Designing a Sports and Performance Analytics Center at the Collegiate Level

(Technical Paper)

The Ethics of Collecting Biometric Data in Collegiate Sports

(STS Paper)

A Thesis Prospectus Submitted to the

Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree

Bachelor of Science, School of Engineering

Peter L. Worcester

Spring, 2020

Technical Project Team Members: Aniket Chandra, Rishab Iyer, Rachel Kreitzer, Maryanna Lansing, Jacob Leonard, Benjamin Metzger, Sarah Nelson, Carl Rhodes, and Daniel Ungerleider

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

Signature _____ Date _____

Peter L. Worcester

Approved _____ Date _____

William T. Scherer, Professor, Department of Systems and Information Engineering

Approved _____ Date _____

Kathryn A. Neeley, Associate Professor of STS, Department of Engineering and Society

Introduction

Charles Barkley once stated, “I’ve always believed analytics were crap. They’re just some crap that some people who are really smart made up to try to get in the game ‘cause they had no talent.” (Baerg, 2017, para. 1). Despite naysayers such as Barkley, sport analytics and its connection to Big Data have blossomed in the past several decades. Even as traditionalists such as Barkely dismiss its value, sports organizations on the professional level have fully adopted analytics and its ability to form competitive advantages on the field or court. Collegiate athletics, while lagging behind, have begun implementing analytics and are starting to see tangible results. This paper will examine the feasibility of designing a sports analytics center, namely at the University of Virginia, as well as the ethics behind the collection and storage of our student athletes’ biometric data.

Technical Topic: Designing a Sports and Performance Analytics Center at the Collegiate Level

Under the guidance of Engineering Systems and Environment Professor William Scherer, Systems Engineers: Aniket Chandra, Rishab Iyer, Rachel Kreitzer, Maryanna Lansing, Jacob Leonard, Benjamin Metzger, Sarah Nelson, Carl Rhodes, Daniel Ungerleider, 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 (Scherer, 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 analyze the data efficiently with dedicated resources to provide a competitive advantage.” (2019). The accepted ideal framework for sport analytics is illustrated in Figure 1. Data management will feed into both predictive analytics and information systems to support the relevant decision-makers. Elite data management is essential to the success of the center because missing or incomplete data will render all analytics useless.

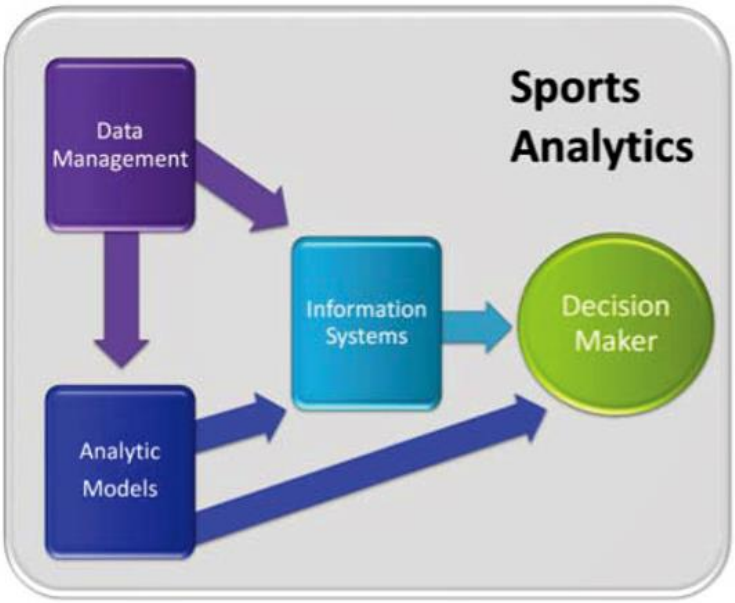

Figure 1: A framework of Sports Analytics (Alamar & Mehrotra, 2011)

Ted White remarked on the current state of our data management, “The Football and Basketball teams, at this point in time, have rudimentary data management capabilities.” (White, 2019).

