A Proposal On Leveraging Workflow Technology For Building Process Aware Visual Analytics System

Sharmin (Tinni) Choudhury
Interaction Design Center
School of Engineering and Information Sciences
Middlesex University, England
t.choudhury@mdx.ac.uk

ABSTRACT
Workflow analysis, conducted using both cognitive workflows and process workflows, has been employed to build and improve visual analytics systems. However, workflows and the visual analytics system have to date remained computationally separate. In this paper, we propose that workflow technology be leveraged to create process aware visual analytics systems. We argue that a process aware visual analytics system would be better able to support users, collect provenance information on user activity and track user decision pathways. This will enable visual analytics systems to become process.

Categories and Subject Descriptors
H.H.m [**Miscellaneous**]: Deals with the potential application of workflow technology to visual analytics..

General Terms
Design, Human Factors

Keywords
Visual Analytics, Workflow Technology, Process Aware Systems

1. INTRODUCTION
Within the domains of human-computer interaction and even visual analytics, the concept of workflows is not unknown. Hitherto, workflows have been used to model and analyze information processing capabilities of visual analytics systems, both from a computer process aspect and from the human cognition aspect. For example, workflow analysis has been used in Geographic Visual Analytics to identify bottlenecks when processing large datasets [1]. An example of cognitive workflow analysis, the work of Mirel et al, who conducted a case-based cognitive task analysis of a biomedical specialist's exploratory workflow as a step towards developing better visual analytics systems for biomedical research [2].

What is yet to be explored, however, is how workflow technology can be leveraged to improve and build visual analytics systems and support the analytical process by making visual analytics systems aware of the process to which it is being applied to. Therefore, in this paper, we first explain what is meant by workflow technology. We then explore potential avenues of leveraging workflow technology for building process aware visual analytics systems, before outlining some research challenges, which is followed by the conclusion.

2. WHAT IS WORKFLOW TECHNOLOGY
In typical application, workflow technology involves the modeling of relevant processes of an entity and executing that process with the aid of a workflow engine [3]. The purpose of the approach is to specify, execute, monitor, and coordinate the flow of work within a distributed environment [3]. In short, workflow technologies consist of the workflow itself, which models relevant processes, and the engine that executes the workflow to support information processing within a given environment [3].

In order to be computer readable and executable, the workflow is expressed in a workflow language - a formal expression for computational purposes [3]. There are a number of workflow languages that comprise of particular XML notation representing the inter-task dependencies [4]. The best known and the most widely used workflow language is the IBM backed Business Process Execution Language (BPEL) [5] and its extensions such as BPEL4People [6].

However, Yet Another Workflow Language (YAWL) [7] has been widely used in research and has been successfully used to introduce process modeling to various industries, such as the Film Industry with YAWL4Film [8,9], that fall outside the big business with enterprise systems that usually employ BPEL and are the traditional customers of IBM.

Both BPEL and YAWL have the potential to be used in visual analytics and further research needs to be conducted to find which is most suitable for application within a visual analytics system. However, workflow languages that render workflows machine readable are just one half the workflow technology equation, the other half is the engine that executes the workflow.

Workflow engines are usually designed to execute workflows written in specific workflow languages. IBM's Websphere can be used to execute BPEL-based workflows [10]. BPEL workflows can also be executed by open source engines such Apache Orchestration Director Engine (ODE) [11].

YAWL has its own engine [7] and indeed, with the exception of BPEL, a workflow language usually has a specific engine that goes with it. BPEL is an exception as IBM has promoted it as a standard and as a result, multiple commercial and non-commercial engines are capable of executing workflows written in BPEL. However, lack of choices in workflow engines is not a mark against a workflow language and workflow languages must be chosen based on their expressive power in regards to the processes needing to be modeled [12].

Embedding a workflow engines within a system forms a type of process aware system that manages and executes operational processes involving people, applications, and/or information sources on the basis of process models, i.e. workflows expressed in a workflow language [13,14]. Process aware systems tend to be more robust, more spatio-temporally aware, more adaptive to changes and generally more aware of human actors and agents who interact with the system and thus better able to support them and their interaction with the system [13].

3. APPLYING WORKFLOW TECHNOLOGY TO VISUAL ANALYTICS
As mentioned in section 1, Researchers involved in visual analytics are already conducting workflow analysis [1,2] but they

are not leveraging workflow technology. Merely using workflows to analyze and model processes. The workflow and the visual analytics system remain computationally separate. However, ideally, the workflow should be embedded within the visual analytics system.

Embedding workflow technology within a visual analytics system, would by definition make the visual analytics system process aware [13,15], i.e. the visual analytics system would be aware of the workflow the user is following. This would allow a visual analytics system to have two key capabilities that it is now missing: the ability to collect provenance information and to track decision pathways.

Provenance is term that comes from the domain of art and is a term applied to a collection of information that provides details of the history of an object. In the domain of visual analytics, "provenance" in its most basic form is the provenance of the data used for analysis. However, provenance can also extend to analytical techniques used on the data and can also reflect how the data plus analytical techniques were used by the analyst. This is important information to important to collect because without a complete provenance picture, it is difficult to build confidence in the results of an analytics process.

An process aware system with embed workflow engines can execute provenance workflows in the background, gathering metadata on the data and the analysis techniques being used and tracking the user through the process. The latter specially is difficult to do if the system is now aware of the process.

