Content Analysis Protocol — Capturing People in Articles

Table of Contents

DATA PROCESSING PROCEDURES	1
COLLECTING INFORMATION ON PEOPLE IN ARTICLES	2
READ THE ARTICLE	
Coding Authors	3
Name	3
Person Type	4
Organization	4
Lookup Name and Process Person	4
CODING SUBJECTS	6
Name	6
Person Type	6
Organization	8
First Quote	9
Lookup Name and Process Person	10
HOW TO LOOKUP A PERSON	11
ONCE YOU ARE DONE, SUBMIT YOUR CODING	13
ILLUSTRATIVE CODING EXAMPLES	13
CODING APPLICATION	14
APPENDIX: FIRST QUOTE EXAMPLES	14

Data processing procedures

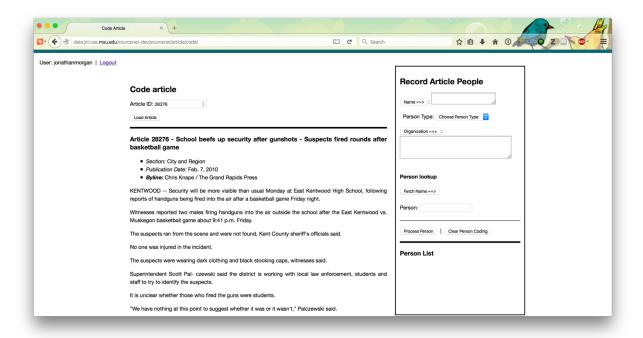
A list of locally implemented news articles will be provided to each coder. The articles have been loaded into a database and will be viewable from within our web-based coding application. Coders will open each article by entering its ID, capture author, subject, and sourcing information, and then save their coding to a database.

Each time a coder begins a new round of coding, he or she should:

- 1) Read through this protocol, to refocus on what we are capturing.
- 2) Look through the last two or three articles coded in the most recent previous session of coding to refresh one's memory of how we code.

For each article, the coders will:

- 1) Load the coding application: http://data.jrn.cas.msu.edu/sourcenet/sourcenet/article/code/
- 2) In the coding form, enter the ID of the article to be coded in the "Article ID" field at the top of the page, then click the "Load Article" button to load that article. The coding page will then display the article text on the left with a coding area on the right:



- 3) Once the article is loaded, read through the story.
- 4) Record the author(s) of the story.
- 5) Go through the story again to look for subjects per the operational definitions outlined below, recording information on each that is present in the story.
- 6) Submit coding for the story to the database.

While coding, you should try to:

- **Focus** Try to code in a quiet place that minimizes distractions and give yourself enough time that you aren't distracted by what comes next.
- Pace yourself Do not rush through coding, take enough time to be thorough and do a good job.
- Avoid fatigue Don't code more than a couple of hours at a time, and if you feel your mind starting to flag or wander, take a break.

Collecting information on people in articles

For each article, the coder will read the article through in its entirety, then use this coding application to identify and categorize the following:

- The <u>author(s)</u> either mentioned in the byline or in a section on "additional reporting provided by" at the end of an article. Every distinct <u>author</u> gets a separate entry.
- <u>Subjects</u> in the story any named person in the story, whether or not they are quoted, including their organization. Each distinct *subject* should get its own separate entry.
- For subjects who are quoted or otherwise specifically cited as a source of information (*sources*), the coder will categorize them as such, then also capture their first quotation.

As you code, part of your job as a coder will be to try to match the name in the article to that person's record in the database. Each distinct person we detect and record in an article is given a unique ID that can be referenced across articles, so we can more readily tie people together across stories. Each name you record in an article will be compared to a database of people already coded from other articles, and if a person with a matching name is found in the database, we tell the coding program to tie this mention of that person to their overall profile.

NOTE: If there is any doubt about the person with whom a quote in an article should be associated, don't link the name with a person from the database. It is easier to go back later and merge than it is to figure out when quotes from multiple people have been attributed to a single person in our system.

Read the article

Before you begin coding, make sure to read the article through in its entirety.

Coding Authors

<u>Author(s)</u> are either mentioned in the byline or in a section on "additional reporting provided by" at the end of an article (example: http://www.nytimes.com/2005/08/29/us/powerful-storm-threatens-havoc-along-gulf-coast.html - this is relatively rare). Every distinct *author* or *reporter* gets a separate entry.

