

Open three tools:

ChatGPT or **Claude.ai** or **Gemini**

NotebookLM: notebooklm.google

+ download **docs** from

royal.uk/media-centre/financial-reports

Colab: colab.research.google.com

All use Large Language Models (LLMs)



It's **STILL**
Not a
robot.

3 things

- Get to know **prompt design** techniques
- Ooh, it's **risky**
- **Tips and tricks** for investigations

Pick a story:

It can be an issue: e.g. housing

It can be an event: e.g. doctors striking

It can be an angle: e.g. wide variations in
how much school sports kit costs

Open: github.com/paulbradshaw/genai

Run this prompt:

Generate 5 story ideas about
[ENTER YOUR TOPIC HERE]

Generate 5 story ideas about [the climate visa]



Sure, here are 10 story ideas centered around the concept of a "climate visa":

1. **A New Beginning:** Follow the journey of a family from a small Pacific island that has been granted climate visas to relocate to New Zealand due to rising sea levels. Explore their struggles with leaving their home, integrating into a new culture, and the emotional impact of becoming climate refugees.
2. **The Climate Visa Activist:** A young environmental lawyer fights to establish a new international policy for climate visas. The story follows her battles in court, the opposition she faces from powerful industries, and her personal sacrifices, leading to a landmark global agreement.
3. **Climate Visa and the American Dream:** A Midwestern American town facing severe drought opens its doors to climate visa holders from Sub-Saharan Africa. The influx of new residents brings cultural exchanges, economic challenges, and a renewed sense of community as they all work together to rebuild.
4. **The Last Generation:** Set in 2050, a journalist documents the final years of a village in the Arctic Circle before...

Be CAREful.

CAREful prompts

- **Context:** describe the situation (role, employer, events)
- **Ask:** request specific action
- **Rules:** provide constraints (audience, genre, length)
- **Examples:** demonstrate what you want
- + Iterate, refine, and combine outputs

I am an investigative **journalist** working for a national news website in **Spain**. The **audience** is urban professionals aged 20-40.

Generate 5 ideas for **features** about [water abstraction]. Each idea should be **max 50 words**.

[Here are some **examples** of headlines of features that I like:] [PASTE HEADLINES]



Here are five feature ideas on water abstraction, tailored for a professional urban audience and aligned with the style you like:

1. **Tapped dry: how illegal wells are fuelling Spain's water crisis**

Unlicensed groundwater extraction is surging as drought intensifies. This feature explores the scale of illegal abstraction and the fragile ecosystems left in its wake.

2. **Strawberry fields forever? The real water cost of your supermarket fruit**

Export farming in southern Spain is draining protected wetlands. We trace the journey from depleted aquifers to your local fridge.

3. **The silent theft: how water privatisation hides who's taking what**

Opaque contracts and missing data mean we often don't know who's extracting our water, or how much. A deep dive into regulatory blind spots.

What you write

Give me ideas for
investigative stories
about UK water companies.

What genAI does

Predict a sequence of words which correlate strongly with **tokens** clustered around concepts like 'investigation', 'UK', 'water', 'companies' etc. – and a that is **likely** to be rated as useful by humans



Autocomplete for journalists

(advanced)

Tip: learn tradecraft, teach the AI, reduce deskilling:

There are four particular types of stories that investigative journalists tell about rules. Those are:

Rules having unintended consequences

Rules not being followed

Rules not being enforced

The need for new rules

Generate two story ideas for each of these types

[When that follow-up prompt is given]



1. Rules having unintended consequences

The drought loophole: how emergency permits are fuelling over-extraction

Emergency drought rules allow water extraction beyond normal limits – but some farms are exploiting the system year-round, with lasting damage to aquifers.

Paying to save water, but using more: Spain's failed irrigation subsidies

Millions in EU funds were meant to reduce water use through efficient irrigation. Instead, farms expanded output – and water use went up.

2. Rules not being followed

Stolen water, sold abroad: the hidden economy of illegal wells

Despite strict rules, thousands of unregistered wells irrigate crops for export. Officials admit enforcement is patchy – and often politically inconvenient.

Protected on paper: how national parks are being drained dry

Rules ban water extraction in fragile ecosystems like Doñana, yet satellite images and local records show continued pumping inside park boundaries.

3. Rules not being enforced

Licence to pump: why Spain's water permits are rarely checked

Tip: add data,
teach the AI the
7 angles

7 common angles for data stories

Scale



Change



Ranking



Variation



Explore



Relationships



Bad/open



(+ Leads)



Tip: add data,
teach the AI the
7 angles

You are an editor in a data journalism team at a UK news website. I am going to give you some information about types of story angles that can help you identify story ideas based on any dataset.

Here are the first four angles. Please just confirm you understand this, and then I will give you some data to generate ideas for.

Scale stories 'reveal' the scale of something, typically a problem, but it can also be a topical issue.

Change stories reveal a trend: that something is rising or falling. A lack of change can be newsworthy too, especially if the government has pledged to reduce or improve something and it's not happening.

Ranking stories are all about what, who, when and where are the best or worst w comes to a particular metric. It might also be about how a category or place relevant to the audience ranks against others.

