Journal of Data Science 23 (1), 149–170
January 2025
DOI: 10.6339/24-JDS1141
Data Science in Action

Visual Analytics for NASCAR Motorsports

Kornelia Bastin[1] and Christopher G. Healey[1,*]
[1]Department of Computer Science, North Carolina State University, Raleigh, NC, USA

Abstract

The National Association of Stock Car Auto Racing (NASCAR) is ranked among the top ten most popular sports in the United States. NASCAR events are characterized by on-track racing punctuated by *pit stops* since cars must refuel, replace tires, and modify their setup throughout a race. A well-executed pit stop can allow drivers to gain multiple seconds on their opponents. Strategies around *when* to pit and *what* to perform during a pit stop are under constant evaluation. One currently unexplored area is publically available communication between each driver and their pit crew during the race. Due to the many hours of audio, manual analysis of even one driver's communications is prohibitive. We propose a fully automated approach to analyze driver–pit crew communication. Our work was conducted in collaboration with NASCAR domain experts. Audio communication is converted to text and summarized using cluster-based Latent Dirichlet Analysis to provide an overview of a driver's race performance. The transcript is then analyzed to extract important events related to pit stops and driving balance: understeer (pushing) or oversteer (over-rotating). Named entity recognition (NER) and relationship extraction provide context to each event. A combination of the race summary, events, and real-time race data provided by NASCAR are presented using Sankey visualizations. Statistical analysis and evaluation by our domain expert collaborators confirmed we can accurately identify important race events and driver interactions, presented in a novel way to provide useful, important, and efficient summaries and event highlights for race preparation and in-race decision-making.

Keywords *analytics; natural language processing; speech-to-text; visualization*

1 Introduction

This paper describes a collaboration between our research group and members of NASCAR's Richard Childress Racing team. The National Association of Stock Car Auto Racing (https://www.nascar.com) is a stock car racing sport in the United States and currently the most popular closed-wheeled racing series in the world. A typical NASCAR race involves approximately 40 cars competing against one another on oval tracks ranging from 0.25 to over 2 miles.

Currently, NASCAR is one of the most lucrative sports in the US, generating an estimated $3 billion per year (ZoomInfo, 2023). Teams are well-funded and interested in any area where a competitive advantage might be gained (Hernandez, 2019; Stoll et al., 2013).

A NASCAR event is a combination of on-track racing and *pit stops* where drivers exit the track to replace tires, refuel, and modify the setup of their vehicles to better fit the current characteristics of the race track. Where millions of dollars in engineering research can produce a few tenths of a second improvement per lap, a well-executed pit stop can produce a multi-second

*Corresponding author. Email: healey@ncsu.edu.

Received February 4, 2024; Accepted June 7, 2024

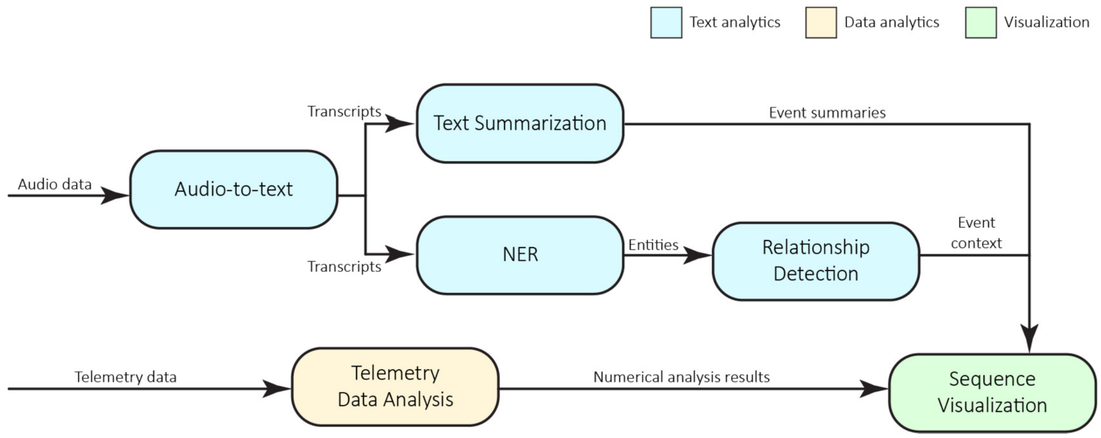


Figure 1: Analytics pipeline for audio and telemetry data conversion to sequence visualization via audio-to-text transcription, event detection and summarization, named entity recognition, relationship detection, telemetry data analysis, and an interactive web-based dashboard containing sequence visualizations of driver events during a race.

improvement over competing teams. Because of this, strategies around *when* to pit and *what* to do during a pit stop are topics of intense study (Heilmeier et al., 2018).

