

Celebrity Sports Twitter Commentary

Assessment 3 – Group Project

MARCH 17

University of Aberdeen
Authored by: Group Sports



UNIVERSITY OF
ABERDEEN

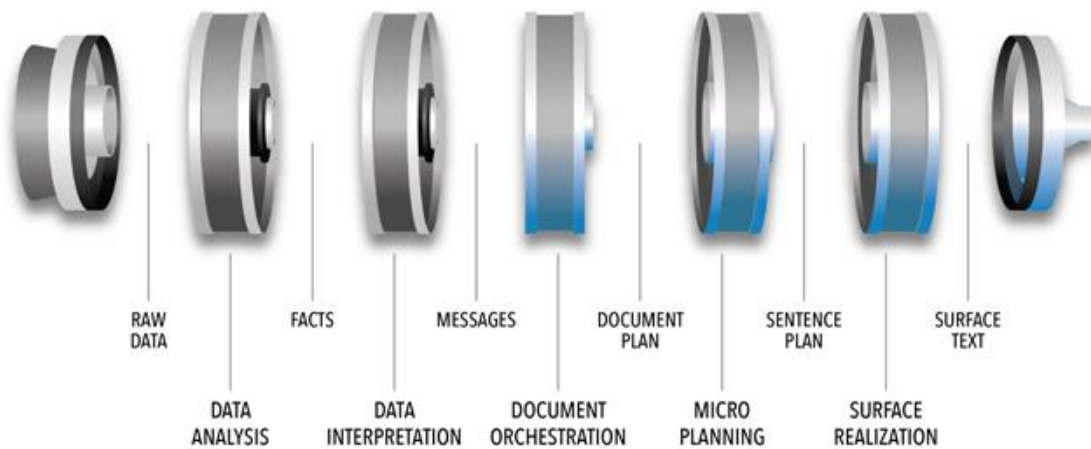
Contents

Introduction	3
1. Vision	4
2. Market.....	5
3. Competitors	6
4. Cost	7
4.1 Office & Workspace	7
4.2 Salary/ Staff	7
4.3 Marketing.....	7
5. Why will people buy this product?	8
6. Development.....	9
Reference	10

Introduction

Natural Language Generation (NLG) is the is language technology that focuses on generating text from structured data or representations. To simplify, NLG can think as a translator, taking the numerical data and representations that exist in everyday business, such as sales figures and produce a narrative by generating natural language.

THE ARRIA NLG ENGINE



At the Heart of the Most Advanced Natural Language Generation Platform

Figure 1: Arria NLG Pipeline

What do we mean by natural language?

A natural language is the language spoken by humans, meaning a language we can read and understand. By generating a natural language, we can bring meaning to data and numbers. Companies such as Wordsmith, offer services to clients to turn data into text in any scale or format. Selling a service to create and gather insights from their data by generating text from their data in multiple spoken languages (*Word Smith, 2019*). Though NLG has been studied for more than 30 years, recently its commercial potential has only recently started to emerge. Our product is based around this technology, demonstrating the capability of its potential in area of sports.

“Arria’s Natural Language Generation (NLG) is revolutionizing how Industries are digitizing their business processes. Using NLG organizations are building new customer experiences, monetizing information assets, introducing new offerings and streamlining operational costs.” – Arria, 2018

