Visual Analytics Education

James Foley
Georgia Tech
(Chair & Author)

Stu Card
PARC

David Ebert
Purdue

A. MacEachren
Penn State

Bill Ribarsky
UNC-Charlotte

ABSTRACT

Visual Analytics is a newly evolving field that spans across several more established disciplines. This panel will discuss how VA system developers and researchers are best educated at the MS and PhD levels. This paper describes several ways in which VA can be characterized – with the goal of using these characterizations to identify knowledge domains that can be used to define VA curricula. Also, a digital library of VA educational resources is described

CR Categories and Subject Descriptors: H.5 Information Interfaces and Presentation; K.3.2 [Computers and Education]: Computer and Information Science Education.

Keywords: Visual Analytics Education, Visual Analytics Curricula

1 INTRODUCTION

The purpose of this panel is to begin a community dialogue on graduate education in the emerging and broadly interdisciplinary field of Visual Analytics (VA). Where does VA fit into the broader set of academic disciplines on which it draws? What is at its intellectual core? How should MS and PhD students be educated to become productive VA developers and researchers?

We begin with an overview of several ways one might conceptualize the body of knowledge that is Visual Analytics, drawing on *Illuminating the Path* [1] and an initial taxonomy developed by the UNC-C and GT Regional Visual Analytics Center.

Next, we pose some of the questions that the panel is expected to address. But, we note that as with any fledgling discipline, the answers to these questions will be, at best, tentative. Only time and experience and experimentation will provide us with some confidence as to the adequacy of these tentative answers.

Finally, we describe the Visual Analytics Digital Library, being developed by the Southeastern RVAC, designed to help educators teach and students learn Visual Analytics.

2 WAYS OF THINKING ABOUT VISUAL ANALYTICS

2.1 Illuminating the Path

Perhaps the most straightforward way to think about the intellectual content of VA is by using the structure developed in *Illuminating the Path* [1], a report representing the thinking of a group of nearly 30 researchers. Figure 1, adapted from [1], depicts the four major report chapters as four quadrants. I have added the "?" to "Core" to reflect one question our panel will address: Is there a set of core VA knowledge, and if so, what is it?

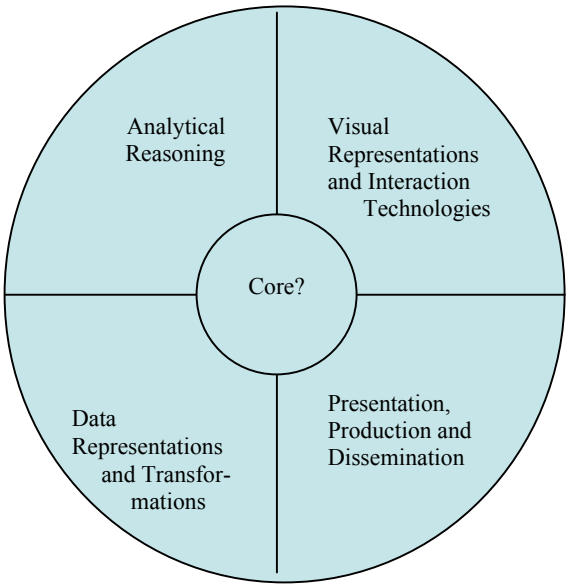

Figure 1. The 'body of knowledge" of Visual Analytics, based on the titles of chapters 2 to 5 of *Illuminating the Path* [1].

2.2 A Starting Taxonomy of Visual Analytics

Education is one emphasis of the Southeastern RVAC at UNC-Charlotte and Georgia Tech. In the process of developing our VA Digital Library (Section 4), we developed a VA taxonomy in a two-step process. The first step used the research agenda of *Iluminating the Path* as a starting point, restructuring some of the material to take more of a pedagogical approach. This ultimately led to the seven top-level categories in the taxonomy (Figure 2). At the same time, we developed a list of about 300 keywords drawn from *Illuminating the Path* and from a VAST Symposium keyword list developed by the program committee. We went through a card-sorting exercise to cluster those keywords into meaningful chunks that became the second-level and third-level categories in the taxonomy.

```
1 Overview/Courses
    Introductory/General
    Uses of Visual Analytics
    Courses
2 Analytical Reasoning & Processes
    Introduction/General
    Models of Analytical Reasoning Processes
    Decision Sciences
    Uncertainty management
    Collaborative reasoning
3 Cognition, Perception & Social Processes
    Introductory/General
    Knowledge Representation & Ontologies
    Perception
```

foley@cc.gatech.edu, card@parc.com, ebertd@ecn.purdue.edu, maceachren@psu.edu, ribarsky@uncc.edu

