Improving Human/Autonomous System Teaming through Linguistic Analysis

Erica L. Meszaros

Eastern Michigan University and NASA Langley Research Center
May 13, 2016

Background

An area of increasing interest for the next generation of aircraft is autonomy and the integration of increasingly autonomous systems into the national airspace. Such integration requires humans to work closely with autonomous systems, forming **human and autonomous agent teams**. The intention behind such teaming is that a team composed of both humans and autonomous agents will operate better than homogenous teams. Procedures exist for licensing pilots to operate in the national airspace system and current work is being done to define methods for validating the function of autonomous systems, however there is no method in place for assessing the interaction of these two disparate systems. Moreover, currently these systems are operated primarily by **subject matter experts**, limiting their use and the benefits of such teams. Providing additional information about the ongoing mission to the operator can lead to increased usability and allow for operation by non-experts. Linguistic analysis of the context of verbal communication provides insight into the intended meaning of commonly heard phrases such as "*What's it doing now?*" Analyzing the semantic sphere surrounding these common phrases enables the prediction of the operator's intent and allows the interface to supply the operator's desired information.

Methods

To provide quantitative evaluation metrics, a proof-of-concept was carried out for the application of Latent Semantic Analysis (LSA) to evaluate human/chatbot teaming. By running the University of Colorado at Boulder's online LSA tool [1] over data from the 2014 and 2015 Loebner Prize competitions [2] of human/chatbot teams, numerical scores were obtained that could be correlated with scores provided by human judges during the competition. Using linear regression over this correlation, formulae were obtained for predicting the score of this interaction. These formulae were then tested over the 2013 Loebner Prize transcripts [3] to determine predictive power. Current research is also being carried out to gather data from verbal interactions between humans and drones that allows for better defined semantic space and provides topic space-specific lexica for analysis. Human communication was observed and recorded during interaction with drones in the Autonomy Incubator at NASA Langley Research Center. A new LSA tool was developed in the statistical computing language R [4] to analyze the observed communication and develop a map of the semantic space of the Autonomy Incubator.
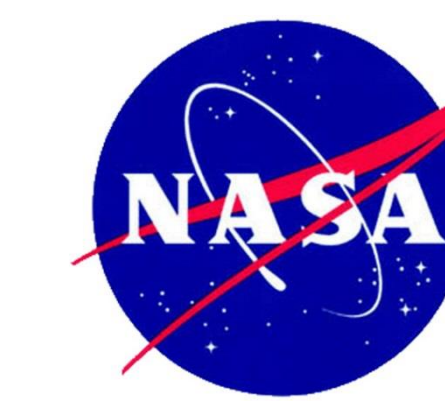
Results