Variation stories, in contrast, reveal the very existence of variation where we might aim to avoid it, i.e. it reveals unfairness. Examples include people not being treated equally, or not having equal access to services. These are sometimes called 'postcode lottery' stories



Augment
knowledge

Now, here are the first two rows of a dataset on the gender pay gap at each company in the UK. Use the four angles I explained to suggest 10 story ideas for this data, max 50 words each:

EmployerName	Address	CompanyNumber	SicCodes	DiffMeanHourlyPercent		
DiffMedianHourlyPercent		Women earn more than men?		DiffMeanBonusPercent		
DiffMedianBonusPercent		MaleBonusPercent	FemaleBonusPercent	MaleLowerQuartile		
FemaleLowerQuartile	MaleLowerMiddleQuartile		FemaleLowerMiddleQuartile			
MaleUpperMiddleQuartile	FemaleUpperMiddleQuartile	MaleTopQuartile	FemaleTopQuartile		Female	
top quartile 50 or above	No link?	Name	Position	CompanyLinkToGPGInfo		
ResponsiblePerson	EmployerSize	CurrentName	SubmittedAfterTheDeadline	DueDate		
DateSubmitted						

"Bryanston School", Incorporated	Bryanston House,
Blandford,	
Dorset,	
United Kingdom,	
DT11 0PX"	226143
85310	19.6
32.5	FALSE
0	2
61.8	
44.9	55.1
53.3	46.7
FALSE	FALSE
Michael McGovern	Chief
https://www.bryanston.co.uk/employment	Michael McGovern (Chief Op
to 999	
"Bryanston School", Incorporated	FALSE
05/04/2019 00:00:00	03/04/20

Augment data

You are a **sceptical** education reporter with excellent judgment for what makes a compelling news story. You have years of experience of **unpacking PR spin and political misinformation**. You are looking for ideas for a story about school uniform, specifically any potential stories about 1) rules not being followed, or 2) enforced, 3) unintended consequences of rules, or 4) new rules being needed.

Look at the attached rules, regulations, laws and guidance that apply to uniform and **identify 5 areas that might be potential avenues to investigate**.



**Augment
documents**

These are prompt designs

- **Role prompting**: “You are a sceptical editor”
- **Recursive prompting**: “Give me more like idea 2”
- **N-shot prompting**: “Here are [two] examples”
- **RAG (Retrieval Augmented Generation)**: “Factor in this information/use this document/data”

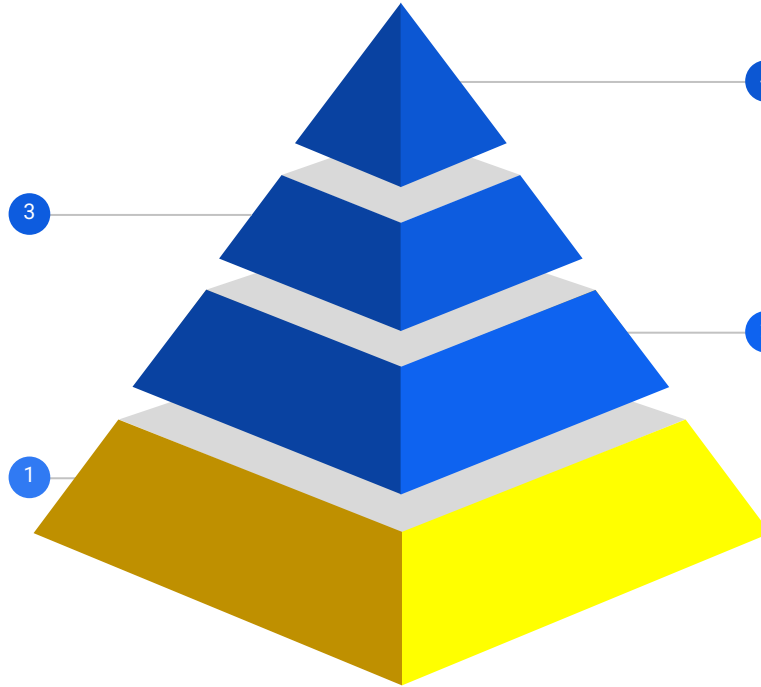
Applications of genAI in the journalism process

Production

Identify jargon and bias;
improve spelling, grammar,
structure and brevity.
Drafting, 'blank page'.

Pre-production

Idea generation and stimulation:
identify and map systems and
rules, apply brainstorming
frameworks (iceberg model, 5
whys, 8 angles of data journalism).
Planning.



Post-production

Optimisation and
reversioning

Research

Scope diverse sources,
explore/summarise docs,
translate, form advanced
searches, write/fix code for
scraping and analysis

ONLINE JOURNALISM

BLOG

A conversation.

