

Data Pipeline for Digitizing Perioperative Flowsheets from Low Middle Income Countries
(Technical Report)

How Data Analytics is Changing the Sport of Football
(STS Research Paper)

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

Darren Klein

November 1, 2021

Technical Team Members:

Christos Chen

Mariam Guirguis

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.

ADVISORS

Bryn Seabrook, Department of Engineering and Society

Don Brown, School of Data Science

Introduction

The hospital system in Rwanda is growing and hospitals are being built at a rapid rate. As a result of this, hospitals are often understaffed in terms of both nurses and doctors. Additionally, all of the Rwandan medical records are written on sheets and forms in the hospital rather than being electronic or digitized. Since the medical professionals in the Rwandan hospital networks have more patients than they can reasonably provide sufficient care for, 5-10 million patients die every year as a result of post-procedure complications. Most of the procedures performed by Rwandan surgeons are deemed low-risk procedures by hospitals in high income countries, such as the United States, the European Union, and Japan, yet they have a high postoperative mortality or complication rate. These deaths could be easily prevented if doctors and nurses in Rwanda and other under-developed nations had access to the data archives and resources that are present in the United States and the other aforementioned high-income nations. The final technical deliverable, a data pipeline between the University of Virginia Health System and the Rwandan Hospital Network specifically focusing on the digitization and analysis of perioperative surgical sheets, will reduce the amount of preventable deaths in the Rwandan Hospital Network through providing Rwandan doctors and nurses with real-time updates on patient risk and possible conditions.

In a different type of data analysis, sports analytics is changing the way professionals view their field of study and make real-time decisions. In the sport of football, sports analytics are influencing both in-game and personnel decisions and an exponentially increasing rate. The NFL and college teams that have utilized sports analytics to the highest extent have been extremely successful in their pursuits of championships, including perennial college football dynasties such as the University of Alabama and Clemson University, as well as NFL teams the

New England Patriots and the Baltimore Ravens. 96% of NFL teams surveyed stated that they utilized player data to build and use metrics to influence decisions (Walder, 2020) The increasing utilization of sports analytics in football has entirely changed the sport, affecting recruiting and personnel management as well as in-game decisions made by the head coach, and will be paramount to the future success of the sport.

STS Topic

Football is one of the oldest sports in America. First starting to be played in 1869(Staff, 2013), the sport of football has captivated the attention of many generations. Continuing to be played throughout chaotic American time periods such as the Jim Crow era, the Great Depression, both World Wars, the Civil Rights Movement, and the Vietnam War, the sport of football has persisted and passed the test of time. Since its' creation, the sport has changed in many different ways. Popularity, equipment, rules, compensation, and many other components have changed over the course of the last 152 years, but football has remained a presence and a staple in American sports and entertainment. Football is the most popular sport in America (Gallup, 2020) with 37% of Americans claiming it to be their favorite sport in 2017. As society has progressed since 1869, the sport of football has progressed with it. Inventions such as radio broadcasting, the jet engine, the internet, computers, developments with the materials plastics, rubbers, and metals, and innovations such as improved camera, audio, and video quality have all increased the popularity and success of the sport of football. However, one of the greatest changes to the sport of football over the past decade and the one that has the capability to revolutionize the sport as a whole is the influence and implementation of data. Data analytics are used by all 32 professional NFL teams, and over 75% of all NCAA football teams. Data

analytics are also used by the NFL to attract the attention of fans and viewers alike and improve their gameday experience, whether that be in person attending games at the stadium or watching the game through a multitude of different methods at home. How has data analytics impacted the sport of football and how will it continue to change the sport through game decisions, team management, and player health and safety?

With a substantial amount of research and studies being done in the current fields of player safety and in-game coaching decision making, there is less being done to project the future impact of this data analytics onto the future of the sport. These wide-ranging studies and research to predict how data analytics will change the sport of football, specifically in the fields of in-game decision making, player health, and team/roster management and acquisition.

Research Question: How has data analytics impacted the sport of football and how will it continue to change the sport through game decisions, team management, and player health and safety?

The STS methods of technological determinism and paradigm shift theory will primarily be used to establish what the current status quo is as well as create predictions and projections as to what data analytics will do to football. The technological determinism framework will set up an analysis and projection of how the field of data analytics can and will completely change the sport of football. The paradigm shift theory framework will help me to set a background for what the current and past methodologies are and were, and then analyze and project what the status quo will be in the future. Some criticisms of these frameworks are that they are conducive to a natural growth environment in society, and that progress and shift changes will occur naturally without the presence of outside influence. These frameworks will be used to analyze the past and present usage of data analytics in the sport of football as well as project and predict how data

analytics will dynamically change football moving forwards, specifically in the fields of player health and safety, game management, and personnel management and acquisition. I will collect my data and information through both primary and secondary sources, from various different academic journals, previous capstone projects, articles, and some interviews with individuals that I have personal relationships with that are experts in their respective analytical fields. To gather my research on player health and safety, I am looking at medical journals and studies that specifically utilize GPS and player biometrics tracking. To gather information on in game decision making, I am using interviews with NFL data analysts, polls of data analytics departments in the NFL, research projects on statistical analyses, and I am hoping to interview the head data analyst of the University of Virginia Football team. To gather information on team management, I am utilizing previously created research models, interviews with personal management staff, articles written about player acquisition, and I hope to interview the director of Player Personnel for the University of Virginia Football team. I have organized my data and resources by theme, and have a wide range of different mediums of conveying information and data.

