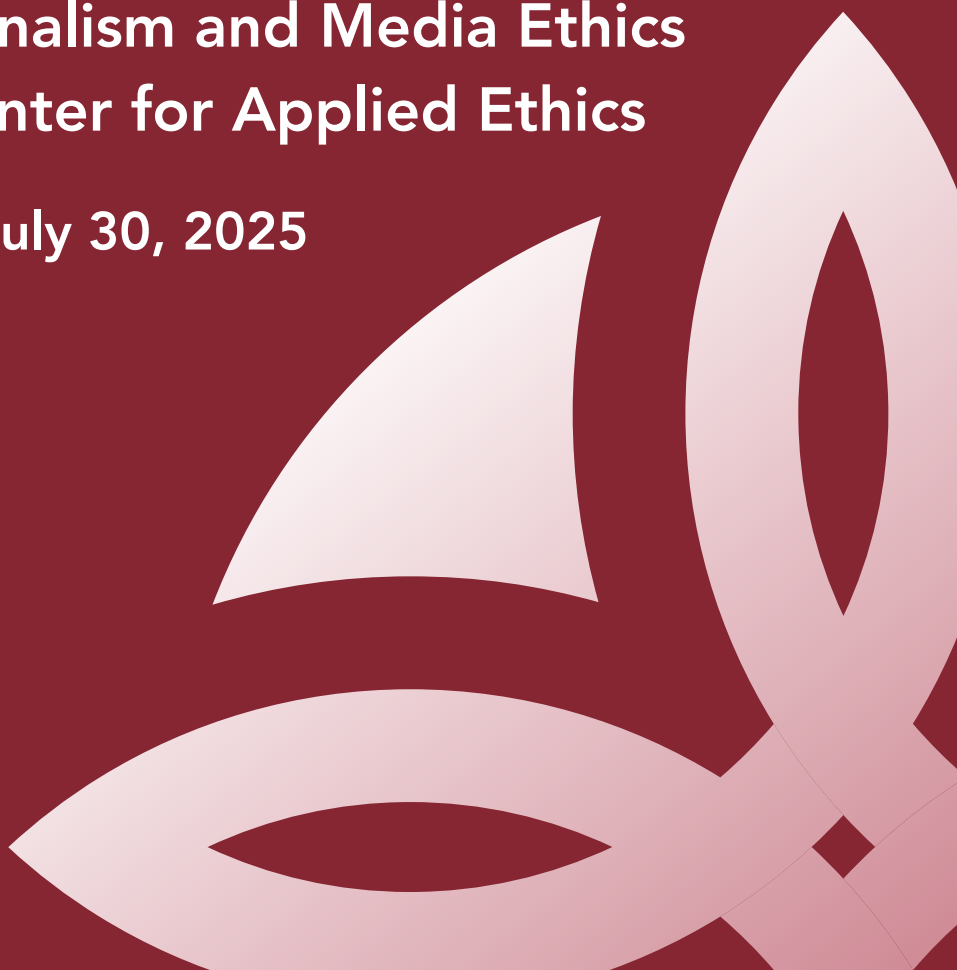# Santa Clara
Markkula Center
for Applied Ethics

1851

Teaching Media Sourcing Literacy Through Annotations

Subramaniam Vincent
Director, Journalism and Media Ethics
Markkula Center for Applied Ethics

July 30, 2025

Teaching Media Sourcing Literacy Through Annotations

At the core of storytelling for news is sourcing and attribution. They are like the two sides of the journalistic-reporting coin. Journalists attribute statements, claims, findings, conclusions, and broadly the content in their stories to sources and justify their inclusion. The justifications and characterizations of sources are often unstructured and embedded in the communication itself, be it writing, audio or video. Without people, organizations, footage, documents, and data serving as sources, journalists would not be able to report out stories from the realities they interrogate.

But how might college students–undergraduate or graduate–systematically observe the use of sourcing and attribution that are present in everyday news? One way is to annotate all the sources and attributions in news reports, using a vocabulary (terms and definitions) and a datasheet.