Term Comparison Predicted vs. Actual Score

Document Comparison Predicted vs. Actual Score
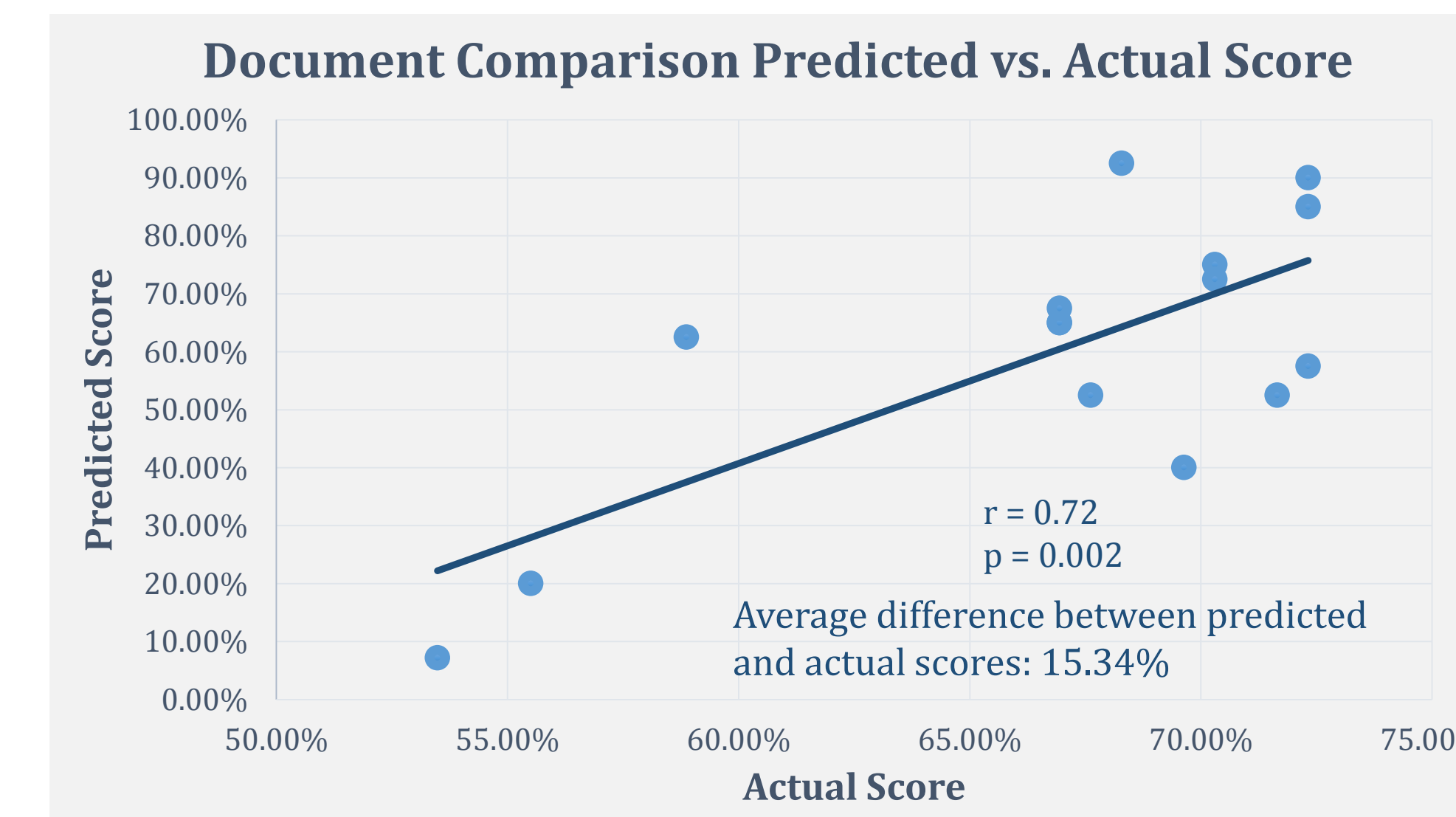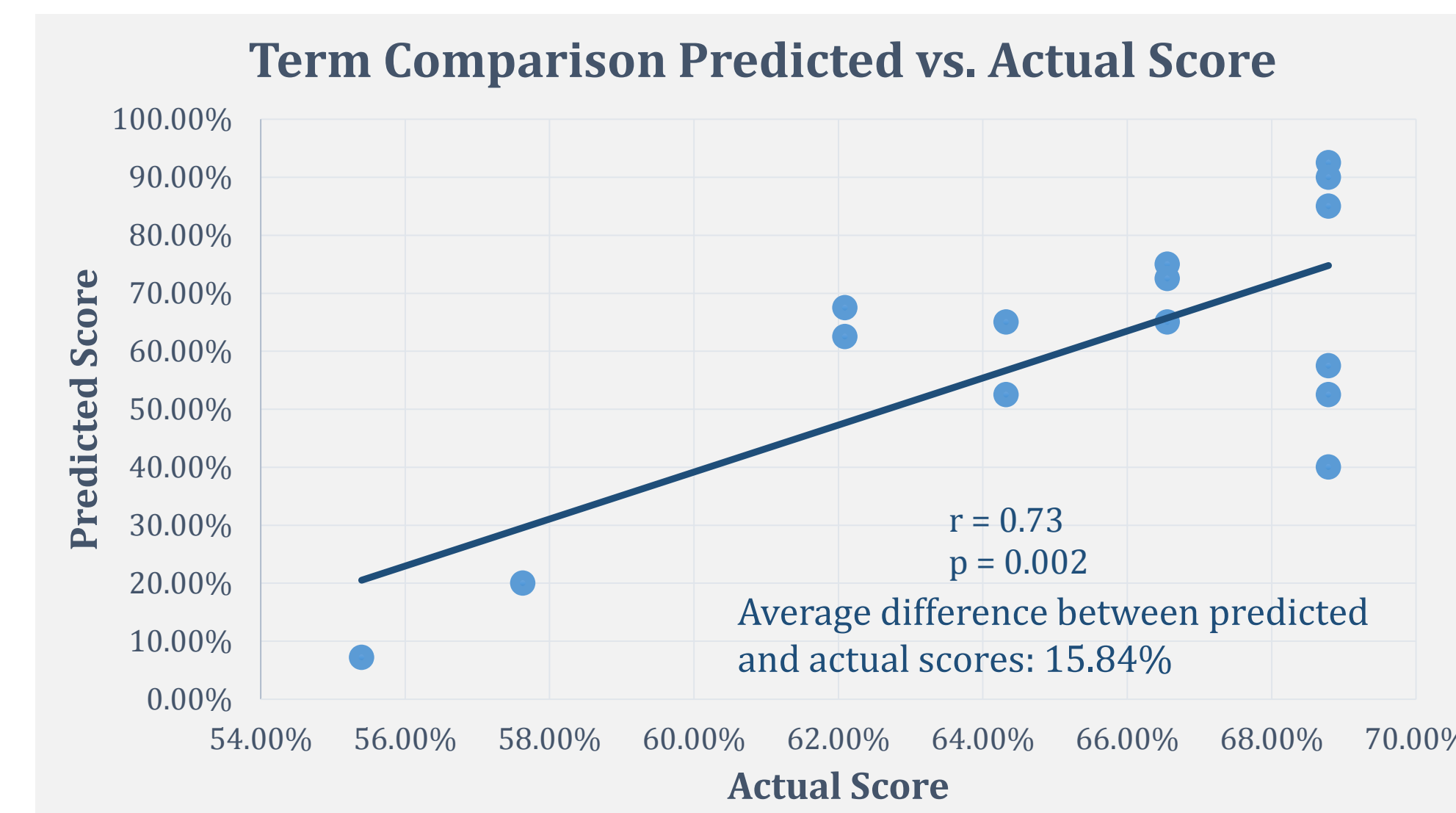
Both term comparison and document comparison LSA provide high correlation between predicted and actual scores, with term comparison producing an r=0.73 and document comparison producing an r=0.72. However, these categories have a similarly high absolute difference between predicted and actual scores, 15.84% and 15.34% respectively. Though there is strong correlation between the predicted and actual scores, their absolute difference suggests that the predictive success of this method is not strong. The limited amount of data and reliance on stock LSA tools may contribute to the limited success of this analysis. The use of specialized topic spaces and lexica, as well as larger data sets, to enhance linguistic analysis may improve the predictive power of these metrics.

