Preparing High School Students for College with Informatics

Stephen J. Zilora
Information Sciences and Technologies Department
Rochester Institute of Technology
Rochester, NY 14623

Session: All other topics relevant to engineering education

ABSTRACT

As we begin the 21st century, a concern among many educators is whether we are providing our students with the skills they will need in this brave new world. Discussion often centers on the issues of globalization and the demands it brings for cultural awareness and appreciation. While these factors do distinguish the 21st century, another equally important factor is the utilization of computing, specifically digital information. This skill is no longer the sole purview of backroom analysts and computer programmers. From health care professionals and educators to scientists and engineers, every 21st century professional will be working with digital information. The utilization of this information is studied within the field of informatics (information + automatic). While several informatics degree programs and areas of study are emerging at universities around the country, no such training exists at the secondary school level. This is a serious omission. Just as 21st century professionals will need to utilize digital information, so will students studying for those professions. We teach all high school students to write research papers even though relatively few of them will ultimately become researchers. The reason is that they need to be able to conduct and write about their research in order to be successful in college regardless of their major. Similarly, 21st century college students need to be able to collect, analyze, and present information in order to be successful in college regardless of their major. This paper describes an informatics course suitable for upper level high school students. The course introduces students to skills that will allow them to work comfortably and effectively with digital information during their college studies. Successful completion of such a course would empower college-bound students and help them succeed in the 21st century.

What is Informatics?

While hardly as mature a discipline as engineering, computer science has evolved significantly in the past half century. This evolution has given rise to several sub-domains and specialties. One such specialty is informatics. The term "informatics" was first proposed by Karl Steinbuch in his 1957 paper [1], "Informatics: Automatic Information Processing", to refer to the general field of computer science. In Europe and many other parts of the world, this meaning of the word informatics has remained. In the United States, however, the term has taken on a more specific meaning. Here, informatics focuses less on the theory and fundamental technology of computer

science and more on the application of such work to a specific domain. “Informatics, in general, studies the intersection of people, information, and technology systems.”[2]

These three factors are critical to defining informatics. We start by noting that the emphasis is on information, not data. Colloquially, we may think of these as synonyms, but there are important differences. [3] A number on a page is an example of data. It is a symbol that represents a quantity, but a quantity of what? And is it a large quantity or a small quantity? Without context, we cannot answer these questions. Information, on the other hand, is data in context. The context might be the rate of temperature increase in a particular reactor or the number of touchdown passes thrown in a Super Bowl game. Regardless, it is the context that makes the data usable. Informatics is about working with information—data in context.

People are consumers of information. A famous riddle asks the question, “If a tree falls in a forest and no one is around to hear it, does it make a sound?” We can paraphrase that to, “If information is presented and no one can understand it, does it have any value?” Informatics addresses this issue that the point of developing or presenting information is so that people can understand and utilize it. A central concern to informatics is identifying key components and indicators within an information set so that the consumer can easily assess the information.

In the world of informatics, it is technology that brings these two aspects together. While the technology often involves computing, it is equally often far from the traditional computing discipline of programming. Instead, the technology is more about working with information. It involves assembling new sets of information from disparate sources, combining it in value-added ways, and presenting it. Fundamentally, in informatics the focus of technology is delivering information to people.

Informatics in Industry

In industry, this intersection of people, information, and technology manifests itself in two similar but distinct manners. In the Research and Development units, it continues to exist as the field of informatics described above. The role of the informatician is to understand the business needs as well as the nature of the information being sought and serve as the mediator between the two. For example, the bio-informatician might work with various protein data banks (of varying formats) and format and present the necessary informatician to a biologist developing a vaccine for some particular disease. The materials-informatician might work with various data sources regarding metal fatigue and format and present the necessary information to an engineer who is trying to determine the best choice of material for a particular service. In these and many similar cases, the informatician needs to have knowledge of both the particular domain and information science and technology.

It is also necessary to consider another situation. It could be called business operations informatics, but it is generally referred to as Business Intelligence. Business Intelligence functions “…range from reporting, query and analysis, and dashboards and visualization, to

intuitive discovery and advanced predictive analytics capabilities, as well as data quality and extract, transform, and load functionality."[4] These are essentially the same functions as informatics, but now applied to business operations. Oracle, SAP, and many other enterprise-focused technology companies are developing platforms that will allow all businesses to carry out this functionality.

