LEVERAGING MULTIDISCIPLINARITY IN A VISUAL ANALYTICS GRADUATE COURSE

NIKLAS ELMQVIST, DAVID S. EBERT*
PURDUE UNIVERSITY

ABSTRACT

There is a growing demand in engineering, business, science, research, and industry for students with visual analytics expertise, but teaching visual analytics is challenging due to the multidisciplinary nature of the topic matter, the diverse backgrounds of the students, and the corresponding requirements on the instructor. We report some best practices from our experience teaching several offerings of a visual analytics graduate course at Purdue University where we leveraged these multidisciplinary challenges to our advantage instead of attempting to mitigate them.

Keywords: visual analytics, education, cross-discipline, multidiscipline, best practices

INTRODUCTION

The emerging field of *visual analytics* is defined as the science of analytical reasoning aided by visual interfaces, and was coined as a scientific term as late as 2005 (Thomas and Cook 2005). It is a rapidly growing area within several branches of society, including academia, where the annual *IEEE VAST* conference is in its seventh year and has been steadily growing since its first offering in 2006; in government, where an increasing number of governmental agencies across the globe are adopting and funding visual analytics; and in industry, where self-proclaimed visual analytics companies, such as Tableau, Oculus, and i2, are gaining increasing success in selling their data analysis software to an ever-expanding market of companies across the range of markets from business to health to manufacturing.

However, visual analytics being a hot topic is a double-edged sword: while stakeholders within academia, government, and industry alike are expanding and, therefore, desperately looking for recruits with visual analytics expertise, institutions of higher education are struggling to respond to this sudden need. This problem is compounded by the intrinsic multidisciplinary nature of visual analytics as a field, which gives rise to three main challenges:

1. The subject matter of what constitutes visual analytics is very wide and encompasses much more material than can comfortably be taught in a standard 15-week graduate course (Figure 1), yet the burden on teaching all this material usually falls onto the visual analytics course since there exists few other complementary courses in traditional academic departments;
2. Students with an interest in taking a visual analytics course typically come from across the academic spectrum, and therefore have very diverse and non-overlapping backgrounds; and
3. The requirements in teaching such a wide and disparate set of topics falls to the instructor, who cannot realistically be expected to be an expert of the full subject matter.

* = corresponding author
Niklas Elmqvist, School of Electrical and Computer Engineering, Purdue University
465 Northwestern Avenue, West Lafayette, IN 47907-2035, USA. E-mail: elm@purdue.edu

In this paper, we report on some best practices for teaching visual analytics in the graduate curriculum in spite of these challenges. Based on our three years of experience from teaching a visual analytics experimental graduate course at Purdue University, we show how to actually transform these multidisciplinary challenges into strengths that will both produce competent visual analytics researchers, as well as attract potential new students to the field.

BACKGROUND

Our experimental graduate course, ECE 695D, is offered at the School of Electrical and Computer Engineering at Purdue University, and has been taught twice: in Fall 2009 (12 students) and in Fall 2011 (14 students). The course is a standard 3-credit, 15-week graduate course that is open to any Purdue graduate student with an interest in visual analytics. We have no specific prerequisites other than experience in a topic such as statistics, data analysis, or visualization. Significantly, even though the course is offered in an ECE department, we do **not** ask that students have programming expertise specifically because we want to attract a multidisciplinary array of students that want to apply visual analytics to their own research.

ECE 695D is mainly designed around a semester-long research project, and is taught in the Fall semester (odd years) so that strong student projects can be polished and improved for submission to one of the *IEEE VisWeek* conferences, which have deadlines at the end of March. The course project itself is designed to teach students every step of how to conduct a successful research project, and thus has regularly spaced deadlines for inception, literature survey, alpha and beta release, final paper, and peer review throughout the semester. These projects can be done individually or in teams of up to three students. We, as instructors, give continual feedback on the submitted material, ensuring that students stay on track and make progress in a worthwhile direction. In general, student feedback on the course project component has been very positive, and they seem to particularly enjoy working on novel research projects that may eventually result in a peer-reviewed publication.

In addition to the course project, we also ask students to perform an analytical exercise with the Atlantic Storm (Hughes 2005) dataset (approximately 50 police and intelligence documents containing a hidden threat) during the first two weeks of the course. In the first week, students try to identify the hidden threat using pen, paper, and standard software tools, whereas in the second, they do the same using the Jigsaw (Stasko et al. 2008) visual analytics tool. The purpose of this exercise is to give students first-hand exposure, as early as possible, to the challenges of investigative analysis, and was inspired by the same exercise used to enlighten the panel who created the visual analytics research agenda (Thomas and Cook 2005). In particular, this exercise often gives rise to confirmation bias because students tend to simply use Jigsaw to confirm what they thought they knew from the first pen-and-paper phase, which students seem to find instructive for the remainder of the course.