Why this guide?

Using a vocabulary that defines a range of terms involving sourcing, annotations require the task of combing through all instances of attribution and their related sources in a single, whole story. In the process, the detailed linguistic and cultural work involved in the way stories expose sources to the world becomes evident to the annotator. There is language that names and introduces sources with titles or links to original documents. There is also language that further adds characterizations, justifications, and context for why the sources are in the story. All of this tends to go unnoticed if one is reading the news casually to make quick sense of the claims and facts.

This guide to teaching sourcing literacy (to students in particular) breaks down sources and their inclusion in news stories into structured parts and shows anyone how to identify the parts separately and populate a spreadsheet. Each row in this spreadsheet would be an actual sourcing statement involving some kind of attribution. Each column is a type of data about the sources without whom that attribution would not have been possible. The rows and columns together become the sourcing annotations for the news story.

One benefit to teaching sourcing annotations is sourcing literacy. It helps build an appreciation for when the different types of sources–from named to unnamed to documents to organizational–become important. But it goes beyond literacy. Annotations, when done for a range of stories in the same news cycle, will show how sourcing decisions drive the framing and the interaction between the different parts of the story itself. Which sources and claims are used

to lead a story, which to corroborate, which to refute, which deepen or add context or offer more information, and so on.

A different benefit is for data-based studies. They enable the building of ground truth datasets which in turn help computational studies of sourcing. Ground truth or ground truth data, refers to verified, true data used for training, validating and testing artificial intelligence (AI) models. In our case, ground truth data for sourcing annotations are useful for evaluating the accuracy of AI tools and models in detecting and annotating sources in stories at scale.

Outline

This guide has five sections, laid out in a step-by-step approach. The actual execution of a training session is explained in the two checklists included. It is best to review the sections in the order below before jumping into an exercise.

1. Vocabulary: This section defines all the terms needed for an annotator to learn about and understand sourcing, its meanings, and applications
2. Annotations dataset template: Our data schema for any spreadsheet you use to capture sourcing annotations
3. Examples
4. Pre-annotation Checklist for Facilitators
5. Pre-annotation Checklist of Annotators (students)
6. Step-by-Step Annotation Instructions
7. Common Errors
8. Credits
9. Citations

Vocabulary

This section introduces the comprehensive set of terms and meanings contextual to news reporting, with examples. There are two sets of terms involved and they emerge from how sourcing and attribution practices manifest in language.

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

2

Caption: News language has signals of sourcing. Slide: Subbu Vincent, 2025 AI talks at Stanford, Baltimore and 2024 at Syracuse.

Diagnostic terms

(to apply meanings and identify the sourcing in a story)

Source
Types of sources
     Named Person
     Named Organization
     Unnamed Group of People
     Anonymous Source
     Document

Dataset terms

(to represent the annotated data in a spreadsheet for a given story)

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

3

Sourced Statements
Name of Source
Type of Source
Title of Source
Source Justification

Source: A source in journalism is a person, organization, document, or another news article from which a journalist takes viewpoints, experiences, claims, expertise, positions, insights, knowledge, data, or documents. Reporters may directly quote their sources or use indirect speech to paraphrase a source's views or claims.

Sometimes the source is the person who sent sensitive material such as emails or documents or other internal organizational correspondence to the reporter. The source is a person who may have been present at meetings where sensitive deliberations or discussions took place, and that person then shared material from the meetings with the journalist.

For e.g., when a reporter attributes a claim or statement to a person or a group of people with words such as, "according to people familiar with..", or, "according to people who were present at the meeting …, " it means that person or those people are the source. When the reporter attributes claims or a statement using words like, "according to a copy of emails reviewed by this newspaper," or, "according to a copy of the document …," it means the source could be a document. If the person who sent the emails or documents to the reporter is granted anonymity by the report, then that person is an anonymous source. (See also, our definition of Anonymous Source.)

A source is an organization which the reporter cites n by contacting its spokesperson or other representative or official. The reporter may cite such spokespersons or officials from a direct interaction or from a press release or social media or corporate blog.