Technical

Hospitals in Rwanda are growing and being built at a rapid rate. As a result of this, hospitals are often understaffed in terms of both nurses and doctors. Additionally, all of the Rwandan medical records are written on sheets and forms in the hospital rather than being electronic or digitized. Since the medical professionals in the Rwandan hospital networks have more patients than they can reasonably provide sufficient care for, 5-10 million patients die every year as a result of post-procedure complications. Most of the procedures performed by Rwandan surgeons are deemed low-risk procedures by hospitals in high income countries, such

as the United States, the European Union, and Japan, yet they have a high postoperative mortality or complication rate. Cesarean sections are the most common low-risk procedures where complications and deaths often occur in Rwanda. These complications occur because the nurses have more patients than they can handle, to the point that family members are tasked with providing care, including food, for their loved ones post-operation.

The project was initiated a couple years ago as a collaboration between doctors in Rwanda and the University of Virginia Anesthesiology department. The goal is to create a method of converting the current method of operation to a more digitized and technologically-friendly recording approach. Currently, perioperative surgical sheets and notes are taken by hand, on paper, by Rwandan anesthesiologists. The goal is to provide a system where doctors can upload their paper surgical sheets and almost instantaneously receive feedback about the patient's risk for postoperative medical issues. This feedback can assist both doctors and nurses in providing critical care to at-risk patients, as well as being used as a learning material to help train residents identify risk and provide patients with the sufficient care necessary, in an effort to alleviate complications and death in the Rwandan Health System. The ultimate goal of the project is to implement and adapt the approach used for Rwanda to other hospitals across the world, specifically in South Africa.

Project Objectives/Goals

The SE Capstone Group will focus on improving the current data pipeline to ingest and read surgical indicators of perioperative surgical sheets sourced in Rwanda. Currently, Rwanda is using a web application to digitize the perioperative surgical sheets. However, an android application has been built in previous years without much testing. Hence, the first major scope of this project is to perform software testing to determine and resolve the bugs. In addition, direct

integration between the mobile app and the analytics system will be developed. Then, user testing will also be conducted to determine how the android app can be modernized and how to create the app to be as simple and intuitive as possible. This should build and complement previous work. Once the app is modernized, the final functional requirement that needs implementation is the capability of sending reminders to the user to check on certain patients given the result from the UVA analytics system. Hence, time permitting the next step of the project would be to allow for this implementation and perform testing on the implementation. Once this is completed, the goal is to provide training to those in Rwanda to help integrate this android application in the hospital.

Technical Approach

The scope of the project has 5 main objectives, listed below:

Software Testing

The current android application has not undergone intensive testing. Hence, the first steps will be to perform functionality testing and implementation testing. Since the current SE capstone group has no familiarity with the app, they will provide an unbiased opinion with regards to the functionality and implementation of the android app. Integration testing will also be performed to ensure the longevity of the android application.

Integration

Direct integration between the mobile app and the analytics system will also be needed. First the SE Capstone Group will test the app to determine the current state of the integration. Then, the SE team will build direct integration between the android application and the UVA analytics

system. In addition, another integration between the mobile application and the cloud database will be developed. A literature review will be conducted to determine the best method.

User Testing

After completing the software testing and integration, acceptance testing will be conducted. Meetings with doctors, nurses, and hospital workers will be scheduled to allow the SE team understanding of pain points of the Rwandan health care workers with regards to the android application. After receiving feedback, the android application will be updated accordingly.

Risk-Based Notification Functionality Addition

As mentioned multiple times by Dr. Marcel Durieux, the goal of the application is to provide real-time notifications to doctors and nurses to check on patients at certain times based on the results of the analysis of the perioperative flowsheets. Currently, the android app lacks this functionality. Hence, a comprehensive literature review will be performed to understand the best way to add this functionality. Based on the research and time, the functionality will be added to the android application and testing will be performed to improve the functionality. (The technical approach has not been decided as a literature review has not been performed yet.)

Flowsheet Viability Scoring Functionality Addition

After the scanning of the surgical flowsheet via the mobile application, analytical techniques such as computer vision(i.e. darkness of ink) and edge detection will be employed. The scoring will ideally set an acceptance threshold based on criteria that will assess the quality of the data that is able to be pulled from the sheet. The viability score will likely be assigned from 0-100.

