Instructions

Overall guidelines

General guidelines

- You will be asked several questions that evaluate a type of similarity between two articles on this scale: Very Dissimilar (VD), Somewhat Dissimilar (SD), Somewhat Similar (SS), Very Similar (VS).
- First, there are four sub-aspects: Geography (GEO), Named Entities (ENT), Time (TIME), Narrative Schemas (NAR). Each question evaluates similarity in one aspect and must be answered *independently*. E.g., your reasoning for the similarity of NARrative Schemas should not depend on the GEOgraphy covered being similar.
- The fifth, main question considers the Overall Story (OVERALL) similarity. It encompasses the previous four aspects and everything else that is not covered by these aspects. It must ignore the style and tone of articles.
- The sixth question provides *auxiliary* information about style and tone. Important: style and tone shall NOT play a role in any of the first five questions.

- When evaluating the given aspect it is crucial to determine the *primary focuses* of the articles and distinguish them from any *secondary focus*es (if present). In most cases, the similarity between the **primary focuses** of each article shall determine the similarity label.

  Sometimes the name of a location or named entity is actually not written explicitly in the articles, even though the articles are clearly about the same locations or named entities. They need to be still considered as such in these cases. E.g., the articles do not mention India explicitly, but the news outlet is clearly India-based and mentions Indian cultural events happening all over the country without specifically saying its name. Or the articles are about a citizen who was the first covid victim in a country, without specifying their identity, yet they clearly are a named entity (it's just that their name was not explicitly mentioned).

The main way to decide similarity labels for a given aspect

- Very Similar
  - The articles are primarily focused on the same or very similar instances of the given aspect. For example: they both focus on the same or closely related people (Very Similar Named Entities) or on identical or overlapping/close locations (Very Similar Geolocation).
- Somewhat Similar
  - The articles are focused on significantly overlapping or related instances of the given aspect. This includes the case where overlapping locations are in primary focus in one article and in a secondary focus in the other article, e.g., one article focuses on the USA and briefly mentions NYC, while the other focuses on New York City and New York State and mentions other states (Somewhat Similar Geography aspect).
- Somewhat Dissimilar
  - The articles are focused on *partially* overlapping or related instances representing the given aspect, e.g., one article focuses on the British queen, while the other focuses on the British prime minister (Somewhat Dissimilar Named Entities aspect) and neither mentions the other article's name entities.
- Very Dissimilar
  - No or minimal overlap between instances of the given aspect. For instance, one article is about Sweden, while the other is about Spain, and the common territory of the EU is not evoked as a substantive focus; or Poland and Senegal are covered respectively, without a common geographic reference (Very Dissimilar Geography aspect)

Tie-breaker rules

- If primary focus does not determine the similarity label for a given aspect, e.g., you are unsure whether to assign the Somewhat Similar or Somewhat Dissimilar label, then you should consider the secondary focus(es) of each article. To decide the final label, you can compare and evaluate in your mind both primary and secondary focuses of articles A and B, namely:
  - primary focus of A with primary focus of B (most important),

- ○ primary focus of A with secondary focus of B (medium important),
  - ○ secondary focus of A with primary focus of B (medium important),
  - ○ secondary focus of A with secondary focus of B (least important).
- In cases where the label is not clear from the comparison of primary focuses alone, a label can be determined by averaging the above four similarities, while giving the most importance to the similarities that involve primary focuses.

Rules of thumb

- If it is difficult to judge primary and secondary focuses or you wonder what the thresholds are on overlaps between sets of entities or locations or narratives, then you can use the following rules of thumb. These rules, however, shall not replace the above considerations and can be used only as an approximate guideline:
  - ○ "Very Similar" - At least *half of the pertinent content* of each article overlaps with the other article.
  - ○ "Somewhat Similar" - At least a *paragraph of pertinent content* overlaps. In addition, if one of the articles is a direct follow-up to the specific other article, then the two articles are usually at least Somewhat Similar.
  - ○ "Somewhat Dissimilar" - The topic is similar and at least a *few sentences of pertinent content* in the two articles contain the same information.
  - ○ "Very Dissimilar" - Almost no overlap
- **Pertinent content** is defined as all content directly or indirectly referring to the dimension in question, i.e., all text that can be used to infer something about locations if we consider GEO, or all mentions of entities if we consider ENT, or all components of narrative schemas if we consider NAR, or all text if we consider OVERALL. "Overlaps" == "says the same thing".
- If some things are mentioned more often than others, then this needs to be taken into account when estimating the overlaps. For instance, Article 1 mentions **A**, A, A, **B, C**, while Article 2 mentions **A, B, C**, D, E, F. Then the overlap w.r.t. A1 is 3/5 = 60%, while the overlap w.r.t. A2 is 3/6 = 50%, so they are Very Similar.