IEEE Symposium on Visual Analytics Science and Technology 2006
October 31 - November 2, Baltimore, MD, USA
1-4244-0592-0/06/$20.00 © 2006 IEEE

```
    Social Processes
4 Data & Knowledge
    Introductory/General
    Data Representations
    Databases
    Data Transformations
    Knowledge Discovery, Representation &
    Management
        Introductory/General
        Fundamentals
        Specific Types of Data/Application
        Systems
    Test & Sample Databases
5 Visualizing Data, Information & Knowledge
    Introductory/General
    Design Principles
    Methods & Techniques
    Interacting with Visual Representations
        Introductory/General
        Taxonomies of Interaction Techniques
        Uses & Methods of Interaction
    Visualization Types
        Introductory/General
        Taxonomies of Types of Visualizations
        Geospatial (Mapping)
        Graphs & Networks
        Sequenced & Time Series Data
        Statistical (Multidimensional) Data
    Visualization Domains
        Introductory/General
        Text & Documents
        Financial Data
        Networks
        Reasoning, Argumentation & Knowledge
    Automatic & Semi-automatic Generation of
    Visualizations
6 Production, Presentation & Dissemination
    Introductory/General
    Data/evidence collection
    Analytical sandbox
    Developing a story within audience context
    Production technologies & management
    Dissemination contexts & dynamics
7 Human Concerns in Use of Technology for Visual
Analytics
    Introductory/General
    Integration within analytic work methods
    Collaboration
    Evaluation of Effectiveness
    Privacy, Security & Legal Issues
8 Software & Hardware for Visual Analytics Systems
    Introductory/General
    Architectures & Systems
    Interaction Devices
    Display Technologies
```

Figure 2. A Visual Analytics Taxonomy

2.3 Leveraging HCI

A common Mantra in Human-Computer Interaction is "*People* perform *Tasks* using *Computers*." The pedagogical implication is that HCI developers and researchers need to know about *people*, the *tasks* the people perform, and the *computer* technologies used to provide a User Interface through which users can perform their tasks. Visual Analytics, to my mind, is a specialized HCI domain, implying that the same Mantra might be applied here. This would lead to a structure along the lines of Figure 3.

```
Task
    Analytical Processes
    Applications of Visual Analytics
People
    Cognition, Perception
    Knowledge Representation & Ontologies
    Social Processes
    Human Concerns in the use of Technology for
    Visual Analytics
Computers
    Data and Knowledge
    Visualizing Data, Information and Knowledge
    Software & Hardware for Visual Analytics
    Systems
```

Figure 3. Structure of Visual Analytics based on an HCI-oriented decomposition

3 Typical Questions the Panel will Address

- Why is / is not VA a field ready for its own MS degree, distinct from a "standard" degree in one of its supporting disciplines, such as computing, HCI, or geography?
- Can a VA M.S. student gain appropriate preparation by appropriate specialization in one of these or some other discipline?
- Is there a core of knowledge that every VA M.S. students should know? Ph.D. student?
- What courses might a VA minor or certificate or specialization at the MS or PhD level include?
- Defining the intellectual content of VA is hard because it draws on so many disciplines. What are some simple criteria to help decide what topics should be explicitly considered as part of a VA curriculum, as opposed to being pre-requisite to VA studies?

4 VA Digital Library

The Southeastern RVAC has started a Digital Library of educational material to help teachers develop courses that are relevant to VA, and to help students learn about VA. The library can be accessed at `vadl.cc.gatech.edu`. It is patterned after, and uses the same software infrastructure, as the Human-Centered Computing Education Digital Library at `hcc.cc.gatech.edu`. The library contains course syllabi, videos, PowerPoint lectures, lectures created with Microsoft Producer [2], exams, assignments and some reference material.

Content is organized according to the taxonomy of Section 2.3; a content item can be placed under several categories in the taxonomy. The user can browse through the taxonomy to get the "big picture" of what types of information are present, or can use a standard text content search to obtain a ranked (be relevance) list of items, Google-style. Submissions to the library are encouraged; the home page links to an easy-to-use submission interface. Submissions are encouraged!