A source can also be another news organization itself. When the journalist has attributed something in the text to another news organization that reported it earlier, then that organization is the source. Sometimes, the reporter will name that organization directly in the text. At other times, the text may simply carry a web link directly to that news article, similar to a document link.

Primary or Secondary source: When a reporter has contacted and talked to or interacted with a source directly before citing their views, that is a ***primary source***. Primary sources know that they are talking to a journalist directly when they communicate to the journalist, usually in response to questions.

When a group of reporters in a press conference ask officials questions, this is a case of collective or shared primary sourcing. Sometimes reporters take views from a person or official in a public place like a meeting or on the road. This is also primary sourcing. If the reporter took the claims or views or assertions of a source from another article or social media, or from the public domain like a speech, that is a ***secondary source***.

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

4

Types of Sources: a) Named Person b) Named Organization c) Document d) Anonymous Source and e) Unnamed Group of People.

Named Person: A source who is a named human being in the story. Note that documents like emails, audio, video, meeting minutes, or other material internal to an organization sent by a human source may contain named people expressing views, taking actions, decisions, etc. Those named people are not sources, even if the reporter included those names and reported their viewpoints or actions in the news article. The person sharing the emails or documents with the journalist is the source, or the documents themselves, if they are authenticated, are a source. (See also, our definition of Document as a type of source.) Also note: When a named person is quoted (directly or indirectly) citing or referring to the contents of a document, the named person is the source, not the document.

Named Organization: An organization named as a source. This includes cases where reporters attribute a statement or statements to unnamed "officials" or a "spokesperson" of a named organization.

Document: An original or authentic document issued by an authoritative organization or person for that field, jurisdiction or expertise. It is usually accessible by the public online or may be retrieved in return for a public records request. The document may itself be a full webpage or PDF or other portable format. It may be attributed through a URL inline in the article. However, when documents sent to a reporter are part of internal material in an organization, such as emails, correspondence, presentations, and data, the person sending the documents is the source.

Also, the reporter may have included statements from a person who is citing or referring to the contents of some document or study or research. That does not make the document the source. The person is the source. The document is the source only if the author of the story has accessed the document itself and used its contents directly for the story.

Anonymous Source: An anonymous source is a person, known by name and identity to the reporter and often the reporter's editor, but the name is being withheld from inclusion in the article as a named source before its publication. You may see the reporter disclose this in the story, that they spoke to the person "on condition of anonymity," or that they were offered anonymity to discuss sensitive matters. This source may have been present at meetings where sensitive deliberations or discussions took place then shared material from the meetings with the journalist. Or this source may be a key player or witness in internal decisions or proceedings at an organization.

For e.g., when a reporter attributes a claim or statement to a person or a group of people with words such as, "according to people familiar with..," or, "according to people who were present at the meeting…," it means that person or those people are the source. This source may have sent sensitive documents, emails, audio, video, meeting notes, or other internal correspondence from inside an organization to the reporter, and the person's identity needs to be protected. So they are not named.

In journalism ethics, the reporter may offer anonymity to the source if the reporter believes or

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

5

concludes that the source's life or career/job or family members lives, etc., might be in jeopardy or that they might face retaliation or retribution. An unnamed spokesperson or official of a named organization is not an anonymous source.

Unnamed Group of People: Sometimes a reporter may attribute a statement or statements to a group of people who are saying or expressing or advocating for something as a group. We call this type of source an Unnamed Group of People. This source is a group the journalist has witnessed on the ground or online or at a meeting or otherwise had access to the group of people. For example, players, teachers, children, protestors, attendees, advocates, activists, participants, onlookers, commenters, etc. The reporter does not name them individually and refers to them as a group, and may quote or paraphrase what the group of people is saying using words like, "the protestors said this," or "the teachers chanted…"

Note 1: Do not confuse them with the Anonymous Source type. Here an individual actually asked for anonymity from the journalist. An Unnamed Group of People-type source is not a group who asked for anonymity from the reporter.