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 (White, 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). Rachel Dawson, the field hockey assistant coach, expressed an interest in such analytics: “The major challenges our team faces are balancing travel, exam schedules, sleep, and even menstrual cycles with performance on the field. I would love for an easy way to optimize and track the health and fitness of my players.” (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 my team’s technical project is to design and develop a proposal for a Center of Performance Analytics Research, here at the University of Virginia. Downey (2005),

argues that a major characteristic of problem definition and solution is that “the process of generating technical solutions includes the non-technical work of assessing the implications of alternative solutions for stakeholders.” That’s why my capstone team conducted interviews with each of the twenty-seven varsity sports as well as eight academic departments at UVA. This center will encompass both athletic and academic initiatives, and be integrated as a “pan-university resource” (Scherer, 2019) for all stakeholder departments at the University of Virginia. After this capstone project, we hope to provide several effective design options. This will include having a centralized database for all teams to have access to, including analysts and coaches. This center will help improve the overall performance of sports teams at UVA. This, in turn, will develop a good reputation among other school’s sports analytics programs. The center will also enhance relationships between athletes and students interested in analytics. It will provide students both the opportunity to learn the data analytics skills and to practice them with real teams. Outside of UVA, this center will engage the Charlottesville community and other local schools. Finally, the center will be self-sustainable and profitable financially. 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. We will conduct interviews with several stakeholders and research the established programs at U.Va’s peer and competitor schools. In addition to the alternatives and proof of concept for this center, my capstone team will identify specific models and pieces of equipment that would play an essential role in recruiting, play-calling, and other developmental aspects for our sports teams, all of which would be housed within the center. Specifically, I will be working with the Women’s Rowing team to develop models based off of their ergometer training data to provide

insightful findings to the coaches regarding player performance that will ideally lead to success on the water.

While 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 (Groves, 2019). My team and I 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 the findings of a feasibility study.

STS Topic: The Ethics of Collecting Biometric Data in Collegiate Sports

"Big Data" has existed for decades but has just now recently gained public attention in a "big" way. This boom in attention is largely attributed to the exponential rise in data volume in both how it is collected and accessed. Data analytics, and its extension into Big Data, have become one of the premier developments in 21st-century sport. Athletes' biometric data comprises a new and valuable category of big data, with roughly 400 million wearable smart devices expected to be sold in 2020. (Arnold, 2017). Historically, teams have used a wide variety of biometric data such as heart rate, biomechanical measurements, reaction time, and self-reported wellness information. (Osborne, B. & Cunningham J., 2017). Whether athletes are ready for Big Data, it does appear to stay. The NCAA recently approved the use of some wearable technologies in football games for the purpose of health and safety. (Dodd, 2014). The rapidly increasing rate of collection in a wide variety of athletes' biometric data will inevitably lead to issues of legality and ethical consequences in the near future. The remainder of this STS research

will examine some of the possible ethical concerns surrounding the collection of this biometric data and how athletes might respond.

The first concern is the validity of our student athletes' biometric data. Polar sensors and Catapult monitors are used pervasively throughout our student athlete population yet the results are improperly studied and not all information is disclosed. Bioethicists Katrina Karzakis of Yale Law School and Jennifer R. Fishman of McGill University pointed out that the perception of biometric wearable devices is that they are precise and objective measures of an athletes' performance. However, the reality is that the device accuracy and validity of the conclusions drawn is largely unknown. Karzakis commented, "There is currently no standard for evaluating the technology or the leagues' and teams' adoption of new methods." (Zagger, 2019). It's up to investigators to conduct research on the validity of biometric data and disclose all information regardless if it is positive or negative.