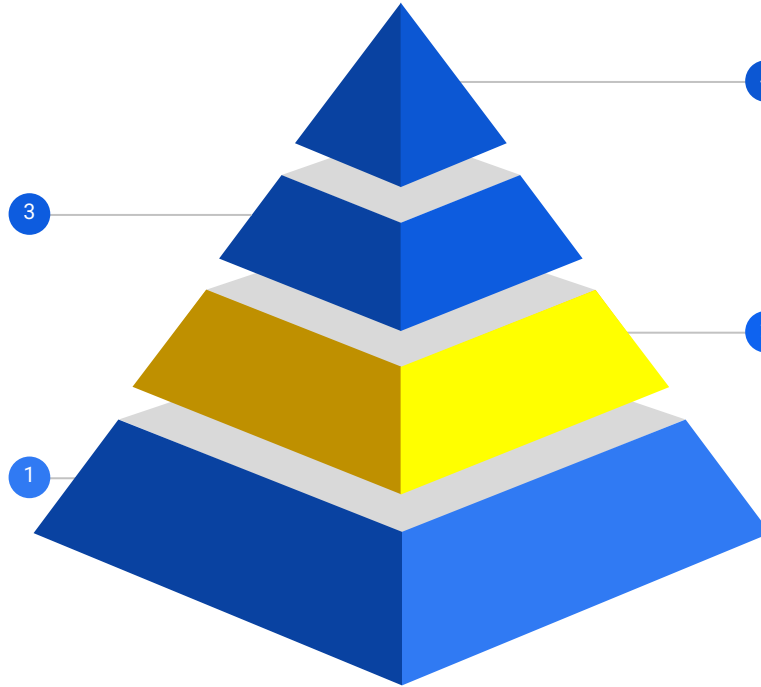
Applications of genAI in the journalism process

Production

Identify jargon and bias;
improve spelling, grammar,
structure and brevity.
Drafting, 'blank page'.

Pre-production

Idea generation and stimulation:
identify and map systems and
rules, apply brainstorming
frameworks (iceberg model, 5
whys, 8 angles of data journalism).
Planning.



Post-production

Optimisation and
reversioning

Research

Scope diverse sources,
explore/summarise docs,
translate, form advanced
searches, write/fix code for
scraping and analysis

ONLINE JOURNALISM

BLOG

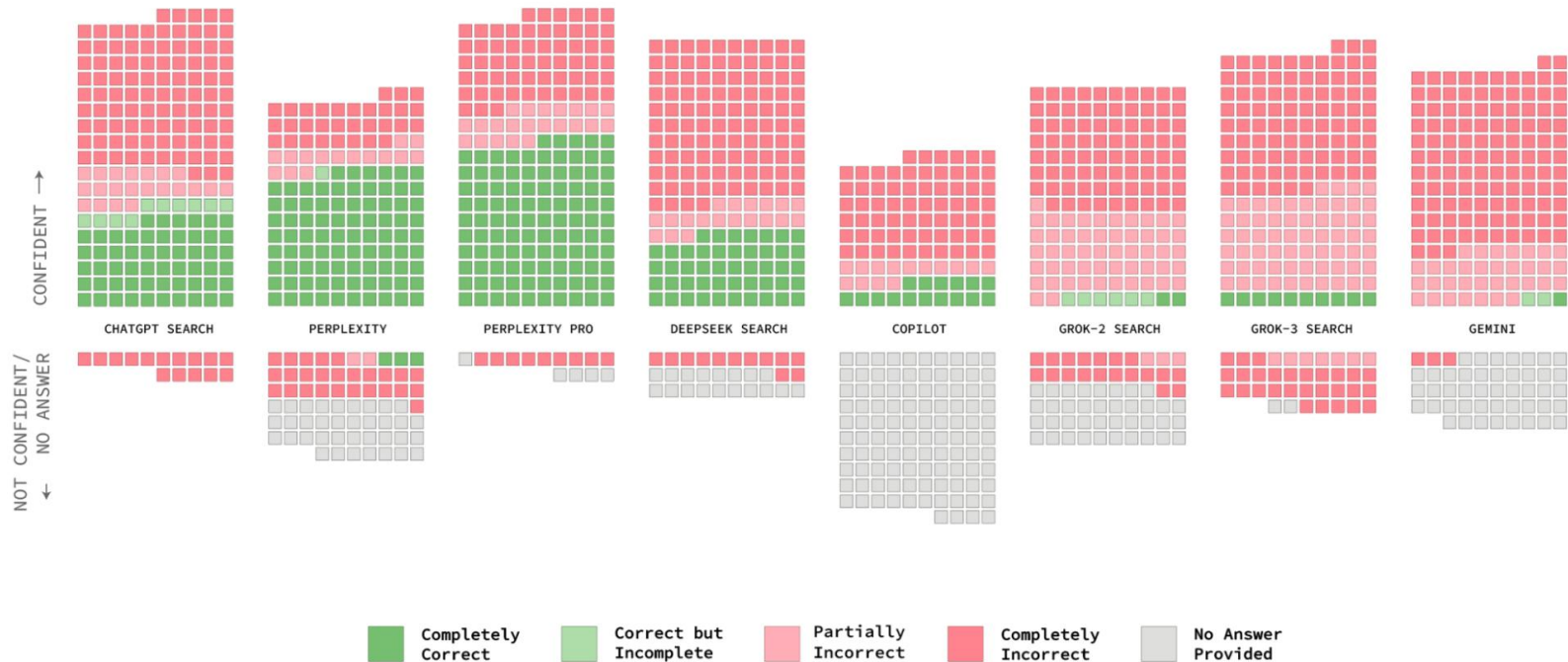


DANGER

ahead

Generative search tools were often confidently **wrong** in our study

The Tow Center asked eight generative search tools to identify the source article, the publication and URL for 200 excerpts extracted from news articles by 20 publishers. Each square represents the citation behavior of a response.