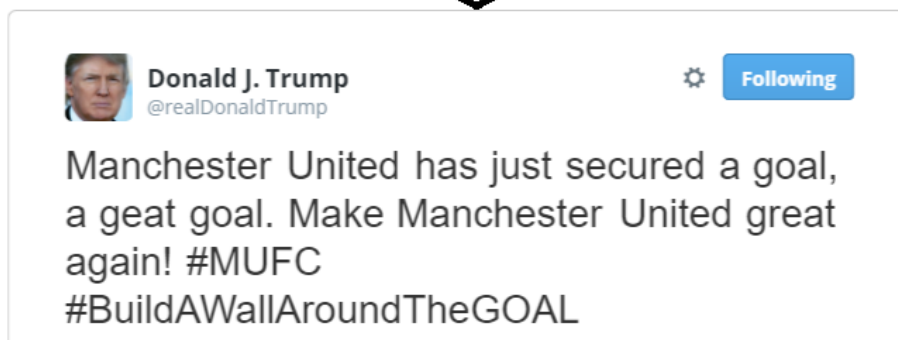
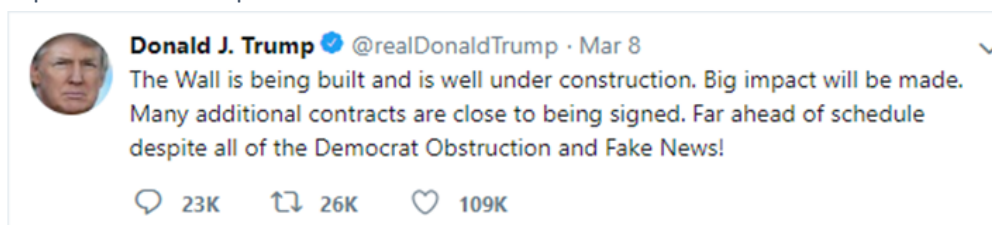
1. Vision

Sports is one the largest consumers market in the world with the British Premier League being the ‘most-watched sports league in the world, broadcast in 212 territories to 643 million homes and a potential TV audience of 4.7 billion people’ (*British Council, 2018*). Many sports have existed for over centuries and has always seemed to evolve and adapt to the modernity of its time, from improvement in broadcasting to goal line technology, sports have become of the most accessible pass time in our world. With the emergence of **Social Media**, the game is more up to date and quick than it has ever been before, platforms such as Twitter giving all teams, fans and organizations a voice.



Source: Manchester United Twitter

However, platform(s) such as Twitter bring more than just high pace up to date information to those in sports. It is controversial platform by nature used by many all around the world for many reasons and purposes, to share thoughts, create discussion and much more. Controversial figures such as Donald Trump use Twitter to **broadcast** political opinions and updates across the world. People such as Trump are controversial figures, not just based around politics but even by the way they write. Celebrities such as Morgan Freeman, known for the cinematics personality and voices all have a unique way of writing and presenting their own literature, which itself is within human nature. Our **vision** is to capture uniqueness, reproduce it in the form of a twitter bot posts that keep fans up to date about sports matches in real time.



Donald Trump Tweet transformation example.