Whether we call it informatics or business intelligence, it is about working with information and presenting it in a manner that people can consume. Businesses value this expertise and are actively seeking it. A simple search at www.monster.com reveals approximately 2,400 advertised positions for business intelligence or informatics [5].

Informatics in College

This popularity of informatics (and business intelligence) is not lost in higher education. Several universities such as Drexel, Baylor, and Rochester Institute of Technology offer an undergraduate degree in bioinformatics. There are also degrees in medical informatics, nursing informatics, and health informatics. Indiana University recently began an Informatics Degree program that allows students to choose a "cognate" area, or specific domain, in which to apply their informatics courses [6].

The value of using a cognate area is rooted both in the essence of informatics and the benefits of interdisciplinary education. The popular benefit of such courses is that it helps students broaden their perspectives. This is not disputed; however, an additional benefit is derived when students see a context and an immediate application of otherwise abstract skills. In the case of informatics, students are able to apply technology to specific domain areas, or cognates [7].

For the past several decades, computing (including programming, networking, and other computer-related activities) has been taught in isolation. The recent explosion of social networking has demonstrated quite vividly that computing now pervades nearly every part of everyone's lives. To teach computing to just a few would-be "geeks" is no longer sufficient. Nor is it sufficient to teach it in isolation [8]. While there remains a need to deeply educate a number of students in the intricacies of computer programming and usage, there also exists a need to educate all in the principles of informatics.

While Facebook®, the iPhone®, and the Internet in general are the most obvious examples of pervasive computing, there is also a nearly ubiquitous use of computing in any data-driven realm. Whether you are a social scientist examining economic factors, or a chemist researching the effectiveness of a particular catalyst, or an engineer studying the durability of insulating materials in extreme environments, your work is data-driven. The same is true for college students studying in these fields. Whether we are discussing professional workers or college students, there is a great need to work with digital information. Workers and students both need to be able to collect, analyze, summarize, and present digital information.

Beginning in high school, we teach our students how to write a research paper. We teach them how to conduct research, construct arguments, and format their papers. What we do not teach them is informatics: how these young *people* can use *technology* to work with *information*. They are left to manually collect and analyze their data. This approach served us well in the 20th century, but we need to do more in the 21st century. Our students need to know about informatics. To be computer literate is no longer good enough; computer fluency (understanding, not just skill) is now a necessity. [9]

Informatics in High School

In order to present an informatics course to high school students, we first need to understand their frame of reference. Depending on their background, students will have a varying degree of literacy and fluency. [10] Many will understand how to send an email, but not the appropriate use of email. But the desire to understand is there. [11] Whether to further their understanding of the tools they use with their peers, position themselves for a better job, or for the study of technology itself, many students want to know more.

It is proposed that an elementary course in informatics can help to satisfy this interest and also better position the students for their college studies. There are several key questions that need to be addressed. First, who should take this course? Because informatics is about applying technology to a particular domain area, it is important that the students have sufficient experience in a domain in which they can apply the technology. Similarly, if the students are to analyze the information, they need to have basic analytical skills. This does not mean a course in statistical analysis is a prerequisite, but rather a developed sense of logic and reasoning. Therefore, the first recommendation is that the course be offered to high school juniors and seniors who have demonstrated strong math and logic skills. We can also state that students should be comfortable using technology, though this is not expected to be a limiting criterion.

A second question is what domain should be used. If we present informatics in the context of economics, health, or engineering, students will be exposed to real information and real problems. Further, the course would be of great benefit if they continue their studies in these areas. But what if they choose to pursue other careers? Or, more problematic, what if they do not have any experience in these areas? To deal with this conundrum, we first introduce the neologism, teen-informatics and define it to be the application of technology to the information that teens typically deal with. Examples of this information include college selection information, iTunes® playlist information, class schedules, and concert schedules. Utilizing this domain leverages their existing worldly knowledge and provides a source of appealing use cases.

A final question is what should be taught. A complete study of informatics would include the Collection, Storage, Analysis, and Presentation of information (as well as the study of what constitutes "information"). As described above, entire degree programs are dedicated to this effort. A single course at the secondary school level needs to be very selective.

A High School Informatics Course

Following is a framework for a high school course on teen-informatics. The course consists of four modules that build upon one another. The modules present tools that the students can use both in this course and their future work (in and after college) to assist them with their information handling. Modules 2 through 4 would involve extensive hands-on exercises within the domain of teen-informatics.

Module 1—What is Informatics?

In this module students will learn about the breadth of the informatics field; how informatics fits alongside computer science, information technology, and other related fields; and how they will be able to put their learning to immediate use.