When you code authors in a story, you will capture the following:

Name

- Capture the <u>Name</u> of the author. You will select the complete text of the author's <u>name</u> in the byline or additional reporting paragraph of an article, then click the "Name ==>" button to store that <u>name</u>.
 - For authors, you will mostly be capturing names from the story's byline, at the top of the article just below the headline.
 - If there are multiple people in a byline, each person should be coded separately. For example, if the byline is "Jeffrey Kaczmarczyk and Rachael Recker / The Grand Rapids Press", first code "Jeffrey Kaczmarczyk" and then "Rachael Recker" as two separate people.
 - Select all of the name as it appears in the byline, to make it easier to authors together across articles. Make sure to include leading initials if they are present ("B. Patrick Murphy", for example).
 - Bylines can include additional information that indicates the affiliation of the reporter, usually separated from reporter names by a forward slash or some kind of dash (" / The

- Detroit News" or "/ Special to the Grand Rapids Press"). Do not include this identifying information with the name. That information goes in the "Title".
- If you accidentally capture the wrong text, just select the correct text and click "Name
 ==>" again.
- o To empty out the Name field, click the "Name ==>" button with no text selected.

Person Type

- Select the "Author" Person Type.
 - Select "Author" for any reporters in the article, either captured from the byline of the
 article or from elsewhere if there is another place where reporters are credited. For
 example, if reporters are credited at the end of an article as having contributed
 reporting, they too should be included as authors for the purposes of this study.

Organization

- Capture the <u>Organization</u> of the author.
 - For this study, the *Organization* of a given reporter is any additional information included in the byline that that indicates the affiliation of the reporter.
 - This is usually separated from reporter names by a forward slash or some kind of dash ("
 / The Detroit News" or "/ Special to the Grand Rapids Press"). Do not include characters
 like slashes or dashes with the title text.
 - See "Example bylines" below for more specific examples.
 - O The Organization field is more flexible than the Name field. Like the Name field, when you select text and click the "Organization ==>" button, that text will be placed in the Organization field. If there is already something in the Organization field, however, the selected text will be appended to the end, rather than replacing the existing text. You can also edit the text in the Organization field.
 - o If you accidentally capture the wrong text, you can just select the incorrect text in the Organization field and delete it.

Lookup Name and Process Person

- Use the Lookup Name field to see if the name from the article refers to any people already entered in the database, and if so, to associate this reference to the existing person. For more details on how the "Lookup Name" field works, see "How to lookup a person" below.
- Once you are done capturing a given person's information, click the "Process Person" button at the bottom of the form to capture the person's information and empty the form for the next person. Each time you process a person, a status message above the person list will tell you the processing is complete and the person will be added to the "Person List" at the bottom of the coding area. See "Bob Dukesherer" example image below.



- To remove a person from the "Person List", click the "Remove" button at the right end of the row of the person you want to remove in the "Person List".
- To edit their information, click their name to load their coding back into the coding form. When
 you are done, you can click the "Process Person" button again to save your changes. If you want
 to discard your changes, click the "Clear Person Coding" button. Your existing coding will remain
 unchanged, but any changes you made since you loaded the coding back into the form will be
 cleared.

Example bylines

Some example bylines, and how you'd code them:

- "Leonard Fleming"
 - o Author *name* is "Leonard Fleming", no *title* information present.
- "Greg Chandler / The Grand Rapids Press"
 - o Author name is "Greg Chandler", title would be "The Grand Rapids Press".
- "Nardy Baeza Bickel and Nate Reens / The Grand Rapids Press"
 - The two author names are "Nardy Baeza Bickel" and "Nate Reens", each coded as a separate person. Each would get the title "The Grand Rapids Press".

- "Christine MacDonald / Special to the Detroit News"
 - The author *name* is "Christine MacDonald". The *title* text would be "Special to the Detroit News".