MULTIDISCIPLINARY TOPICS

A course on visual analytics by necessity becomes something of an umbrella course that introduces the basics of many different topics (Figure 1), yet is unable to go into depth on any of these topics. This is unfortunately a fact of life because the courses that would teach each of these topics in depth are spread across multiple traditional academic departments such as computer science, statistics, psychology, informatics, and engineering, and few students are able to take all of these courses separately. In fact, in some cases, a single university may not even offer all of the required graduate courses.

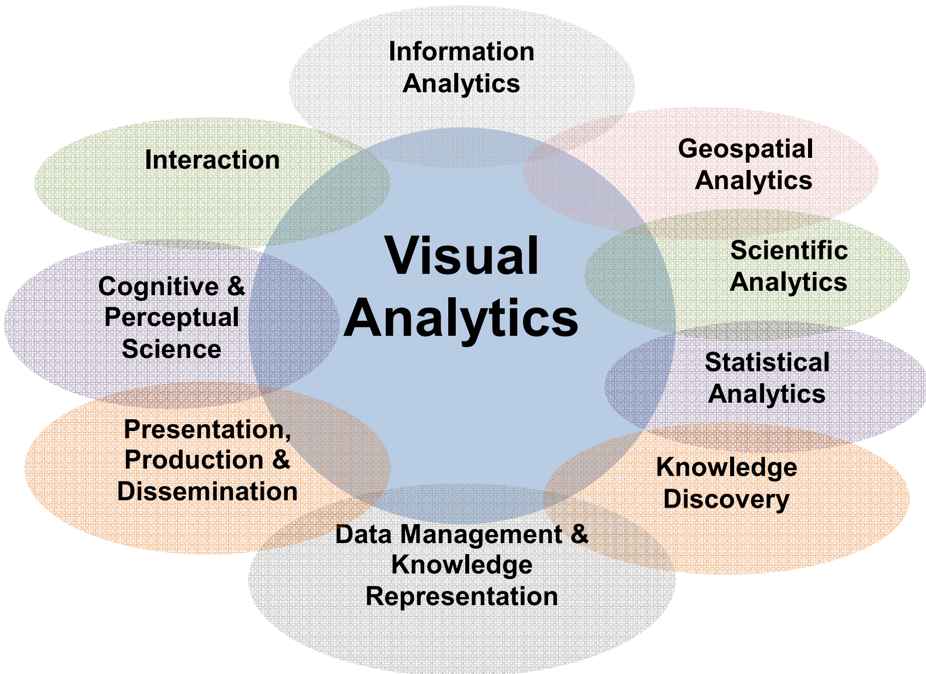

Figure 1: The multidisciplinary scope of visual analytics (created and refined by Jim Thomas, Daniel Keim, David Ebert, Niklas Elmqvist, etc).

Designing, organizing, and teaching a visual analytics course, therefore, becomes a problem of describing the bare essentials of each topic in an order that makes sense to the students while charting the visual analytics landscape to inspire future exploration by the students. Table 1 shows our current 15-week schedule, including a brief overview of the contents of each topic. In particular, note the heavy emphasis on the human analytical process in the first five weeks of the course; most of this material is new to the students, yet introduce them to topics that are central to designing effective visual analytics systems.

Week	Topic	Contents
1	Introduction	Analytical exercise
2-3	Analytical Reasoning	Analysis process, critical thinking, sensemaking, situation awareness
4	Perception	Human perception, preattentiveness, color, shape, texture
5	Cognition	Cognitive theory
6-7	Data Management	Representations, trasnformations, statistics (temporal and spatial)
9	Visual Representations	Visualization techniques
11	Interaction	Interaction techniques
12	Communication	Production, presentation, dissemination
13	Collaboration	Collaborative visual analytics
14	Evaluation	Evaluating visual analytics
15	Advanced Topics	Conducting VA research, novel computing platforms, mobile VA

Table 1: 15-week syllabus for ECE 695D (Introduction to Visual Analytics).

Of course, a visual analytics course such as this would be even better if it could rely on an array of advanced graduate courses that went into detail on each of these topics, allowing students to specialize in a particular area. At Purdue University, this is really only possible for statistics, cognitive science, human factors, and visualization; most other existing graduate courses in other departments do not cover the subject matter that ECE 695D introduces to the students. This naturally puts additional strain on the instructor expertise; see below.