Note 2: A document sent by a source to a journalist may contain references to groups of people saying something. That is not an example of an Unnamed Group of People source. The document is the source. Note 3: When reporters attribute a statement or statements to unnamed "officials" or "spokesperson" or "lawyers" of a named organization, the type of source is Named Organization, not Unnamed Group of People.

In any news story, you may find one or more of these types of sources.

Sourced Statement: Every statement in a story that the reporter would NOT have been able to put in without drawing or receiving the content or part of the content from one or more sources, is a sourced statement. These are actual text lines in the article that reporters have written with attribution to their source.

Sourced statements may contain quotes or indirect speech. They may include viewpoints, experiences, criticism, questioning, advising, support or other expressions from a source.

Sourced Statements also include all other attributions the reporter has made in the story, referring to the source's conduct, position, or attitude. For example, lines that state that a person criticized, supported, questioned, or decided something, may be present in addition to quotes from that person.

Name of Source: Name of Source applies to the following types of sources: Named Person, Named Organization and Document. For Named Person and Named Organization types, use the person or organization's full name. This is identifiable as a proper noun in the article. When a source is referred to using a generic designation such as, "spokesperson" or ",representative" or ",officials" from a named organization, the name of the source is the name of the organization.

When the source is a public document, the name of the source is the document's publishing organization or person, if available. For Anonymous Sources, the Name of Source does not apply, i.e. it has a "null" value. For Unnamed Group of People source, the name of source is the

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

6

term being used to refer to the group. For e.g., it might be "teachers", "participants", "parents", "attendees", "rallygoers", "protestors" etc.

Title of Source: The words defining or designating the formal position of power, authority, or leadership held by the Named Person source. For example, Director, Mayor, Vice President, President, Secretary, Treasurer, Board Member, Professor, Provost, Principal, Congresswoman, Senator, Chief of Police, General Counsel, Spokesperson, etc. If the person is a legislator, the title includes the constituency. Experts are often Named Person sources and their expertise is signaled by their title, designating their position or role usually held at some organization, and specialization. Sometimes people in leadership roles will simply be referred to as "leader". That is also a title.

Note 1: Characterizations or designations like team member, player, senior, sophomore, junior, freshman, activist, protestor, attendee, participant, eyewitness, etc. are NOT titles. Occupations in the trades like carpenter, plumber, janitor, installer, coal miner, etc., are not titles. They represent people in a type of trade, life journey, or activity.

Note 2: For anonymous sources, the title may or may not be in the story and only the association or justification for inclusion in the story is mentioned. Also, journalists may include the voices of everyday people and people who are NOT in formal positions of authority or power. Such Named Person sources may not have a title. Finally, for Named Organizations and Document sources, Title of Source does not apply.

Source Justification: This refers to any additional source characterization or explanation or context that justifies to the reader why the source is in the story or that section of the story, how they are connected to the story and/or to other sources in the story.

Any of the five defined types of sources may have such justifications and explanations present. It is not the same as the title of source, which is the previous definition above. It may be a few words, a part of a sentence, multiple sentences, or a full paragraph. The reporter will usually offer a justification in the story when they introduce the source. Source justification may be a part of the sourced statement itself. Sometimes the source justification comes later or earlier in the article where the source or some situation involving the source is referred to.

While source justification is NOT the same as title of the source, it may include the title of the source.

For named persons, the source justification text may narrate the lived experience of the source. When sources are people who are stakeholders to the issue being reported on, who witnessed something happen, or have a lived experience related to the issue of the story, or a co-litigant in a lawsuit, etc., narrating this demonstrates their significance in the story for readers. For example, someone who went through a period of homelessness may be quoted for their lived experience and opinion about solutions.

Someone else may have spent four years waiting to get a job or to get their voting rights back because of a prior felony conviction. Remember that Source Justification is not the same as title of the source, but may include both the title of the source and the explanations for the source's role in the story or relationship to other sources.

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

7

Named Persons or Anonymous Sources without a title may still have source justification present in the text.