Coding Subjects

Each person explicitly identified by name in a story is captured as a <u>subject</u>. When you code subjects in a story, you will capture the following:

Name

- Capture the <u>Name</u> of the subject. You will select the complete text of the person's <u>name</u> in the article, then click the "Name ==>" button to store that <u>name</u>.
 - Select all of the name as it appears in the article.
 - Make sure to include leading initials if they are present ("B. Patrick Murphy", for example).
 - If there is a formatting problem with the name, just select the full name as it appears in the article, and you can correct the problem when you try to lookup the person in the database. For example: "Scott Pal- czewski". Select the entire name including the hyphen and space as the name, then after you fetch the name for lookup, delete the hyphen and space so you are looking for the proper name in the database (for more details on Person lookup, see "How to lookup a person" below).
 - o Do not include titles or other related identifying information.
 - o If there is confusion as to whether certain words are part of a subject's name, just capture the subject's last (or "given") name.
 - For example, with the name "DJ Kalendr", if you are unsure if "DJ" is part of the name, just capture the last name: "Kalendr".
 - Do not capture general, non-specific references to people.
 - For example, in "commissioners said", even though "commissioners" are people, they are not explicitly identified as individuals, and so would not be considered a subject or a source for this study. Individual commissioners who were named would be subjects, and sources if they were directly quoted.
 - If you accidentally capture the wrong text, just select the right text and click "Name ==>"
 again.
 - To empty out the Name field, click the "Name ==>" button with no text selected.

Person Type

• Select the appropriate *Person Type* for the subject:

- Subject (not quoted) Select this person type if the person is mentioned in the article, but not quoted in any way.
- Source (quoted) Select "Source (quoted)" if the person is quoted, either directly or indirectly, in the story. To be considered a source, the source of information must be explicitly named and directly quoted, which usually means being the subject of a verb that depicts direct or indirect communication. Direct communication in news stories is indicated by verbs of attribution such as "said," "reported," "stated," and "noted," known collectively as "said" verbs. Indirect communication is indicated by verbs related to mental states such as "hopes," "feels," and "believes." These verbs could be in past or present tense. Phrases like "According to" are also considered evidence of attribution. Examples:
 - Direct quote: information from the person is contained in quotation marks, attributed using a "said" verb.
 - Example direct quote of Pete Goodell:

"We went from green grass to a lot of snow," said Pete Goodell, a manager at Pando. "We still had to make a lot with the machines, but it was a great start.

- Indirect or paraphrased quote: information from the person is paraphrased, attributed using a "said" verb.
 - Example indirect quote of Bob Dukesherer (note present tense "says"):

Bob Dukesherer, a National Weather Service meteorologist, says it's a near certainty that will happen by Tuesday, when a storm will settle in and hang over the region through Saturday.

- Mixed quote: a mix of paraphrase and direct quotation, attributed with a "said" verb.
 - Example mixed quote of Greg Holzman:

"It really is a supply-and-demand issue," said Dr. Greg Holzman, the Department of Community Health's chief medical executive, noting the agency feels it already has reached many of the people who would have been hard hit by the illness.

- Implied attribution: could take many forms, but in general, the person is introduced, then a quote directly follows. The quote is implied to be attributed to the person, but no said verb is present.
 - Example implied quote of Gary Solis (no said verb):

Gary Solis is a law professor at Georgetown University and a former Marine prosecutor. "I think there are a number of cases that never make it to the reporting stage, and in some that do make it to the reporting stage, there has been a reluctance to pursue them vigorously."

Organization

- Capture *Organization* information for the subject.
 - o For this study, we will capture the name of the primary organization with which the subject is affiliated in the *Organization* field.
 - You should aim to capture as complete and accurate a name of the individual's organization as possible, but only capture information that is contained within the article. If a person is said to be a member of the "state Senate", for example, even if you know the state in which the subject serves, just capture "state Senate" as the Organization.
 - If a given source has multiple organizational affiliations represented in the article and the primary affiliation is not obvious, capture all of the related organizations in the *Organization* field.
 - In some instances, you'll need to pull together information from throughout the article to make a complete *Organization*:
 - for legislative bodies, include name of locality represented if known (Michigan state House, Michigan Senate, Park Township Board, etc.).
 - for people affiliated with geographical governments, include the type with the name if known (so "city of Lansing" rather than just "Lansing", "Park Township", etc.).
 - for names of companies, include as complete a company name as is referenced in the article. For example, the company "Integrated Fabric Resource" is referred to later in an article as "Integrated Fabric". For Organization, you should capture "Integrated Fabric Resource".