This functionality will allow real time feedback on the quality of the flowsheets, and it will ensure that the data quality and accuracy will be improved.

Conclusion

This paper covers an investigation into the usage of data analytics in the sport of football as well as how data pipelines and analytics can save lives and ensure hospital patients in Rwanda receive necessary care on time. The team will create a data pipeline between the University of Virginia Health System and the Rwandan Hospital Network to allow for perioperative flowsheets to be digitized and sent back and forth, along with necessary information pertaining to patients present and future medical risks and conditions.

On a similar note of data analytics, this paper explores the effective usage of data analytics in the sport of football, both professional and college football, with an emphasis on the NFL. Specifically, this paper targets the fields of player health and safety, game management, and personnel management and acquisition. As a whole, society is turning towards more data-driven decisions and data-centric metrics, and the sport of football is following those same trends. This change of methodology is important to notice and study because it is indicative of the future, in both society and in football. This paper will analyze the current usage of data analytics in football and utilize that to project and predict what the future of football will look like.

References

Beckwith, J. *et al.*, "Systems Analysis for University of Virginia Football Recruiting and Performance," *2019 Systems and Information Engineering Design Symposium (SIEDS)*, 2019, pp. 1-6, Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8735611>
doi: 10.1109/SIEDS.2019.8735611.

This article is actually another Capstone project at the University of Virginia, this time from 2019. The project dives into the usage of data analytics in recruiting as well as on-field performance. The on-game performance component studied sub-optimal play calling areas as well as opponent scouting, while the recruiting component targeted overperforming athletes and a commitment likelihood predictor.

I will use this article as evidence that data analytics impacts the success of football programs, and will reference statistics and records to show this. The project utilizes paradigm shift theory and technological determinism, and I will use the project to show both theories.

Corscadden, J. *et al.*, "Developing analytical tools to impact U.Va. football performance," *2018 Systems and Information Engineering Design Symposium (SIEDS)*, 2018, pp. 249-254, Retrieved from <https://ieeexplore.ieee.org/abstract/document/8374746>
doi: 10.1109/SIEDS.2018.8374746.

This paper is actually a Capstone project from 2018 that analyzes the use of data analytics in the University of Virginia Football program. The project focuses on the correlation between practice workload and in game performance, as well as looking into a

4th down decision predictor model. This project utilizes examples of both paradigm shift theory and technological determinism, and I will use it to show both theories.

Ferguson, Daniel David Reid, "Utilizing Analytics in American Football to Improve Decision Making on Fourth Down" (2018). *Graduate Theses - Engineering Management*. 7.

Retrieved from

https://scholar.rose-hulman.edu/engineering_management_grad_theses/7

The article by Ferguson details data analytics, both predictive and retrospective, specific to 4th down decisions in football. The article utilizes a formulated strategy from statistical inputs to determine whether or not a coach should decide to punt, kick a field goal, or go for it on 4th down.

I will use this article in my exploration of how data analytics influences coaching decisions. This source provides evidence to the types of processes that are used in formulating data-driven decisions. This article provides examples of both technological determinism and technological momentum.

Fox, L. (2021, August 12). How The NFL Uses Analytics, According To The Lead Analyst Of A Super Bowl Champion. Retrieved from

<https://www.forbes.com/sites/liamfox/2021/08/12/how-the-nfl-uses-analytics-according-to-the-lead-analyst-of-a-super-bowl-champion/?sh=7d8f26aa424e>

This article explores the usage of data analytics in sports, especially the role of Ryan Paganetti, who was a data analyst for the Philadelphia Eagles and who made a large impact on the team when they won the Super Bowl in 2017. The article along with quotes from Paganetti, looks into the statistics of Expected Points Added (EPA) and EP (Expected Points). The article also dives into the copycat nature of the NFL, and how

Paganetti and the Eagles' success with analytics led to the rest of the league mimicking and adopting similar ideologies towards data analytics. This source will be used as evidence that data analytics directly affects success, and provides examples of both paradigm shift theory and technological determinism.

Gallup. (2020, March 30). Sports. Retrieved from <https://news.gallup.com/poll/4735/sports.aspx>

This article is a poll that summarizes what sports are most popular in the United States. I will use this information as background to state that football is the most popular sport in America.

Ghasem, W., Valenzuela, J., & Saxon, L. A. (2021). Player Tracking Technology and Data for Injury Prevention in the National Football League. *Current Sports Medicine Reports*, 20(9), 436-439. Retrieved from: https://journals.lww.com/acsm-csmr/Fulltext/2021/09000/Player_Tracking_Technology_and_Data_for_Injury.3.aspx?cohtml=LatestArticles doi:10.1249/jsr.0000000000000873

This article speaks to the recent increase in attention towards sports analytics, specifically citing Amazon's "Amazon Web Services", as well as the usage of GPS and biosensor technology to analyze injuries and injury prevention in the NFL. The article also discusses the transition from data as a tool for fan engagement to a tool for team doctors and personnel departments to make informed and better decisions.