Side notes

- Contents of videos, advertisements, or other audiovisual sources are not to be coded. We label only textual content of the articles, including tables and infographics/figures with text.
- If you notice any issue or weirdness in the text of the following instructions and examples, e.g., a wrong link to a news article, please comment it directly in this Google Doc.
- You may find that some articles are behind reader registration/subscription paywalls. The following steps usually help in such cases:
  - You could try opening the article in a Private/Incognito browser window and/or to remove cookies
  - It may also be convenient to install a plugin that is able to bypass such obstacles. You can install Bypass Paywall:

- - as a Chrome extension: Go to https://github.com/iamadamdev/bypass-paywalls-chrome and follow the instructions (you can also unpack the ZIP first and then click on "install from unpacked folder" on the Chrome developer extension page
  - if you use Firefox via this link/file (also on the github site)
  - Note that only specific sites are supported: https://github.com/iamadamdev/bypass-paywalls-chrome#bypass-the-following-sites-paywalls-with-this-extension
  - Combining Bypass with an Ad Blocker works even better. uBlock Origin seems to be recommended by the Bypass creator
- If an article has multiple pages, please only code the first page.
- We welcome your feedback to the annotation interface and questions. If you have some suggestions, e.g., for a modification in the interface or a new question/feature, please let us know via Slack.

The Comment/Feedback textbox

- Apart from the four questions, you will have an option of providing a free-text comment to a given pair of news articles. **When writing comments, please use the following four main keywords: LANG, URL, PAYWALL, NOT_RELEVANT, MANY**:
  - Most important (it's as important to use them as to answer the similarity questions)
    - A pair links to the same text on two different websites (possibly with minimal modifications to the layout or text) — please annotate all aspects as "Other" and use the keyword DUPLICATE.
    - You see again the same or nearly the same pair of articles as a pair that you labeled before — please please annotate all aspects as "Other" and use the keyword REPEAT or REPEATED.
    - If both the original and the Internet Archive links for a given article do not lead to an article of societal or political relevance, e.g., they lead instead to a unrelated question on a forum or an unrelated social media post or an article about sports or entertainment, then please annotate all aspects as "Other" and leave a comment in the text box starting with the word "NOT_RELEVANT article 1" or "NOT_RELEVANT article 2".
  - Less important (please use them, but currently they have lower priority -- this may change in the future)
    - If a link leads to a page that is unavailable, please try again. If this does not help, try the Internet Archive link that is placed below the original link. If the issue persists, please annotate all aspects as "Other" and leave a comment in the text box starting with the words "URL article 1" or "URL article 2".
    - If one of the articles is behind a paywall, then please annotate all aspects as "Other" and write "PAYWALL article 1" or "PAYWALL article 2" in the comment.
    - If one of the articles is in a language that you do not understand, please annotate all aspects as "Other" and leave a comment reporting the issue starting with the words "LANG article 1" or "LANG article 2".
    - If a URL leads to what seems like two distinct articles with distinct headlines, please try opening it again. If it does not help, please check the Internet Archive link. If the issue persists, only code the *first one*

and leave a comment reporting the issue starting with the word "MANY".
- ■ If one article is a translation from another one: "TRANS"
- **If you want to discuss a given pair, please mark the checkbox "Flag survey item for further discussion"**. In such cases, we encourage you to use the following template as your comment and write your comments after the capital letters representing each aspect. Writing comments is very helpful when you are uncertain about the labels and you'd like to discuss it later with someone. The discussions of pairs will happen typically after a few days since you labeled the pair, so your comment will help you in recalling your thought process for the given pair. **Leaving such detailed comments is particularly important for difficult pairs and pairs which you labeled very differently than other labelers.**

```
GEO: Please write your explanation for the GEO aspect here, e.g., "A1
is all Queensland, while A2 is Queensland and several other  states,
but Queensland appears as a primary state"
ENT:
TIME:
NAR:
OVERALL:
STYLE:
TONE:
```

Question-specific guidelines and examples

GEO: How similar is the geographic focus (places, cities, countries, etc.) of the two articles?

Notes:

- GEO similarity can be judged by (a combination of)
    - a. the fact that locations are placed in the same (administrative, cultural, etc.) areas. E.g., two cities in the same federal state, country, continental area (like Western Europe)
    - b. the intersection of areas (e.g., the EU vs. Italy)
    - c. the set overlap of the same places mentioned in both articles (places A,**B,C** in article 1, places **B,C**,D,F,G in article 2). If you are unsure about the thresholds on overlaps between sets of locations, please use the aforementioned rules of thumb as an approximate guideline.
    - d. ...while weighing the locations that are in (primary, then secondary) focus more

- Try to *detach from your personal life context* and take a "global look" at the similarity. E.g., you live in the UK; so you know UK cities well and articles on two different cities may seem very different to you. But: they are both in the UK, so from the global perspective they may be Somewhat Dissimilar or Somewhat Similar, depending on the overall focus of the article. If both articles focus primarily on the UK as a nation,

that might push the rating further towards Very Similar, as now, the same geographic instance of the whole of the UK is more in focus in both.

- The place of *publishing* (as often stated at the beginning of articles or inferred from a newspaper name) is irrelevant by itself. However, looking at this place and/or the regional focus of the medium can sometimes help to decode references such as "in our state", "the city", etc.
- If something with the name of a location is described/acts like a person ("Washington intervened", "London responded") it can still indicate the location should also be considered as a named entity in the next question.
- Do not code fictional locations, just infer the "real world" location
- If there is no clear location in one or both articles, then you should code GEO as "Other". In such a case, please note in the comment with the "GEO:" prefix what was the case (both or one article without GEO).