Examples:

• In the following:

"The authority also appointed John Nunneley as its legal counsel. Nunneley headed the bond practice of Detroit law firm Miller, Canfield, Paddock and Stone PLC."

 Nunneley is a lawyer, and so could have many clients through his law firm, with those clients likely changing over time.
 Because of this, his primary affiliation would be with "Miller, Candield, Paddock and Stone PLC", not with the authority for which he is legal counsel.

- Pete Goodell works for the Pando Winter Sports Park and is referred to in an article as "a manager at Pando." The good Organization value for this person is "Pando Winter Sports Park" (the full name of the organization with which the person is primarily associated). To capture this:
 - select "Pando Winter Sports Park" and click "Organization ==>".
- To minimize the potential for typing words wrong, please select text and either use the "Organization ==>" button or copy and paste whenever possible.
- The Organization field is more flexible than the Name field. Like the Name field, when you select text and click the "Organization ==>" button, that text will be placed in the Organization field. If there is already something in the Organization field, however, the selected text will be appended to the end, rather than replacing the existing text. You can also edit the text in the Organization field, so you can manually add details if the organization name is broken up in such a way that you need to add information to make it specific or accurate.
- If you accidentally capture the wrong text, just select the incorrect text in the Organization field and delete it.

First Quote

- If the person is a source, capture the first sentence of their First Quote.
 - Capture a single entire sentence in which the source is quoted by name. This can be either a direct or indirect quote, and it can be a sentence that contains other information, as well.
 - To start, try to find and capture the entire sentence of a person's first direct quote (a quotation where the person's exact words are enclosed in quotation marks and attributed to the person using a "said" verb see more details above in the "Subject: Person Type section").
 - If there is no direct quotation of the person, then look for the person's first indirect
 quote (where information from the person is paraphrased and attributed using a "said"
 verb see more details above in the "Subject: Person Type section")
 - o Example:

On the other hand, Ottawa County Sheriff's Sgt. Keith Koeman, who commands the force's marine patrol, urges boaters to familiarize themselves with the law for drinking while operating a boating

"One beer on the water is equal to three on the shore," Koeman said.

- The verb "urges" in the first sentence could be considered an attribution verb, making that sentence an indirect quote. But, in this case, the second sentence is a direct quote, and so you would capture the second sentence instead.
- If the only quotation is an ambiguous one like the sentence above with "urges" and you
 can't otherwise tell if the reporter spoke with the person, mark them as a subject, not a
 source.
- See "Appendix: First Quote Examples" for more examples.

- The field to hold the source's first quote, First Quote Text, appears when you select
 "Source (quoted)" as the Person Type. Select the entire text of the person's first quote,
 then click the "First Quote Text ==>" button to store it.
- If you accidentally capture the wrong text, just select the right text and click "First Quote Text ==>" again.
- To empty out the First Quote Text field, click the "First Quote Text ==>" button with no text selected.

Lookup Name and Process Person

- Use the Lookup Name field to see if the name from the article refers to any people already entered in the database, and if so, to associate this reference to the existing person. For more details on how the "Lookup Name" field works, see "How to lookup a person" below.
- Once you are done capturing a given person's information, click the "Process Person" button at
 the bottom of the form to capture the person's information and empty the form for the next
 person. Each time you process a person, a status message above the person list will tell you the
 processing is complete and the person will be added to the "Person List" at the bottom of the
 coding area. See "Bob Dukesherer" example image below.



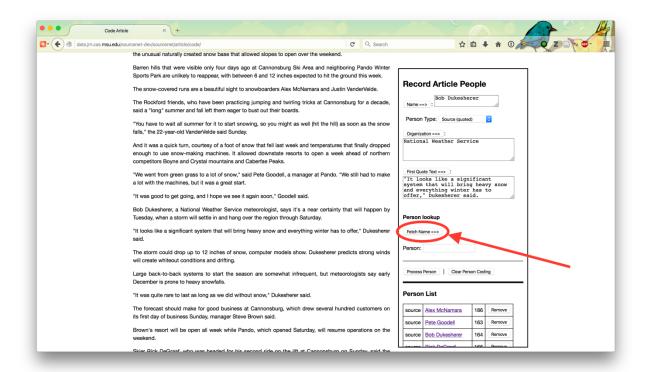
- To remove a person from the "Person List", click the "Remove" button at the right end of the row of the person you want to remove in the "Person List".
- To edit their information, click their name to load their coding back into the coding form. When you are done, you can click the "Process Person" button again to save your changes. If you want to discard your changes, click the "Clear Person Coding" button. Your existing coding will remain unchanged, but any changes you made since you loaded the coding back into the form will be cleared.