The gathering of provenance data is also a first step in the second important capability an embedded workflow technology can bring to visual analytics systems – the tracking of decision pathways, which essentially involves generating and audit trail for the analyst that can be followed at a later date to cross-check decisions made by the analyst.

The capability to generate an audit trail of user interaction with the visual analytics system is especially important for intelligence visual analytics systems as those decisions often have to be presented in the court of law, where it is not only important to show the final path that led to the decision but also paths that were explored by abandoned for some reason or another.

By being process aware, a visual analytics system is better able to track decision pathways; including the paths the user explored but ultimately abandoned as well as the final path that lead the user to the decision. This is because, a process aware visual analytics system would have a reference model, which would typically be a cognitive workflow of the users mental model, and thus have some awareness of the cognitive process of the user.

In other words, the broad research question that is yet to be addressed within the field of visual analytics is how both cognitive and information process workflows can be expressed in a formalized, computer-readable workflow language such as BPEL or YAWL and how these formalized, computer-readable can be executed by a workflow engine, embedded within a visual analytics system, to build process aware visual analytics systems with capabilities such as the ability to collect provenance data and track decision pathways.

4. RESEARCH CHALLENGES

There are a number of research challenges in this area. Firstly, it needs to be experimentally determined which workflow language is best suited for expressing the kinds of workflows needed for visual analytics systems.

Secondly, there is the question of modeling the workflows and expressing them in the chosen workflow language in a manner best suited to maximize the benefits to a visual analytics system.

Thirdly, what needs to be investigated is the desired affect when the workflow is executed by a workflow engine. For example, BPEL workflows are generally used to string together distributed software processes that come together to facilitate a task, e.g. booking a holiday – which involves purchasing tickets, booking hotels and maybe even pre-purchasing attraction tickets. However, while we may desire this kind of behavior in a visual analytics system, this is not all that workflow technology can do for a visual analytics system.

Lastly, what also needs to be investigated is how workflow technology be exploited to confer upon visual analytics systems the ability to gather provenance information and track user decision pathway by keeping an audit trail of user interaction with the visual analytics system.

5. CONCLUSION

This paper gave a brief introduction to workflow technologies and the concept of process aware information systems and suggested how workflow technology can be leveraged to build process aware visual analytics systems. We believe that this is a promising direct for visual analytics systems research that would allow visual analytics system to better support users and acquire new capabilities such as the ability to collect and compile provenance information and track the pathways a user explores as well as identify the pathway that leads to the final decision. This area is ripe for research and should be explored as part of the ongoing work on visual analytics systems.

6. ACKNOWLEDGEMENTS

We would like to acknowledge the EPSRC and the Home Office INSTINCT programme for their continued support of the Making Sense Project; this paper and its content forms part of the Making Sense Project SAR work package.

7. REFERENCES

[1] M. Kramis and C. Gabathuler, "Streamlined Workflow for Large-Scale Interactive Geographic Visual Analytics," *Geospatial Visual Analytics Workshop GIScience*, 2008.

[2] B. Mirel, F. Eichinger, B. Keller, and M. Kretzler, "A cognitive task analysis of a visual analytic workflow: Exploring molecular interaction networks in systems biology," *Journal of Biomedical Discovery and Collaboration*, vol. 6, 2011, pp. 1-33.

[3] C.A. ELLIS, "Workflow Technology," *Computer Supported Cooperative Work*, John Wiley & Sons Ltd, 1999, pp. 29-54.

[4] the G.W. Forum, "Workflow Description Languages."

[5] R. Khalaf, N. Mukhi, F. Curbera, and S. Weerawarana, "The Business Process Execution Language for Web Services," *ProcessAware Information Systems*, John Wiley \& Sons, 2005.

[6] M. Kloppmann, D. Koenig, F. Leymann, and G. Pfau, *WS-BPEL Extension for People – BPEL4People*, 2005.

[7] W.M.P.V.D. Aalst and A.H.M.T. Hofstede, "YAWL: yet another workflow language," *Information Systems Journal*, vol. 30, 2005, pp. 245-275.

[8] C. Ouyang, A. ter Hofstede, M.L. Rosa, M. Rosemann, K. Shortland, and D. Court, "Camera, Set, Action: Automating Film Production via Business Process Management," 2008.

[9] C. Ouyang, M.L. Rosa, A. ter Hofstede, M. Dumas, and K. Shortland, “Towards Web-Scale Workflows for Film Production,” *QUT ePrints Archive*, 2008.
[10] M. Kloppmann, D. König, F. Leymann, G. Pfau, and D. Roller, “Business process choreography in WebSphere: Combining the power of BPEL and J2EE,” *IBM Systems Journal WebSphere Application Server*, vol. 43, 2004, pp. 270-296.
[11] T.A. Foundation, “Apache ODE,” 2008.
[12] J. Jang, A. Fekete, P. Greenfield, and D. Kuo, “Expressiveness of Workflow Description Languages,” *ICWS*, 2006, pp. 104-110.
[13] M. Dumas, W.M.P.V.D. Aalst, and A.H.M.T. Hofstede, “Process-Aware Information Systems: Bridging People and Software through Process Technology,” 2005, pp. 235-255.
[14] W.M.P. Van Der Aalst, “Process-Aware Information Systems: Lessons to Be Learned from Process Mining,” *Transactions on Petri Nets and Other Models of Concurrency II*, vol. Lecture No, 2009, pp. 1-26.
[15] N. Russell, “Foundations of Process-Aware Information Systems,” *Language*, 2007, pp. 1-419.