

CSCI E-2A Assignment 9. Due December 7 at 11:59 pm EDT. Please email your answers to cscie2a@googlegroups.com Please attach your work as a file, either a pdf (preferred) or a .doc file. Please name the file as follows: Lastname_Firstname_Assignment_5.ext (e.g., Moore_Tyler_Assignment_9.pdf).

1. (9 points, 3 for each edit description) Wikipedia has 3 core content policies which “determine the type and quality of material that is acceptable in Wikipedia articles”. They are:

1. Neutral point of view (<http://en.wikipedia.org/wiki/NPOV>)
2. No original research (http://en.wikipedia.org/wiki/Wikipedia:No_original_research)
3. Verifiability (<http://en.wikipedia.org/wiki/Wikipedia:Verifiability>)

These editorial policies are enforced regularly by Wikipedia's many editors. They are especially important for maintaining Wikipedia's standard of quality and for dealing with controversial topics. Find one such controversial topic on Wikipedia, and view the edit history (by clicking on the 'History' tab at the top of the page). Find three edits that violated each of the core content policies (one edit for each policy). Describe the edit, why it violates the policy, and state how long it took before the edit was corrected to reflect the core policies.

2. Given Wikipedia's success, it is no surprise that many other applications are trying to enlist the help of Internet users to carry out collective tasks. One such enterprise is called PhishTank (<http://www.phishtank.com/>). Phishing websites impersonate banks in order to carry out identity theft. PhishTank asks people to submit suspected phishing websites to their site. It also lets anyone vote on whether they believe submitted websites are phishing. Once the list has been vetted with enough votes, it is passed along to web browsers and used to stop users from visiting phishing websites. Visit the website and try it out.

A. (3 points) Wikipedia has to protect against inaccuracies and potential bias from its authors. What threats does PhishTank need to protect against?

B. (3 points) Are Wikipedia's content policies suitable for PhishTank? If so, why? If not, what policies would you advocate?

3. (3 points) For those that did not attend the class on December 1 (either in person or watching live online), please answer the following question.

According to its Wikipedia page, the Wikiscanner (<http://wikiscanner.virgil.gr/>) “consists of a publicly searchable database that links millions of anonymous [Wikipedia](#) edits to the organizations where those edits apparently originated, by cross-referencing the edits with data on the owners of the associated block of [IP addresses](#).” The Wikiscanner has been used, for instance, to find that someone from the New York Times edited George W. Bush's Wikipedia entry by adding the word 'jerk' 12 times. Wikiscanner has been praised as a tool for increasing the transparency of Wikipedia edits.

Do you think the transparency afforded by Wikiscanner is a good thing, and if so, what are the limits? Should a similar tool be developed to identify the location of people who post comments to websites and blogs? What would the advantages and disadvantages of such an approach?