To lookup a person in our database, you'll use a lookup field that accepts names, full or partial, searches for them in the database, then returns a list of people who match.

You'll start by fetching and looking up the name exactly as it appears in the article. If there is a match, you'll check to make sure it looks correct, and if so you'll confirm the match and move on. If there is no match for the full name as it appears in the article, then you'll try a few basic searches for parts of the name (last name only, just first and last name if there was a middle name, etc.). If these searches don't result in a match, you'll just process the reference without choosing an associated Person, and a new Person database record will be automatically created when the coding is submitted to the server.

To lookup a person:

• To start, once you have selected and captured the name of an author or source, click the "Fetch Name ==>" button in the "Person lookup" area of the coding sidebar to pull the person's full name from the Name field above down into the Person lookup field:



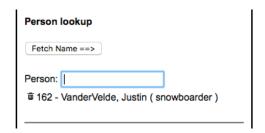
• The name from above will be pulled down into the Lookup field and submitted to the server. If an exact match is found, a box will pop up below the Person lookup field with an ID, a name, and any title for the person that matched. It will look like this:



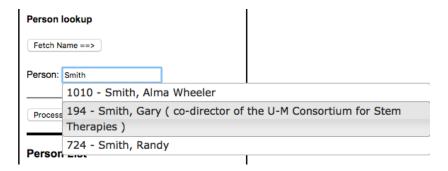
• If this person's details look correct, click that box that popped up (containing the ID, name, and title) to associate the name in this article with that person:



• Once you click to choose a person, the person will appear below the "Person" lookup field, as in the image below:



- If you select a name in error, you can click the trash can icon at the left of the person and the person will be un-selected and removed.
- If the name of the person as it appears in the article doesn't result in a match (no box will pop up at all), then you can edit the text in the "Person" lookup field to try to find a match:
 - Start with just the last name in the "Person" lookup field. This will bring up all the people with that last name. For example, "Smith":



o If there are matches, see if any look like matches for the rest of the name parts (if there is a middle name in the story, for example, perhaps just first and last are present). As

- you move your mouse over the results, the person the mouse is on will turn gray (see above). If one is a match, move your mouse over it, then click to associate the reference with that person.
- You can also search for other complete name parts (just the first name, just the last name).
- o If there is a middle name (for example, "Alexander Graham Bell"), try searching for just the first and last name ("Alexander Bell").
- o If there is the potential for a spelling error, you can try different spellings. For example, change "Smirth" to "Smith".
- o If you've tried all the above and not found a match, move on.
- If you do not find a match, just don't select anyone from the lookup, then once you are done with all the other coding, click "Process Person" to record the person and move on. When a person is submitted with no person from the database associated, a new person will be automatically created in the database.
- NOTE AGAIN: If there is any doubt about the person with whom a quote in an article should be associated, don't link the name with a person from the database. It is easier to go back later and detect and merge multiple Person records for the same actual person than it is to figure out when quotes from multiple people have been incorrectly attributed to a single person.

Once you are done, submit your coding

• Once you are done coding a given article, scroll to the very bottom of the coding box on the right side of the page and click the "Submit Article Coding" button. The page will reload and at the top, you will see a status message that indicates whether your coding was saved or not.

Illustrative coding examples

• If there is a quote where more than one person "said" something, that quote can be used as an indication that each of the people is a source, and could be the first quote for any person who has no other individual quote of their own (we are looking for the first place where a given person is quoted by name). In the example below, a single quote indicates that both Alex and Justin are sources. It could be the first quote for both IF there are no subsequent quotes attributed to just one or the other. If there is a subsequent quotation that is attributed to just one or the other by name, use it as that person's first quote. The example:

The snow-covered runs are a beautiful sight to snowboarders Alex McNamara and Justin VanderVelde.

