Putting the "I" in Interaction: Interactive Interfaces Personalized to the Individual

Interactive data exploration and analysis is an inherently personal process. One's background, experience, interests, cognitive style, personality, and other sociotechnical factors often shape such a process, as well as the provenance of exploring, analyzing, and interpreting data. This viewpoint posits both what personal information and how such personal information could be taken into account to design more effective visual analytic systems, a valuable and under-explored direction.

Exploratory data analysis necessitates user interaction. Users interactively examine different aspects of data in order to derive insights. Furthermore, such a process is almost certainly unique for each user. Such a personal exploration process is influenced by their educational background, experience with both the system and the data/domain, cognitive style of reasoning, the portion of the data that they are interested in exploring, and even their personality.

Consider two users with very different personalities who wish to explore the same dataset for the first time, using a new analytical system that they have never seen before. One user is adventurous and happy to begin clicking through the system menus and buttons to learn how the system works, while the other is cautious and methodical, and would prefer to first read the tutorial and learn the basics of the system. This sole difference in the personality of these users drastically changes their approach to using the system, and further will affect how they explore the data and how they interpret the results of their exploration.
Understanding, quantifying, and responding to such user personality traits is a relatively new and open research area in the context of visual analytics and exploratory data analysis. The end goal of such research is to create adaptive visual analytic systems that can suit the needs of each user. To reach this goal, additional research is needed to identify which user tasks and characteristics may influence the behavior of users and to what degree in a visual analytic process. Further, there is little active research in how a system can effectively learn about the personality traits of users from their interactions with a visualization. Each of these steps are necessary precursors to creating truly responsive, personalized analytical systems.

In this Viewpoint, we examine existing research on understanding users and their characteristics pertinent to interactive visual analytic systems. In doing so, we wish to highlight the role that provenance data can play in inferring the characteristics of users from their interactions, pertinent to performing interactive visual analytic tasks. Provenance data, in the context of our work, means any data collected during the analysis process --- steps taken, items examined, views generated, etc.

While similar work has been undertaken in the HCI community for building personalized systems, personalized visual analytic systems must consider both data and analytical challenges. In particular, accomplishing visual analytic tasks often requires that users interact with large amounts of data, select appropriate analytic approaches, and synthesize the results from different explorations. Unlike building other personalized systems (e.g., personalized movie recommender systems), personalized visual analytic systems face much greater challenges because of the greater task complexity and larger user cognitive burdens. Many existing visual analytic systems are already personalized, considering user tasks, user preferences, and even computing devices, such as desktops, mobile phones, and large screens. Our goal with this Viewpoint is to encourage existing research along this direction by considering additional characteristics, especially inherent user characteristics (e.g., cognitive style or personality).

We discuss the rationale behind incorporating user characteristics into visual analytical systems, and we speculate that such user characteristics can be learned dynamically. We outline user characteristics and personality traits pertinent to visual analytic tasks, relating these characteristics to active challenges in visual analytics. Additionally, we discuss the benefits and risks for varied intervention designs during the life cycle of visual provenance based on user intent and individual differences, and we indicate future research directions in these areas.

TOWARDS HYPER-PERSONALIZED VISUAL ANALYTICS

Considering the uses of personality and characteristics to support visualization and analytical processes, we note that users can be modeled from three key aspects. The first aspect is to understand user intent—based upon their interactions, can a model determine the interests or the tasks of a user? The second aspect is to anticipate user behavior based upon their previous interactions or inferred interests, such as the approach proposed by Bors et al. to iteratively build a provenance-based task abstraction framework. A third aspect is to infer users' innate characteristics, which may be most useful to assist with post-hoc analysis of system behavior. The correlational patterns that connect user characteristics to their behavior could be learned to improve future versions of a system.

We propose that all three of these aspects could be modeled together to establish a more comprehensive and accurate understanding of user characteristics—a hyper-personalized system. For example, the intent of a user may be inferred more accurately from their interactions and their personality. Likewise, predicting future user behavior could be based on both user intent and user personality traits. Not only could a system predict the next step a user is likely to take, but it could also predict a potential sequence of interactions. As a result, such systems could adaptively personalize based upon a better understanding of users, including their intent, behavior, knowledge acquired in the process, and innate characteristics.

In a similar vein, deriving intent or future actions from personality could be complemented by learning personality based upon the actions of a user. The semantic interaction approach serves as a useful model here; rather than interpreting user intent through captured interactions, a learning module could infer user characteristics based upon those same interactions. To motivate this discussion, we next describe how user intent
modeling is performed using semantic interaction, and how such modeling must be enhanced to support the nuances of personality characteristics.