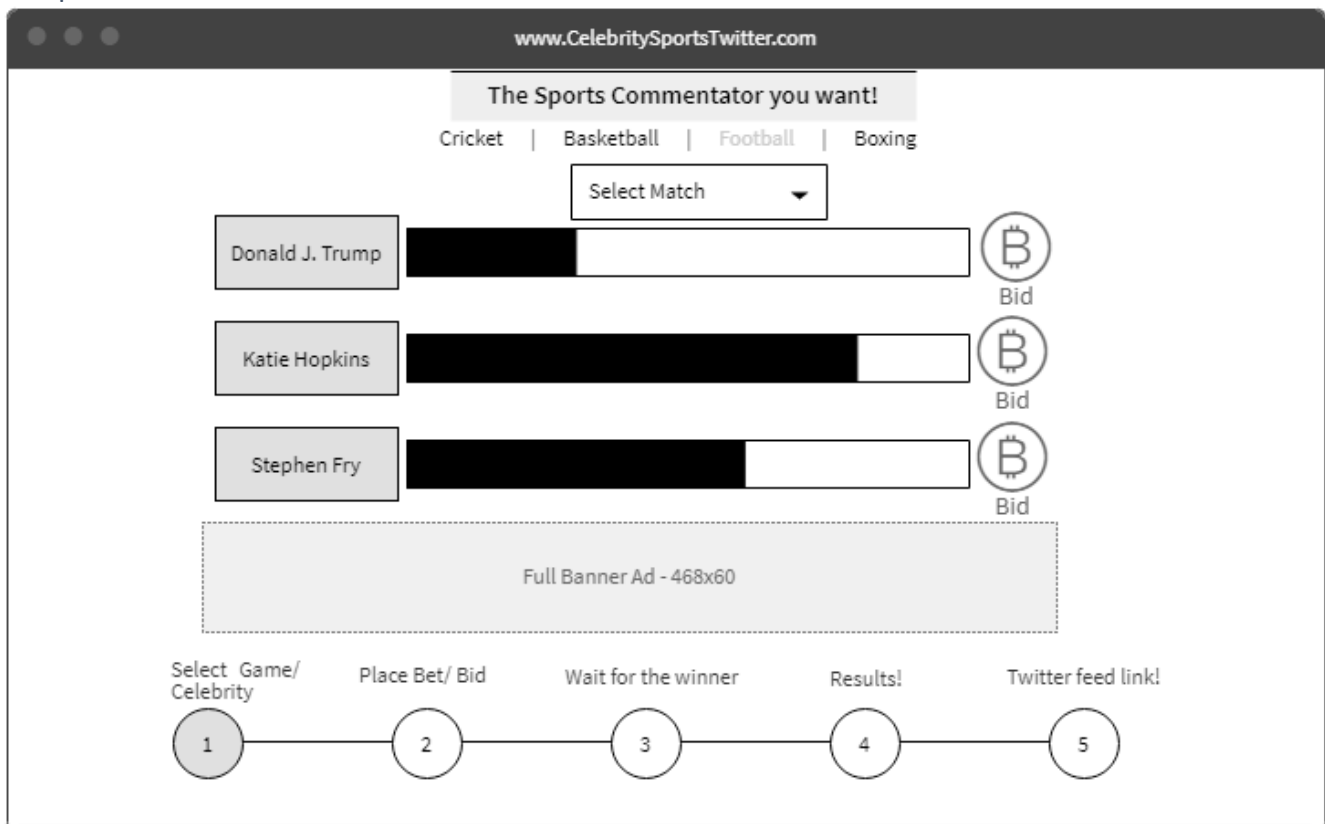
2. Market

As stated the premier league is the most watch football league in the world, with its revenue reaching over £4.5 Billion (*Bill Wilson, 2018*) it is without doubt, an area with huge market potential. We also look to platform we utilize for our service, Twitter. With a user base of over 260 million and roughly 46% of the current user base using it daily we have the potential audience of over 100 million users (*Omni Core Agency, 2019*). The numbers clearly indicate a user base clearly indicates a huge potential market size using a platform allowing access to a huge audience.

The product we propose is a novelty-based product, designed based on satire and viral trends, we look to develop digital product available to a wider audience. Our approach to this unique product; is to develop a web platform to sell the product and use a hybrid-gambling model of which we later discuss our justification based of the competitor market (3).

We consider our target market space to be very fine, in that we are targeting a novelty market. A customer looking for sports updates will traditionally go with a company based on speed and accuracy of the product and how it is delivered. Our market is that of satire creating a novelty service open to as many people as possible and we treat as such.

To understand the market better and why we have adopted our approach we need to first look at competitors.



Wireframe of web platform

3. Competitors

Due to the uniqueness of our product and goal, we widen the scope of our competitor research. We focus on 2 key aspects, **Sports** and **Social Media**. In terms of finding an exact competitor we find none, nobody attempts the service of our nature let alone monetize it. First, we look at spoof twitter accounts which yes already exist, these twitter feeds are not however created by companies.

These tweets while novelty based such as ours, are not direct competitors as their tweets do not tie the goal of commentating sports matches. However, there are twitter accounts that do tweet updates of sport matches, from the clubs of teams participating themselves, companies and even journalists. These we would see as our more direct competitor.



Example Spoof Tweet of Queen Elizabeth

An example competitor as such would be ESPN, a sports company with huge revenue shares and part of Disney offers many services to sports fans for free sports updates. However, these are from their journalists and company twitter accounts and do not utilize NLG techniques to mimic celebrity behavioral tweets (ESPN, 2019). To further our research, we look at companies that offer NLG products and services, once again we see companies do not offer the novelty service that we do directly as a platform. If a consumer would want to achieve our product they would have to engage a potential competitor and develop a solution from scratch.