Real-world examples:

Pair	Article 1	Article 2	Geographic similarity
S1, L8	National guard helping covid efforts in several US states. No focus on one state, except short focus on pentagon Direct | I-Archive	US Army converting buildings for covid use, mainly in NY state, with additional reference to other states Direct | I-Archive	Somewhat Similar A: the primary focus is on the US as a whole talking about multiple states with a brief mention of New York City B: the primary focus is on New York state, with a secondary focus on other states
S2, L13	Stock market troubles with focus on US stock markets and measures Direct | I-Archive	Central bank regulatory measures in UK, US, EU Direct | I-Archive	Somewhat Similar/Dissimilar A: the primary focus is the US, a marginal secondary focus is global. B: the primary focus is global, the secondary focus is the UK, US and EU. This pair is borderline Somewhat Similar/Dissimilar, since the primary focuses are different. Note: the Overlap of the publishing medium, TIME and main ENT (El-Erian) *must not* play any role in determining this dimension.
S1, L15	Cheerful Covid masks in Bulgaria Direct |	Temporary lockdown in Jamaica Direct |	Very dissimilar No geographic overlap at all

	I-Archive	I-Archive	
S2, L10	Electric bills in Karachi Province, Pakistan Direct | I-Archive	Lockdown in Karachi Province, Pakistan Direct | I-Archive	Very Similar Both articles focus on the province first and foremost, with minor mentions of federal topics
S2, L5	Prime minister of the UK gets covid (London, UK) Direct | I-Archive	Youngest victim of Covid in the UK (Bristol, UK) Direct | I-Archive	Somewhat Similar Different cities in the same country Both evoke the frame of the UK as a whole, but with focus on different regional contexts
S2, L12	Economic relief in India (as in: the whole nation) due to covid Direct | I-Archive	Lockdown in India (as in: the whole nation) Direct | I-Archive	Very Similar, whole nation of India is referenced in both
S1, L12	Covid testing in/around city of Chandigarh, India (only local frame) Direct | I-Archive	Community transmission in India as a whole (no regional/local focus on Chandigarh) Direct | I-Archive	Somewhat Dissimilar India is the common location, but is not referenced as the whole nation/country in both, A1 doesn't reference it at all
S2, L16	Chinese firm donates something to Serbia Direct | I-Archive	Serbia-based expert forum comments on "global vaccine race" Direct | I-Archive	Somewhat Similar A1: The main focus is Serbia and China. A2:There are 3 of the 8 paragraphs mentioning Serbia, while the main focus is global.

Summarized explanations of labels for the above examples using the tie-breaker rules:

Pair	Similarity between Primary and Secondary focuses of the two articles				Label
	PrimaryA vs PrimaryB	PrimaryA vs SecondaryB	SecondaryA vs PrimaryB	SecondaryA vs SecondaryB	

S1, L8	SS (US vs NY state)	VS (US)	SD/SS (NYC vs US)	SD/SS (NYC vs US)	SS
S2, L13	SS/SD (US vs global)	VD (US vs EU/UK)	VS (global vs global)	SS (global vs EU/UK/US)	SS
S2, L5	VS (UK vs UK)	SS (UK vs a city in the UK)	SS (UK vs a city in the UK)	SD (two different cities in the UK)	SS
S1, L12	SD/SS (Chandigarh vs India)	- (no Secondary focuses)	- (no Secondary focuses)	- (no Secondary focuses)	SD
S2, L16	VD/SD (Serbia/China vs Global)	VS (Serbia/China vs Serbia)	- (A1 does not have Secondary focus)	- (A1 does not have Secondary focus)	SS

ENT: How similar are the named entities (e.g., specific people, companies, organizations, products, living beings), <u>excluding previously considered locations</u> appearing in the two articles?

Notes:

- Named entities are specific entities that typically have a proper name. E.g. "a dog" or "a house" are not named entities, while "Lassie" or "The White House" are. A specific person will nearly always have a unique name, but that name may not be mentioned in the article or may appear in non-explicit ways, e.g., the president of the United States instead of Donald Trump.
- Two different named entities can be related even if they are not the same, e.g., the prime minister and the queen of the UK are members of the UK government. I.e., even if the entities in both articles do not overlap in terms of exact matches, if all entities in article 1 are closely semantically related to those in article 2, this allows for an SS or SD rating, for example.
- If you wonder about the thresholds on similarity between sets of entities (by semantic similarity or set overlap), use the aforementioned rules of thumb.
- Sometimes the name of an entity is actually not written explicitly in the articles, even though the articles are clearly about the same named entities. They need to be still considered as such in these cases. E.g., articles are about a citizen who was the first covid victim in a country, without specifying their identity, yet they clearly are a named entity (it's just that their name was not explicitly mentioned).
- The news outlet publishing the two articles should not be considered, unless it is mentioned explicitly as part of a sentence in an article and is relevant to it, e.g., an article is focused on describing views of a Fox News collaborator.
- The "excluding previously considered locations" rule is mainly meant to avoid double-consideration of locations (that often have proper names) in the GEO and

ENT questions.

