

External Representation of Provenance in Intelligence Analysis

Ashley J. Wheat

1 INTRODUCTION

Visual Analytics systems allow users to gain insight and understanding of often large and complex datasets by coupling visualisation, interaction and computational power. Throughout the process of an analysis it can be important to keep a record of interaction with data and the analyst's reasoning process, to show how an insight was reached [1]. *Provenance* in this context refers to the historical account of the process of an analysis.

Key to the design and analysis of visual analytics systems with the embedded ability to track provenance, is an understanding of the way in which provenance is already represented by analysts through the use of external resources such as notes, whiteboards and computer tools. By gaining this understanding from the analyst's perspective, researchers and developers can design better interfaces [1].

We believe this is a fundamental element of our work, which focuses on generating an understanding of the way analysts elicit the use of external resources in support of sensemaking throughout an analysis. Through the study of analysts carrying out intelligence exercises, we aim to provide a conceptual framework of *Distributed Sensemaking* grounded in Distributed Cognition [3, 4, 7, 2] and sensemaking theory, communicated through elements of The Resources Model [15, 16].

In this workshop our wish is to stimulate a discussion—driven by a number of questions outlined at the end of this paper—centred around analysts' use of external resources in support of sensemaking, and how the visual analytics and HCI communities can approach research and design in this area.

2 BACKGROUND

There has been recent shift in the way we think about human cognition moving beyond the scope of the individual actor to a view of cognition involving the interaction of internal and external representations [17]. Work by Zhang & Norman [17], Scaife & Rogers [11] Larkin and Simon [5], and Hutchins [3, 4], among others, has sought to understand and describe the interplay between internal and external representations in human cognition.

2.1 Distributed Cognition

In his work on Distributed Cognition, Hutchins “extends the reach of what is considered *cognitive* beyond the individual to encompass interactions between people and with resources and materials in the environment” [2]. In so doing, Hutchins' work endorses the notion that the human cognitive system is ‘distributed’ in nature, and intelligent action comes as a result of an interplay of social interaction between actors and their interaction with external resources [7, 2, 9]. In their 2000 paper Hollan et al. unpack the Distributed Cognition approach into three “tenets”:

Socially Distributed Cognition which is concerned with the cognitive processes distributed across a social group.

Embodied Cognition which is concerned with the division of cognitive processes internally and in the external environment.

Culture and Cognition which concerned with the way in which cultural ecologies and social practices shape cognitive processes.[2]

It is argued that central to human intelligence is our ability to recognise our cognitive shortcomings and design our environment to overcome them. Therefore any inquiry into human cognition must include a theory outlining the extension of our cognitive performance through the appropriation external resources in our environment [17, 15]. However, according to Wright, Fields and Harrison, despite its relevance, there has been little adoption distributed cognition in HCI [15, 16]. This may be as a result of its lack of a set of categorical features to look for in an analysis, making it ineffective as a “quick and dirty” approach [15, 16, 8, 10, 14]. In response to this, The Resources Model proposes an approach to HCI research grounded in Distributed Cognition concepts [15, 16].

2.2 The Resources Model

The aim of The Resources Model [15, 16] is to provide a language and a set of concepts for HCI practitioners to study the interaction between people and resources by modelling Distributed Cognition in a way that can be practically applied to research in HCI. The Resources Model describes external cognition as *resources* that are drawn upon during user interaction. Resources can be represented internally (e.g. a set of memorised procedures) or externally (e.g. written instructions), informing action [8, 10]. Resources are classified as the following:

Plans are sequences of actions, events or states that to be carried out.

Goals are states of the world the user wishes to bring about.

Possibilities are the next possible action that can be taken according to the state of a system.

History which is comprised of the set of actions, events or states already achieved.

Action-Effect Relations are the causal effect of some event or action being carried out.

State refers to the to the values or information pertaining to the objects featuring in an interaction, at a given point.

The composition of resources comes as a result of *interaction strategies* such as *plan following* and *goal matching*. Furthermore, resource configurations are changed as a result of action in a cyclic process: when an action is taken, the configuration of resources is changed, which in turn informs the next action.

3 DISTRIBUTED SENSEMAKING

In ongoing work we are seeking to understand and describe the way in which intelligence analysts appropriate external resources in service of sensemaking. We argue that during the process of an analysis the analyst will leverage a number of resource configurations—represented through various external instruments—as a means of computational offloading. The aim of our work is to be able to identify the taxonomy of resources in the domain of intelligence analysis. Through the mechanism of The Resources Model and sensemaking theory, we will carry out an analysis of subject matter experts carrying out intelligence analysis exercises, discerning the role of external resources in the

sensemaking process, and outline the characteristics of resources in this context. As a result of the analysis of these studies we envisage the formation of a conceptual framework, which we call *Distributed Sensemaking*.