Annotations Template Data Sheet

Here is a template that an annotator can use to represent a story's sourcing data.

NEWS STORY SOURCING ANNOTATION TEMPLATE						
URL: URL of story annotated						
Title: Title of story						
Sourcing Manifest or Source Data Table		**Types of Sources: a) Named Person b) Named Organization c) Unnamed Group of People d) Anonymous Source e) Document**				
No	**Sourced Statements**	**Type of source, e.g. "Named Person"**	**Name of Source**	**Title of Source**	**Primary or Secondary Source (P/S)**	**If, Secondary Source, put link here**	**Source Justification (includes addition source characterizations beyond the title, additional context about the source)**
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
15							
16							
17							
18							
19							
20							

Examples

<screenshot 1>

No	Sourced Statements	Type of Source	Name of Source	Title of Source	Source Justification
7	Steve Smith of Civic Nebraska, part of a large coalition of groups supporting the measure, said the wait creates a group of taxpayers who can't choose their representatives.	Named Person	Steve Smith	Civic Nebraska	Steve Smith of Civic Nebraska, part of a large coalition of groups supporting the measure
8	"You're civically dead and you can't vote for the people who are levying those taxes," he said	Named Person	Steve Smith	Civic Nebraska	
9	At the time Nebraska was in step with other states. Now, while a few states require wait times for specific offenses or define completion of a sentence as including things such as fines and restitution, Nebraska is alone in requiring a general waiting period beyond imprisonment and release from parole or probation, said Margaret Love, co-founder and director of the Collateral Consequences Resource Center, which keeps a 50-state database on restoration of rights.	Named Person	Margaret Love	co-founder and director of the Collateral Consequences Resource Center	which keeps a 50-state database on restoration of rights.
10	The bill's author, Democratic state Sen. Justin Wayne, said he was going door to door in his first election in 2016 and was told by would-be constituents that they could not vote. Much of the reason was confusion over the law's waiting period, he said.	Named Person	Justin Wayne	Democratic state Sen.	The bill's author
11	Wayne, who represents parts of Omaha with strong minority populations, said reconnecting people to the voting process is integral to successful reentry.	Named Person	Justin Wayne	Democratic state Sen.	His bill advanced this past week from a committee to the full Legislature.
12	"When people get out of our system, they've got to feel engaged in their community, and the number one way for a person to feel engaged in their community is to be able to vote for the leadership of that community," he said.	Named Person	Justin Wayne	Democratic state Sen.	
13	Kathy Wilcot, a member of the University of Nebraska Board of Regents, was the lone dissenter from among the nearly 20 witnesses who spoke on Wayne's bill. Wilcot stressed she was speaking as an individual and not on behalf of the university.	Named Person	Kathy Wilcot	a member of the University of Nebraska Board of Regents	was the lone dissenter from among the nearly 20 witnesses who spoke on Wayne's bill.
14	"I do think that hopefully the waiting period reinforces the fact that voting is something very special, and hopefully that will be part of the things that an individual would consider if they're tempted to break the law again," she said.	Named Person	Kathy Wilcot	a member of the University of Nebraska Board of Regents	
15	Three of the witnesses with criminal records who spoke in favor of the legislation said in later interviews the waiting period is not a deterrent to future crime, but rather a barrier for those who have served their sentences.	Unnamed Group of People			

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

8

ANNOTATIONS

Pre-annotation Checklist for the Facilitator/Faculty

1. Create a shared drive for students, and post into it an empty data template spreadsheet for the students.
2. Identify a set of reported stories from diverse news outlets, and save their URLs into a different assignment spreadsheet. If you have 10 students for example, and each student annotates two to three stories a week, a 10-week quarter would need at least 20 stories X no.of.students for your annotations corpus.
3. Best practices for identifying stories:
   a. Take major news cycles visible using trending tag/label/topic names on news aggregators. For each topic, hundreds of stories are reported. If you take 10 topics, you can easily find 5-10 (non-paywalled) reported stories (URLs) per topic.
   b. Go to specific local and or mission-oriented news sites you and your students may already be monitoring or using to discuss media issues. Directly pull reported stories off their sites.
4. In your assignment sheet, you could put a student name across every URL and a status column.