However, deviating from this rule, sometimes named entities could be used as both locations or actors. For instance, in the sentence "St. Mary Hospital began vaccinating against covid" the hospital is acting both as an organization and location, but in a phrase "next to St. Mary Hospital" it is only a location. Only in the former case, the named entity shall be considered both in GEO and ENT questions. In the latter case, it shall be considered only in the GEO question. Similarly, if a city or a country is mentioned as an actor, e.g., "the city of London has decided…" or "Germany has to react", then it acts as an organization, since it refers to the governmental body or citizens, and it typically also acts as a location. In these special cases, it shall be considered as both ENT and GEO.

- Also, a story could be about conflicts in the middle east and mention "The intervention of Washington and Paris in the Gaza strip did not help" or likewise. Then the GEO focus would be completely on the Middle East/Gaza, not the US and France, since they refer only to governments.
- Likewise, "the U.S. Embassy in Jerusalem" can act as an organization, but in a given article it may serve as a location only. Note that as a location, it is related to Israel and not to the USA. As an organization it is related to both the USA and Israel.

Examples:

Pair	Article 1	Article 2	Entity similarity
S1, L8	National guard helping covid efforts in several US states. Direct | I-Archive	US Army converting buildings for covid use Direct | I-Archive	Somewhat similar Shared entities: Covid, National Guard, Esper, Pentagon, Trump, Several other entities are dissimilar A2 has a focus on the army corps of engineers, and their director, as well as NYC-based ENTs. It mentions the National Guard in a few places. A1 has a focus on the N.G. and their head, plus a much stronger focus on the pentagon and Esper. It mentioned interviewers and some ships. In sum, most entities don't overlap, except for the N.G. However, all entities are closely related. Thus, this is a weak SS (borderline SS/SD)
S1, L15	Cheerful Covid masks in Bulgaria Direct | I-Archive	Temporary lockdown in Jamaica Direct | I-Archive	Very dissimilar Covid is shared, but a few other ENTs are mentioned. Each of these articles focuses on much more specific topics than just Covid and named entities appear related to these specific focuses. A1 mentions Angel Baby and some people, while A2 mentions Jamaican authorities (without writing them in Capital Letters, but it mentions them nevertheless). Covid is a generic topic.

			That's why this is a weak VD (a borderline VD/SD).
S2, L10	Electric bills in Karachi Province, Pakistan Direct | I-Archive	Lockdown in Karachi Province, Pakistan Direct | I-Archive	Somewhat similar, Somewhat dissimilar Shared entities: Covid, Sui Southern Gas Company, K-Electric,Murad Ali Shah A2: DG Rangers Sindh , IG Police, Armed Forces, Criminal Procedure Code, Constitution, QESCO, SEPCO
S2, L11	German authorities say man with virus in critical condition Direct | I-Archive	Authorities fear German tourist brought Covid-19 to Netherlands Direct | I-Archive	Very Similar The two articles focus on the same German citizen, although they don't mention his/her name. Other shared ENTs include: the hospital in Erkelenz, German authorities, German news agencies
S2, L13	Stock market troubles US stock markets and measures Direct | I-Archive	Central bank regulatory measures in UK, US, EU Direct | I-Archive	Somewhat similar Shared: El-Erian, Covid,FOX Business, Allianz Global Investors A1: The Stock Market, S&P, Liz Claman, ETFs, Troubled Asset Relief Program, George W. Bush A2: Maria Bartiromo, Bank of England, Treasury, Washington (as in government), European Central Bank, Christine Lagarde, 2008 Great Financial Crisis, Bloomberg, Federal Reserve
S2, L12	Economic relief in India Direct | I-Archive	Lockdown in India Direct | I-Archive	Somewhat Dissimilar Shared: Modi, Covid, while mentioning a set of other ENTs A1: Independence, RBI, Chicago Booth, Raghuram Rajan, LinkedIn,, Shamika Ravi, Brookings India, IMF A2: Prime Minister, Centre and State Governments
S1, L12	Covid testing in/around city of Chandigarh, India Direct | I-Archive	Community transmission in India as a whole Direct | I-Archive	Somewhat similar Covid19, Indian Council of Medical Research, ILI, several non-shared ENTs

Summarized explanations of labels for the above examples using the tie-breaker rules:

Pair	Similarity between Primary and Secondary focuses of the two articles				Label
	PrimaryA vs PrimaryB	PrimaryA vs SecondaryB	SecondaryA vs PrimaryB	SecondaryA vs SecondaryB	
S1, L15	SD/VD (Bulgarian citizens vs Jamaican government, covid)	VD	VD	VD	VD
S2, L13	SS (El-Erian, Covid, S&P vs Central Banks)	VD	VD	SD (Fox, Allianz, but everything else different)	SS
S2, L5	SD (British PM vs a British child, covid)	VD	VD	VD	SD
S2, L12	SD (Rajan vs Modi, covid)	SD (indian institutions)	SD (indian institutions)	SD (indian institutions)	SD
S1, L12	SS (ICMR, covid, UT administration not shared)	SS (ICMR vs MoHFW and symptoms)	SD (institutions of Chandigarh vs ICMR)	SD (institutions of Chandigarh vs MoHFW and symptoms)	SS

TIME: Are the two articles relevant to similar time periods or describing similar time periods?