**USER INTENT MODELING FROM INTERACTIONS**

Adaptive techniques personalize application inferences for users automatically by adjusting a system's settings in response to user behavior. In such systems, the data displayed within a visualization is modified based upon the interactions of a user, updating as the system gradually learns the data-centric interests and preferences of a user. A benefit of adaptive techniques is their ability to use provenance data in order to recommend future actions, visualizations, and interfaces. There are also systems that automatically infer user intent from user behavior and then recommend system settings based on the inferred user intent. Such systems still let users be the decision maker. For example, Gotz and Wen describe a system that can recommend suitable visualizations based on dynamic user interactions, while it is still up to the users to adopt the recommended visualizations.

The semantic interaction paradigm, often also referred to as demonstrational interaction, is a clear example of an adaptive technique. Systems such as StarSPIRE (Figure 1) permit users to interact directly with the visualization to demonstrate some intent, or alternatively to externalize knowledge from which intent can be discerned. Such systems can then infer the intent of the user from the interaction before updating some underlying parameters to reflect that intent within the visualization. We note that this intent does not necessarily need to be inferred; a system can use active learning to directly query the user in order to disambiguate.

![Figure 1. The StarSPIRE system uses a semantic interaction approach to adapt the layout of documents in the visualization based on user interaction.](image)

Focusing on intent modeling from interaction such as in the case of the semantic interaction approach, a necessary step is the inference phase that maps the user's interactions to the user's analysis goal. Many existing systems assume a one-to-one mapping. In StarSPIRE for example, dragging two documents closer together is assumed to reflect the user's desire to express higher similarity between the documents. However, in systems that allow more free-form and expressive user interactions (e.g., ActiveInk), such assumption of a one-to-one mapping might not always apply. Without clear-cut rules, that one approach to model this complexity is by considering the cardinality of both sides of this interaction/intent relationship.
One interaction implies one intent: This is the trivial case. For example, consider the direct manipulation of a control widget (click the button to submit the form). Performing such an interaction is directly mapped to a single intent.

Many interactions imply one intent: This case is demonstrated in flexible UI design. For example, there are many different interactions and interaction sequences that can be used to bold text in Microsoft Word (e.g., using the font dialog, clicking a button in a toolbar, using the keyboard shortcut).

One interaction implies many intents: This case is often an underspecified interaction: one interaction from a user could be inferred in several different ways, or one interaction from a user could simultaneously express multiple intents. As an example, the semantic interaction of directly repositioning data items in dimensionally reduced projections is clearly overloaded, as such a repositioning could be interpreted with respect to other data items at the source and/or the target of the interaction, or to everything in the projection. If the interaction is with respect to particular data items, it may be with respect to a single other data item or to a collection of items (e.g., a cluster).  

Many interactions imply many intents: This is the most interesting case, because the natural interpretation is flexible interaction design. The user could perform any gesture or interaction, and the system could use a set of meta-rules (or user behavior, or provenance, or any of the above) to infer the user's intent for the interaction. This naturally is the most difficult of these relationships to predict, but recent work in this area is proving its feasibility. For example, the Metatation system (Figure 2) combines a linguistic data-model with the recent sequence of user interactions (free-form annotations) to enumerate a list of user intents, which are then used to recommend next steps in analysis.

![Figure 2. Metatation](image)

Figure 2. Metatation infers multiple intents from reader annotations on a poem (right image), and presents suggested data visualizations, on for each intent, as scrollable tiles (left).

With regards to inferring user personality traits, the latter two relationships are most relevant. A single interaction performed by a user does not provide sufficient detail to infer personality, as a standalone interaction could be performed by an adventurous, tired, or inventive user. Examining a sequence of interactions might be used for this inference, potentially considering the speed, variety, or location of the interaction collection. Likewise, the behavior of the user may change within a single interaction session if the user is temporarily distracted by another analyst, briefly excited by a new insight, or increasingly uncertain of the quality of their analysis. As such, a system attempting to infer personality traits requires the same flexibility.

Additionally, personality traits could be used to disambiguate user intent when existing paradigms like semantic interaction contain overloaded interactions. For example, this additional channel of information could
consider the knowledge (provided or inferred) that a user is adventurous or timid when determining the magnitude of change to make to an underlying model in response to an interaction, which would in turn affect the magnitude of the update to the visualization output.

PERSONALITY TRAITS AND COGNITIVE STATES

User modeling is an active research area, with a rich body of existing research demonstrating how to learn user characteristics, personality traits, and cognitive states. However, an open question remains: which user personality traits would impact visual provenance? We propose a collection of such traits in the sidebar, which we believe are most pertinent to our goal of enabling better visual provenance.