A collection of private and public data is available for each NASCAR race. Public sources of information include real-time SportsMEDIA telemetry feeds (SMT feeds) provided by NASCAR that include vehicle GPS position, speed, acceleration, throttle, brake, and steering position. Additionally, audio communication between drivers and their pit teams is broadcast to allow fans to follow their favorite drivers. Although this seems like a rich source of information, it is currently ignored since the time required to analyze the multi-hour feeds for all drivers is prohibitive. This data may provide two important insights, however. First, it could inform a team of potential issues with a vehicle, suggesting when to pit and what modifications to perform during the pit stop. Second, it could provide details about the strengths and weaknesses of competitor teams' car setups and on-track conditions.

Our high-level goal is to automate identifying and summarizing important pit stop and vehicle balance events within driver–pit crew communications, integrating the events with SMT data, and presenting the results for one or more drivers in an interactive visualization dashboard. To achieve this, we completed the following analytic operations (Figure 1).

1. Transcribe driver–pit crew audio communication into a text transcript.
2. Subdivide text transcripts into "event blocks" to construct event-driven term vectors.
3. Construct a domain-specific dictionary of important terms to re-weight events of importance in each term vector.
4. Perform extractive summarization to compress the text transcript into a compact description of critical events.
5. Apply dependency parsing and relationship extraction to include relevant context in the event summaries.
6. Combine event sequence visualization of important events, excerpts from the race summary, and corresponding SMT data in a web-based interactive dashboard.
7. Validate our results, both statistically and through collaboration with our NASCAR domain experts.

Statistical analysis and anecdotal feedback from our NASCAR domain experts were both positive. We believe our research contributes the following novel contributions to the areas of natural language processing (NLP), analytics, and event visualization.

1. **Audio Summarization.** Audio transcripts are converted to text using audio-to-text services specifically selected to handle the unique characteristics of imperfect audio with rapid back-and-forth communication, poor-quality recording with significant noise, and grammatically incomplete utterances.
2. **Event Detection.** Transcribed audio recordings are converted into "event blocks" to extract important events using a domain-specific dictionary.
3. **Event Summarization.** Extractive text summarization is extended and combined with named entity recognition (NER) and relationship extraction to generate compact, contextual summaries of all important events during a race based on driver–pit crew audio communication.
4. **Sequence Visualization.** Sankey diagrams are modified to generate sequence visualizations of race events, augmented with SMT numeric data, summary text, and important event occurrences.

1.1 Text Analytics

Text analytics involves investigating patterns and properties embodied in text "documents." We briefly review past work relevant to the text analytics task of *summarization.* Three common methods of summarization include *keyword:* a small collection of words (terms) summarize the content of a document collection *extractive:* text from the document collection is extracted verbatim and combined to form a summary; and *abstractive:* unique text is generated to summarize the document collection (Aggarwal, 2018; Zuang and Zhang, 2019).

Preprocessing steps are typically performed prior to most text analytic tasks, including choosing a method to *represent* text (e.g., term vector, *n*-gram); removing *stop words* with no semantic information; conflating words to a common *stem* or *lemma*; weighting the importance of terms based on *term frequency* within a document and *inverse document frequence* across a document collection (TF-IDF (Spaȑch Jones, 1972)).

1.2 Summarization

Summarization compresses a document or document collection into a short, descriptive text summary that highlights the most important details in the original text.

Keyword summarization selects a small set of keywords to summarize a text corpus, for example, selecting the top *n* terms based on their TF-IDF weights. Word cloud visualizations are a visual example of keyword summarization (Coupland, 1995).

Extractive summarization identifies a subset of text from the original corpus to form a summary. Here, the size of the elements extracted and their importance ordering is critical to forming an effective summarization (Gupta and Lehal, 2010; Moratanch and Chitrakala, 2017). Once candidate text is selected, both unsupervised (clustering, graph-based, fuzzy logic, concept-based) and supervised (machine learning) approaches have been proposed to generate the final summary. For example, latent Dirichlet allocation (LDA) converts a term–document matrix (TDM) into a concept–document matrix. The concept–document matrix is used to determine the amount of each "concept" the text candidates contain, allowing the selection of a subset of candidates that span the identified concepts and avoid redundancy.

Abstractive summarization is an open research problem in NLP (Kryściński et al., 2018). Recent work has proposed using deep neural networks and generative adversarial networks (GAN) (Aggarwal, 2018; Rekabdar et al., 2019; Zuang and Zhang, 2019). Other approaches include topic detection, phrase-table machine translation, and quasi-synchronous grammar methods. Many current models apply a `seq2seq` framework of two neural networks to encode input text into fixed-length vectors v, then decode v to predict the input sequence (Xu et al., 2018). Most recently, massive language models like OpenAI's GPT-3 and ChatGPT have provided support for both zero-shot (unsupervised) and trained (supervised) long-form summarization (OpenAI, 2021, 2022; Payne, 2021).