- Note that some articles may focus on longer time periods, e.g., the World Cup, while others may focus on a particular point in time, e.g., the date of a particular soccer match.
- If the news article focuses on specific time periods or dates explicitly, then these times should be considered, as opposed to the time of article's publication. If the news article focuses on recent events without mentioning their specific times, then one can use the publication date of the article as a proxy for the time period of focus.
- We provide the following simplistic rules for deciding time similarity **for events with a clear time point**:
  - Very Similar - The events and time periods described in the two articles happen within a period shorter than a week
  - Somewhat Similar - The events happen within a time period of two weeks
  - Somewhat Dissimilar - The events happen within 6 weeks

- ○ Very Dissimilar - The events happen within a time period of more than 6 weeks
- If several time points of equal importance are mentioned, one can average between them to determine the overall time point.
- To judge if **covered periods** are similar (e.g. "Euro Cup" vs. "the summer of 2021"):
  - ○ Very Similar - There is an overlap between the periods of at least ½ of each period.
  - ○ Somewhat Similar - There is an overlap between the periods of at least ¼ of each period.
  - ○ Somewhat Dissimilar - The two periods must either be consecutive, e.g., January 2020 and February 2020, or they need to have a non-zero overlap of less than ¼ of each period .
  - ○ Very Dissimilar - There is no overlap between both periods and they are not consecutive.
- If one article focuses on a specific date (like a concrete World Cup match) in the first half, but then also speaks about the World Cup in general at a significant length in the second half, it has a focus on a specific time point as well as a period. In this case, if the second article only focuses on the entire World Cup, they are only somewhat similar.
- If a time (period) is only mentioned as a side note, e.g. "since World War 2", but with no further focus on or mention of it, it can be ignored.

Pair	Article 1	Article 2	Temporal similarity
S1, L8	National guard helping covid efforts in several US states (March 19, 2020). Direct | I-Archive	US Army converting buildings for covid use (March 21, 2020) Direct | I-Archive	Very similar The two events happened within 2 days.
S1, L15	Cheerful Covid masks in Bulgaria (March 16, 2020) Direct | I-Archive	Temporary lockdown in Jamaica (March 30, 2020) Direct | I-Archive	Somewhat similar The two events happened within 2 weeks.
S2, L10	Electric bills in Karachi Province, Pakistan (March 23, 2020) Direct |	Lockdown in Karachi Province, Pakistan (March 22, 2020) Direct |	Very similar The two events happened within 1 day.

	[I-Archive](#)	[I-Archive](#)	
S2, L13	Stock market troubles US stock markets and measures (March 13, 2020) [Direct](#) | [I-Archive](#)	Central bank regulatory measures in UK, US, EU (March 11, 2020) [Direct](#) | [I-Archive](#)	Very similar The two events (interviews) happened within 2 days. The fact that one article talks to a small extent about the financial crisis in 2008 is neglectable.
S2, L5	Prime minister of the UK gets covid [Direct](#) | [I-Archive](#)	Youngest victim of Covid in the UK [Direct](#) | [I-Archive](#)	Very dissimilar The two events happened within almost exactly two months.
S2, L12	Economic relief in India due to covid (April 5, 2020) [Direct](#) | [I-Archive](#)	Lockdown in India (March 24, 2020) [Direct](#) | [I-Archive](#)	Somewhat similar The main event of A1 is the blog post, for A2 the Modi address, both of which seem to have occurred immediately before the articles were published. The two events happened within 2 weeks.
S1, L12	Adoption of covid testing protocols in/around city of Chandigarh, India (April 17, 2020) [Direct](#) | [I-Archive](#)	Study release on community transmission in India as a whole (April 10, 2020) [Direct](#) | [I-Archive](#)	Very similar The two events happened within 1 week. If we consider period overlap then this is VS, if time points then it's VS/SS, so in sum TIME is closer to VS than SS.
S1, L14	Initiative for mental support in during Covid-19 in India (launched March 27, started counseling April 2, free session April 14, 2020) [Direct](#) | [I-Archive](#)	Deportivo Alaves confirmed Covid-19 cases (Tests on March 16, 2020) [Direct](#) | [I-Archive](#)	Somewhat dissimilar A1 is a summary of some mental health service offers, where one of them starts from March 27th and one on April 14th. A2 talks about tests on March 16th, but the release of the report could have happened 1-2 days later. On average over the dates in A1 , we are over the 2-week limit.
S2, L8	Virginia sees new coronavirus cases (April 15,	VDH confirms 1st coronavirus-relat	Very dissimilar The two events happened more than six weeks from each other.

	2020) Direct | I-Archive	ed death in Virginia (March 14, 2020) Direct | I-Archive	

NAR: How similar are the narrative schemas presented in the two articles?

- A narrative schema is a sequence or set of *events and actions* that determine the semantic structure of a news article. A NAR could be for example "a criminal is caught and punished for murder" and should also clearly identify first a primary NAR/focus and then -if present- any secondary NAR/focus. It is in turn constituted from several events/actions described in the article.