These user characteristics can be separated into transient and persistent traits -- sometimes referred to as cognitive states and traits. For example, one's level of stress or frustration is an example of transient characteristic. As a user is interacting with a system to explore their data, they may get stuck at a dead end in the analysis, not understand how to use the system to perform the analysis that they hope to achieve, or might even be working against a tight deadline. Being able to detect such frustration or stress and adapt the system to provide guidance or temporarily reduce the cognitive load on the user would be beneficial.

In contrast, a characteristic such as adventurousness is often a persistent personality trait. As discussed in the sidebar, these persistent traits can include characteristics that can be developed over time, as well as more permanent traits that reflect the mean behavior of a user. We suggest in the Introduction that highly adventurous users may be glad to explore a new system to understand its capabilities. In contrast, less adventurous users may be more hesitant and require tutorials or guidance in order to begin to use a system to its full potential. Accommodating the needs of both types of users within an adaptive system leads to a generalizable tool.

USER PERSONALITY AND EXISTING RESEARCH CHALLENGES

Beyond real-time learning of user characteristics, a number of existing challenges in visual analytics have ties to the characteristics of users. When introducing hyper-personalized user modeling into such systems, we are provided with new avenues to both interpret and address some of these open questions. At a high level, we feel that future systems should support the ability to deliver different data in different ways for different types of users.

For example, the detection, assessment, and mitigation of bias is an area of active research in visual analytics. While bias can also result from sampling, data processing, and visual mappings, cognitive bias is most relevant in this space. Human limitations such as working memory capacity, along with analyst-specific considerations such as internal predisposition towards the data or a hypothesis, will affect the conclusions that are drawn by those analysts while they use the system. Identifying and mitigating these biases in an adaptive system by inferring the characteristics of a user will improve the quality of conclusions drawn, so long as the bias detection is accurate.

Concurrently, ensuring complete information and analysis coverage of the data is important to mitigate such cognitive bias. In reality, some forms of bias are necessary in analytical systems. For example, if a user is interested in exploring a hypothesis, then the presented data should be relevant to that hypothesis. While intellectually curious users may naturally tend towards exploring both sides of the hypothesis, other users could potentially be steered towards drawing biased conclusions if the system presents an inadvertent imbalance between information that supports and refutes the hypothesis.

Detecting if a user is in need of help or reassurance could come from a number of sources when interacting with complex systems, including but not limited to issues with understanding the system itself, the data, the visual output, or the effects of an interaction. Assisting users with low need for cognition may prove necessary to assist them in interpreting complex visualizations. Reducing the rate of information flow to a lower
level is one means of mitigating this challenge, giving the user fewer items to focus on until they have recovered. In contrast, a user with high locus of control likely will not wish such severe interventions to occur. A further discussion on intervening to assist a user is provided in the next section.

There are risks to poor intervention selection. For example, another method for responding to a user in need of help is to present some tips for how to continue with the analysis. This heavy-handed intervention may be received poorly by some users. As an example, consider the reputation of Clippy from the Microsoft Office product line. This effect is magnified if the tips are irrelevant to the current issue faced by the user, removing them from their workflow to respond to an agent that is providing more harm than help. As we expand upon in the next section, avoiding "Clippy-like" irritation is crucial to designing useful adaptive systems.

**INTERVENTIONS**

After determining in response to some detected behavior that there exists a need to intervene between the standard system operation and the user, a system designer must determine how best to intercede in order to assist the user. Using the case of frustration detection for an example, the system could respond to detected frustration by scaling back the amount of displayed data, locking out the functionality of some subsystems, providing helpful prompts to the user for next steps, and more. These interventions could be realized by either augmenting the user interface or the analytic models, with an inherent tradeoff in these options.

Interventions that are **handled by the interface** are more transparent to the user, allowing them to identify helpful suggestions or avoid unnecessary actions. These interventions can also take on many levels of visibility, a spectrum ranging from applying subtle visual scent to suggest controls through locking out those controls. Because of their front-facing property, these interventions risk causing an interruption to the user's workflow as they switch from, say, performing data analysis to responding to the intervention.

The ideal solution to this issue is to design non-interruptive interventions in such a way to provide assistance without affecting the workflow of the user. One possibility could involve a dedicated area of the interface that can be populated by the most recent recommendation, which in turn fades from view over time if it remains ignored. A history of such recommendations could also be accessed if the user explicitly requests them at some point in their analysis. Such an interface can be further learned and customized for individual users via training. For instance, SUPPLE® learns a new interface layout from the user's "offline training" with the interface, producing a layout that is optimal for the user.