To provide such improvements, observations of human communication during drone operations were made over six days, totaling 1589 utterances. These utterances were analyzed using a new R-based LSA tool, which uses singular value decomposition to generate eigenvalues which were then scaled into two-dimensional space using multidimensional scaling. The resulting data points were then clustered using k-means clustering (k=5) into general semantic categories. See the below graph for the clusters and suggested categories.
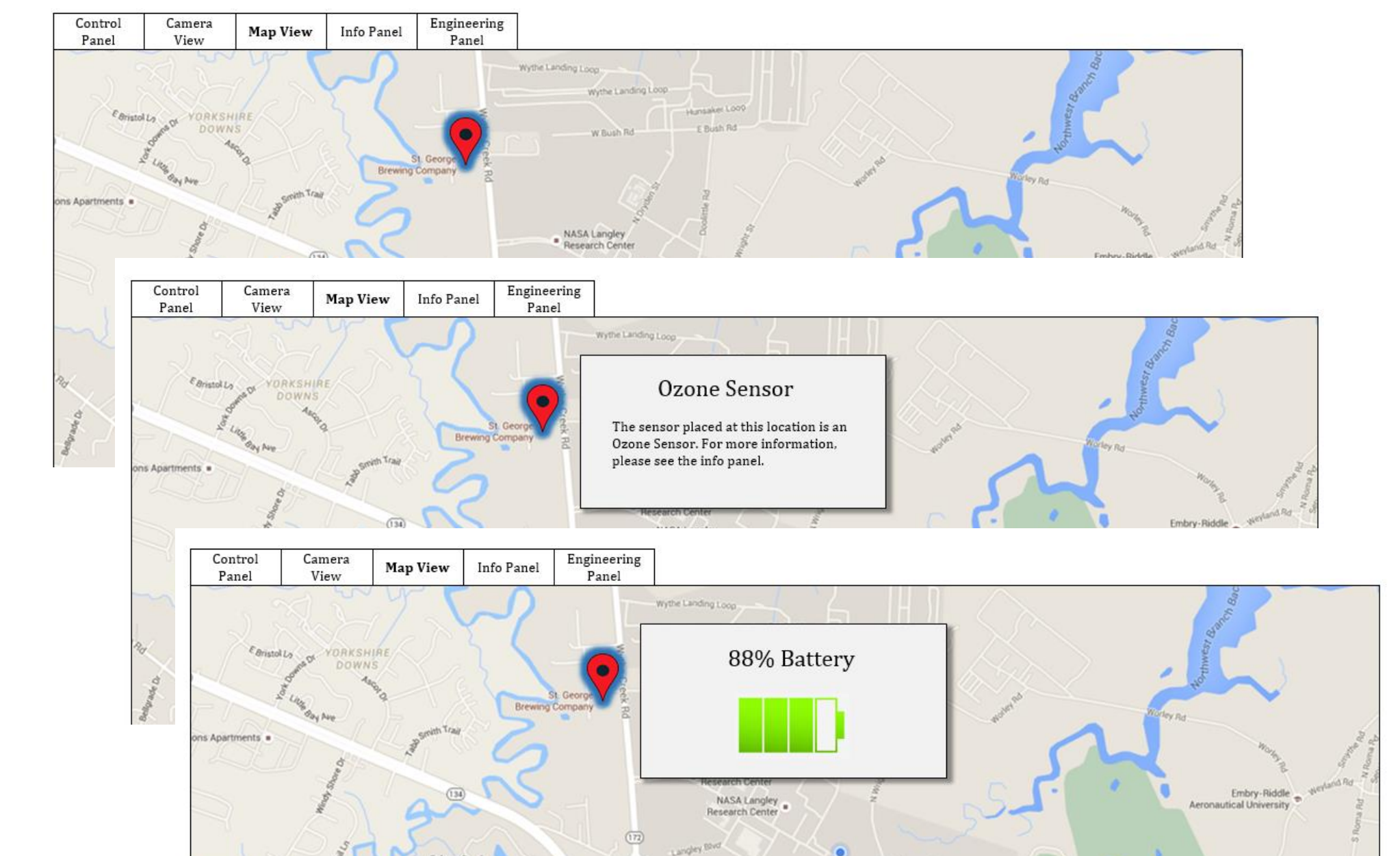
Semantic Map of the Autonomy Incubator
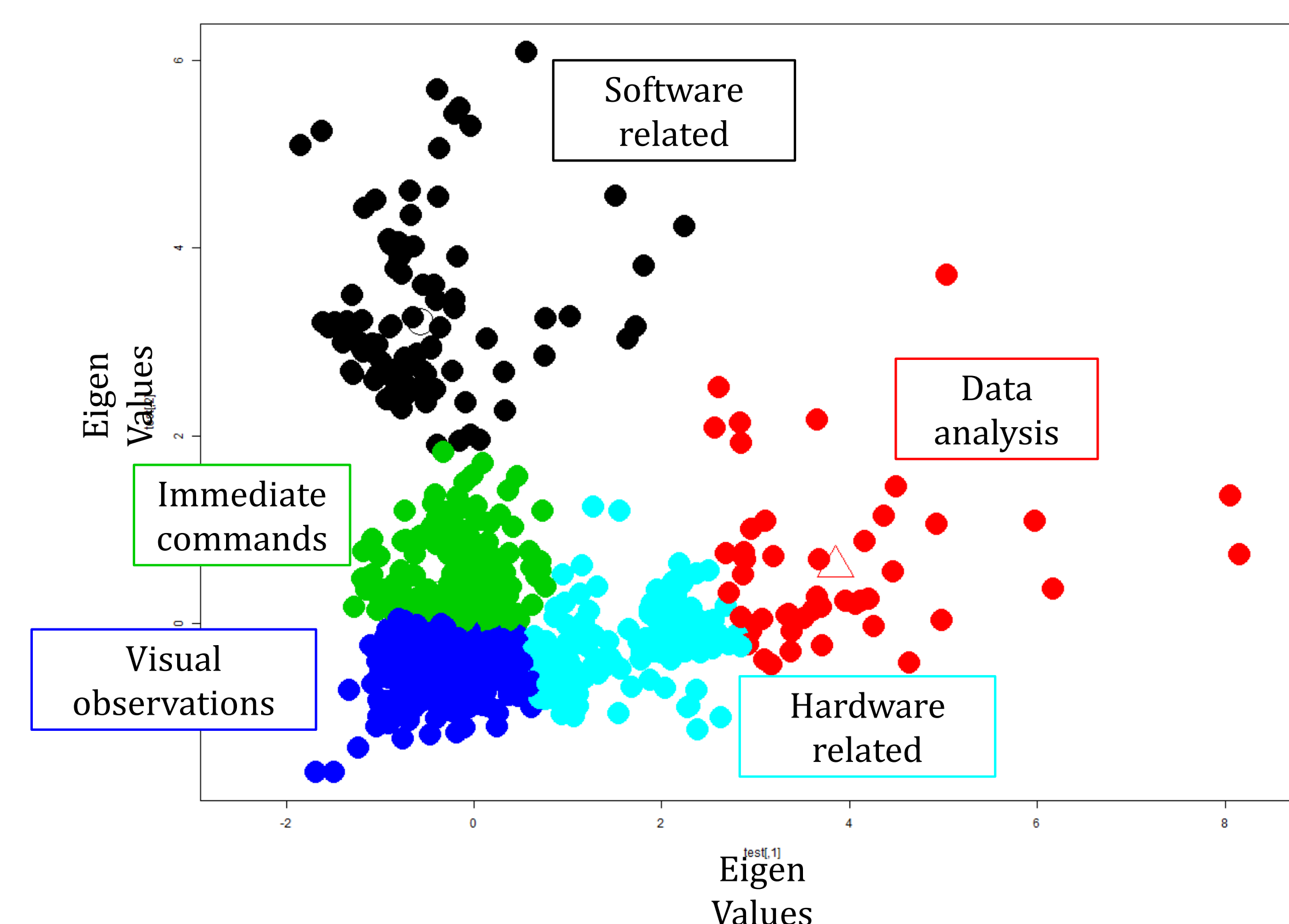
Keywords from Each Category

want, see

work, turn

right, now

script, run

check, data

Continuing Work

Having established a map of the semantic space of the Autonomy Incubator and determined the contained semantic categories, the semantic meaning of new utterances can be determined by mapping them to this existing space. By including this functionality with the drone interface, information can be provided to the user based on generic questions such as "*what is it doing now?*" The above image displays a scenario where a user has selected a portion of the display showing where a drone has placed a sensor, then queries "*Which sensor is that?*" followed by "*How's its battery life?*" Information on user queries and information provided can be saved to continuously improve the background semantic map.

Conclusions

This research presents a novel use for standard natural language processing techniques, leveraging tested methods of communication evaluation in autonomy and human factors research, fields that are currently seeking quantitative metrics. Beyond the use of LSA for evaluating human/autonomous system interaction, linguistic tools can also serve to **increase trust** in the system by enabling transparency and mitigating frustration. Linguistic analysis can also serve to **increase the usability** of an autonomous system by enabling natural language communication and feedback, allowing these systems to be used in more fields and by more people.

Works Cited

[1] http://lsa.colorado.edu/
[2] http://www.aisb.org.uk/events/loebner-prize
[3] http://www.loebner.net/Prizef/loebner-prize.html
[4] https://www.r-project.org/about.html