It is worth briefly discussing the relation between visualization and visual analytics graduate courses.We designed ECE 695D to be a complement to our existing graduate course on visualization, and we expect virtually all graduate students who take one of the courses to also take the other (although in an unspecified order). For this reason, we do not spend much time talking about visual representations in ECE 695D, and analogously, we do not spend much time talking about visual analytics in our visualization course. We would recommend that all universities that teach one of these courses would also teach the other, but if a choice has to be made, visual analytics is a better choice due to its wider scope, applicability, and general interest. Students can always individually read up on the relevant visualization techniques to use for their projects provided that instructors give them the foundation and references during the class.

Diverse Student Backgrounds

In our two offerings of ECE 695D (a total of 25 students), we have taught an astoundingly diverse pool of students, including students from not only traditional engineering departments such as computer, mechanical, and industrial engineering, but also from forestry, psychology, physics, geomatics, education, economics, and technology. We explicitly market the course as one to take if you have a data analysis problem to solve for your own research, regardless of the field. Accordingly, we do not ask students to have any previous experience in the subject matter besides some data analysis knowledge; significantly, we do not ask that they have software development expertise, which is otherwise common in visualization courses.

Of course, this gives rise to a significant challenge when it comes to the project course, where we ask students to make a contribution to the field of visual analytics that conceivably be published in a peer-reviewed visual analytics conference. Such contributions—for example, those published at the *IEEE VAST* conference—typically involve a non-trivial amount of software development in order to build the implementations of the novel techniques presented in the papers. However, such software development is clearly out of reach for most the students with a non-CS/CE background, and asking them to learn not only visual analytics but also programming during a single semester is clearly unfair and unrealistic.

To date, we have dealt with this challenge in ECE 695D in two different ways that actually leverage the diverse student backgrounds instead of treating it as a problem:

- **Multiple project types** – Visual analytics contributions these days do not necessarily have to be new software systems; in fact, the *IEEE VisWeek* call for papers explicitly identifies contribution types (evaluations, design studies, and models) (Munzner 2008) that do not necessarily have to involve building new software systems. Accordingly, we allow students to work on any of these project types, enabling them to harness their own expertise in a specific field towards their project.
- **Multidisciplinary teams** – In several instances, we were able to form project teams consisting of both technical and non-technical students, thereby providing complementary expertise that

significantly multiplied the capabilities of the team as a whole. In one example, a student team consisted of an industrial engineer with human factors expertise grouped with two computer engineers with strong software development expertise; their work is currently under review for a prestigious visualization journal.

In other words, our recommendation is not to see diverse student backgrounds as a problem, but rather to embrace it as an opportunity for reaching new heights in student learning and success.

INSTRUCTOR EXPERTISE

Finally, the fact that visual analytics spans such a wide array of topics does not only lead to challenges for the course schedule and diverse student background, but also puts an inordinate strain on the expertise of the instructor teaching the course. Simply put, it is impossible to be intimately familiar with all of these disparate topics that a visual analytics course typically should contain, and thus an instructor will easily get overwhelmed in preparing for and teaching these topics to students.

We have found that the best remedy for this challenge is to offload this burden between a team of instructors instead of a single one. Both of the authors are co-instructors for ECE 695D, and have been intimately involved in lecturing, writing notes, evaluating, grading, and giving feedback on different components of this course throughout both offerings. In our particular case, we have similar yet complementary expertise in different topics of the course, and have taken turns in lecturing on these topics. This is also beneficial in expanding one's own visual analytics expertise. Furthermore, we have also been very open to bringing in guest lecturers from our colleagues across campus, from the pool of senior students in our research groups, or even from different universities (typically in conjunction with research visits). Several times we have also recommended that students attend relevant seminars across campuses in lieu of traditionally scheduled lectures. These guest lecturers have been able to lend a different perspective and different expertise on aspects of the course that we feel can only widen the students' horizons on visual analytics, which is the whole purpose of ECE 695D.

Unfortunately, there currently exists no comprehensive textbook on visual analytics that could serve as the backbone of a visual analytics course, and this is a major barrier against adoption in graduate curricula. While visualization textbooks are now fairly common—such as those by Ward et al. (2010), Ware (2004), and Spence (2007)—these tend to focus on visual representations only and therefore generally fail to cover the entire scope of visual analytics. Until a VA textbook is written, we hope that the course schedule in this article (Table 1) will serve as a guideline on which topics to cover in a visual analytics course. Furthermore, the Visual Analytics Digital Library (VADL) at http://vadl.cc.gatech.edu/ contains many additional resources, syllabi, slides, lecture notes, and schedules that a new lecturer in visual analytics can adopt and extend for his or her own purposes.