Audio summarization typically starts with audio-to-text conversion followed by text summarization. Characteristics of audio transcripts often produce poor results when fed directly to a traditional text summarizer, however. Lower levels of detail, fragmented sentence structure, grammatically incorrect utterances, interruptions, repetition, and transcription errors due to noise challenge text summarizers, since they were not developed to address these issues. These issues were considered during our investigations.

1.3 Sports Analytics

Analytics and visualization in sports have gained interest in recent years. With more data being collected, domain experts, fans, and sports professionals need ways to efficiently explore, identify, and understand this data. Perin et al. (2018) subdivides sports data into three categories: (1) *box score data:* discrete in-game event data (*e.g.*, scores); (2) *tracking data:* continuous spatio-temporal motion data; and (3) *metadata:* data related to a sporting event (*e.g.*, weather, team colors, or winner predictions). Different sports, including cricket, soccer, basketball, and hockey, have all been visualized in different ways (Chen et al., 2016; Fu and Stasko, 2022; Perin et al., 2013; Pileggi et al., 2012; Tharoor and Dhanya, 2022). This includes motorsports (Stoll et al., 2013), which uses simulation visualizations to model expected vehicle performance, or track visualizations to allow fans to follow drivers throughout a race.

1.4 Sequence Visualization

Sequence visualization displays an ordered sequence of events The most well-known sequence visualization is a line chart showing events and the continuous change between them.

A second example is storyflow visualization. Initially conceived to display characters' actions and interactions in a movie, the idea has been extended to optimized layout algorithms. and movies that contain non-linear timeflow (Padia et al., 2019). Figure 2 shows the interactions between the main characters in the 1977 movie Star Wars. Each line represents a character

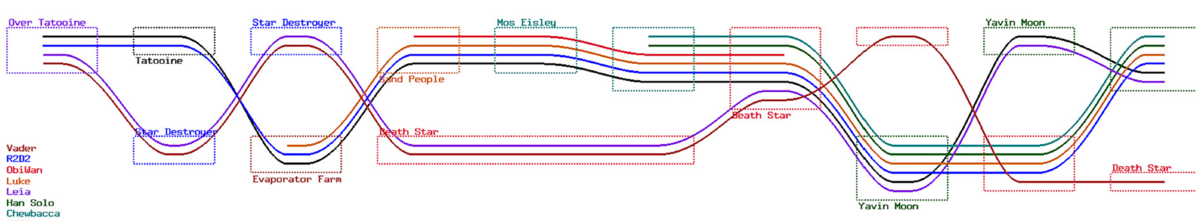


Figure 2: Storyflow sequence visualization following actions between the main characters in the 1977 movie Star Wars.

progressing through the movie from left to right. Nodes where lines meet represent character interactions. If we consider drivers in a NASCAR race as characters, events as actions, and events involving multiple drivers as interactions, it suggests storyflow could be extended to sports visualization.

Another method for visualizing sequence progressions are flow visualizations. Initial "components" move left-to-right over time, splitting or combining to produce intermediate forms. A related method, and the one we use as the basis for our visualization design, is a Sankey diagram. Entities are shown as edges progressing through a common set of stages (events) from left to right. Each step is represented as a vertical axis. The width of an entity's edge between corresponding stages represents an attribute value, an attribute amount relative to all other entities, or both. Captain Henry R. Sankey presented the original Sankey diagram in 1898 to visualize thermal efficiency in a steam engine (Sankey, 1902). An early but very well know use of a Sankey diagram appears in Charles Minard's famous "Napoleon's Russian Campaign of 1812" infographic (Brinch, 2019). Sankey diagrams are generally meant to represent many-to-many relationships. In our case, this corresponds to multiple drivers and multiple events during a race.

2 Audio Summarization

One primary data source for our research was driver–pit crew audio communication for all drivers and races in 2021. Automated voice transcription is an active research area. Many excellent tools are available. However, challenges like accents, background noise, and speaker differentiation still exist (Bokhove and Downey, 2018; Johnson, 2011; Wienrich et al., 2021).

We compared four leading audio transcription systems: IBM Watson (IBM, 2022), Google Cloud Speech-to-Text (Google, 2022), Amazon Transcript (Amazon, 2022), and Microsoft Azure Speech to Text (Microsoft, 2022). IBM Watson and Google were eliminated because they could not manage the speed of communication, at times omitting entire sentences. Amazon and Microsoft did not suffer from this problem, and both had comparable transcription accuracy, functionality, and speed. However, customizing Azure to handle domain-specific text is less demanding, requiring as little as 400 domain-specific example term uses. Amazon requires a minimum of 10,000 examples and prefers 100,000 or more. The minimal requirement for Azure allowed us to improve our base model quality and raise accuracy by multiple percentage points with a limited training set. The final accuracy for the retrained Microsoft model approached 90% versus Amazon's baseline 80%. Table 1 shows examples of Azure's base model performance (left) and domain-trained performance (right), where highlighted words show errors in the baseline model that are corrected in the domain-trained model.