One likely response athletes will have to the Big Data boom is the issue of privacy, specifically the increased surveillance of players and the associated risks regarding data security. Standards need to be developed on how to best collect, store, and use biometric data to both optimize its usefulness while also respecting the privacy rights of our student-athletes. (Arnold & Sade, 2017). Where Big Data monitoring could prove to be problematic is through health data. This data is primarily collected through off-field monitoring. Teams are concerned with what their athletes are doing away from official competition. Their goal is that the players are getting enough sleep, eating healthy, and avoiding detrimental habits. In the past, there was no clear way of monitoring this information. However, with technological advances in athlete biometric data it is now possible to quantify each of these activities and form metrics. Pete D'Alessandro, current Denver Nuggets front office executive, expressed his desire for this form of data, "We need to be

able to have impact on these players in their private time." (Baerg, A., 2017). Data collected in this manner would become part of the athlete's health record. Ultimately, teams are responsible for the data. A representative from a major wearable technology company reported that no injury-related data is stored but in fact solely de-identified movement-based data. (Osborne, B. & Cunningham J., 2017). This data is stored in the cloud and only used by the company to develop new algorithms. The team, on the other hand, owns the raw data files. This is why the current status of athletes' biometric data with regard to privacy relies on the management of the team to regulate the protection of the player data and the extent to which the data is disclosed.

Additionally, there is the ethical concern of conflicts of interest. Universities often get roped into projects without fully understanding the implications. For example, the University of Michigan recently signed a $170 million contract that allows Nike to harvest the personal data of their student athletes from their wearable technology. (Arnold & Sade, 2017). This contract lacks the protection of our student athletes' data and there are concerns whether the university even fully understood the nature of the contract. Karzakis commented, "I wonder if Michigan knows what they've just done." (Tracy, 2016). Other corporations are also looking to take advantage of our student athletes' data. This is especially true at the NCAA level where historically student-athletes have been manipulated for monetary concerns and brings up the ethical concern of coercion. Coaches and fellow teammates may apply peer pressure to reveal personal data. Recently, the NCAA has been prioritizing the monetization of player statistics. In 2017, the NCAA announced the launch of a new initiative: monetizing college athletes' game statistics. The NCAA signed a 10-year partnership with the U.K.-based Genius Sports to centralize the data. (Edelman, M., 2018, May 14). The worry is that the NCAA will sell this personal data to gambling-related ventures. Acts such as these have raised contention for years and have finally

culminated in the recent decision to allow college athletes to get paid for use of their names, images, and data. (Dwyer, C., 2019, October 29). In this surprise move, student-athletes are now able to profit from their images and performance on the field. While the timeline for implementing the changes is still unclear, this is a blockbuster move for the NCAA and data analytics in collegiate sports.

When examining such intricate problems as the ethical considerations of data collection in sports, it can be helpful to utilize an analytical framework to uncover all the actors responsible for the problem's development. Looking at the impact of Big Data from a sociotechnical perspective, one begins to see why biometric data is of ethical concern. Technically, data analytics is advancing at an unprecedented rate, with the amount of available data nearly doubling every two years. (Lohr, 2012, para. 8). Organizationally, there are huge economic benefits that businesses stand to benefit from consumer analytics. Finally, culturally, consumer data manipulation has several potential negative impacts and ethical issues. Application of this framework in my STS research paper will aid in the identification of the background forces behind the data privacy concerns, underlying a field where data analysis and technology certainly take place. Furthermore, the monetization of player data gives an economic perspective to this sociotechnical system of sports analytics. Combining a sociotechnical investigation with the previously mentioned economic analysis should result in a cohesive look into what data privacy means for our players at both the collegiate and professional level.

Conclusion

Continuing advancements in technology, coupled with the rise of Big Data, provide the potential for data analytics to thrive at both the professional and collegiate level of sports. The

technical portion of my thesis project will culminate in the spring with a report submitted to Ted White and Jim Ryan with a conceptual design for a sports performance analytics center. The STS research paper will explain some of the implications that a sport analytics could have on our athletes and summarize some of the key concerns athletes have with their private data, namely biometric health data.