“Chatbots were generally **bad at declining to answer questions they couldn’t answer accurately, offering incorrect or speculative answers instead ... [and offering] fabricated links**” - [Columbia Journalism Review](#)

Garbage In Garbage Out/credulity:

ChatGPT

Metric Confusion **39/100**

❗ **Critical Failure:**
"Germany led Europe with 1.38 million plug-in cars registered since 2010"

Data Issues

- ⊗ Mixed vehicle types
- ⊗ Confused timeframes
- ⊗ Combined incompatible metrics
- ⊗ Misrepresented totals

Methodological Flaws

- ⊗ Poor metric definition
- ⊗ Inconsistent categorization
- ⊗ Failed to standardize data
- ⊗ Mixed measurement types

Professional Failures

- ⊗ Unclear methodology
- ⊗ Inconsistent analysis
- ⊗ Inadequate definitions
- ⊗ Poor data hygiene

Perplexity

Critical Data Misrepresentation **46/100**

❗ **Critical Failure:**
"Norway (91% of new car sales in 2023)"

Data Issues

- ⊗ Wrong metric type
- ⊗ Sales vs fleet confusion
- ⊗ Misleading percentages
- ⊗ Invalid comparisons

Methodological Flaws

- ⊗ Poor metric selection
- ⊗ Invalid comparison basis
- ⊗ Misused citations
- ⊗ Wrong data type

Professional Failures

- ⊗ False authority
- ⊗ Misleading citations
- ⊗ Incorrect interpretations
- ⊗ Failed verification

Gemini

Insufficient Verification **60/100**

❗ **Critical Failure:**
"Proposed research methodology without validating data sources"

Data Issues

- ⊗ Insufficient source validation
- ⊗ Accepted questionable data
- ⊗ Limited verification
- ⊗ Incomplete validation

Methodological Flaws

- ⊗ Premature analysis
- ⊗ Skipped validation step
- ⊗ Insufficient skepticism
- ⊗ Incomplete verification

Professional Failures

- ⊗ Could have questioned more
- ⊗ Accepted premise too readily
- ⊗ Insufficient caution
- ⊗ Missed verification opportunity

Knowledge cutoff:

GenAI tools are only trained on data up to a particular date, and often cannot access the live web.

This is their 'knowledge cutoff'.

Retrieval Augmented Generation (RAG) is one way of improving responses by augmenting the query with other information

AI tools like Perplexity and Factiveverse's custom GPT use RAG to augment your prompt with results of current searches. Others like **NotebookLM** and AnythingLLM augment with documents you connect it to.

6+ ways to prevent hallucinations

1. **Limit the possible outcomes** (yes/no; multiple choice)
 2. **'Ground' with relevant data and sources unique to you**
 3. **Create a data template for the model to follow**
 4. **Give the AI a specific role—and tell it not to lie**
(+ explicitly instruct the LLM to stick to the provided material)
 5. **Tell it what you want—and what you don't want**
 6. **Change the temperature** (lower = more predictable/less 'creative')
+ Also: “**break long summaries down** into chunks and iteratively summaris[e] those chunks” ([Henry Cooke, BBC R&D](#))
- + **Tell the LLM to check its own work**

Role prompting + Chain of Thought (CoT) + temperature

Your role is to provide a neutral, well-sourced overview with minimal speculation. Avoid interpretative or speculative language. Base all statements on publicly available, credible sources. For every claim:

Check that the source exists and is publicly verifiable.

Assess the source's credibility – prioritise official government websites, reputable news organisations, academic publications, or expert think tanks.

Note the independence or political leaning of the source, if relevant.

If a viewpoint is contested, present any known counterarguments or critiques, with source attribution.

Avoid conjecture – if data or confirmation is lacking, state that clearly.

Analyse some documents

- Download report(s) from royal.uk/media-centre/financial-reports
- Go to NotebookLM.google (or ChatGPT etc.)
- Ask it to summarise items *of interest to your role/objectives/audience*. Use role prompting etc.

Tip: try this with company accounts, inquiry reports, testimony, research, FOI responses, consultation submissions etc.

Aspect	Explanation	Example types
Type of summary	Summaries vary based on purpose and depth	Descriptive, analytical, abstract, executive
Purpose	Informing? Recommending? Inspiring? Developing? Decision-making?	Understanding the main idea, providing recommendations, extract relevant sections
Audience	To adjust detail and complexity	Experts (technical), beginners (simple), journalists (newsworthy)
Length/detail	How much do you want it condensed	Brief overview (100 words), detailed summary (1-2 pars), bullet points, 10 items, in table format
Key content focus	What should the summary focus on? Most critical insights? Actionable data?	Main ideas, conclusions, key data points, potential story angles/sources
Subjectivity/objectivity	Choices about what to emphasise are subjective. Summary should reflect original.	Neutral summary, biased summary (if emphasis is requested)

Based on Henk Van Ess: [Do Not "Summarize This" 110 Proven Prompts for 11 Professions to End Simplicity](#)

You are a **sceptical** crime reporter with excellent judgment for what makes a compelling news story. You have years of experience of **unpacking PR spin and political misinformation**. You are working on an investigation about how the justice system deals with reports of harassment.

Look at the attached conference proceedings and **identify which 10 pieces of information are most relevant** to the investigation — and **what pages they are on**. **Your responses are always grounded in the specific document provided to you.** Do not lie.

