Regarding the sentiment analysis performed to determine which platform would be best to build a suite of apps for aid workers in remote locations, I've found some interesting conclusions.

To review, I first performed a sentiment analysis for the phones and operating systems in question on the common crawl of the world wide web. While the common crawl is currently at 5 billion pages and counting, I performed a sentiment analysis on a smaller section, a still large number of approximately 150,000 web pages. This is a large enough sample to ensure that I didn't have a biased sample by accident to give misleading results. Remember: a sentiment analysis analyzes the word content of sentences of web pages to determine if a sentence is “positive” or “negative” pertaining to the subject in question, in this case a variety of cell phones and mobile operating systems but specifically iPhone and Galaxy handsets.

First, I trained a model on known data to achieve the highest accuracy possible, so that I could be confident that the sentiment determined from the large data matrix that we gathered from the common crawl was accurate. Next I compiled the data mined from the common crawl into a clean matrix before applying the model and reaching a conclusion. Both the iPhone and the Galaxy sentiment analysis were performed using the same process.

The result of the sentiment analysis was that both the iPhone and the Galaxy handsets have a positive sentiment on the internet. The graph below illustrates our result, where anything above a “3” on the scale is positive (though a “5” would not be more positive than a “4” in this analysis), and anything below would be negative, with 0 and 3 being and indeterminate sentiment.

As far as confidence, the models for determining sentiment were 77 and 79% accurate for the iPhone and the Galaxy, respectively, which represents a very accurate estimation. Thus were are confident in the sentiment conclusion reached by the model and suggest moving forward with its result. One caveat is that some of what was captured would have certainly been advertisement, but that is anticipated to be incredibly small in relation to all of the unbiased opinion found within 150,000 web pages.

Now we can ask better questions: given that the iPhone and the Galaxy have similar sentiment, what else might be considered when deciding who to partner with? Two considerations are cost and ubiquity, of both handsets and the operating system. While the iPhone and iOS might be the big name item in North America, the iOS software development kit is $99 per year. Further, the Android operating system is the dominant mobile platform around the world, accounting for 85% of the phones shipped. Thus in building a suite of apps for aid workers, availability of replacement phones and compatible operating systems must be considered, especially given the remote places these workers will be stationed. With that in mind, given the goals of this app suite and the target users, I would strongly suggest considering partnering with Samsung and the Galaxy instead of Apple and the Iphone.