

AN INTERDISCIPLINARY APPROACH TO SPORTS ANALYTICS IN A UNIVERSITY SETTING

WHAT ARE THE RISKS ASSOCIATED WITH PERFORMANCE ANALYSIS FOR INDIVIDUALS?

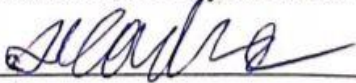
A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Systems Engineering

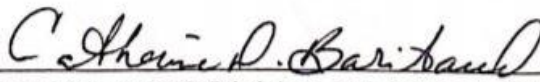
By
Aniket Chandra.


October 31, 2019

Technical Project Team Members
Iyer, R., Kreitzer, R., Lansing, M., Leonard, J., Metzger, B., Nelson, S., Rhodes, C., Ungerleider, D. & Worcester, P.

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed:  Date: 12/05/19

Approved:  Date: Dec 12, 2019
Catherine D. Baritaud, STS Division, Department of Engineering and Society

Approved:  Date: 12/12/19
W.T. Scherer, Department of Engineering Systems and Environment.

According to Alamar and Mehrotra (2011), the world of sports has experienced an explosion in the use of analytics over the past few years (p. 1). The idea of integrating data science and statistics into sports management can be traced back to over 60 years ago (Lindsey, 1959, p. 197), but it only started to take notice after the 2005 release of Michael Lewis' book: "Moneyball: The Art of Winning an Unfair Game". Currently, there's an increasing number of collegiate athletic departments and individual teams employing data analytics to improve performance. As sports analytics focuses on team decision-making and athletic performance of individuals, the same technology can be applied to learning and productivity analytics for students and teachers. Learning analytics is the collection and analysis of data generated during the learning process in order to improve the quality of both learning and teaching. (Siemens, Dawson & Lynch, 2013, p. 8).

However, According to experts, the present technological advances to collect and analyze massive sets of data is likely to lead to revolutionary but also disadvantageous changes in business, and society (Kumar R, 2016). The technical research and loosely coupled STS research will look to provide an understanding of what can be achieved through a foray into performance analytics and the likely outcomes. For the technical report, my team and I will research and design a Sports and Performance Analytics Center for the University of Virginia, that will be integrated with various departments including both academic and athletic. Loosely coupled, my STS research will focus on the potential negative impacts of the increasing use of data analytics and digitization in our society.

SPORTS AND PERFORMANCE ANALYTICS CENTER

Under the guidance of Engineering Systems and Environment Professor William Scherer, Systems Engineers: Rishab Iyer, Rachel Kreitzer, Maryanna Lansing, Jacob Leonard, Benjamin Metzger, Sarah Nelson, Carl Rhodes, Daniel Ungerleider, Peter Worcester, and I aim to create a proposal to be presented to President James Ryan. The 2018-2019 collegiate athletics season for the Virginia Cavaliers was considered one of the most successful years for any college sports program (McElroy, 2019). The Virginia Cavaliers were awarded “National Champion” status in two sports: NCAA Division 1 Men’s Basketball and Men’s Lacrosse. However, for the Virginia Athletics department to continue to operate at the highest level they need to adapt their approach to current trends. Bradley Smith, a sports performance analyst at Northwestern University, stated “I think that very soon we’re going to see that the teams that don’t use analytics — they’re going to be left behind.” (2019).

Currently, there are many teams in the Virginia Athletics department that have engaged in data exploration and analysis to obtain a competitive advantage. For example, engineering students have developed analytics tools to support decision-making on and off the field for the Virginia Cavaliers football team (Hayhurst, 2019, para. 14). There have also been capstone projects for Systems Engineering students that have created analytics frameworks for the Men’s and Women’s Golf, Field Hockey and Softball programs (W. Scherer, personal communication, August 2019). Other sports in the program have collected little to no data whatsoever. As of now, all the teams that have integrated data analytics with their operations have done so independently of the Athletics department.

Ted White, the Assistant Athletic Director and our capstone project’s sponsor, argued; “We need a structure that would allow all the teams to collect, process and analyse the data efficiently with dedicated resources to provide a competitive advantage.” (T. White, personal communication, September, 2019). The accepted ideal framework for sport analytics is illustrated in Figure 1. “The Football and Basketball teams, at this point in time, have rudimentary data management capabilities.” (T. White, personal communication, September, 2019).