Events	Roles
A search B	A = *Police*
A arrest B	B = *Suspect*
B plead C	C = *Plea*
D acquit B D convict B	D = *Jury*
D sentence B	

- These events and actions should be interpreted at a fairly general level, e.g., the concrete statement "Boston Police arrests John W." should be abstracted to "Law enforcement ARRESTS Citizen", to be comparable over articles. I.e. concrete ENTs are rather understood as their *roles* in such an event or action, as ENTs are already covered in the ENT question.
- An article can contain several such events and actions for example: "A SEARCHES B, A ARRESTS B, D CONVICTS B", where the specific roles can be filled in by various subjects/objects, e.g., A = "a police officer" , B = "a suspect" D = "a federal judge". The table on the right shows an extended version of this example.
- What concretely stands in for A, B, D in the text in the above example could either be the actual strings listed (e.g. "a suspect") or more specifically A="NYPD", B= "Peter Smith", C="European Court of Justice". Since we are abstracting, these concrete ENTs would not be considered in this question, but only their semantic *roles*.
- The NAR question is focused on comparing sequences and sets of events .
- Another example: consider the narrative schemas of two hypothetical articles: A ATTACKS B and B ATTACKS A, where A = protester Alice and B = policeman Bob. Then, ENT is Very Similar (both articles focus on Alice and Bob), while NAR is only Somewhat Similar, because the actors - and more importantly their roles - in the two narrative schemas have been swapped (protester ATTACKS policeman vs policeman ATTACKS protester).
- To compare narrative schemas, first we determine what sequences/sets of events and actions each article primarily focuses on. Often, the primary narrative schema is at least hinted at in the title, while secondary narrative schemas are covered in the text. Then, we compare whether these sequences/sets are similar.
- For instance, an article focusing on "US President becoming ill with coronavirus" has a very similar narrative schema to an article focusing on "UK Prime Minister becoming ill with the deadly new 2022 flu", even if concrete events and people making up the schemas are different. Both share a common narrative schema not

only of a person becoming ill, but more specifically of a head of state contracting a contagious virus in the times of pandemia, even though the articles are about different specific people and countries. In fact, the specific instances of people and country (and even type of sickness) should not play a role in deciding the NAR similarity (this will matter for the OVERALL question, but it does not matter for the NAR question), as we want to abstract from the concrete instances to find and compare more general narrative schemas!

Examples:

Pair	Article 1	Article 2	Narrative Schema similarity
	U.K.'s Boris Johnson leaves hospital where he was treated for COVID-19 Direct | I-Archive	Trump leaves hospital but 'not out of woods' – video Direct | I-Archive	Very Similar Both share a main narrative schema about "Head of State released from hospital but not yet fully recovered" that is in primary focus. A1 features some secondary narratives that are however marginal and only serve to provide context for the primary narrative.
	Coronavirus: Boris Johnson admitted to hospital over virus symptoms Direct | I-Archive Note: only consider A1, up until the second header (Analysis)	Trump Covid: US president flown to hospital, says 'I think I'm doing very well' Direct | I-Archive	Very Similar Both share a main narrative schema about "Head of State hospitalized". A1 additionally features a "Fellow politicians wish head of state well" schema. A2 has an additional secondary schema of "head of state reassures citizens in message". In total, these are still Somewhat Similar and the total rating is therefore Very Similar.
	Italy's Health Care System Groans Under Coronavirus — a Warning to the World Direct | I-Archive	India Covid: Hospitals overwhelmed as deaths pass 200,000 Direct | I-Archive	Very Similar The primary narrative schema in both articles is the overload of the healthcare system due to Covid19, including a secondary narrative of "war" against the virus. A1 focusses as well briefly on the possible abandonment of a democratic principle. A2 focusses secondarily on the schema of international help to India, Vaccinations as well as that of political

			responsibility in the wake of elections
S1, L8	National guard helping covid efforts in several US states. Direct | I-Archive	National guard converting buildings for covid use in NYS and other states Direct | I-Archive	Somewhat similar On the general primary schema of "armed forces are helping during pandemic" both articles agree. But A1 is discussing help options in several states much more generally while the specific primary focus of A2 is on conversion of buildings into hospitals throughout, and in particular in one city.
S1, L15	Cheerful Covid masks in Bulgaria Direct | I-Archive	Temporary lockdown in Jamaica Direct | I-Archive	Very dissimilar Production and use of masks is unrelated to complaints about failing restriction of movement.
S2, L10	Electric bills in Karachi Province, Pakistan Direct | I-Archive	Primary focus: Lockdown in Karachi Province, Pakistan Secondary: Government request for utility bill relief Direct | I-Archive	Somewhat dissimilar The Primary focus of A1 is related only to the secondary (or even tertiary) focus of A2, which cover the same schema of "extended utility bill payment deadlines for consumers". However the main NAR of A2 about lockdowns and another NAR about Army help are VD. The secondary NAR about electric bills is SS in that certain actions/events of A2 are repeated (the request by the government figure) but new ones are added on top (the refusal to the request, proposing an alternative plan). Also by our rule of thumb, this is a follow-up article, thereby suggesting SS as the label. In total, these considerations balance out to SD.
S2, L13	Stock market troubles US stock markets and measures Direct | I-Archive	Central bank regulatory measures in UK, US, EU Direct | I-Archive	Somewhat similar Both deal primarily with the narrative schema of uncertainty in the financial system due to Covid. However, A1 deals very concretely with stock markets, while A2 covers the actions by central banks.
S2, L5	Prime minister of the UK gets covid Direct | I-Archive	Youngest victim of Covid in the UK Direct | I-Archive	Somewhat dissimilar Shared: A serious infection of a noteworthy individual. But A1: focusses on the recovery of the PM and secondarily on other positive aspects like congratulations and on the workings of government. A2: narrative is more on children being targets of infection and general danger of the illness spreading.
S2,	Economic relief	Lockdown in	Somewhat dissimilar:

