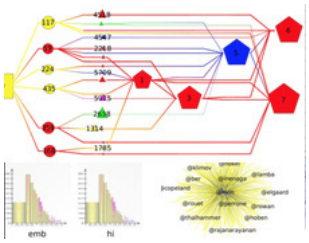
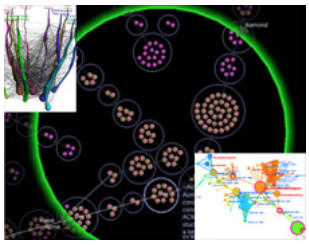
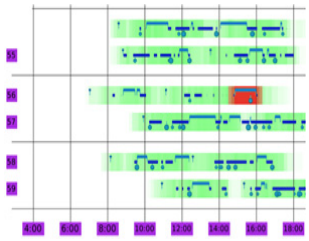
IEEE VAST 2010 Challenge

JUDGING CRITERIA







Judging Process and Criteria for Judging (DRAFT)

Submissions will be reviewed by *judges both from the analytic community and from the visualization community*. All individual reviews will be collected and provided to participants.

Both *qualitative and quantitative* scores will be computed and provided to each team individually.
All ratings will be based solely on the materials provided by the participants i.e. short and detailed answers, video and - for the Grand Challenge - debrief.

Accuracy metrics will evaluate the correctness of the data table answers based on the known "ground truth" embedded in the dataset. Participants will be given points for correct answers and penalized for incorrect answers. The correct answers will be based on the "ground truth" embedded in the data. For example when we asked for the participants in a certain activity, they will be given points for finding those who did participate and penalized for missed participants or for identifying irrelevant people. If participants report additional suspicious elements that were not part of the known ground truth, the analysts who created the datasets will review these elements for legitimacy. If they are deemed to be legitimate (i.e. supported by evidence), those elements wll be added to te ground truth and accuracy recomputed.

Due to the increased number of submissions, a two step judging process will be followed.

1st level peer review

Entries will be sent out for an initial peer review (similarly to paper reviews).
Reviewers will be asked to first view the video, read the materials provided by the teams, and then provide the following ratings using a 1-7 scale - and providing comments:

1. Clarity (Does the submission clearly explain how Visual Analytics tools were used to analyze the data and scenario?)
2. Analytic process
3. Visualizations
4. Interactions
5. For Toolkits only: If the tool used is a toolkit, rate the effort needed to customize the toolkit for this specific analysis. Consider such things as programming ability required, amount of time needed.
6. Novelty (How much novelty do you see in this submission (data processing, visualization, interaction, hypothesis generation or evaluation, overall process, etc…)
8. Overall rating

Remember that during these assessments the judges will not be able to ask the participants any questions so the clarity of the explanations provided in the text is critical. The judges cannot correctly assess something they do not understand.

2nd level review

After the 1st round of reviews, the top entries will be reviewed in more detail by a team composed of the challenge committee and professional analysts. This team will make the final decision about awards.

The list of awards is not determined in advance but most likely will resemble those given in the past. Examples of secondary criteria which could be used in award nominations if a submission particularly stands out with a novel approach includes:
• Scalability (i.e. are some aspects of the analysis automated? Are the results of the automatic processing understandable and believable? Are there mechanisms to guide the analysis or the use of the tools?)
• Versatility (i.e. did the team demonstrate that the tools could be used for multiple data types and situations?)
• Data integration (can data from various sources be integrated smoothly?)
• Handling of missing data and uncertainty
• Support for collaboration
• Learnability
• Reporting (how is the preparation of the debrief supported)
• Other features such as a history mechanism, the ease of importing and exporting data, innovative features in general, and others.

Questions? **Send email to challengecommittee AT cs.umd.edu**