Interventions that are handled by the underlying models may not be as obviously-presented to the user, which may mean in some cases that the user sees changes to the interface and/or functionality without understanding their cause. Returning again to the example of detecting frustration, an action by the system to reduce the amount of information displayed could be interpreted by the user as a global absence of information supporting a hypothesis rather than a temporary, artificial limitation imposed by the system. Further, it is possible that the user may not actively notice the system's attempt at an intervention at all, making the operation ineffective or inconsequential if the user remains frustrated after the system reduces the quantity of presented information. However, these interventions are also less likely to interrupt the user's workflow, potentially making them lower risk to introduce.

Making the model interventions more apparent to the user will make their existence more clear, but again with a greater risk of workflow interruption. Techniques from Explainable AI research with regards to visual model transparency may prove beneficial to optimizing the intervention information conveyed without interruption (i.e., introducing non-intrusive ways to communicate the inferences of the system to the user). Similar to explaining the "black box" of machine learning systems in AI, the same methods can be used to explain what user visual interaction behavior is used to predict user characteristics and make recommendations and how reliable the results might be. Moreover, the rationales can also be explicitly represented (e.g., as a set of rules) and presented to the users, which can then be tuned or reconfigured to perform future interventions.

Interventions can also be risky, especially if there is a lack of information about user tasks and/or user behavior. In the absence of such information, it is even more important to explicitly explain the rationales behind the interventions as described above. Users can then decide whether they wish to accept the intervention and what level of interventions might be used. For example, certain users prefer to be guided during their tasks while other users are more independent and do not wish to be interrupted during their task processes. When
user characteristics are not known, we need a means to quickly gain the necessary knowledge to inform a model. Early work shows that it is feasible to learn from a brief conversation in the absence of a user's demonstrated behavior.\textsuperscript{13}

Note that it is not always the case that these model interventions have negative consequences for user understanding. For example, Battle et al.\textsuperscript{3} use predictive prefetching to support large-scale data exploration. If the prediction was correct, the user received the benefit of a "faster" response from the system; otherwise, they experience the default (not-prefetched) response time. In this sense, making use of past user interactions can only improve the user experience.

**FUTURE RESEARCH OPPORTUNITIES**

With all of these characteristics, opportunities, and challenges in mind, we now consider the path forward towards introducing user personality modeling into visual analytics systems. Given the large number of possibilities, we focus this discussion on the need for contributions in five broad topics.

**Training and Model Longevity:** Determining the appropriate role of model training is critical to the process of inferring user characteristics. At one end of the spectrum, a single pre-trained model is applied to a diverse pool of users, with their characteristics inferred from this sole source. While this is the most generalizable approach, it may not be the most effective means of responding to each individual user. At the other end, either no trained model exists and all learning about the user happens in real-time, or alternatively a model exists that can be continuously updated or adapted to the actions of the user. This second approach is much more customizable but is simultaneously much more difficult, as all information needs to be determined from the actions of the user.

**Signal Detection:** As a user is interacting with a system, they will inject a large amount of information into the system. Identifying which components of this information are useful for understanding the personality of the user and which are simply noise is essential for accurately and efficiently capturing the relevant user characteristics. While research exists in this space to address specific tasks like the recommendation of visualizations\textsuperscript{5} and models exist to extract higher-level intent from sequences of low-level interactions,\textsuperscript{2} no generalizable signal extraction technique currently exists.

**Balancing Instruction and Freedom:** When mitigating bias, there is the obvious risk that the system is also stifling the creativity and freedom of the user to explore data in the manner that they feel is best. Determining when to intervene and when to let the user continue along their current exploration trajectory is a difficult but important threshold to identify. Ideally, "quiet" interventions that do not affect the workflow of a user, but that can still provide assistance if necessary, can be presented.

**Impacts on User Experience:** Though initial steps in this space have been made, creating a complete hyper-personalized analytical system remains a future endeavor. When such systems begin to appear, evaluating their effectiveness with respect to today's state of the art will be necessary for understanding their impact on user experience and performance (e.g., correctness of analytic output, degree of user comprehension, or magnitude of accuracy improvement), as well as for addressing personality, accuracy, or over-correction issues. Determining the optimal means of performing this evaluation (and of measuring the accuracy of a personality model) is an open challenge.

**Ethical Implications:** The acts of inferring the personality of a user and introducing that information into software systems present numerous ethical challenges that must be addressed as such systems become available. Beyond the parallels to current AI training issues involving race and facial recognition, the security of storing personality information, the act of sharing personality information between systems, and the marketability of personality information represent just a few of the ethical concerns facing this emerging field.