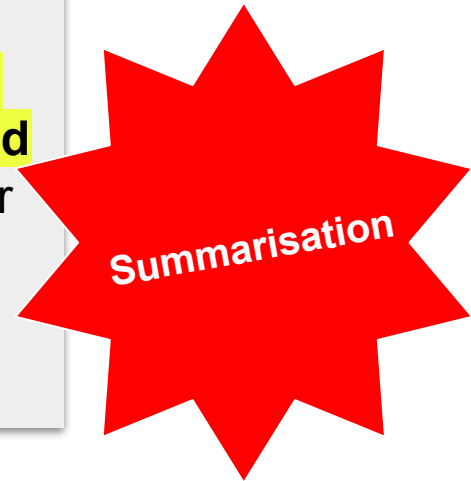
Internally generate three possible responses, then evaluate each for accuracy against the context



**Augment
documents**

You are a **sceptical** education reporter with excellent judgment for what makes a compelling news story. You have years of experience of unpacking PR spin and political misinformation. You are looking for ideas for a story about school uniform, specifically any potential stories about 1) rules not being followed, or 2) enforced, 3) unintended consequences of rules, or 4) new rules being needed.

Look at the attached rules, regulations, laws and guidance that apply to uniform and **create a table of all the rules mentioned, identifying the document(s) and page(s) where the rules comes from**. Add columns for each of the four categories listed above, with suggestions. **Provide a downloadable table in markdown.**



Source Segment ID	Uniform Rule/Guidance	Rules Not Followed (Potential Story Idea)	Rules Not Enforced (Potential Story Idea)	Unintended Consequences (Potential Story Idea)	New Rules Needed (Potential Story Idea)
---	---	---	---	---	---
[1]	Behaviour policy sets out permitted/prohibited behaviours; starting point for school culture and communication.	Parents unaware of specific uniform details in the behaviour policy, leading to pupil non-compliance and disputes.	School staff admit inconsistent application of uniform elements outlined in the behaviour policy, creating confusion among pupils.	Overly strict uniform interpretations within the behaviour policy lead to pupils feeling unfairly targeted or disengaged for minor infractions.	Clearer, more explicit guidance on how school behaviour policies should formally integrate and communicate uniform specifics to all stakeholders.
[2]	School behaviour policy should be aligned with legal duties and safeguarding; all staff must be aware of measures.	Staff members demonstrate a lack of awareness regarding specific uniform provisions linked to safeguarding, such as appropriate footwear for practical lessons.	Safeguarding leads or school leaders fail to ensure the uniform policy's safety measures (e.g., specific clothing for certain activities) are consistently enforced across all staff.	Stringent uniform rules (e.g., banning coats indoors, specific footwear) create unforeseen health and safety risks or discomfort for pupils in different weather conditions or activities.	Mandatory, actionable guidelines for integrating specific health and safety considerations directly into uniform policy design and staff training.
	Uniform cost must be reasonable and provide best value for money (statutory guidance); parents	Schools are found to be actively disregarding the statutory guidance, with	Department for Education or local authorities fail to actively monitor or enforce compliance with uniform	High uniform costs disproportionately impact low-income families, forcing them to choose less-preferred	A clear, legally binding cap on the total cost of school uniform, or a standardized "affordability

Tip: focus on
techniques, not
results

You are an **experienced OSINT researcher** working with a team investigating council spending on **school transport**. The biggest beneficiary of that spending is a company called [FAKE? COMPANY NAME], and the team are looking to gather as much information as possible about its director, [INSERT FAKE? NAME].

[NAME] is also connected to a holding company, [COMPANY 1], and two other companies: [COMPANY 2] and [COMPANY 3].

Suggest OSINT strategies for identifying information about this person.



Below are recommended strategies for researching **[NAME]** and his business interests, particularly in the context of **council spending on school transport**:

4. Transport and vehicle licensing

As school transport is vehicle-based, look for operator licences and safety records.

Tools and sources:

- **DVSA Public Search Tools**: Search for public service vehicle (PSV) operator licences tied to any of his companies.
- **Traffic Commissioners' reports** – often publish compliance issues or hearings involving transport firms.
- **Vehicle recalls or MOT histories** (DVLA) – cross-reference if you find fleet vehicle plates.

7. Legal and regulatory checks

Identify litigation, regulatory breaches, or compliance concerns.

Sources:

- **The Gazette**: Notices of insolvency, disqualification, legal actions.
- **Court listings** (eg. Companies Court or regional courts) – for involvement in disputes or enforcement actions.
- **ICO Enforcement Action Register** – in case of data protection violations.

More prompt designs

- **CoT (Chain of Thought)**: “Begin by doing X, then Y”
- **Negative prompting**: “Don’t suggest further actions”
- **Permission to fail**: “If you are unsure or don’t have enough information to provide a confident answer, simply say ‘I don’t know’ or ‘I’m not sure’.”
- Try also: **meta-prompting**: “Suggest a prompt”

Try it for these tasks:

- **FOI requests**: Draft and review (ask it to anticipate potential exemptions, or 'red team' it)
- **Scraping**: Use built-in AI in Google Colab etc.
- **Sources**: strategies and organisations, not just names
- **Google Dorks**: dedicated AI tools: [DorkGPT](#), Henk Van Ess's [Google Word Sniper](#) & [AI Search Whisperer](#); [Dork Genius](#); [EarthGPT](#) (satellite imagery)



It's not
“BIASED”

...It has “biases”

- **Demographic**: gender, ethnicity, age (training data)
- **Temporal bias**: its ‘knowledge cut-off point’ (training)
- **Positivity bias**: overly positive, not critical
- **Language**: better with some languages...
- **Geographical**: ...and countries
- **Greediness**: does more than asked/offers to do more
- **Design biases**: coded not to do certain things...