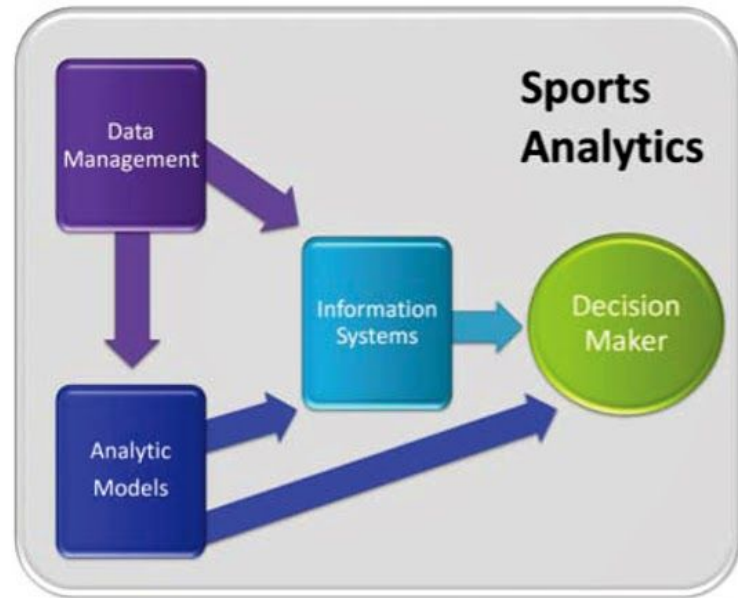


Figure 1: A framework of Sports Analytics: The key components in the transformation of data to knowledge. (Alamar & Mehrotra, 2011)

Illness Prevention and Injury Prediction analysis is also an unexplored area within the athletics department. For athletes, the state of their bodies is equally important for their success as their level of skill in the sport (T. White, personal communication, September, 2019). Technological advances have gone some way towards solving the issue of size—sensors, microprocessors and memory devices for processing and collecting the data are now miniature (Armstrong, Jovanov & Kerwin, 2007). Mark Allen, the Men’s and Women’s Squash Head Coach, expressed an interest in such analytics: “I would love an easy and accurate method to track the health and fitness levels of my players, as the data I have now has proven to be less than useful.” (M. Allen, personal communication, September 2019).

The University of Virginia's Strategic Plan or The 2030 Plan is an outline of goals and initiatives the university's administration has set out to be achieved by the year 2030. One of the aims of the plan is "To become the best public university in 2030, and one of the very best in the world, whether public or private." "To support this work, we will invest in our data and analytics capabilities." (U.VA. Office of Communications, 2019). One proposed way to utilize data and analytics and raise U.Va's reputation is through Learning Analytics. "It is envisaged that education systems that do make the transition towards data-informed planning, decision making, and teaching and learning will hold significant competitive and quality advantages over those that do not." (Siemens, Dawson & Lynch, 2013, p. 2).

The objective of the team's technical project is to design and develop a proposal for a Center of Performance Analytics Research, here at the University of Virginia. This center will encompass both athletic and academic initiatives, and be integrated as a "pan-university resource" (W.Scherer, personal communication, September 2019) for all stakeholder departments at the University of Virginia. This project is a two-semester long research capstone that will cover a range of areas starting with initial outscoping and ending with the hopeful outcome of a presentation to the Office of the President. The team will conduct interviews with several stakeholders and research the established programs at U.Va's peer and competitor schools.

Whilst working with the Athletics Department on designing the sports analytics aspect of the project, we aim to collaborate with various academic departments on the learning or education analytics side. As with most academic initiatives, funding for the project is a critical factor in development (A.Groves, personal communication, October 2019). The project team

will research the various fundraising mechanisms the university already has at its disposal, including the donation programs and alumni outreach. We aim to write a group paper in the form of a proposal or findings from a feasibility study.

THE POTENTIAL NEGATIVE IMPACTS OF BIG DATA ANALYTICS

Although Big Data is a trending buzzword in both academia and in industry, its meaning is still shrouded by much conceptual vagueness. The term is used to describe a wide range of concepts: from the technological ability to store, aggregate, and process data, to the cultural shift that is pervasively invading business and society, both drowning in information overload (De Mauro & Greco, 2015) . Big data is created every day by the interactions of billions of people using computers, GPS devices, cell phones, sensors and medical devices, data-intensive areas such as atmospheric science, genome research and astronomical studies. Now realizing the great importance of big data, many analytical companies are engaged in finding hidden information in consumer data. This process offers a way to better understand and meet their customers' demands. By analyzing customer behavior, as well as vast troves of reviews and feedback, companies can nimbly modify their digital presence, goods or services to better suit the current marketplace (Uzialko, 2018). According to experts, the present technological advances to collect and analyze massive sets of data is likely to lead to revolutionary but also disadvantageous changes in business, and society (Kumar, 2016) .