Due to our lack of competitors for this type of product, we do not have enough research to help decide the monetization strategy for our product. Since there are similar services in terms of twitter sports commentary and satire accounts that are free our selling point is the trained model of mimicking a chosen celebrity. This will not be a set price product. We conclude that a better solution is a hybrid gambling model.

As shown in (2) [the wireframe], we propose a web-based platform using a pay model that incorporates a gambling element. Potential customers come to the website find the football match or sport game they wish to be commentated by the celebrity twitter of their choice and place a bid of £5 or higher. Potential customers can see a progress bar of how the current bidding for each celebrity is going, once bided the timer will expire 3 hours before game time. The celebrity who receives the bids is chosen to commentate, customers who bided on a losing celebrity receive 50% refund of their initial bid. From their all participants are linked to the twitter bot feed to see the tweets in action.

We believe this is the best model to not only create fair profit but will fit our marketing strategy that relies upon the virality of our product. Our tweets are a key part of our advertisement strategy, we wish to include as many participants apart from our product for it to expand and capitalize on our lack of competitors.

4. Cost

For cost analysis, we segment our approach, using research to determine cost estimates and strategy to dictate our budget and spending strategy. Since this is a start up product we consider other elements outside development also, including overheads. We base our findings on research through various, recruitment and housing websites cited.

4.1 Office & Workspace

For office space we believe for our initial development stages and product launch we will only require a small amount due to having low amount of staff and needs. For this assume we will require 1500 square foot of office space, for initial setup potential minor expansion. This space will development desks and conference room.

- £25.00/sq. ft. pa
- Location: Aberdeen
- Estimated total cost pa: £37,500

(Zoopla, 2018)

4.2 Salary/ Staff

Our current team will be comprised of 4 developers, as a small start-up project we do not require huge staffing needs. In terms of requirements for our developers we need programmer with a wider skill base and not just language specialty. For this cost estimation we base it of the job title *Software Developer*.

- £30,000 pa
- Estimated total cost: £120,000 pa

(PayScale, 2019)

4.3 Marketing

For our product/ service to succeed we depend on marketing, an entirely different skillset to that of development needs. Due to the stark differences in skill set we decide to *outsource* our marketing needs. Our cost research is based on the running cost service twitter marketing/ social media marketing rates.

- £2,500 pm
- Total: £30,000 pa

(contentfac.com, 2019)

Name	Cost	Total per annual cost
Office Space (Not including over-head)	• £25.00/sq. ft. pa	£37,500
Software Developer (Staff)	• £30,000 pa	£120,000
Marketing (out-source)	• £2,500 pm	£30,000
		£187,500 total

5. Why will people buy this product?

Our product is that of novelty, a key challenge in our marketing strategy and battle against potential NLG competitors is it justify the use case. Our idea does not create *convenience*, address real world *problem*, instead we focus on novelty. We base our justification of purchase for our product based on the landscape of trends found on the market. Going viral has become more and more key to a lot of marketing strategy, videos, tweets and among other things on the internet can gain record breaking exposure overnight that is just possible through other mediums such as TV (*VeritcalRail.com, 2016*).

Does viral marketing work? One thing to consider this is a high-risk strategy in marketing that does not have guaranteed success but to demonstrate the possibility of viral marketing we refer to the **Egg**. On January 4th, 2019, Chris Godfrey, a gentleman from a marketing agency based in London created a twitter account called the world record egg and posted a plain picture of an egg. The message read: “Let’s set a world record together and get the most liked post on Instagram. Beating the current world record held by Kylie Jenner (18 million)! We got this.” The picture broke world records becoming the most liked picture with over 52 million likes and the account gaining over 9 million followers (*Standard, 2019*).