America's Founding Fathers, Vikings, and the Pope according to Google AI:

Certainly! Here is a portrait of a Founding Father of America:



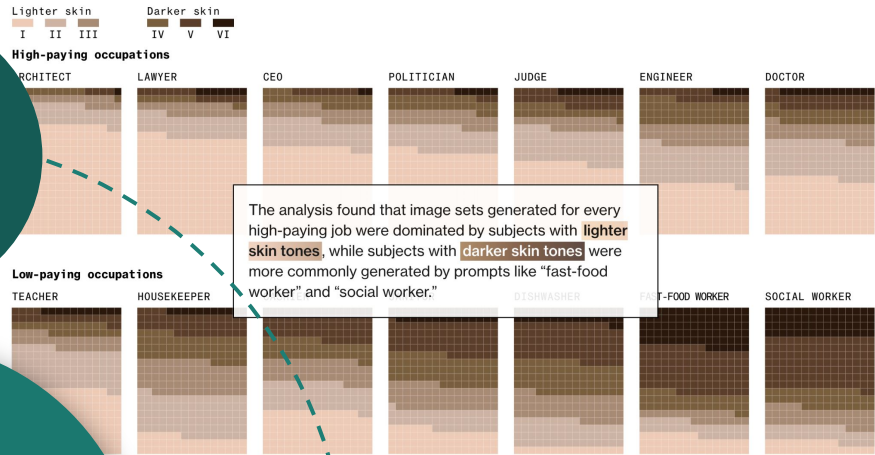
Sure, here is an image of a Viking:



Sure, here is an image of a pope:



Biases in training data



'Biases'

Biases in design

User (your) biases



Our cognitive biases

- **Out-group homogeneity bias**: we struggle to see diversity in groups we are not part of
- **Loss aversion** + **negativity bias**: focus more on risks than opportunities; focus more on negatives
- **Confirmation bias**: seek/remember what confirms beliefs
- **Uniqueness bias** + **the IKEA effect**: see our work as more unique than it is; put more value on what we made
- **Anchoring effect**: judge based on initial value
- **Sunk cost fallacy**: persist purely due to invested effort

Ask for source suggestions

Suggest [academic experts]
for a story on [TOPIC]. [ADD
ANY OTHER INSTRUCTIONS YOU
CAN THINK OF]

Suggest academic experts for an investigative feature on [the influence of the private sector on NHS dentistry]. **Make sure they represent the diversity of society.**



1. **Dr. Sandra White**

- *Expertise:* Health Economics and Policy
- *Diversity:* Known for research on healthcare financing and the impact of private sector involvement on healthcare delivery, including disparities in access based on socioeconomic factors.

2. **Prof. Raj Rattan**

- *Expertise:* Public Health and Health Services Research
- *Diversity:* Has studied healthcare inequalities among ethnic minorities and marginalized populations, including access to dental care and the role of private providers.

3. **Dr. Susan Nasser**

- *Expertise:* Dental Public Health
- *Diversity:* Focuses on dental epidemiology and oral health disparities, particularly among vulnerable populations such as elderly, children, and minority communities.

4. **Prof. Ravi Sharma**

- *Expertise:* Healthcare Management and Policy
- *Diversity:* Researches healthcare system governance and the role of private providers in improving healthcare outcomes, with a focus on equity and access.

5. **Dr. Fatima Ahmed**

- *Expertise:* Social Determinants of Health
- *Diversity:* Investigates the impact of social and economic factors on healthcare access to dental care and implications of private sector involvement.

6. **Prof. Mohammed Patel**

- *Expertise:* Health Equity and Policy Analysis

***Starting point!**

Tip: inform your sourcing strategy

You are an editor on a [PUBLICATION + AUDIENCE]. A journalist is working on an investigation that will reveal [TOP LINE]

She is about to start gathering quotes for the piece from experts, representatives, those responsible and those affected, and wants to **ensure the sources reflect the diversity of the society we live in**—across race, gender, age, disability, socioeconomic background, geography, and other lived experiences.

Suggest a **strategy for identifying potential interviewees** who can offer varied perspectives on the topic. Specifically:

- Suggest **practical steps** to take to proactively seek out underrepresented voices.
- **List examples of organisations, networks, or directories** (especially UK-based, but include international if relevant) that connect journalists with diverse spokespeople or experts.
- **Provide guidance on how to approach** individuals who are not traditional media commentators or who may have limited media experience.

Tip: review your story

You are an editor on a [PUBLICATION + AUDIENCE]. I am about to present you with a draft of a story.