Pre-annotation Checklist for the Annotator (Student)

1. Start with a blank Annotations Template (Xlsx downloadable file) provided with this guide.
2. Copy the link and title of the story you are reviewing (assigned or self-identified) into the top left-hand side cells.
3. Read the article end to end.
4. Follow ALL the steps under, "Annotation Instructions," below to extract (annotate) the following data from the article. Sourced Statements, Name of Source, Type of Source, Title of Source, and Source Justification.

Step-by-Step Annotation Instructions

This section will help the annotator go through the story scanning for sourcing statements by type of source. It goes in the following order.

1. Annotate for Anonymous Sources
2. Annotate for Unnamed Groups of People
3. Annotate for Document

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

9

4. Annotate for Named Organizations
5. Annotate for Named People

(An alternative way to annotate is to use the full worth of the vocabulary to find all statements involving any attribution, enter into a row, map it to one of the five source types, and finish the annotation column entries for that row. Both methods ought to result in the same data tabulation.)

Anonymous Sourcing Annotation

1. Using the definitions provided above in the Vocabulary section, identify all the Anonymous Sources in the article. Remember that for Anonymous Sources, Name of Source has a 'null' value.

2. For each Anonymous Source in your list above, find all the associated Sourcing Statements. For sourced statements, extract the exact full sentences from the article whether they are quoted or indirect speech as the data. Do not paraphrase or summarize. Retain them as they are. Copy and paste the exact wording used in the article for each statement. And pay special attention to multiple statements from the same source. Remember that for Anonymous Sources, Name of Source has a "null" value.

3. Applying the definition of Title of Source provided above, for each Anonymous Source in your list, extract the Title, if present in the text. Remember that Anonymous Sources may or may not be included with a title, since the title itself may be sensitive information. But if a title is present, extract it.

4. Now apply the definition of Source Justification given, and find the source justification for each anonymous source in each sourcing statement. Extract only the text that matches this definition directly from the story. Remember that source justification refers to words or sentences doing additional characterization or justification for inclusion of that source in the story. Do not synthesize or generate or summarize in your own words. Extract the actual text. If more than one source justification is present in the text, concatenate them as one single string separated by a semicolon.

5. Fill out the row for that sourced statement in your annotation spreadsheet. The row starts with the "Sourced Statement" column and ends with the"Source Justification" column.

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

10

Unnamed Group of People Sources Annotation

6. Next, apply the definition of the Type of Source "Unnamed Group of People" and find all such sources in the article. Remember that the type of source "Unnamed Group of People" is different from "Anonymous Source" so apply the definition carefully.

7. For each source of the type "Unnamed Group of People" in your list above, find all the associated Sourcing Statements. Look only for statements where groups of people are quoted directly or indirectly. Remember that none of the Sourced Statements in your Anonymous Sources list must appear on this list.

8. And remember that for Unnamed Group of People, the Name of Source is the term the reporter has used to refer to the group. For example, it might be "teachers", "participants", "parents", "attendees", "rallygoers", "protestors", "advocates", "activists", "students", etc. If not present, set Name of Source to the -"null" value.

9. For each "Unnamed Group of People" source in each sourcing statement, find the Source Justification, by applying the definition given. Extract only the text that matches this definition directly from the story. If the same Unnamed Group of People source is attributed into multiple sourcing statements, you may copy the Source Justification from the first sourced statement to the others for that source. Remember that source justification refers to words or sentences doing additional characterization or justification for inclusion of that source in the story. Do not synthesize or generate or summarize in your own words. Extract the actual text. If more than one source justification is present in the text, concatenate them as one single string separated by a semicolon.

10. Fill out the row for that sourced statement in your annotation spreadsheet. The row starts with the "Sourced Statement" column and ends with the "Source Justification" column.