The overall aim of Distributed Sensemaking is not only to provide a set of concepts outlining the way in which intelligence analysts employ the use of resources in aid of sensemaking, but to deliver a framework that can be used in the analysis of HCI research and the design, development and evaluation of novel technologies supporting sensemaking.

4 PROVENANCE IN DISTRIBUTED SENSEMAKING

We argue that the external representation of provenance information is a key component in support of the sensemaking process. By eliciting resource configurations, provenance information can be represented externally, supporting “reflection-in-action” [12, 13]. This can support analysts by maintaining an external account of the history of an analysis, serving as an element of the Distributed Sensemaking paradigm. By achieving an understanding of the resource configurations at play in representing provenance, and through a description of the affordances provided by them, we can gain vital insight in the design and analysis of visual tools, better supporting naturalistic interaction.

5 POINTS TO DISCUSS

Following the position outlined above, we propose a discussion in this workshop centred around the external representation of provenance and its service of sensemaking in intelligence analysis. We wish to stimulate this discussion based on following questions:

1. Based on the notion that provenance can be examined on three levels, how might analysts elicit configurations of resources to represent provenance:
 - (a) **At a data level** taking into account that all data will have some source, and a path between this source and its use in analysis.
 - (b) **At the analysis level** accounting for the actions performed and techniques used in the analysis at a given point.
 - (c) **At the reasoning level** dealing with the way the conclusions in analysis have been reached. [6]
2. What impact do these external representations and resource configurations have on sensemaking in an analysis?
3. How might we approach a programme of research rooted in concepts outlined by The Resources Model to understand and describe the characteristics of configurations of resources representing provenance in support of sensemaking?
4. How can we design and develop visual tools which embody affordances for representing provenance supporting the work of intelligence analysis?

REFERENCES

[1] D. Gotz and M. X. Zhou. Characterizing users visual analytic activity for insight provenance. *Information Visualization*, 8(1):42–55, 2009.

[2] J. Hollan, E. Hutchins, and D. Kirsh. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, 7(2):174–196, June 2000.

[3] E. Hutchins. *Cognition in the Wild*, volume 262082314. MIT press Cambridge, MA, 1995.

[4] E. Hutchins. How a cockpit remembers its speeds. *Cognitive Science*, 19(3):265–288, July 1995.

[5] J. Larkin and H. Simon. Why a diagram is (sometimes) worth ten thousand words. *Cognitive science*, 99:65–99, 1987.

[6] J. Roberts and D. Keim. From Ill-Defined Problems to Informed Decisions. *EuroVis Workshop on ...*, pages 1–5, 2014.

[7] Y. Rogers. A brief introduction to distributed cognition, 1997.

[8] Y. Rogers. New theoretical approaches for human-computer interaction. *Annual Review of Information Science and Technology*, 38(1):87–143, Sept. 2005.

[9] Y. Rogers. Distributed cognition and communication. *The encyclopedia of language and linguistics*, 2:181–202, 2006.

[10] Y. Rogers. HCI Theory: Classical, Modern, and Contemporary. *Synthesis Lectures on Human-Centered Informatics*, 2012.

[11] M. Scaife and Y. Rogers. External cognition: how do graphical representations work? *International journal of human-computer studies*, pages 185–213, 1996.

[12] D. Schön. The Architectural Studio as an Exemplar of Education for Reflection-in-Action. *Journal of Architectural Education*, 38(1):2–9, 1984.

[13] D. Schön. Educating the reflective practitioner: Toward a new design for teaching and learning in the professions. *Australian Journal of Adult Learning*, 50(2), 1987.

[14] A. J. Wheat. A Framework for Distributed Sensemaking in Investigations. In *The 28th British HCI Conference*, 2014.

[15] P. C. Wright, B. Fields, and M. D. Harrison. Distributed information resources: A new approach to interaction modelling. In *Proceedings of ECCE8: Eighth European Conference on Cognitive Ergonomics*, pages 10–13, 1996.

[16] Wright, Peter C., Robert E. Fields and M. D. Harrison. Analyzing human-computer interaction as distributed cognition: the resources model. *Human Computer Interaction*, 1(15):1–41, 2000.

[17] J. Zhang and D. Norman. Representations in distributed cognitive tasks. *Cognitive science*, 22, 1994.