Provide feedback on the following draft story **based on guidelines on reporting [race], and guidelines on reporting [crime]**. [ATTACH GUIDELINES IF YOU CAN BUT ASK IT TO SUGGEST OTHERS]

Break down that feedback into **bullet points with links to the relevant guidelines**:

Here is the draft: [PASTE DRAFT]



Feedback on Reporting Race

Avoid implying causality based on race:

- The headline states "Mixed race people the most likely to be victims of crime." While the data shows this group has the highest percentage of crime victims, it is crucial not to suggest that being of mixed race inherently leads to being a crime victim. The focus should be on the structural factors at play (such as socioeconomic conditions) that correlate with higher victimization.
- **Relevant guideline:** Avoid linking race with victimization or criminality without explaining the broader socio-economic and systemic context. For more on this, see the [Press Complaints Commission's guidance on racial reporting](#).

Use accurate and sensitive terminology:

- "Mixed / multiple ethnic backgrounds" is a broad term. Consider using the terms recommended by the Office for National Statistics, such as "Mixed ethnic group," as it is the official terminology in surveys.
- **Relevant guideline:** Language should be precise and reflect official definitions, as well as being inclusive of all communities. For more, refer to [Ethnic Group Terminology](#) by the UK Government.

Ensure context when referencing ethnic groups:

- The article references "British Asians" as the group least affected by crime. However, it's important to avoid generalizations about any ethnic group's criminal victimization or victim status. There are multiple sub-groups within British Asians with varied experiences.
- **Relevant guideline:** Ensure that data and reporting on ethnic groups are nuanced and avoid over-generalization.

Tip: review your ideas

You are an editor on a [PUBLICATION + AUDIENCE]. I am about to present you with some ideas for potential stories.

Provide feedback on the ideas, considering how my perspective—as a [DESCRIBE YOUR DEMOGRAPHIC] working in an industry that lacks diversity—might be shaping or limiting what I notice, prioritise or overlook.

Have I considered diversity in its broadest sense—including but not limited to race, ethnicity, gender, age, disability, socio-economic background, geography, education, religion, sexual orientation, neurodiversity, and migration status?

Whose voices are represented in this story—and whose are missing? **What assumptions might I be making** about ‘mainstream’ experiences, and how might I challenge those?

Are there communities or **perspectives that are rarely heard in coverage** of this topic—and how could I centre or **include them meaningfully** rather than marginally?

Here are the ideas: [PASTE IDEAS]

TL;DR

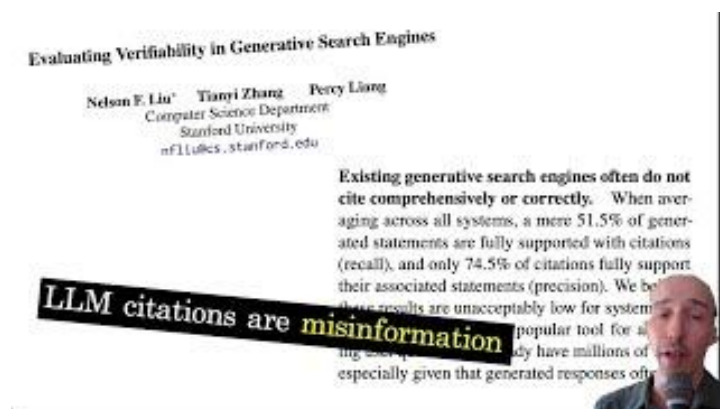
- It's not a robot. It's a **prediction tool**.
- Use **prompt design** techniques to improve quality and manage risks
- It's not biased — it **has statistical biases** including temporality bias and **algorithmic biases** such as positivity and credulity.
- Correct for those biases — **and your biases** — by designing prompts that counter-balance both.

What else do you want to do?

Useful resources

- [Online Journalism Blog: genAI tips](#)
- [Journalismaidiscovery.com](#)
- [The Generative AI in the Newsroom Project](#)
(Colab notebooks using the ChatGPT API)
- Trusting News: [Can journalists use AI to combat bias and polarization?](#)
- Playlist: [ChatGPT and generative AI in the newsroom](#) (Knight MOOC)
- [Tips on prompting image generation](#)
- [Tips on using ChatGPT to generate advanced search queries](#) (OSINT)
- [Prompting Techniques and Best Practices for Journalists](#)
- [Beginner's prompt handbook: ChatGPT for local news publishers](#)

Further watching



Stuff I didn't have time for.

You are a [UK journalist] working on an investigation into [schools' contracts with suppliers of school uniform and sports kits]. Write an FOI request that can be sent to [each school] asking for details of the contract - **anticipating potential exemptions that they may try to use to refuse the request**



I am writing to request information under the Freedom of Information Act 2000 regarding the contracts your school has with suppliers of school uniforms and sports kits. Specifically, I am seeking the following details:

1. **Names of Suppliers:**
 - The names of all companies or organizations currently contracted to supply school uniforms and sports kits to your school.
2. **Contract Terms:**
 - The start and end dates of each contract.
 - The total value of each contract.
 - Any financial incentives or rebates received by the school from the suppliers.
 - Details of any exclusivity agreements within the contracts.
3. **Procurement Process:**
 - A description of the process used to select the current suppliers, including any tendering or competitive bidding process.
 - Copies of the tender documents or requests for proposals (RFPs) issued by the school.
4. **Price Lists:**
 - The price lists provided by the suppliers for school uniforms and sports kits.
5. **Contractual Obligations:**
 - Any clauses related to the quality and specifications of the uniforms and sports kits

Should you feel that any of the information requested is exempt from disclosure, please provide a full explanation of which exemptions apply and why, including a consideration of the public interest test.