2.1 Transcript Preprocessing

Although text preprocessing in NLP is well-defined, audio transcripts are not the same as written text. They include redundant information, repetitions, repairs, word fragments, incomplete or incorrect grammar, and linguist inaccuracies. More problematic is the lack of reliable indicators of where text elements like sentences or paragraphs start and end (Hori and Furui, 2003). We developed and evaluated a number of novel methods to preprocess the speech-to-text transcripts to account for these issues.

Event Blocks. Summarization depends on clearly identified boundaries between text entities.

Table 1: Audio transcription of NASCAR driver–pit crew communication: (left) Microsoft Azure Speech to Text, baseline model; (right) domain-trained model, highlighted text shows incorrect terms in the baseline model that were corrected in the domain-trained model.

try two more times when you get **this** two more times leader into one got the fifty two in front of five two **pride** next time be green checker for us next time **we'll** be outside pole order **player player player** crashing crashing crashing check your temps tweleve in the same areas there again about three maybe four times through the center **wanted to** the extra two twenty three four **pine trees** are alright maybe a little loose into **water** OK ten four year tyler's report and there's quite a bit of guys scrambling loose right now so we'll put a one wedge in this thing yeah eric **we can throw** a little bit of wedge in this thing the eight did that on the last stop and they liked it a lot so we'll go from there **aye regularly** straight shot in and the one is the lucky dog that's in front of us so should be fine yes sir thanks go ahead anbd cut your brake **fan** off their dog **hi period** open open security yellow switch down six is two in front of us here ninety nine i think it **lap** down then you got to seventeen seventeen at the other end of pit road i believe in the **sixties that** you're ten away get the eighteen behind us twenty one **'s** going to be at your one blue two in front of us	try two more times when you get **the** two more times leader into one got the fifty two in front of **you** five two **alright** next time be green checker for us next time **will** be outside pole order **clear clear clear** crashing crashing crashing check your temps tweleve in the same areas there again about three maybe four times through the center **one and two** the extra two twenty three four **pine trees** are alright maybe a little loose into **one** OK ten four year tyler's report and there's quite a bit of guys scrambling loose right now so we'll put a one wedge in this thing yeah eric **we're gonna put** a little bit of wedge in this thing the eight did that on the last stop and they liked it a lot so we'll go from there **alright rick it'll be** straight shot in and the one is the lucky dog that's in front of us so *it* should be fine yes sir thanks go ahead anbd cut your brake **ban** off their dog **alright pit road** open open security yellow switch down six is two in front of us here ninety nine i think it **lapped** down then you got to seventeen seventeen at the other end of pit road i believe in the **sixes if** you're ten away get the eighteen behind us twenty one **is** going to be at your one blue two in front of us

Normally, punctuation like periods or carriage returns is used to separate sentences or paragraphs. However, these are not reliably transcribed from speech due to poor recording quality, background noise, short unstructured speech snippets, repetitions, and lack of formal grammar (McKeown et al., 2005).

We investigated a number of different approaches to divide our transcripts into *event blocks:* blocks of text that discuss a single event. Each event block represents a document, and the collection of blocks from a single driver–pit crew audio transcript represents a document collection.

1. **Traditional.** The transcript was split by sentence. Each sentence represented an event block. This approach was ineffective since punctuation and sentence structure were not reliably assigned. Also, due to the nature of the communication, some suggested sentences consisted of only one or two words.
2. **BERT.** Punctuation was removed, and bidirectional encoder representations from transformers (BERT (Kenton et al., 2019)) was applied to identify sentences representing event blocks.

This approach had similar drawbacks to the traditional method. BERT's punctuation assignment was poor due to grammatical incoherence in the transcript.

3. **Word Count.** Artificial sentences with a pre-determined word count were created to define event blocks.
   This approach was inaccurate since the actual sentences consisted of a wide variety of words.
4. **Fragments.** Text fragments representing event blocks were constructed using silence of a threshold length in the audio communication.
   This approach proved successful. Combining text between periods of silence provided more context and was more informative. It also reduced the number of event blocks in an audio transcript from 500+ to ~200, making the documents longer. Longer sentences eliminated the issue of short, non-informative documents.