From the start of the 21st Century, consumers have volunteered volumes of personal data through social media services. Google has amassed vast quantities of data about users of its search facilities, and progressively of other services. The company's acquisition, retention, and

exploitation of all Gmail traffic have allowed it to build archives of its users and their correspondents' communications (Wigan & Clarke, 2013).

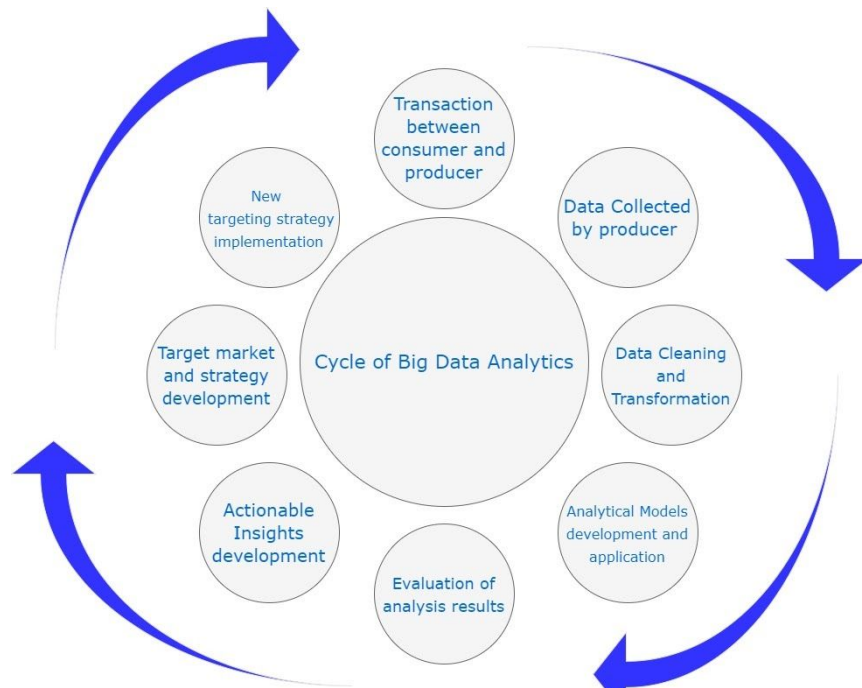


Figure 2: The Cycle of Big Data Analytics: How businesses take advantage of consumer data. (Chandra, 2019)

As seen in figure 2, the process starts from the initial transaction between the customer and the business. The type of transaction doesn't affect the process as data can be collected from any sort. For example, data can be collected from a user after they sign up for an account on Facebook, or when a customer pays for groceries using cash. In both cases, the business or producer instantly receives both quantitative and qualitative data from the consumer. The data cleaning and transformation is inherent in every analysis program; the models may be specific to the producer in question. Using the results from the program, insights and strategy can be developed through simple evaluation. For example, if the models show that men in the age group

18-25 are the largest consumers of a Cola drink, then the producer can launch a targeted marketing campaign on Instagram (which stores gender and age data).

Issues arise concerning users' data, including informed consent for use and disclosure, retention (even after the account is closed), access, and the adequacy of the consideration provided (Wigan & Clarke, 2013). The majority of the debate on Big Data's downsides today addresses threats to personal information privacy. Loss of privacy can in turn result in crimes such as identity theft or cyberstalking. The field of information security overlaps somewhat with privacy, because when security is breached, privacy can be compromised. Another concern is the threat of illegal discrimination in areas such as housing, education, finance, or insurance or unacceptable classification and labeling of individuals as part of target marketing (Markus, 2015).

There's been many cases in recent years when big corporations, whose primary business is the collection and trade of data, have come under scrutiny in the public light due to their data management practices. In August 2011, Facebook faced criticism when it was thought to be exposing the names and phone numbers of everyone in the contacts of mobile devices that used the "Contacts" feature on the Facebook mobile application. Earlier in the year, two security researchers announced that iPhones were regularly recording the position of each device to a hidden file (Davis, 2012). It is also possible for even the most restrictive of data management mechanisms to have accidental breaches. There are two widely publicised cases of information leaks occurring through two individuals associated with the US government—Bradley Manning leaking US diplomatic cables to the Wikileaks site, and Edward Snowden leaking classified NSA data to various media organisations (Nunan & Di Domenico, 2015). These weren't accidents in

the traditional sense and they were supposedly for the benefit of society. It does, however, raise the issue of whether similar data breaches and theft can be done for criminal activities.

