*The WhatEvery1Says (WE1S) project uses digital humanities methods to study public discourse about the humanities. The WE1S project team is assembling an English-language primary corpus of publications (including newspapers, magazines, and television and radio transcripts) to collect full-text digital articles that will be ingested for text analysis.*

*Corpus scoping work has identified publications and databases that will be useful to building the primary WE1S corpus. Each RA is responsible for one or several “areas of focus,” to ensure that WE1S is achieving representativeness in its main corpus (these areas, for example, include global nations and regions, alternative/indie sources, and news sources targeted toward various racial, ethnic, and gender groups.)*

*In addition to this primary corpus, WE1S will also be producing in the future smaller sub-corpora, which will likely include historical newspapers, scholarly journals, speeches and mission statements, government and political documents, and Spanish-language publications.*

*As we have been completing research for our primary corpus areas of focus, the WE1S RAs have brought up thought-provoking questions and reflections about representativeness, canonicity, source availability, and the source collection process. This form is meant to provide a space for RAs to keep track of their reflections, which may in the future provide a foundation for more in-depth research reports.*

*Please answer the following questions about your area of research focus as you proceed with your research, and feel free to add additional reflections to section 6 below. This report will eventually be published on the WE1S public Web site.*

*Please avoid fancy formatting (special fonts, colors, tables, spacing adjustments, etc., since this Word document will eventually be converted into Markdown before being posted online.*

*This document: version 1, 19 Feb. 2018*

## Collection Area of Focus Report Template

### 1. Overview

#### What is your area(s) of focus?

#### Why is this area of focus important to the WE1S corpus?

### 2. Source Scoping Process

#### How have you been selecting sources for the WE1S corpus? (e.g. collecting from particular databases, using “impact” lists, etc.)

#### If you are using external lists to guide your selection of sources, include links here and indicate who produced them, for what purpose the list was produced, and any potential bias issues involved.

### 3. Corpus Representativeness

#### How representative do you think your corpus is? (“Representativeness” can be interpreted and addressed in a number of ways, so tailor it to be most productive for your area.)

#### What challenges in achieving representativeness have you encountered?

#### Provide a tally breakdown of the various facets of sources in your area of focus that WE1S is considering as possible measures of overall corpus "representativeness" (for example, by source or media type, nationality, region, political orientation, identification with specific racial, ethnic, and gender audiences, etc.). (The following is an example):

Distribution Method

Print: 277

Broadcast: 33

Online: 474. Reflections

#### What challenges or difficulties have you encountered in the source selection or collection process? Do you anticipate any challenges emerging from your work going forward?

### 5. Research Scan

#### Conduct some preliminary research on the questions or challenges that you provided in sections three and four.

#### Have other scholars reflected on these issues? Are there publications that address these problems? Has research been conducted on how to overcome these challenges or at least acknowledge them productively?

### 6. Additional Comments/Reflections

#### Include any other issues or questions that you have encountered that may not fit into any of the above categories.