**CONCLUSION**

This Viewpoint provides a discussion of topics from the broad range of possibilities that exist when introducing user personality modeling into visual analytics systems. By combining an understanding of the intent of a user, anticipating future user actions, and modeling users' innate characteristics, researchers have the ability...
to create hyper-personalized analytical and visualization systems that are more responsive to the needs and
goals of users. A wide variety of both transient and persistent user states and traits can be modeled in such
systems, but incorporating these user characteristics in such systems comes with many challenges that need to
be addressed by future research. We hope that the details and speculation contained within this Viewpoint
serve as inspiration and points of reflection for the future of these fields.

**SIDEBAR: CHARACTERIZING USERS**

This sidebar provides a sample list of user characteristics and their effect on the use of analytical systems.
This is not intended to be a complete list, but rather demonstrates the breadth of possibilities in this space.
Broadly, some of these characteristics change often, such as emotion-related characteristics, some character-
istics are more persistent over a person's lifetime, while some characteristics, such as cognitive capabilities,
can be developed over time with proper training.

**Transient Characteristics**

**Fatigue** -- A tired or fatigued user is more likely to make errors and is slower to react to a changing environ-
ment.\(^17\)

**Frustration** -- Frustration (and similar emotional states like confusion and anger) can affect productivity and
learning rates.\(^18\)

**Level of stress** -- Stress causes physiological changes which, among other effects, may cause changes in typ-
ing and mouse movement behaviors as a result of increased muscle tension, and can be affected by cognitive
load.\(^19\)

**Persistent Characteristics**

**Adventurousness** -- An adventurous user may be more likely to explore a system, being less likely to feel
disorientation during such exploration.\(^20\)

**Intellectual curiosity** -- A curious user is more likely to investigate information gaps in data presented to
them, leading to opportunities for thorough investigation.\(^21\)

**Inventiveness** -- Inventive users can adapt to changing data, produce innovative solutions to problems, and
favor exploratory processes.\(^22\)

**Locus of control** -- Users who feel "in control" of events have been shown to explore visual representations
differently from those who blame external factors.\(^23\)

**Need for cognition** -- Users with a low need for cognition do not often engage in or enjoy tasks that require
thinking, and may need assistance in interpreting visualizations.\(^24\)

**Can Be Developed**

**Perceptual speed** -- The speed with which a user can compare and locate figures and symbols can have an
effect on which visualizations they interpret most accurately.\(^24\)

**Spatial ability** -- Users with high spatial ability can produce, manipulate, and interpret visual representations
more accurately.\(^25\)

**Tolerance of uncertainty** -- Users are better at understanding uncertainty if it is clearly presented; otherwise,
they are likely to misinterpret both causes and implications of uncertain data.\(^16\)

**Working memory capacity** -- Users with high visual working memory have preferences for visualization
types, those with high verbal working memory benefit from faster response times, and those with high spatial
memory show increased performance.\(^26\)
ACKNOWLEDGMENTS

This Viewpoint is the result of discussion at Dagstuhl Seminar 18462: Provenance and Logging for Sense Making.

REFERENCES

15. E. Wall, L. M. Blaha, L. Franklin, and A. Endert, “Warning, Bias May Occur: A Proposed Approach to Detecting Cognitive Bias in Interactive Visual Analytics,” in


---

### ABOUT THE AUTHORS

**Remco Chang** is an Associate Professor of Computer Science at Tufts University. Contact him at remco@cs.tufts.edu.

**Christopher Collins** is an Associate Professor of Computer Science and Canada Research Chair in Linguistic Information Visualization at Ontario Tech University. Contact him at christopher.collins@ontariotechu.ca.

**Michelle Dowling** is a PhD Candidate in the Computer Science Department at Virginia Tech. Contact her at dowlingm@vt.edu.

**Alex Endert** is an Assistant Professor at the School of Interactive Computing at Georgia Tech. Contact him at endert@gatech.edu.

**John Wenskovitch** is a Visiting Assistant Professor in the Computer Science Department at Virginia Tech. Contact him at jw87@cs.vt.edu.

**Kai Xu** is an Associate Professor in Data Analytics with the Department of Computer Science at Middlesex University. Contact him at K.Xu@mdx.ac.uk.

**Michelle Zhou** is co-founder and CEO of Juji, Inc. Contact her at mzhou@juji-inc.com.
Contact department editor Theresa-Marie Rhyne at theresamarierhyne@gmail.com.