Interface Design for Analysts in a Data and Analysis-rich Environment

PRESENTED BY
Karin Butler, Human Factors
Improving Analysis, Search, and Visualization of ‘Big Data’

Analysis of complex, multivariate time series data from tests of device performance

Stevens Adams, Haass, Matzen, & King (2016)
## Allocation of Function

<table>
<thead>
<tr>
<th><strong>Human Strengths</strong></th>
<th><strong>Human Weaknesses</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Creative</td>
<td>Memory</td>
</tr>
<tr>
<td>Adaptive</td>
<td>Declarative</td>
</tr>
<tr>
<td>Goal Driven</td>
<td>Retrospective</td>
</tr>
<tr>
<td>Multipurpose</td>
<td>Prospective</td>
</tr>
<tr>
<td>Multimodal</td>
<td>Procedural</td>
</tr>
<tr>
<td>Teachable</td>
<td>Computation</td>
</tr>
<tr>
<td>Efficient</td>
<td>Vigilance</td>
</tr>
<tr>
<td>Empathetic</td>
<td>Endurance</td>
</tr>
<tr>
<td>Intuition</td>
<td>Distractible</td>
</tr>
<tr>
<td>Will-to-Live</td>
<td>Speed</td>
</tr>
</tbody>
</table>

Allocate computational functions to automated systems to be more quickly and accurately performed.
Full Autonomy vs. Human Computer Integration

I, for one, welcome our new computer overlords. -Ken Jennings (2011) on being beaten by Watson on Jeopardy!
Challenges of Partial Automation

Characterization of Human Interactions with Automation

- Use/disuse is influenced by mental workload, trust, and risk
- Misuse caused by violations of the assumption of the automation
  - Lack of understanding of automation requirements
- Abuse: Use of automation negatively affects human user, redefining their role as a by-product of the automation

Parasuraman & Riley (1997)
Growing Collections of Data and Analyses
## Incorporating Statistical Learning and Other Data Science Techniques

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Neural Nets</th>
<th>SVM</th>
<th>Trees</th>
<th>MARS</th>
<th>k-NN, Kernels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural handling of data of “mixed” type</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>Handling of missing values</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>Robustness to outliers in input space</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>Insensitive to monotone transformations of inputs</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>Computational scalability (large N)</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>Ability to deal with irrelevant inputs</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>Ability to extract linear combinations of features</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>Interpretability</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td>▼</td>
</tr>
<tr>
<td>Predictive power</td>
<td>▼</td>
<td>▼</td>
<td>▼</td>
<td></td>
<td>▼</td>
</tr>
</tbody>
</table>

From *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* by Hastie, Tibshirani, and Friedman (2016)
Binary Code Reverse Engineers

Simulating how a computer system would process instructions and inputs.

Workflow
◦ is hypothesis-driven and discovery-based
◦ demands high cognitive load
◦ demands integration of information across days/weeks
◦ could utilize more automated tools for asking specific question about code

Mental models are complex and include uncertainty.
Growing Collections of Data and Analyses
Human Factors Consideration: Selection and Use

How are computational partners integrated into a workflow?

◦ What knowledge do users require to select computational tools?
◦ How cognitively demanding is selection?
◦ How cognitively demanding is use?
◦ How are outputs used to understand the system or situation?
Situational Awareness

Three components
1. Perception of the elements in the environment within a volume of time and space
2. Comprehension of their meaning, and
3. Projection of their status in the near future.

Complex, dynamic domains (military command and control, emergency services, aviation)

Does not fully capture the need to integrate information across time to make decisions.
20 Guidelines for the Design of Human-Autonomy Systems (Endsley, Bolte, & Jones, 2003) focusing on three areas 1) support of human understanding of autonomous system 2) minimize the complexity of the autonomous system and 3) support of situation awareness.
Sensemaking

Sensemaking is the ability to give meaning to an ambiguous situation, one with high complexity or uncertainty, in order to make a decision.

Seven types of sensemaking processes in data-frame theory

- Connect data to a frame
- Elaboration of the frame
- Questioning the frame
- Reframing
- Preserving the frame
- Comparing frames
- Seeking the frame: Searching for a new representation of hypothesis

Mental models = Data + Frame

Klein, Phillips, Rall & Peluso (2007)
Demands of Analyst Interfaces in an Analysis-rich world

◦ Use: Influenced by mental workload, trust, and risk
  ➢ Workload
    ➢ selecting the right analysis
    ➢ evaluating analysis results and
    ➢ integrating analysis results
  ➢ Trust
    ➢ using tool correctly, all required parameters
    ➢ uncertainty in output
Demands of Analyst Interfaces in an Analysis-rich world

◦ Misuse: Failures that occur when people inadvertently violate critical assumptions and rely on automation inappropriately
  ➢ Provide feedback on critical assumptions
Heuristics for Trusted Autonomy

Visibility of current systems behavior
Visibility of probable system behavior
Accessibility of system rationale
Visibility of system capabilities & limitations
Awareness of latencies & delays
Transparency of failure
Fit with users and operations.

Jackson, Prasov, Vincent, & Jones (2016)
Using Computational Tools for Sensemaking

Outputs of tools are mapped to mental representations or hypotheses that are consistent with those data.

- Which mental representations are appropriate?
- How can we support easier mapping?
How Can the Work of Computational Partners be Communicated to Human Users?

Hypothesis-Driven Knowledge Sharing

Knowledge-Transfer Graph

Zhao, Glueck, Isenberg, Chevalier, & Khan (2018)
Conclusions

Consideration of how analysis-based automated tools will be used within unique workflows

- What are the frames used to make sense of the workflow?
- What information is needed for hypothesis generation?
- What information is needed for discovery?
- What needs to be remembered in long-term memory?
- What needs to be remembered in short-term memory?
References


