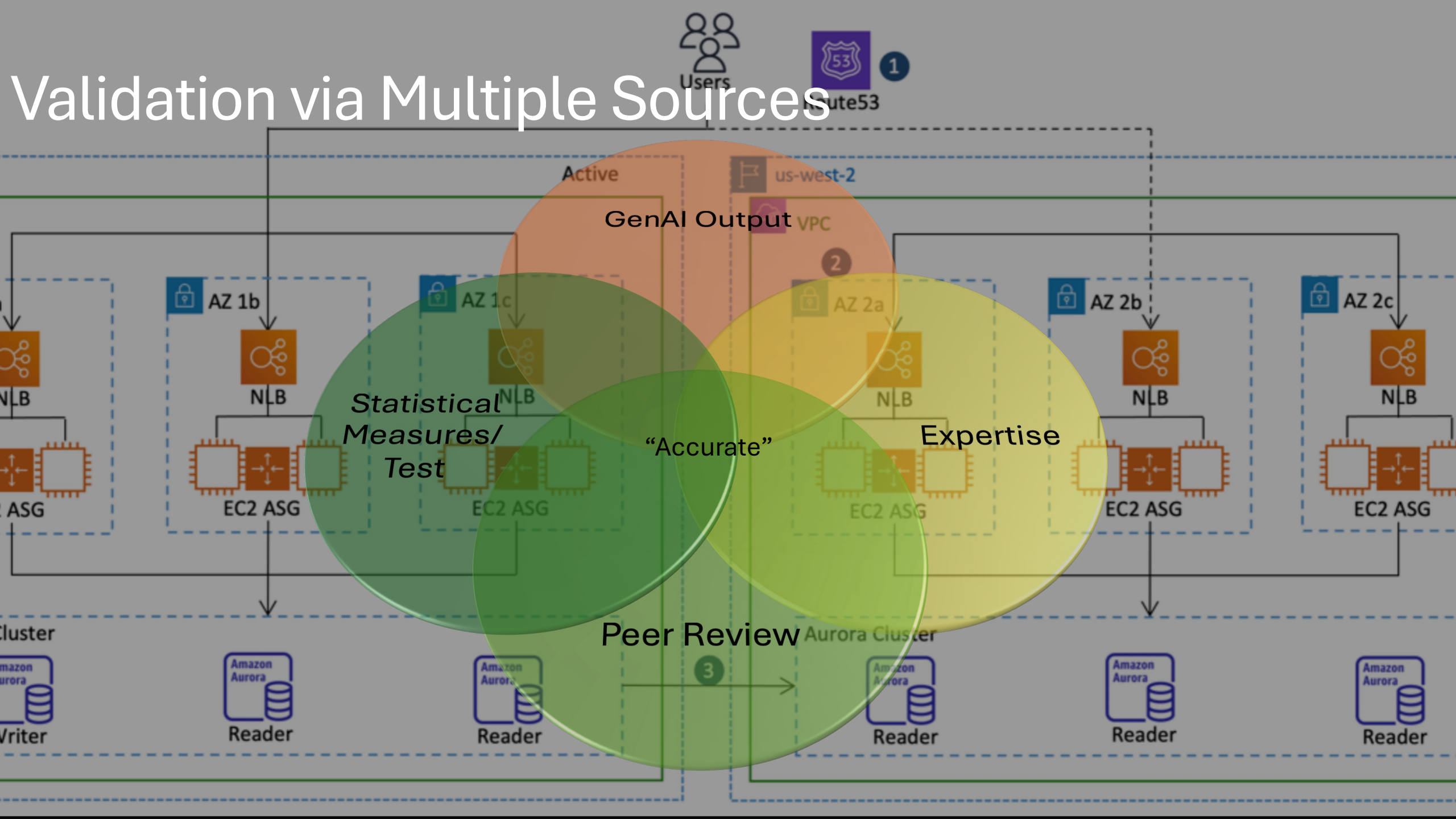
"A lie can travel halfway around the world before the truth has a chance to get its boots on."

Current Best Practices

- **Multi-Source Verification:** Journalists are encouraged to verify information with multiple sources, especially when covering breaking news or controversial topics.
- **Transparency with Audiences:** Show the work behind the reporting including citing sources, explaining the verification process, & acknowledging uncertainties.
- **Collaboration & Fact-Checking Networks:** News organizations collaborate with fact-checking organizations & share resources for verifying information.
- **Adapting to Misinformation:** Training in recognizing & debunking misinformation has become crucial. Journalists now play a key role in fact-checking claims made on social media & in viral content.
- **Ethical Considerations:** There's an increased focus on the ethics of reporting in the digital age, including considerations about the potential harm of spreading unverified information.

PEREF DESIGN DESIGN DESIGN ✓ VALIDATIE VALIFATE VALIFRY

Validation via Multiple Sources



Key Takeaways



LLM's are good,
not perfect



Understand inaccuracy tolerance
(perfect or good?)



Math and
Language skills
will improve



Human
Expertise is
amplified, not
replaced



IV&V has never
been more
important



Embrace a culture of verification & Validation, consider
using journalistic best practices

plé
discutere

discutir

discuter

चर्चा

讨论

keskustella

diskutere

обсуждать

discutir

jadili

話し合う

megbeszél

ræða

Discuta

ناقش

Discussion

Future Trends

- **AI and Automation in Verification:** The use of artificial intelligence to assist in verifying content and detecting deepfakes is on the rise, though it comes with its own set of challenges and ethical considerations.
- **Audience Participation:** Crowdsourcing and audience engagement in the fact-checking process are becoming more common, leveraging the collective knowledge and resources of the public.
- **Global Fact-Checking Initiatives:** As misinformation becomes a global issue, international collaborations and initiatives are becoming more important in developing standards and sharing best practices.