5 Panel Participants

Jim Foley is a professor in the College of Computing at Georgia Tech, and holds the Fleming Chair in Telecommunications. While on Georgia Tech leave from 1996-99, he was CEO and Chairman of MERL – Mitsubishi Electric Research Labs where he was responsible for corporate R&D in North America. He joined Georgia Tech in 1991 as the founding director of the Graphics, Visualization and Usability Center in the College of Computing. Foley is co-author of three computer graphics texts and is a Fellow of AAAS, ACM and IEEE. He received SIGGRAPH's bi-annual Steven Coons award for contributions to computer graphics, and is one of seven inaugural members of the SIGCHI Academy for contributions to Computer-Human Interaction. He was chairman (2001-2005) of the Computing Research Association.

David Ebert is a Professor in the School of Electrical and Computer Engineering at Purdue University and directs the Purdue University Regional Visualization and Analytics Center. His research interests include volume rendering, information visualization, perceptually based visualization, illustrative visualization, and procedural abstraction of complex, massive data. Ebert has been very active in the visualization community, teaching courses, presenting papers, co-chairing many conference program committees, serving on the ACM SIGGRAPH Executive Committee, serving as Editor in Chief of *IEEE Transactions on Visualization and Computer Graphics*, and serving on the National Visualization and Analytics Center's National Research Agenda Panel.

Stuart Card is a Senior Research Fellow and the manager of the User Interface Research group at the Palo Alto Research Center. His study of input devices led to the Fitts's Law characterization of the mouse and was a major factor leading to the mouse's commercial introduction by Xerox. His group has developed theoretical characterizations of human-machine interaction, including the Model Human Processor, the GOMS theory of user interaction, information foraging theory, and statistical descriptions of Internet use. These theories have been put to use in new paradigms of human-machine interaction including the Rooms workspace manager, papertronic systems, and the Information Visualizer. Card is co-author of *The Psychology of Human-Computer Interaction*, co-editor of *Human Performance Models for Computer-Aided Engineering* and of *Readings in Information Visualization*. He has served on many editorial boards, government panels, and university review boards. He has been an adjunct faculty member at Stanford University. He is currently developing a supporting science of human-information interaction and visual-semantic prototypes to aid sensemaking. Card is a Fellow of the ACM, the first recipient of the ACM CHI Lifetime Achievement Award, and the first member of the ACM CHI Academy.

Alan MacEachren is 2004-2007 E. Willard and Ruby S. Miller Professor of Geography and Director of the GeoVISTA Center at Pennsylvania State University. He also directs the new North-East Regional Visualization & Analytics Center. MacEachren's research foci include geovisualization, geocollaboration, interfaces to geospatial information technologies, human spatial cognition as it relates to use of those technologies, human-centered systems and user-centered design. He served as chair of the International Cartographic Association Commission on Visualization and Virtual Environments (1999-2005) and was named honorary fellow of that organization in 2005. He has been a member of the National Research Council Computer Science and Telecommunications Board Committee on the Intersections Between Geospatial Information and Information Technology (2001-2002) and of the National Visualization and Analytics Center R&D Agenda panel (2004-2005). MacEachren is author of *How Maps Work: Representation, Visualization and Design*, and *Some Truth with Maps*, and is co-editor of several additional books (including *Exploring Geovisualization*) and journal special issues (including a recent issue of *IEEE Computer Graphics and Applications* on Geovisualization, and a forthcoming special issue of the *International Journal of Geographical Information Science* on Visual analytics & spatial decision support).

William Ribarsky is the Bank of America Endowed Chair in Information Technology at UNC Charlotte, is founding director of the Charlotte Visualization Center, and director of the Southeastern Regional Visualization & Analytics Center. His research interests include visual analytics, 3D multimodal interaction, bioinformatics visualization, virtual environments, visual reasoning, and interactive visualization of large-scale information spaces. Formerly, he was the Associate Director for External Relations of the Georgia Tech GVU Center. Dr. Ribarsky is the former Chair and a current Director of the IEEE Visualization and Graphics Technical Committee, and chairs the Steering Committees for the IEEE Visualization Conference and the IEEE Virtual Reality Conference. He is an Associate Editor of IEEE Transactions on Visualization and Computer Graphics. Dr. Ribarsky co-founded the Eurographics/IEEE visualization conference series (now called EG/IEEE EuroVis) and led the effort to establish the Virtual Reality Conference series. For the above efforts on behalf of IEEE, Dr. Ribarsky won the IEEE Meritorious Service Award in 2004.

REFERENCES

[1] James Thomas and Kristen Cook, eds. Illuminating the Path: The Research and Development Agenda for Visual Analytics. IEEE Computer Society, 2005.

[2] Microsoft Producer, http://www.microsoft.com/office/powerpoint/producer/prodinfo/default.mspx