I will use this article as evidence that data based analytics in injuries can be used to prevent and increase recovery times. This article uses the STS concept of paradigm shift theory, and I will use it as examples of the theory.

Li, R. T., Salata, M. J., Rambhia, S., Sheehan, J., & Voos, J. E. (2019). Does Overexertion Correlate With Increased Injury? The Relationship Between Player Workload and Soft Tissue Injury in Professional American Football Players Using Wearable Technology. *Sports Health: A Multidisciplinary Approach*, 12(1), 66-73.
doi:10.1177/1941738119868477

This medical journal explores and studies the effects of overexertion and injury. The study utilizes wearable tracking devices to analyze workload and determine if workload affects injuries. I will use this journal as both background and evidence that data analytics can help prevent injuries. This journal provides examples of paradigm shift theory, and I will use it as evidence to show that theory.

Numbers game: Analytics spreading through college football. (2017, August 06). Retrieved from <https://www.usatoday.com/story/sports/ncaaf/2017/08/05/numbers-game-analytics-spreading-through-college-football/104310592/>

The article describes the rise of the presence of data analysts in the world of college football, and how necessary certain coaches deem analytics to be in their decision making. Specifically, companies called Championship Analytics Inc. (CAI) and Zcruit are discussed, as well as the multitude of college coaches that use and endorse their products. CAI focuses on analytic decisions in game, while Zcruit focuses on analytics in recruiting.

I will use this article in my exploration of how data analytics affects both recruiting and in game coaching decisions. This source provides evidence to the impact that analytics has in the world of college football, and provides examples of technological determinism and to an extent, paradigm shift theory.

PFF Team Services. (n.d.). Retrieved from <https://www.pff.com/team-services>

This source is simply a fact of the number of football teams at both the professional and collegiate levels that utilize Pro Football Focus, a pioneering data analytics company. I will use this article as examples of technological determinism.

Raphael, Rudy. "Paul DePodesta and the Power of Analytics." (2020, December 10). Retrieved from <https://www.greenbook.org/mr/brand-impact/paul-depodesta-and-the-power-of-analytics/>

This article looks into the impact of data analytics across all mediums, but especially on Paul DePodesta, an analytical guru in the world of American football. I will use this information as background for the field of data analytics, and to explore how many different fields it can affect. The article does not utilize any STS theories, but since I will be using it as background, I will not be using it as evidence to support any theory.

Staff, H. (2013, September 25). Who Invented Football? Retrieved from <https://www.history.com/news/who-invented-football>

This article summarizes the history of football. I will use this article for background information, specifically to cite what year football was invented and started to be played.

Stodden, D. F., & Galitski, H. M. (2010). Longitudinal Effects of a Collegiate Strength and Conditioning Program in American Football. *Journal of Strength and Conditioning Research*, 24(9)

This journal summarizes the effects of data analysis on strength and conditioning training for American football athletes. Specific measurables explored in the article and study are both strength based, using bench press and chin ups, speed/explosivity based, using 40

yard dash, 5-10-5 shuttle, and vertical jump, and build based, using weight, body fat, and lean muscle mass. It was observed that there were measurable results through multiple years of training.

I will use this information as evidence that data-based analytics reduces injuries and can lead to higher performance levels. This source provides examples of paradigm shift theory.

Walder, S. (2020, September 24). 2020 NFL analytics survey: Which teams are most, least analytically inclined? Retrieved from

https://www.espn.com/nfl/story/_/id/29939438/2020-nfl-analytics-survey-which-teams-most-least-analytically-inclined

This article looks into which NFL teams are most and least inclined to use data analytics, as well as what areas of the NFL are most affected by data analytics. I will use this information as background as well as evidence to draw conclusions between data analytics and success in terms of wins and losses. This article does not specifically utilize any STS concepts, but I will use it to support evidence of technological determinism.

Walter, L. *et al.*, "Implementation of a recruit visualization tool for UVA football," *2017 Systems and Information Engineering Design Symposium (SIEDS)*, 2017, pp. 168-173,

Retrieved from <https://ieeexplore.ieee.org/abstract/document/7937710>

doi: 10.1109/SIEDS.2017.7937710.

This project is yet again a former Capstone project, this time from 2017 at the University of Virginia. This project explores a model that is predictive of how well a player will fit into the University of Virginia football program, as well as predicting possible success. I will use this article as evidence that data analytics impacts the success of football recruiting, and will reference statistics and records to show this. The project utilizes paradigm shift theory and technological determinism, and I will use the project to show both theories.