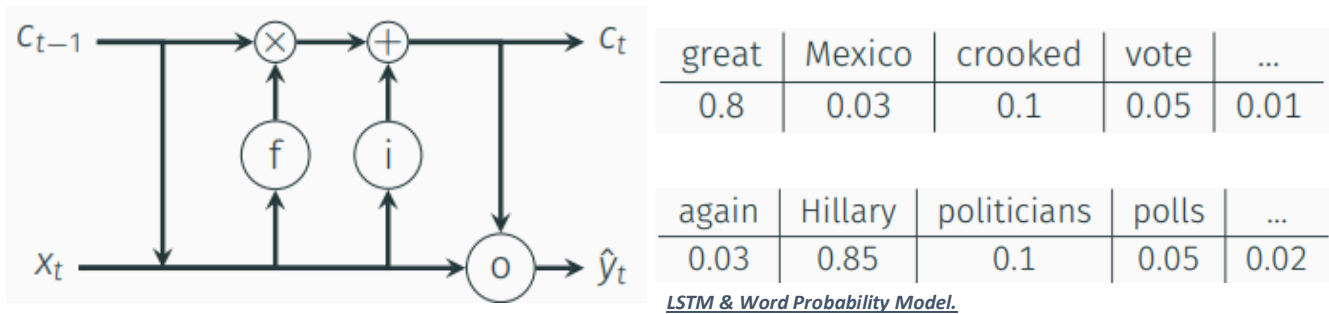
Viral marketing can provide some of the best result and people participate out of trends overcoming our challenges addressing needs and convenience. On top of this when we look at other NLG based companies, none adopt a similar marketing strategy. When identifying the top companies in the space NLG products/ services none adopt a similar marketing technique (*Gauri Bansal, 2018*). Companies such as Arria specialize in certain areas of NLG offering specific products and services that they specialize in; AX NLG Cloud offers service in turn data into any language. Some NLG services only sell in their insights in English. Yet what is common among them is that they do not use our strategy or type of product for marketing.

The closest form in a competitor in this sense of advertisement is IBM Watson. A powerful AI system, that is used in nine languages IBM Watson is one of the world’s strongest AIs available. To help with the promotion and advertisement of Watson, it partook in the famous quiz tv show, Jeopardy! Playing against some of all-time greats of the show Watson was triumphant becoming the winner and over night media hit (*Adam Gabbatt, 2011*). This is the closest that we assess an AI has come to using a similar marketing strategy of using the advantages of going viral.

To summaries we believe will buy our product not out of need, quality, or even convenience, our strategy is to convince potential to partake in a growing viral trend. We believe that when we can produce interesting tweets creating a viral culture people will want to join and take part in our enterprise, something that no NLG or sports twitter-based company has attempted yet.

6. Development

The original concept for this product was based of a presentation given at the University of Victoria, *make neural nets great again*, where academics train a neural network system to produce Trump like tweets (C. Foster, 2019). The presentation demonstrates the academics method using a neural net to establish word probability from Donald trump tweet data for training and producing trump like tweets using a LSTM (Long-short term memory) NLG.



Our System – (Hidden Markov Model)

- To begin our version of this concept we are using similar techniques in acquiring data for our NLG system. We train our model to adopt the chosen celebrity text by using Twitters own API as a scrapping tool to acquire fundamental data. Due to technical difficulties we are not able to apply filter system there for we also include retweets within out data system. We store this data as CSV file.
- Once in CSV we then read and filter empty lines, links and noisy data. We then use POS (Power of speech) tagging for the whole data set.
- From this we establish **hidden Markov model**, where the transitions are the POS tags and the emissions are the words themselves.
- We believe using the Markov model we can take simple sentences and expand them using POS tagging.

Our Platform(s) – Website & Twitter

- We develop a simple HTML website as our prototype to demonstrate how our platform would look the end user wishing to engage with the product.
- Our platform for the actual NLG system will be using Twitter bot accounts using Twitters own APIs.

For our programming language we use Python through **jupyter**, and web development using basic html and simple backend systems.