In summary, our recommendation is to offload the burden of teaching a visual analytics course on several instructors, if possible, and to leverage colleagues, visitors, and senior students as guest lecturers in the course. We have also given some pointers and resources on how to structure a visual analytics course for novice instructors with no expertise in the research area itself.

STUDENT RESULTS AND SUCCESS

Several student projects from ECE 695D have gone on to become full-fledged research visual analytics projects that have been published in peer-reviewed conferences. In a few cases, these publications were the first papers the students in question published in a visual analytics conference. This suggests that ECE 695D has succeeded in being a stepping stone for new students to enter the visual analytics scientific community. Figures 2 and 3 give examples of student work that since have been published; several other student projects are currently under review.

The course website for ECE 695D can be found here:

- https://engineering.purdue.edu/~elm/teaching/ece695d/

CONCLUSION

Designing and teaching a visual analytics course can be challenging, but our experience is that the broad scope of the research field can be harnessed to make the course even more engaging and compelling to the students. So far, our own visual analytics course at Purdue University received top ratings in course surveys, and the course seems popular in the graduate student population. We hope that some of the best practices we report on in this article will help other instructors to improve and refine their own education programs on visual analytics. Furthermore, we would be very open to hearing about your own experiences in teaching visual analytics; please see below for information on how to get in touch with us.

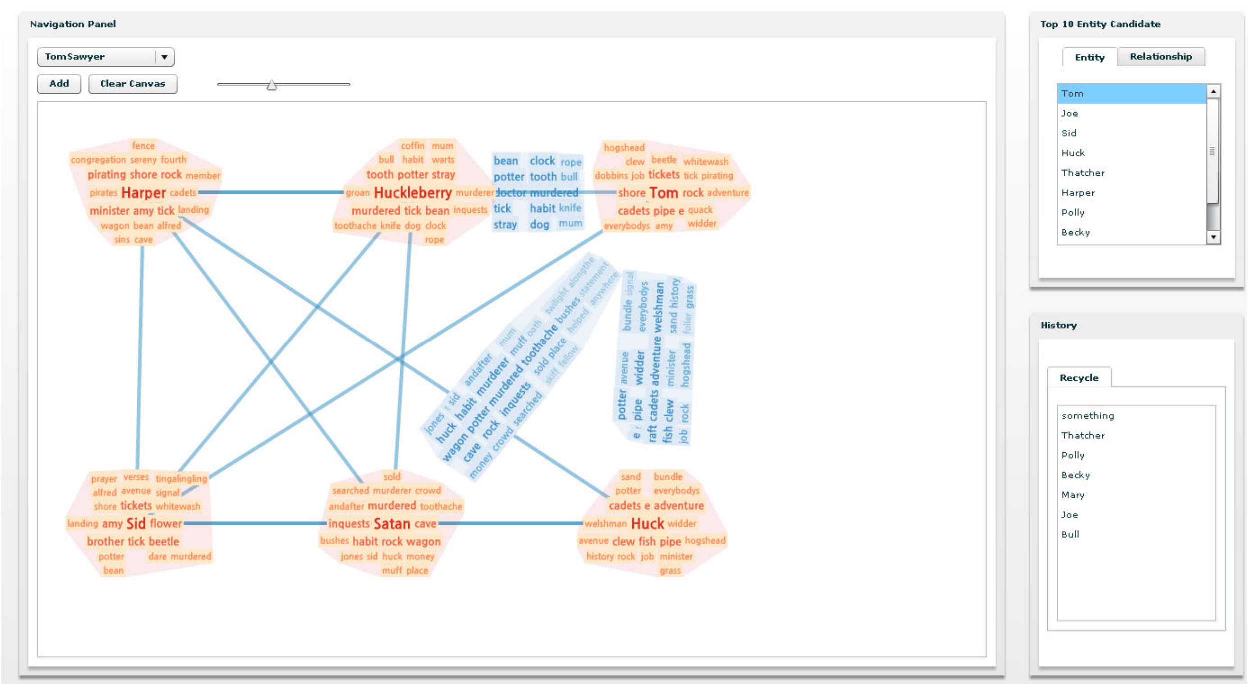

Figure 2: The WordBridge text analytics visualization showing relations in document collections (Kim et al. 2011).