L12	in India Direct | I-Archive	India Direct | I-Archive	Shared: Both primary narrative schemas touch on lockdown measures, but A1 is an opinion on relief measures for the population and heavily focuses on economic impact, while A2 is simply a statement on current lockdown measures
S1, L12	Covid testing and restrictions in/around city of Chandigarh, India Direct | I-Archive	Community transmission in India as a whole Direct | I-Archive	Somewhat Similar: A1: primary focus: concrete testing measures being adopted, including their description, secondary on development of case numbers A2: primary focus: tracking of infection chains remains extremely hard, secondary: officials adapt testing strategy, which is described A1 primary and A2 secondary focus are very similar, A2 primary is unique

Summarized explanations of labels for the above examples using the tie-breaker rules:

Pair	Similarity between Primary and Secondary focuses of the two articles				Label
	PrimaryA vs PrimaryB	PrimaryA vs SecondaryB	SecondaryA vs PrimaryB	SecondaryA vs SecondaryB	
S2, L10	VD/SD (response from utility companies vs lockdown measures, hard judge whether ...)	SS (request to government to pay utility bills vs request to utility companies to relieve utility bills)	VD (request to government vs lockdown measures)	SS/VS (request to government to pay utility bills vs request to the companies to relieve bills)	SD

OVERALL: Overall, are the two articles covering the same substantive news story? (excluding STYLE and TONE)

This category is meant as the most general similarity score that summarizes all relevant sub-aspects of similarity.

- Note that the four aspects that you label first (Geography, Entities, Time, Narrative Schemas) do not cover *series of particular actions or events by specific people at specific times and places*. Concretely: Narrative Schemas cover series of *types* of actions and events, but should not consider their particular instances, such as specific series of actions by Named Entities at specific Geographic locations. Geography, Named Entities and Time, on the other hand, cover their respective aspects, but do not tie them with action and event types. The OVERALL question considers these aspects in unison.

- However, beyond the four aspects that you label first (Geography, Entities, Time, Narrative Schemas) you shall take into account all other information present in the article and use the best of your judgment to weigh that information and the four aspects against each other. For example, articles covering news happening at exactly at the same TIME might be Very Dissimilar overall, since TIME is not by any means a sufficient condition to make two articles OVERALL similar (they can be about very different stories that are just coinciding by chance).
- Proceeding in this way, two stories may end up being Somewhat Similar or Somewhat Dissimilar OVERALL, despite most of the other four aspects being Very Similar. This case may be very uncommon, but understanding it is crucial to accurate labeling. For instance, two articles may focus on Massachusetts (Very Similar GEO) and time periods (Very Similar TIME) and mention similar entities, e.g., state government (Somewhat Similar/Dissimilar ENT), but one describes growing house prices, while the other describes coronavirus restrictions (Very Dissimilar NAR), so they can be OVERALL Very Dissimilar.
- In any case, the OVERALL (or any other) score should **never** include any notion a) of style of writing (e.g. sincere vs. poetic), or b) framing/opinion bias (e.g. a politically conservative vs. a liberal view on exactly the same events). In the most extreme case, we could have two opinion pieces with different views (one positive, one negative) on the exact same event / news story (such as the quality/content of a book being released). This would still be a Very Similar OVERALL rating, as the substantive story is about the quality of the released book and only the framing is different (unless one piece explicitly connects this to some other story, like the authors' political motivations, the general increase in books of this type in the last decade, etc., while the other doesn't).

Pair	Article 1	Article 2	Overall similarity
S1, L8	National guard helping covid efforts in several US states (March 19, 2020). Direct | I-Archive	US Army converting buildings for covid use (March 21, 2020) Direct | I-Archive	Somewhat similar Both are about aid activities of the armed forces, but share different primary focuses. A1 is about the general role of the U.S. military and whether state or the federal governments should direct their work. It mentions hospital ships. A2 is specifically about work the national guard is doing retrofitting buildings as temporary hospitals in NYC.
S1, L12	Covid testing and restrictions in/around city of Chandigarh, India	Community transmission in India as a whole	Somewhat dissimilar Shared: covid, India, ICMR A1: concrete local testing measures A2: a discussion of community transmission in India