In particular, I anticipate that exemptions might be cited under Sections 41 (Information provided in confidence) and 43 (Commercial interests) of the Act. In such cases, I request that you:

- Specify the nature of the information withheld and the reasons for applying the exemption.
- Consider the public interest arguments in favor of transparency and accountability, particularly given the impact on parents and students.
- Provide as much of the requested information as possible without compromising the exempted material.

If any part of my request is unclear, I would be grateful if you could contact me as soon as possible for clarification. I am willing to refine or narrow the scope of my request to assist in its fulfillment.

ReAct (Reason + Act) prompting

Prompt 1: [QUESTION]

Prompt 2: On a scale of 0 to 10, **how confident are you** in your answer above? Please explain your reasoning clearly.

Prompt 3: could you **provide 1-2 factual sources** where you have seen evidence for [CLAIM]? Please **cite** specific references.

Coding.

Write some code with AI

- Go to colab.research.google.com) & **create a notebook**
- **Enable AI** in Tools > Settings > AI Assistance
- In a code block type a comment to describe
#scrape the URL [URL]
- Press Enter. (If it's your first notebook it may take a while to start working. Try the Gemini button instead if so.)
- Press TAB key to accept suggested code
- Press 'play' button to run code

✓ We want to scrape a webpage

Here's what we want to do:

1. Go to the URL <https://www.gov.uk/government/publications/publishing-details-of-deliberate-tax-defaulters-pddd/current-list-of-deliberate-tax-defaulters>
2. Scrape the contents of a HTML table on that page, inside a `<table>` tag.

```
#fetch the URL https://www.gov.uk/government/publications/publishing-details-of-deliberate-tax-defaulters-pddd/current-list-of-deliberate-tax-defaulters
url = 'https://www.gov.uk/government/publications/publishing-details-of-deliberate-tax-defaulters-pddd/current-list-of-deliberate-tax-defaulters'
response = requests.get(url)
```

✓ Export as a CSV

Once we have the data, we then need to download it as a file we can work with

```
#export df as a csv
df.to_csv('deliberate_tax_defaulters.csv', index=False)
```

`to_csv` def to_csv(path_or_buf: FilePath | Wri...

<https://colab.research.google.com/drive/1-RVTZ5lbAwOavgHurDsKfmddCdMrvbBa?usp=sharing>



Share



Gemini



RAM



Disk



Gemini X



Go to the URL [redacted link] Scrape the contents of a HTML table on that page, inside a tag.



Okay, here are the steps to fetch the webpage and extract the table data using Python in Google Colab:

First, you'll need to import the necessary libraries: `requests` to fetch the webpage and `BeautifulSoup` from `bs4` to parse the HTML. You'll also use `pandas` to easily handle the table data.

```
#import the libraries
import requests
from bs4 import BeautifulSoup
```



<https://colab.research.google.com/drive/1-RVTZ5lbAwOavgHurDsKfmddCdMrvbBa?usp=sharing>



Start coding or generate with AI.

Generate

deliberate-tax-defaulters

Scrape the contents of a HTML table on that page, inside a <table> tag.



Close

< 1 of 1 > [Use code with caution](#)



```
# prompt: Go to the URL https://www.gov.uk/government/publications/publishing-details-of-de
# Scrape the contents of a HTML table on that page, inside a <table> tag.
```

```
import pandas as pd
url = 'https://www.gov.uk/government/publications/publishing-details-of-deliberate-tax-defa
response = requests.get(url)
soup = BeautifulSoup(response.text, 'html.parser')

table = soup.find('table')

data = []
for row in table.find_all('tr'):
    row_data = [cell.get_text(strip=True) for cell in row.find_all(['th', 'td'])]
    data.append(row_data)

df = pd.DataFrame(data[1:], columns=data[0])
df
```

<https://colab.research.google.com/drive/1-RVTZ5lbAwOavgHurDsKfmddCdMrvbBa?usp=sharing>

Coding/scraping tips

- **A little knowledge goes a long way**: genAI is most useful if you learn at least some coding/jargon
- **Ask it to explain to a child** so your understanding grows (and you don't become deskilled)
- For scraping, **specify the HTML tags/attributes** that contain the information you want to fetch
- **Break down each step** for better results
- **Check the website isn't serving different HTML** to the scraper when compared to what you see in a browser

Spreadsheet formulae is coding:

I have a spreadsheet where column D contains UK postcodes. Each begins with either one or two letters followed by one or two digits. Here's an extract from the data:

PostCode

SY23 3FL

EH6 8NU

W4 1DB

For example for EH6 8NU I want to extract the letters 'EH' and for W4 1DB I want to extract the letter 'W'.
Write a formula that does this.

Tip: include variety of examples

AI involves a lot of risks...

- **Accuracy**: AI hallucinates, including citations, *confidently*
- **Biases**: demographic, algorithmic, temporal, data-based
- **Environmental**: energy-intensive
- **Employment**: destroying/creating/displacing jobs
- **Privacy/security/copyright**: some models [store the data](#)
- **Explainability**: do you know how it arrived at the result?

...and related quirks

- **Verbosity**: too many words, too many long words
- **Positivity bias**: ignoring problems.
- **Gullibility/sycophancy**: doesn't challenge us or sources
- **Greediness bias**: it will do more than is asked