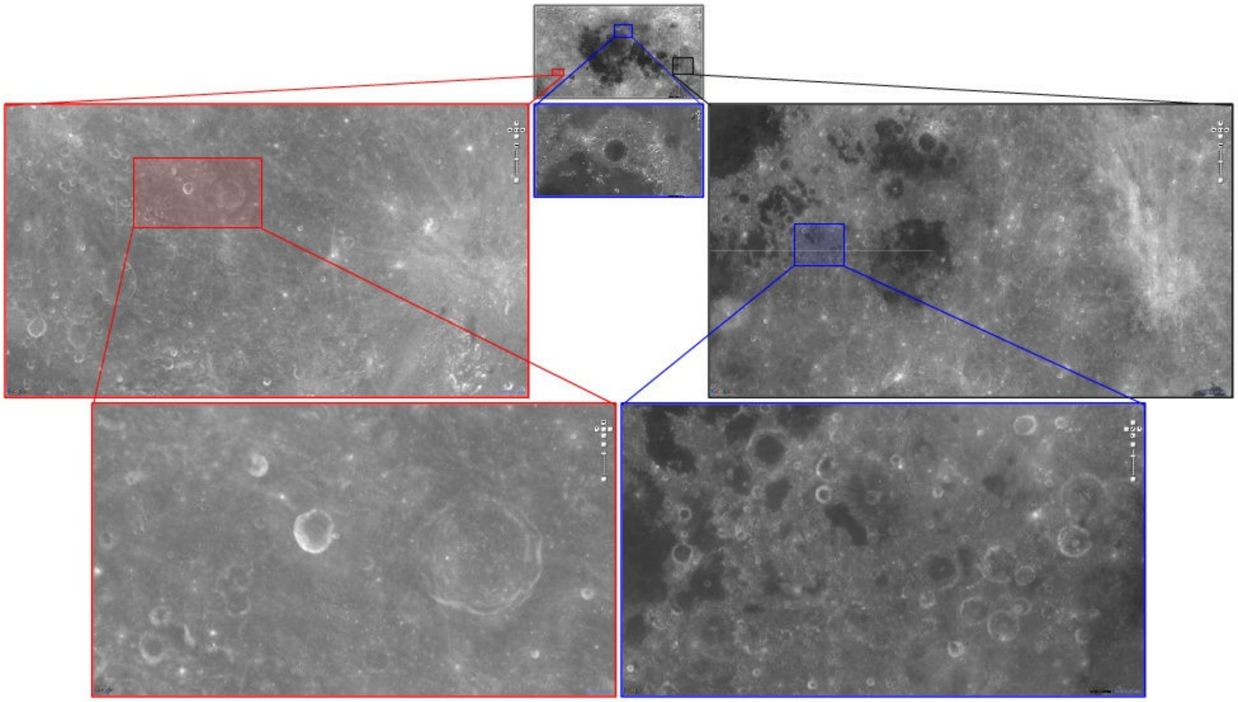

Figure 3: Using the PolyZoom multi-focus technique to explore a Lunar map dataset (Javed et al. 2012).

BIOGRAPHIES

Niklas Elmqvist is an assistant professor in the Purdue University School of Electrical & Computer Engineering. Contact him at elm@purdue.edu.

David S. Ebert is the Silicon Valley Professor in the Purdue University School of Electrical & Computer Engineering. Contact him at ebertd@purdue.edu.

REFERENCES

F. J. Hughes. Discovery, Proof, Choice: The Art and Science of the Process of Intelligence Analysis, Case Study6, "All Fall Down", Unpublished report, 2005.

Waqas Javed, Sohaib Ghani, and Niklas Elmqvist. PolyZoom: Multiscale and Multifocus Exploration in 2D Visual Spaces. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*, to appear, 2012.

KyungTae Kim, Sungahn Ko, Niklas Elmqvist, and David Ebert. WordBridge: Using Composite Tag Clouds in Node-Link Diagrams for Visualizing Content and Relations in Text Corpora. In *Proceedings of the Hawaii International Conference on System Sciences* (Visual Analytics Minitrack), 2011.

Tamara Munzner. Process and Pitfalls in Writing Information Visualization Research Papers. In *Information Visualization: Human-Centered Issues and Perspectives*. Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, Chris North. (editors) *Lecture Notes in Computer Science*, Volume 4950,Springer, pp. 134-153, 2008.

Robert Spence. *Information Visualization: Design for Interaction*. Prentice Hall, 2007.

John Stasko, Carsten Görg, Zhicheng Liu, and Kanupriya Singhal. Jigsaw: Supporting Investigative Analysis through Interactive Visualization.Proceedings of the IEEE Symposium on Visual Analytics Science & Technology, pp. 131-138, 2007.

James J. Thomas and Kirsten A. Cook. (eds.). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*.IEEE Computer Society, 2005.

Matthew O. Ward, Georges Grinstein, and Daniel Keim. *Interactive Data Visualization: Foundations, Techniques, and Applications*. A K Peters, 2010.

Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2004.