Topics:

- General overview of informatics
- Careers as informaticians
- Informatics in the sciences
- Informatics in a research paper
- Introduction of Teen-Informatics

Module 2—Excel for Informaticians

Excel is one of the most widely used tools for simple data analysis. In this module students will learn both the basic features of Excel and also its advanced analysis and charting capabilities.

Topics:

- (Review) basic spreadsheet principles
- Formula usage
- Basic statistical functions
- Advanced referencing techniques
- Data lookup functions
- Charting capabilities
- When to use each chart type

Module 3—XML for Informaticians

XML has emerged as the lingua franca for exchanging information between applications. In this module, students will learn how to read and write XML. They will also learn how to transform the XML data file (via XSLT) into reports or web pages for others to view.

Topics:

- Overview of XML purpose and features
- XML syntax
- XML files as a source of information
- Overview of XSLT
- Using XSLT to format an XML file
- Overview of XPath
- Advanced XSLT with XPath

Module 4—Mashups

Mashups are combinations of two or more applications (and their data) that produce new usable information or functionality. A good example is the website, www.weatherbonk.com. This website combines the Google map functionality with the Weather Channel's forecasts. The result allows a user to click on a map and view the weather forecast for that area.

In this module students will learn about the appropriate use of mashups and be able to create their own mashups using a common framework that automatically handles much of the underlying coding.

Topics:

- What is a mashup?
- Why use mashups?
- Introduce BIRT
- Developing mashups with BIRT

Summary

We live in an information age that might also be called an information-overload age. Everywhere we look, we see sources of information. The problem we face is how to make use of it. This is the central question in the field of informatics and it is also a key concern of businesses, researchers, and students. If we are going to leverage all this information, we need to know how to access it, analyze it, and present it. A high school course in informatics is a first step toward this goal. It will prepare our students for their college studies and their employment.

Acknowledgements

This work has been supported by the following members of the Informatics Research Team in the School of Informatics at Rochester Institute of Technology: Chintan Mota, Karakoz Serikbayeva, and Devavrat Sud. Mr. Terrance McCarthy, LOTE/Tech/Business Curriculum Supervisor for the Webster Central School District (Webster, NY), has also provided valuable assistance.

Biography

Stephen Zilora (stephen.zilora@rit.edu) received his M.S. in Computer Science from the New Jersey Institute of Technology in 1996. He has designed or developed numerous applications and managed software development operations for companies ranging from small firms to Fortune 50 firms. He is currently an associate professor in the Information Sciences and Technologies Department, School of Informatics, Rochester Institute of Technology. His research is focused on the storage, consumption, and presentation of information.

References

[1] Steinbuch, K. Informatik: Automatische Informationsverarbeitung. Berlin: SEG-Nachrichten. 1957.

[2] Groth, D., and J. MacKie-Mason. "Education Why an Informatics Degree?" Communications of the ACM 53.2 (2010): 26-28.

[3] Ackoff, R. "From Data to Wisdom." Journal of Applied Systems Analysis 16 (1989): 3-9.

[4] SAP AG. "All Information, All People, One Platform." SAP AG, 2008.

[5] "Google Search." Google. Retrieved Mar. 2010. <http://www.google.com>.

[6] "Requirements - Undergraduate Programs - Academics - School of Informatics and Computing." Indiana University School of Informatics and Computing. Indiana University. Retrieved Mar. 2010. <http://www.informatics.indiana.edu/academics/undergrad_requirements.asp>.

[7] Deremer, D., and K. Herbert. "An Interdisciplinary Undergraduate Science Informatics Degree in a Liberal Arts Contex." SIGCSE '06. Houston, TX: ACM, 2006. 179-183.

[8] Jurema, A., and M. Costa Lima. "Towards a Pedagogy or Informatics." CSCL '95 Proceedings. 1995. 187-190.

[9] Williams, K. "Literacy and computer literacy: Analyzing the NRC's 'Being Fluent with Information Technology.'" Journal of Literacy and Technology 3.1 Spring 2003.

[10] Hoffman, M., and D. Vance. "Computer Literacy: What Students Know and From Whom They Learned It." SIGCSE '05. St. Louis, MO: ACM, 2005. 356-360.

[11] Schulte, C., and J. Magenheim. "Novices' Expectations and Prior Knowledge of Software Development - Results of a Study with High School Students." ICER '05. Seattle, WA: ACM, 2005. 143-153.