Document Sources Annotation

11. Next, apply the definition of the Type of Source "Document" and find all such sources in the article. Remember that Document sources are different from Anonymous Sources and Unnamed Groups of People, which you already identified.

12. For each Document source in your list, set the Name of Source to the "null" value. But if the document is a public document, set the Name of the Source to the document's publishing organization or publisher, if available.

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

11

13. For each Document source in your list above, find all the associated Sourced Statements. Remember that you already saved the sourced statements Anonymous Sources and Unnamed Group of People sources. The sourced statements for the Document sources, if any, will be different. Extract the exact full sentences from the article whether they are quoted or indirect speech as the data. Do not paraphrase or summarize. Retain them as they are. Copy and paste the exact wording used in the article for each statement. And pay special attention to multiple statements from the same source.

14. Now apply the definition of Source Justification given, and find the Source Justification for each Document source in each sourcing statement. Extract only the text that matches this definition directly from the story. If the same Document is attributed into multiple sourcing statements, you may copy the Source Justification from the first sourced statement to the others for that document. Remember that source justification refers to words or sentences doing additional characterization or justification for inclusion of that source in the story. Do not synthesize or generate or summarize in your own words. Extract the actual text. If more than one source justification is present in the text, concatenate them as one single string separated by a semicolon.

15. Fill out the row for that sourced statement in your annotation spreadsheet. The row starts with the "Sourced Statement" column and ends with the "Source Justification" column.

Named Person Sources Annotation

16. Next, apply the definition of the Type of Source "Named Person" and find all such sources in the article. Remember that all named persons in the text of the news article are not Named Person sources. And this type of source is different from Named Organizations.

17. For each Named Person source in your list above, find all the associated Sourced Statements.

18. Applying the definition of Title of Source provided above, for each Named Person source in your list, extract the Title, if provided. Remember that Named Persons in positions of power, formal authority, leadership, or expertise will usually be quoted with a title. But when the voices of everyday people and people without structural power are included in the story, they will also be Named Person sources and may not have a title.

19. Now apply the definition of Source Justification given, and find the Source Justification for each Named Person source in each sourced statement. Extract only the text that matches this definition directly from the story. Remember that source justification refers to words or sentences doing additional characterization or justification for inclusion of that source in the

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

12

story. Do not synthesize or generate or summarize in your own words. Extract the actual text. If more than one source justification is present in the text, concatenate them as one single string separated by a semicolon.

20. Fill out the row for that sourced statement in your annotation spreadsheet. The row starts with the "Sourced Statement" column and ends with the "Source Justification" column.

Named Organization Sources Annotation

21. Next, apply the definition of the Type of Source "Named Organization" and find all such sources in the article. Note: All named organizations in the text of a news article are not sources. And this type of source is different from Named Persons.

22. For each Named Organization source in your list above, find all the associated Sourced Statements.

23. Now apply the definition of Source Justification given, and find the Source Justification for each Named Organization source in each sourced statement. Extract only the text that matches this definition directly from the story. Remember that source justification refers to words or sentences doing additional characterization or justification for inclusion of that source in the story. Do not synthesize or generate or summarize in your own words. Extract the actual text. If more than one source justification is present in the text, concatenate them as one single string separated by a semicolon.

24. Fill out the row for that sourced statement in your annotation spreadsheet. The row starts with the "Sourced Statement" column and ends with the "Source Justification" column.

25. For any of the rows you are adding, if you have a question or concern or there is ambiguity, put a note on the Student Comment column (last column).

Final steps

26. Scan the full article text again from top to bottom–double check you did not miss anything.

Common errors

1. E.g. of something that is NOT Title(+Affiliation),
- *"the two people familiar with the board's deliberations,"*
- *"Anduril promotional video,"*
- *"oversees the program coaches at Strength-Based Community Change"*

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

13

The above are good examples of **Additional Source Characterization or Justification For Inclusion In Story**

2. For anonymous sources, there is no entry needed in the Name of Source column. It intended to be blank.

3. For organizations named as sources, leave the Title of Source column blank. Title is only named person sources.

4. Names of organizations like "Customs and Border Protection" need to be in the Name (of Source) column, not in the Title.