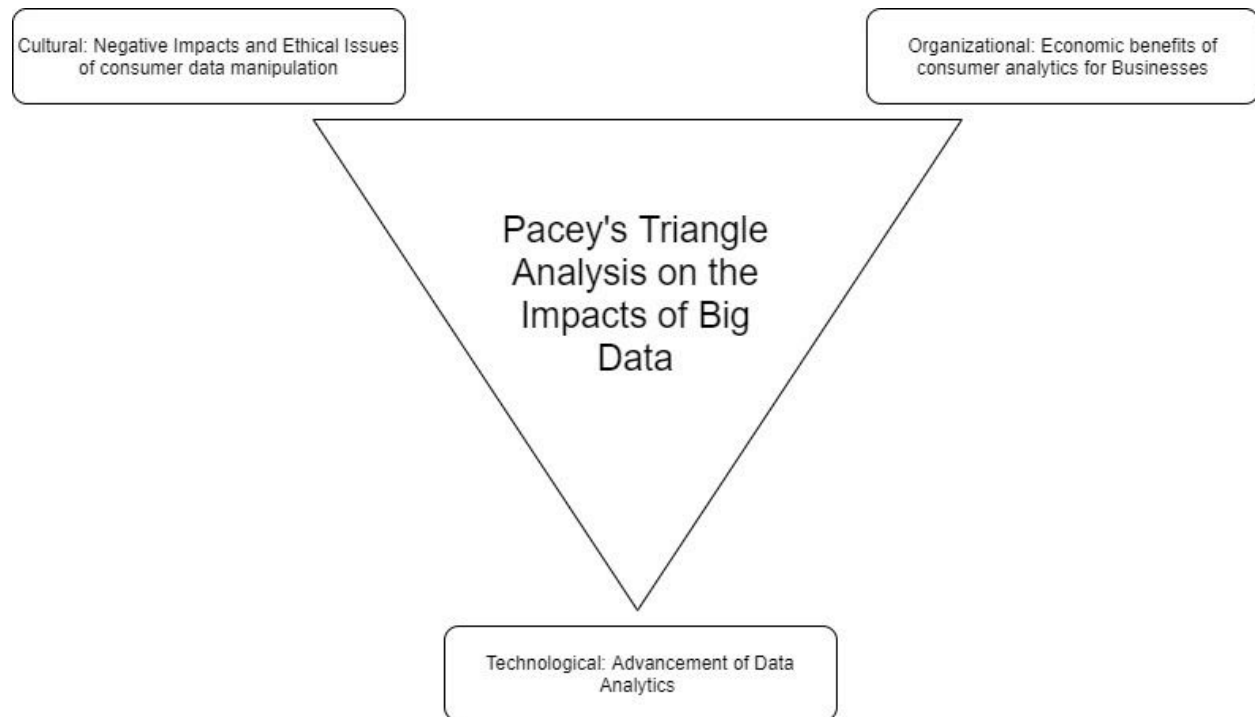


Figure 3: Pacey's Triangle: Analysis of the cultural, technological and organizational aspects of Big Data. (Chandra, 2019)

My STS research project will result in a scholarly article that aims to analyse the negative impacts of Big Data in our society today. Using the general social construction of technology theory and Pacey's Triangle analysis as seen in Figure 3, I hope to evaluate the current and future benefits of data analytics versus the associated harmful consequences that may arise. Through the research and analysis of the current cultural perception of Big Data, I hope to present a proposition that will spark further debate into the foray of analytics development, regulation and acceptance.

WORKS CITED

Armstrong, S., Jovanov, E., & Kerwin, D. (2007, April). Wireless connectivity for health and sports monitoring: a review. *British Journal of Sports Medicine*. Retrieved from

<https://bjsm.bmj.com/content/41/5/285>

Hayhurst, C. (2019, August). Data Analytics Helps College Coaches and Athletes Optimize Training and Performance. *EdTech Magazine*. Retrieved from: <https://www.richmond.com/>

Lindsey, G. R. (1959, April). Statistical Data Useful for the Operation of a Baseball Team.

Operations Research, Vol. 7, 197-207.

McElroy, W. (2019, June). UVA has had the best athletic year of any college sports program.

Richmond Times-Dispatch. Retrieved from: <https://www.richmond.com/>

Siemens, G. Dawson, S. & Lynch, G., (2013, December). Improving the Quality and Productivity of the Higher Education Sector. *Society for Learning Analytics Research*. Retrieved

from <http://www.solaresearch.org/>

Smith, B. (2019, August). A Conversation with Chris Hayhurst. *EdTech Magazine*. Retrieved

from: <https://edtechmagazine.com/higher/>

University of Virginia: Office of Communications. (2019). *Great And Good: The 2030 Plan*.

Charlottesville, VA: Author.