The Rockford friends, who have been practicing jumping and twirling tricks at Cannonsburg for a decade, said a "long" summer and fall left them eager to bust out their boards.

• In the following indirect quotation:

He said Gov. Jennifer Granholm has said if districts do not seek ways to cooperate and consolidate services, then there could be a 5 percent reduction in state aid in the 2006-07 budget.

Even though there is a "said" verb next to Jennifer Granholm, the "He" is the one who is quoted here, describing what Jennifer Granholm said. She is a subject, but not the source in this case.

More to come as we code and find examples that should be mentioned.

Coding Application

The coding web application is online at

http://data.jrn.cas.msu.edu/sourcenet/sourcenet/article/code/. Each coder must have a username and a password to access the coding page. Your username and password will be provided to you with your list of articles to code.

Appendix: First Quote Examples

As an example, we'll look at first quotes in the following article:

CANNON TOWNSHIP -- A winter storm bearing down on West Michigan this week is expected to pump up the unusual naturally created snow base that allowed slopes to open over the weekend.

Barren hills that were visible only four days ago at Cannonsburg Ski Area and neighboring Pando Winter Sports Park are unlikely to reappear, with between 6 and 12 inches expected to hit the ground this week.

The snow-covered runs are a beautiful sight to snowboarders Alex McNamara and Justin VanderVelde.

The Rockford friends, who have been practicing jumping and twirling tricks at Cannonsburg for a decade, said a "long" summer and fall left them eager to bust out their boards.

"You have to wait all summer for it to start snowing, so you might as well (hit the hill) as soon as the snow falls," the 22-year-old VanderVelde said Sunday.

And it was a quick turn, courtesy of a foot of snow that fell last week and temperatures that finally dropped enough to use snow-making machines. It allowed downstate resorts to open a week ahead of northern competitors Boyne and Crystal mountains and Caberfae Peaks.

"We went from green grass to a lot of snow," said Pete Goodell, a manager at Pando. "We still had to make a lot with the machines, but it was a great start.

"It was good to get going, and I hope we see it again soon," Goodell said.

Bob Dukesherer, a National Weather Service meteorologist, says it's a near certainty that will happen by Tuesday, when a storm will settle in and hang over the region through Saturday.

"It looks like a significant system that will bring heavy snow and everything winter has to offer," Dukesherer said.

The storm could drop up to 12 inches of snow, computer models show. Dukesherer predicts strong winds will create whiteout conditions and drifting.

Large back-to-back systems to start the season are somewhat infrequent, but meteorologists say early December is prone to heavy snowfalls.

"It was quite rare to last as long as we did without snow," Dukesherer said.

The forecast should make for good business at Cannonsburg, which drew several hundred customers on its first day of business Sunday, manager Steve Brown said.

Brown's resort will be open all week while Pando, which opened Saturday, will resume operations on the weekend.

Skier Rick DeGraaf, who was headed for his second ride on the lift at Cannonsburg on Sunday, said the day's moderate temperatures made for a nice day on the hill.

"It makes Michigan living a good thing," the 59-year-old from Rockford said.

First quotes:

- Alex McNamara (compound "The Rockford friends... said" here is this person's only quote):
 - The Rockford friends, who have been practicing jumping and twirling tricks at Cannonsburg for a decade, said a "long" summer and fall left them eager to bust out their boards.
- Justin Vandervelde (this is VanderVelde's first direct quote):
 - o "You have to wait all summer for it to start snowing, so you might as well (hit the hill) as soon as the snow falls," the 22-year-old VanderVelde said Sunday.
- Pete Goodell (direct quote):
 - o "We went from green grass to a lot of snow," said Pete Goodell, a manager at Pando.
- Bob DukeSherer (direct quote):
 - o "It looks like a significant system that will bring heavy snow and everything winter has to offer," Dukesherer said.
- Steve Brown (indirect quote, with partial title):

- o The forecast should make for good business at Cannonsburg, which drew several hundred customers on its first day of business Sunday, manager Steve Brown said.
- Rick DeGraaf (direct quote):
 - o "It makes Michigan living a good thing," the 59-year-old from Rockford said.