Requirements

REQ-1.1.1 The NLP model will accept a source of text to be analyzed. Ambiguity: what form may the text take (link, tweet, plaintext, news page, etc)?

REQ-1.1.2 The NLP model will identify colonialist language within a given body of text.

REQ-1.1.3 The NLP model will report the results of its analysis to the user. Ambiguity: How does this report occur? What is displayed?

REQ-1.1.4 Optional: The NLP model will suggest decolonialized alternatives to colonialist language uses within an analyzed body of text.

REQ-1.2.1 The web scraper will collect the requisite data from a given website.

REQ-1.2.2 Requisite data (as described in REQ-1.2.1) may take the form of a news article, peer-reviewed study, or social media post.

REQ-1.2.3 The web scraper will store the data collected from sources for use in training the NLP model.

REQ-1.2.4 The web scraper will pass the data collected from sources for analysis by the trained NLP model.

REQ-1.2.5 The web scraper will accept a target to be scraped. Ambiguity: how flexible must the scraper be in what it can collect data from?

REQ-1.2.6 The web scraper will parse a web page via traversing the displayed HTML to collect requisite text. Ambiguity: should the scraper traverse the DOM or only parse raw HTML? Ambiguity: what tags are relevant to the scraper's collection process?

REQ-1.2.7 The web scraper will collect data from a given web page upon activation from the user interface by a given user.

REQ-1.2.8 The web scraper will complete its collection of data for a given website within 2s of user initiation.

REQ-1.2.9 The web scraper will provide visible feedback to the user upon completion of its processing.

REQ-1.3.1 The NLP model will accept input of text from the web scraper application.

Requirements that will likely need to change to reflect changes in our project: 1.2.2, 1.2.8