An *audio transcript* and its corresponding *fragments* are analogous to a *document collection* and its corresponding *documents*. We will use the terms transcript and fragment to emphasize our audio-based analysis, but once the transcript is converted to text-based fragments, traditional document and document collection algorithms act as starting points for our fragment and transcript analysis.

Different gap lengths of silence were tested to generate fragments. NASCAR races are characterized by a rapid pace, resulting in quick conversations. Even a few seconds of silence normally indicate the end of a conversation. Gaps from 1 second to 15 seconds were investigated. Based on manual inspection of the results, a gap of 10 seconds produced the best event block fragments. Shorter gaps split conversations, and longer gaps grouped conversations.

Once the text was divided into event blocks, standard text preprocessing was applied, including stop word removal, stemming, and weight-adjusted TF-IDF.

Weight-Adjusted TF-IDF. In addition to short sentences and repetitions, spoken language is often domain-specific, containing terms that may only be familiar to individuals participating in a conversation. Abbreviations and shortened versions of words are also common. Important terms may be used sparingly, making them difficult to manage using standard text analytic approaches. In our transcripts, certain important issues and significant race events were mentioned infrequently, making them hard to label as significant using standard TF-IDF. To address this, we constructed a small set of important terms in collaboration with our NASCAR collaborators and overweighted them in the initial term–document matrix (Table 2). Initial weights of $w = 1$, 1.5, 3, and 4 produced final summary accuracies of 60%, 64%, 50%, and 40%, respectively. Based on the decreasing accuracy as w increased, we did not pursue larger w, instead choosing to overweight important term frequencies by 1.5×.

Table 2: Important NASCAR-specific terms that were overweighted when constructing the term–document matrix.

****	adjustment	bad	balance	blown	bottom	bouncing	break	broke	bumper
bust	crack	damage	engine	entry	exit	fix	flat	free	front
fuel	grip	handling	hitissue	loose	low	mess	middle	need	nose
pit	rear	save	steering	tight	tire	trouble	wedge	wrong	

2.2 Fragment Similarity

Given fragment term vectors weighted using our TF-IDF approach, we next wanted to cluster similar fragments. Clustering allows us to identify key topics in a transcript and extract a subset of fragments from each topic to generate a final summary. Most clustering algorithms use a pairwise document similarity matrix to determine when a pair of documents are "close enough" to belong in a common cluster. Numerous methods exist to measure the similarity between documents. We investigated both term-based and concept-based approaches.

Cosine Similarity (Term-Based). TF-IDF-weighted term vectors $\overrightarrow{d_1}$ and $\overrightarrow{d_2}$ are normalized. Their similarity is the cosine of the angle between them, $\cos\theta = \overrightarrow{d_1} \cdot \overrightarrow{d_2}$ Intuitively, term vectors pointing in similar directions are similar, producing $\lim_{\theta \to 0} \cos(\theta) = 1$.

Latent Semantic Analysis (Concept-Based). Latent semantic analysis (LSA) occurs between TF-IDF weighting and cosine similarity. It is used to: (1) identify *latent* concepts within the document collection, converting the term–document matrix X into a concept–document matrix X_k; and (2) reduce the number of unique terms n to a set of concepts k, $k << n$ such that the k concepts capture the majority of the variance in the original TDM (Deerwester et al., 1990). In the reduced representation $X_k = U_k \Sigma_k V_k^T$ the k columns of U_k represent concepts whose row (term) values define the amount of each term contained in the concept. V_k^T's columns represent documents whose column values define how much of each concept is contained in a document.

Latent Dirichlet Allocation (Concept-Based). Latent Dirichlet allocation (LDA), similar to LSA, attempts to extract latent concepts from a document collection (Blei et al., 2003). To do this, each document d_i forms a probability distribution over k topics (*i.e.*, it defines how much of each topic d_i contains). The topics themselves are probability distributions over all terms in the document collection. An iterative generative algorithm based on Dirichlet distributions for both word-in-topic and topic-in-document probabilities is used to produce the final topic–word and document–topic probabilities. The algorithm uses $p(t_j|d_i)$, the proportion of terms in document d_i assigned to topic t_j, and $p(w_k|t_j)$, the proportion of documents belonging to t_j because they contain term w_k, to determine these probabilities.

2.3 Fragment Clustering

A manual examination of the three techniques showed all worked acceptably, but the concept methods produced more accurate similarity scores versus TF-IDF–cosine alone. We therefore compared LSA and LDA to determine which produced more accurate clusters. Figure 3 uses multidimensional scaling (MDS) to project concept-based fragment similarities from the Richmond Cup race for LSA with $k = 3$ clusters on the left and LDA with $k = 6$ clusters on the right. The axes represent the arbitrary positions selected by MDS, which are rotationally invariant. This highlights LDA's ability to associate fragments more tightly with similar text, reducing the ambiguity of overlapping topic clusters using LSA (Figure 3a). We chose weight-adjusted TF-IDF → LDA → cosine similarity to generate pairwise fragment similarity and a corresponding fragment similarity matrix.