While the proposed design of the sports performance analytics center is only conceptual, future capstone projects will be able to build upon it and ideally begin conducting their own data modeling in time for the new year. The hope is that all varsity athletics at UVA and their data can be consolidated in one location. Teams can learn from each other, athletes can perform better, and students can hone their data analytic skills in an exciting manner.

REFERENCES

Alamar, B., Mehrotra, V. (2011, October). Beyond 'Moneyball': Rapidly evolving world of sports analytics, Part I. (2016, July 14). Retrieved October 31, 2019, from http://analytics-magazine.org/beyond-moneyball-the-rapidly-evolving-world-of-sports-analytics-part-i/.

Armstrong, S., Jovanov, E., & Kerwin, D. (2007, April). Wireless connectivity for health and sports monitoring: a review. *British Journal of Sports Medicine*. Retrieved October 31, 2019, from https://bjsm.bmj.com/content/41/5/285

Arnold, J. F., & Sade, R. M. (2017). Wearable Technologies in Collegiate Sports: The Ethics of Collecting Biometric Data From Student-Athletes. *The American journal of bioethics : AJOB*, 17(1), 67–70. doi:10.1080/15265161.2016.1251648. Retrieved December 4, 2019.

Baerg, A. (2017). Big data, sport, and the digital divide: Theorizing how athletes might respond to big data monitoring. *Journal of Sport and Social Issues, 41*(1), 3-20.

Dodd, D. (2014). NCAA denies ACC use of helmet cams, sideline communications. *CBS Sports*. Retrieved December 9, 2019.

Dwyer, C. (2019, October 29). NCAA Plans To Allow College Athletes To Get Paid For Use Of Their Names, Images. Retrieved October 31, 2019, from https://www.npr.org/2019/10/29/774439078/ncaa-starts-process-to-allow-compensation-for-college-athletes.

Edelman, M. (2018, May 14). The NCAA Is Prioritizing Monetizing Player Statistics Over Protecting Student-Athletes. Retrieved October 31, 2019, from https://www.forbes.com/sites/marcedelman/2018/05/14/the-ncaa-is-prioritizing-monetizing-player-statistics-over-protecting-student-athletes/#17324138655c.

Hayhurst, C. (2019, August). Data Analytics Helps College Coaches and Athletes Optimize Training and Performance. *EdTech Magazine.* Retrieved October 31, 2019, 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. Retrieved October 31, 2019.

Lohr, S. (2012). The age of big data. *New York Times*, *11*(2012). Retrieved December 10, 2019.

McElroy, W. (2019, June). UVA has had the best athletic year of any college sports program. *Richmond Times-Dispatch.* Retrieved October 31, 2019, from: https://www.richmond.com/

Osborne, B. & Cunningham J. (2017). Legal and Ethical Implications of Athletes' Biometric Data Collection in Professional Sport, *28 Marq. Sports L. Rev. 37*. Retrieved October 31, 2019, from https://scholarship.law.marquette.edu/sportslaw/vol28/iss1/3.

Siemens, G. Dawson, S. & Lynch, G., (2013, December). Improving the Quality and Productivity of the Higher Education Sector. *Society for Learning Analytics Research.* Retrieved October 31, 2019, from http://www.solaresearch.org/

Smith, B. (2019, August). A Conversation with Chris Hayhurst. *EdTech Magazine.* Retrieved October 31, 2019, from: https://edtechmagazine.com/higher/

Tracy, M. (2016, September 9). With Wearable Tech Deals, New Player Data Is Up for Grabs. Retrieved December 10, 2019, from https://www.nytimes.com/2016/09/11/sports/ncaafootball/wearable-technology-nike-privacy-college-football.html?auth=login-facebook&login=facebook.

University of Virginia: Office of Communications. (2019). *Great And Good: The 2030 Plan*. Charlottesville, VA: Author. Retrieved October 31, 2019.

Zagger, Z. (2019, October 24). States On Deck To Protect College Athlete Biometric Privacy. Retrieved December 9, 2019, from https://www.law360.com/articles/1212731/states-on-deck-to-protect-college-athlete-biometric-privacy.