5. Follow the same spellings in your sheet as the reporter has done in the news story. If the reporter has misspelled it, that's the correction for the newsroom to make. For annotation purposes copy exactly what the story writer has used.

6. For source justification, remember to copy the text from the sentence or graf. It can be a fragment or partial sentence (a set of words), or a whole graph. All of that is valid source justification.

7. Attributions using indirect speech or paraphrasing are valid cases of sourcing. Direct quotes are the easiest noticeable from. But indirect speech is a key part of sourcing.

8. Document vs. organization source type
9. Primary vs Secondary; document source can be primary too.
10. <Name, Title> instead of <Name.. with words.. and Title>

Credits

The sourcing annotations project is an interdisciplinary effort at the Journalism and Media Ethics Program at the Markkula Center, involving students and faculty at the College of Arts and Sciences and the School of Engineering.

SCU Department of Computer Science & Engineering
Dr Yi Fang

       Graduate Students
       2024, 2025
       Zhan Shi (ongoing, LLM accuracy metrics)

       2024
       Phoebe Wang (First version/MS Thesis)

       2021-2023 (Source Diversity/NLP)

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

14

Xiaoxiao Shang
Xuyang Wu
Louise Li

Ground truth annotation dataset creation

Undergraduate students
George Packard (Communication)
Kylie Bennett (Communication, Political Science)
Emily Hofstetter (Communication, Public Health)
Gigi Patmore (Electrical Engineering and Computer Science)
Kelly Perasso (Communication, Italian Studies)
Sarah El Shenawy (Computer Science, Communication)

Graduate students (Engineering)
Haowei Liu
Leo Wei
Shuowei Li
Jinming Nian
Zhan Shi
Xiaoxiao Shang

What students have said

During 2024 and 2025, SCU students have learned and implemented sourcing annotations for news stories as part of an internship class offered by the Markkula Center. At the end of the course, they submit a reflection on their learnings. Here are some quotes of what students have said.

Genevieve Patmore SCU'26

"Going over the ideas of stakeholders within stories made me realize the importance of sourcing and how leaving out key perspectives and people has the ability to bias any article… reading through the different assigned articles made me realize how different the storylines ended up being based on the types of sourcing used."

"Learning about these different types of sources led me to gain a better understanding of journalism as a whole and how incorporating a multitude of sources from different perspectives is important for the overall integrity of the article."

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

15

Kelly Perasso SCU'25

"Because of this internship, I have begun to pay more attention to attribution, especially the quality, quantity, and consistency of sourcing we see in articles daily."

"It is important for journalists to implement their knowledge and context amongst others' opinions and facts in order to build a story. Adding context also involves an ethical issue, for sometimes source characterization can change how a reader interprets the source and their information."

George Packard, SCU'25

"I learned a lot about the conceptual understanding/vocabulary of a legitimate news story, such as the five types of attributions a journalist can use for a sourced statement: named person, named organization, document, anonymous source, and unnamed group of people. Learning the definitions and thresholds for each was essential in order to be able to identify and tabulate all statements from an article."

Kylie Bennett, SCU'27

"..when referencing information the author has gained secondhand, it is their ethical responsibility to make that information's origin also accessible to the reader."

"By giving additional justification on a quote or quoted topic, the author provides the readers with a deeper understanding of what the sourced statement adds to the story."

References/further reading:

Could Quoting Data Patterns Help in Identifying Journalistic Behavior Online?
S. Vincent, X. Wu, M. Huang, and Y. Fang (2023). Could Quoting Data Patterns Help in Identifying Journalistic Behavior Online? #ISOJ Journal, 13(1), 33-64]

The Ethical Distribution of News: Roundtable Use Cases.
Santa Clara University. (n.d.). *The Ethical Distribution of News: Roundtable Use cases*.

Teaching Media Sourcing Literacy Through Annotations | Subbu Vincent, Director, Journalism & Media Ethics
Markkula Center for Applied Ethics, July 30, 2025

16