The similarity matrix is fed to a k-means clustering algorithm. To determine an appropriate number of clusters k, we investigated silhouette scores (Rousseeuw, 1987) and Davies-Bouldin scores (Davies and Bouldin, 1987). The silhouette score measures cluster consistency using a fragment's within-cluster cohesion versus its between-cluster separation, producing a value on the range $-1 \ldots 1$ with larger values indicating a better k (Figure 4a). The Davies-Bouldin

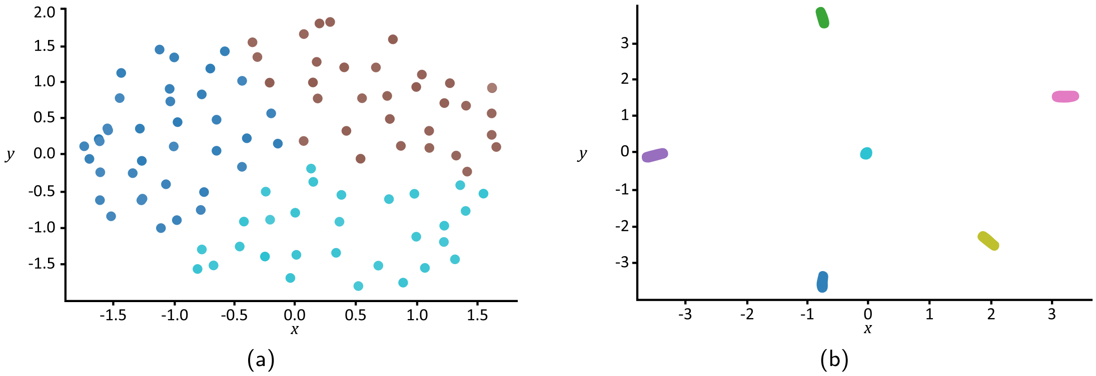


Figure 3: Projection of concept-based fragment similarities: (a) LSA with $k = 3$ clusters; (b) LDA with $k = 6$ clusters.

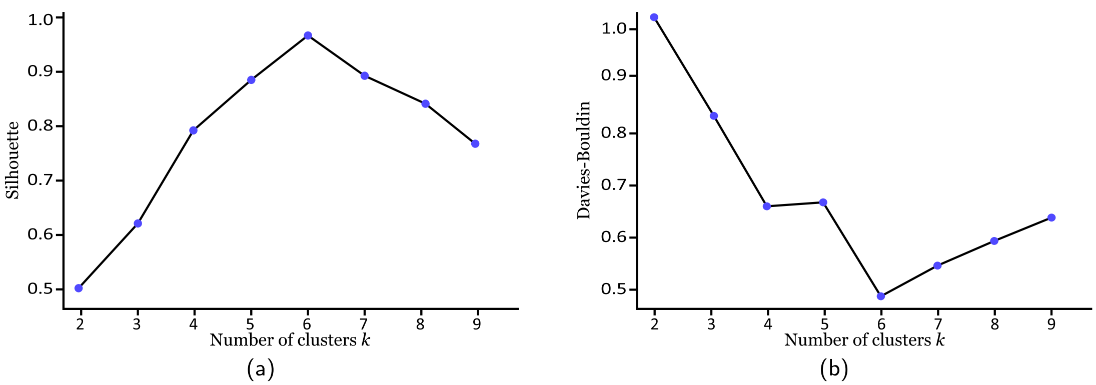


Figure 4: Silhouette and Davies-Bouldin scores for the LDA similarity matrix from the 2021 Richmond Cup: (a) silhouette scores with a maximum at $k = 6$ clusters; (b) Davies-Bouldin scores with a minimum at $k = 6$ clusters.

score identifies cluster separation using a ratio of within-cluster to between-cluster distances. The minimum Davies-Bouldin score is 0, with lower values indicating a better cluster ratio (Figure 4b). A comparison of silhouette and Davies-Bouldin scores showed they were identical in almost all cases. If they were different, we used the smaller of the two scores as our k value during clustering.

2.4 Summarization Results

Once clusters were constructed, we applied extractive summarization by combining 20% of the fragments closest to the centroid of each cluster to form a driver–pit crew communication summary for a given race.