While experiencing development challenges which we log in our README file of our prototype, our goal was to create enough prototype to demonstrate the intended function and operation of our proposed system. We have developed a basic web design in conjunction to demonstrate the front-end vision of our product, while also separately developing an example Markov model to demonstrate how we intended to generate text for our tweets. We feel given the time scale allowed for our preliminary development we have been to produce a basic outline demonstrating the intended system we desire to create and expand upon in future work.

Reference

- Anon, 2019. How Much Does it Cost to Outsource Social Media Marketing? The Content Factory. Available at: <https://www.contentfac.com/how-much-does-social-media-marketing-cost/> [Accessed March 15, 2019].
- Anon, About Overview. Arria NLG. Available at: <https://www.arria.com/about/about-overview/> [Accessed March 15, 2019].
- Anon, Wordsmith. Automated Insights. Available at: <https://automatedinsights.com/wordsmith/> [Accessed March 15, 2019].
- Bansal, G., 2018. Seven Best Natural Language Generation Companies to Watch For. Analytics Insight. Available at: https://www.analyticsinsight.net/seven-best-natural-language-generation-companies-to-watch-for/?utm_campaign=News&utm_source=hs_email&utm_medium=email&hsenc=p2ANqtz--ip6B1KjQPYYfItRGAUSHJnm0ZjIFswMIahkReR5ziCke7LnstcJYCmKQZ2glIZ16sXYST [Accessed March 15, 2019].
- British Council, Playing the game: The soft power of sport. Playing the game: The soft power of sport | British Council. Available at: <https://www.britishcouncil.org/organisation/policy-insight-research/insight/playing-game-soft-power-sport> [Accessed March 15, 2019].
- C. Foster, D. Gibb, C. Seinen, C. Watson., 2016, Make neural nets great again. Available at: <https://fosterelli.co/file/talk/make-neural-nets-great-again.pdf> [Accessed March 15, 2019]
- Eddy, S.R., 1998. Profile hidden Markov models. Bioinformatics (Oxford, England), 14(9), pp.755-763.
- Edmonds, L., 2019. Creator of record breaking Instagram Egg revealed to be London ad man. Evening Standard. Available at: <https://www.standard.co.uk/tech/person-behind-record-breaking-instagram-egg-post-revealed-to-be-london-advertising-creative-a4056636.html> [Accessed March 15, 2019].
- Gabbatt, A., 2011. IBM computer Watson wins Jeopardy clash. The Guardian. Available at: <https://www.theguardian.com/technology/2011/feb/17/ibm-computer-watson-wins-jeopardy> [Accessed March 15, 2019].
- PayScale, Average Software Developer Salary. PayScale. Available at: https://www.payscale.com/research/UK/Job=Software_Developer/Salary [Accessed March 15, 2019].
- Reiter, E. and Dale, R., 2000. Building natural language generation systems. Cambridge university press.
- Ritter, A., Cherry, C. and Dolan, W.B., 2011, July. Data-driven response generation in social media. In Proceedings of the conference on empirical methods in natural language processing (pp. 583-593). Association for Computational Linguistics.
- Salman.aslam.mughal, 2019. • Twitter by the Numbers (2019): Stats, Demographics & Fun Facts. • Twitter by the Numbers (2019): Stats, Demographics & Fun Facts. Available at: <https://www.omnicoreagency.com/twitter-statistics/> [Accessed March 15, 2019].
- Vertical Rail Staff, V.R., 2016. What Does "Going Viral" Mean? - "Gone Viral" Defined. Vertical Rail. Available at: <https://www.verticalrail.com/kb/what-does-going-viral-mean/> [Accessed March 15, 2019].
- Wilson, B., 2018. Premier League club revenues soar to £4.5bn. BBC News. Available at: <https://www.bbc.co.uk/news/business-43820603> [Accessed March 15, 2019].
- Zoopla, Offices to let in Aberdeen. Zoopla. Available at: https://www.zoopla.co.uk/to-rent/commercial/offices/aberdeen/?floor_area_units=sq_feet&q=aberdeen&radius=0&results_sort=lowest_price&search_source=refine [Accessed March 15, 2019].