			OVERALL is more dissimilar than NAR mainly because of GEO.
	Direct | I-Archive	Direct | I-Archive	
S2, L8	VDH confirms 1st coronavirus-related death in Virginia (March 14, 2020) Direct | I-Archive	Virginia sees new coronavirus cases (April 15, 2020) Direct | I-Archive	Somewhat Similar Both articles list the number of covid cases per county. A1 describes much more than this: the first covid death, reactions to it, and prevention measures. A2 focuses only on the list of infections and is more of a situation update. Also, the two articles mention events that are 1 month apart. Arguably, this pair is borderline Somewhat Similar/Dissimilar.
S2, L10	Lockdown in Karachi Province, Pakistan (March 22, 2020) Direct | I-Archive	Electric bills in Karachi Province, Pakistan (March 23, 2020) Direct | I-Archive	Somewhat similar/dissimilar One article talks about the relief request of utility bills and the other talks about the response of an electricity company to the relief. Periods and locations are very similar. A2 is in fact an update on parts of A1. Since most of A1 is focused on other topics than the relief request, this pair is borderline Somewhat Similar/Dissimilar.
S2, L13	Stock market troubles US stock markets and measures Direct | I-Archive	Central bank regulatory measures in UK, US, EU Direct | I-Archive	Somewhat similar One talks about the US stock market while the other is focused on the world economy. Periods are very similar. Main named entity (the interviewee) is the same.
S2, L5	Prime minister of the UK gets covid Direct | I-Archive	Youngest victim of Covid in the UK Direct | I-Archive	Somewhat dissimilar The two articles both talk about a British got covid-19. However, one is the Prime minister whose event may change the global political state, while the other one is just a kid showing the infectivity of Covid-19 virus. The location is somewhat different. Time is borderline Somewhat/ Very Different.
S2, L12	Economic relief in India due to covid Direct | I-Archive	Lockdown in India Direct | I-Archive	Somewhat dissimilar The two articles just both talk about the measures that India adopted to confront Covid-19. Entities are somewhat dissimilar. Periods are also somewhat different.
S2, L11	German authorities say man with virus	Authorities fear German tourist brought Covid-19	Very Similar The two articles both talk about the same german covid victim, his infection and his

	in critical condition Direct | I-Archive	to Netherlands Direct | I-Archive	trip to netherlands.
S2, L9	Key dates in Indiana's path to reopening amid the coronavirus pandemic Direct | I-Archive	Outlining "measured" lifting of lockdown Direct | I-Archive	Very Similar The two articles describe exactly the same event that India is going to reopen. They are different in terms of level of details and styles (the first one includes infographic).

ADD: Additional questions

- STYLE: Do the articles have similar writing styles?
  - This question judges whether the two articles are written in a similar style of writing (e.g., sincere vs poetic, formal vs informal,) and have similar format/structure (e.g., continuous narrative vs bullet points or list-like elements vs tables vs infographics).

- TONE: Do the articles have similar tones?
  - This question asks you if the general sentiment/tone in both articles is similar. That is, are they both positive or negative, optimistic or pessimistic, supportive or indifferent, reinforcing or critical, etc towards something. They do not *need* to have that tone towards a shared target, but *might* have a shared target. However, having or not having a shared target should not influence the rating of TONE.

Pair	Article 1	Article 2	Style and framing similarity
S2, L14	Concor Head rallies SA to fight virus Direct |	Church Leader Defies Coronavirus Measurers Direct |	*Style*: Very Similar Same formal style. *Tone*: Very Dissimilar The negative opinion about the lockdown

			measures of religious leaders vs a positive opinion of a CEO of a company.
S2, L7	Impacts of Trump's Coronavirus Fumbles—and How to Protect Yourself Direct |	Fewer Than 1,900 Americans Tested For Coronavirus, 'Atlantic' Reports, Virus Now Found In Digestive Tract Direct |	*Style*: Somewhat Dissimilar The first article uses renarration/historical reference to start the article and is broken up into sections where the corresponding issues are discussed and recommendations are given.Overall, the style is that of an opinion piece. The second article is more like a brief reporting news article. *Tone*: Somewhat Dissimilar A1 is overly critical and reproaching of the lockdown actions of the US government and its president. A2 has a more neutral tone, although still showing some disapproval.
S2, L8	VDH confirms 1st coronavirus-rela ted death in Virginia (March 14, 2020) Direct | I-Archive	Virginia sees new coronavirus cases (April 15, 2020) Direct | I-Archive	*Style*: Somewhat Dissimilar Article 1 is a standard running text interspersed with several direct quotes, but Article 2 is a list of bullet points. Both articles use formal language. *Tone*: Somewhat Similar Reporting is neutral, but A1 includes sentimental information via statements from people.
S2, L9	Key dates in Indiana's path to reopening amid the coronavirus pandemic Direct | I-Archive	Outlining "measured" lifting of lockdown Direct | I-Archive	*Style*: Somewhat Similar A significant part of Article 1 is an infographic, which does not appear in Article 2, but the style is formal in both articles. *Tone*: Very Similar Both articles have a *neutral* tone.
S1, L5	Covid loans extended in the Seychelles Direct | I-Archive	Pleasant activities people are doing during Covid lockdown Direct | I-Archive	*Style*: Very Dissimilar A1 is a running text, A2 is a list-like format and is much more casually written *Tone*: Somewhat Dissimilar A1 has some positive, but also negative traits. A2 points out mostly pleasant, positive activities. The difference, however, is not big.
S1, L10	Toyota Extends Plant Closures In North America Until May Amid Plunging	Toyota extends shutdown of North American plants through April 17	*Style*: Very Similar The style is generally the same *Tone*: Very Similar Both articles feature a rather negative outlook

	Demand Direct | I-Archive	Direct | I-Archive	
	Adam Nathaniel Furman unveils "monumentally joyful" pyramid at King's Cross Direct | I-Archive	Opinion: 'Best Starts For Kids' Provides Vital Educational Support Direct | I-Archive	*Style*: Very Dissimilar A1 is a regular journalistic article text with interview elements, while A2 is an opinion piece. *Tone*: Very Similar Both articles have a very positive, optimistic tone