Sampling. To test our results, we block-sampled the collection of all 1,383 driver–pit crew files based on track type and driver skill level to manage the time needed for audio transcription. Blocking is an experimental design technique that groups subjects based on their similarity to avoid unexplained variability (Bernstein, 1927; Chaudhuri et al., 2004). NASCAR tracks are divided into four types: (1) *superspeedway*, oval tracks 2+ miles long; (2) *intermediate speedway*, oval tracks 1–2 miles long; (3) *short track*, oval tracks 0.5–1 mile long; and (4) *road course*, non-oval tracks with left and right turns 2–4 miles long. Based on NASCAR driver rating, We categorized the 32 full-time drivers into 11 *good*, 10 *intermediate*, and 11 *poor* drivers. We note that the difference between a good and a poor driver is less than 1 second per lap, so all 32 drivers are considered highly skilled.

We first blocked on four randomly chosen drivers from each skill group, then on two randomly selected races for each race type, producing a total of $(4 \cdot 3) \cdot (2 \cdot 4) = 96$ audio files. Our proposed approach was applied to each audio transcript to produce a final race summary. To evaluate the summary, our three domain experts manually identified fragments that should be included in each summary. Any fragment identified by at least two experts was retained. We then computed precision p, recall r, and an F_1 score for each summary based on the true positive (t_{pos}, fragments we identified that were considered important), false positive (f_{pos}, fragments we included that were not considered important), and false negative (f_{neg}, fragments we discarded that were considered important) rates:

$$p = \frac{t_{\text{pos}}}{t_{\text{pos}} + f_{\text{pos}}}, \qquad r = \frac{t_{\text{pos}}}{t_{\text{pos}} + f_{\text{neg}}}, \qquad F_1 = 2\left(\frac{p \cdot r}{p + r}\right). \tag{1}$$

The average scores for our 96 summaries were $p = 0.59$, $r = 0.72$, $F_1 = 0.64$. Although the precision score was below 60%, it was still considered satisfactory, since our NASCAR collaborators favored longer summaries with potentially less relevant text over summaries that missed important information. Summarization reduced the transcript length by 80% while still retaining the majority of the important fragments. To determine if driver skill or track type affected summarization performance, we performed a Kruskal-Wallis analysis of variance (ANOVA) on both blocking factors. For driver skill, $F(2, 95) = 1.57$, $p = 0.21$. For track type, $F(3, 95) = 0.46$, $p = 0.71$. No statistically significant difference across driver skill or track type was found.

Table 3 shows an example of a summary during the Atlanta Motor Speedway race. The audio transcript for car 18 generated four clusters. Selecting 20% of the fragments closest to the cluster centers produced 14 fragments and 462 words in the summary.

Table 3: Audio summary for car 18 during the 2021 Atlanta Motor Speedway race.

Tires look good, normal Atlanta green, but no courts, no chunks.
the back just feels a little bit like help did not bad but a little bit This doesn't really feel like it travels a whole lot. and uh stays off the platform probably two good and then center office i feel like it's too easy to shoot here sample
i like how they able to finish two steals as many laps on tires yeah it's start to get away from the or working away from the twelve a little bit so it cost traffic you have Two more. still of twenty off
I don't know what you guys did there, but it's just too tight. Need rebalance. or could have back have our how the half around out of the right rear we'll put it back next stop good finish
bright hot right front had a a spot on it might be what the contributing your tight there

3 Event Detection

Our second goal is to extract important event discussions related to pit stops and vehicle balance from our transcripts. We compared dictionary matching and a trained NER for this task. Given the events, we needed to supplement them with context. Here, we compared dependency parsing and relationship extraction, both of which are active research areas in text analytics and NLP.

NER is a supervised technique requiring a training set and a test set. We randomly selected one race from our collection of transcripts, the Atlanta 2021 cup race. Once this was decided, we also selected the Charlotte Motor Speedway race since we wanted both transcripts to be from the same race type (intermediate). We used the Atlanta race as our training set and the Charlotte race as our test set. For dictionary matching, we constructed a dictionary built from the Atlanta race transcript and used it to identify events in the Charlotte race transcript.

3.1 Important Event Detection

In addition to summary overviews of race communication, we used event detection to identify "important" events that occurred during the race. We focused on pit stops and vehicle balance issues since these were highlighted by our NASCAR collaborators as the events they were most interested in. Since we are only looking for mentions of an event, we did not need to build conversation clusters. The raw text was evaluated directly using the two candidates. *Dictionary matching* uses the dictionary of important terms from text summarization. The dictionary was constructed based on input from our NASCAR collaborators. We extended this set with additional terms in the Atlanta race transcript related to pit stops and balance to better locate important events. The additional terms were added verbatim, without any stemming applied. Our NASCAR collaborators requested this, since they wanted *only* the additional terms and not conflations of those terms to match important events. Table 4 shows the additional pit stop and balance terms we included in consultation with our NASCAR collaborators.

Dictionary Matching. Raw text generated from our audio transcripts was examined to locate fragments containing important event terms. Figure 5 shows the number of raw text fragments and the corresponding number of important event fragments after dictionary matching for twelve drivers during the Atlanta race. Twenty to thirty percent of the raw text fragments contained an important term.

Table 4: Additional terms added to the dictionary used during dictionary matching for both pit stop and balance events, *e.g.,* "321" and "54321" are countdowns (3-2-1 and 5-4-3-2-1), "pi rd" is short for pit road, terms are added verbatim to match the speech-to-text conversion.

Pit Stop									
321	54321	come in	for there	fuel	full	now	pi rd	pit	tank
this time	tire	tires							

Balance									
adjust	adjustment	bad	balance	bounce	bouncing	control	drives	feel	feels
free	handle	handling	issue	issues	loose	lose	smooth	steering	tight
trouble	wrong								

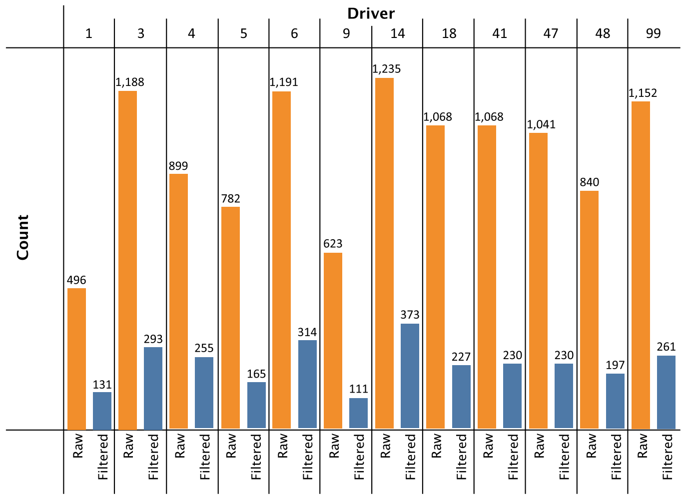


Figure 5: Count of raw text fragments versus text fragments containing important event terms for twelve drivers from the Atalanta cup race.

To evaluate our results, we manually enumerated the number of pit stop events P and balance events B contained in the Charlotte race transcript. Accuracy P_{acc} and B_{acc} was defined by a simple comparison of the number of events detected, P_d and B_d, versus the number of events that occurred:

$$P_{\text{acc}} = \frac{P_d}{P}, \qquad B_{\text{acc}} = \frac{B_d}{B}. \tag{2}$$

Dictionary matching produced accuracies $P_{\text{acc}} = 81\%$, $F_1(P) = 89\%$ and $B_{\text{acc}} = 76\%$, $F_1(B) = 86\%$. The results were considered good but perhaps not as high as expected. This is caused by misspellings or terms that imply a pit stop but are not in the dictionary. This second issue is important since it shows that the generalizability of an exact-match dictionary is unlikely to be perfect due to unexpected terms in new transcripts.

NER. *Named entity recognition* involves annotating entities in the transcript to train an NER model. Similar to dictionary matching, we used the race transcripts for twelve drivers from the Atlanta race to train an NER model. Each transcript was annotated in its entirety to avoid missing misspelled or previously unidentified terms. To improve this process, we used the Prodigy annotation tool (Explosion Inc., 2023a). The NER model was trained using BERT and spaCy (Kenton et al., 2019; Explosion Inc., 2023b). Because the training set contained both pit stop and balance entities, we could only calculate a combined set of accuracy metrics C_{acc} and F_1 when we applied our model to the Charlotte test transcript rather than individual pit stop and balance results (Eq. 2). Our model produced $C_{\text{acc}} = 90\%$ and $F_1 = 87\%$, an improvement over the dictionary method, particularly for accuracy. Given our small training set of 12 driver transcripts for one race, we consider these results promising and open to future improvement.

SMT Data. A final possibility is identifying the time when pit stop and balance issues occurring using the SportsMEDIA telemetry feeds. These times would be cross-referenced with the audio transcripts to identify corresponding text at the time of the event. For pit stops, this is relatively simple. For green flag (racing) laps, there is a stable pattern of brake, speed, steering angle, and throttle position patterns. The patterns show a clear variation when a driver prepares to enter pit lane.

Identifying balance, however, is less straightforward. We can use the gradient of steering angle α to lateral acceleration a to determine how many degrees d of steering are needed to achieve $1g$ of lateral acceleration. Linear regression through a cloud of (α, a) points over time allows us to monitor the change in d to identify when a car is trending in the direction of understeer or oversteer. The problem, however, is that a vehicle's balance is highly subjective based on an individual driver. To determine whether a driver believes there is a balance issue requires locating an occurrence of the driver reporting a balance problem, then examining the slope of the regression line at that time to determine a baseline for understeer or oversteer. Since the audio transcript is already examined to locate the balance event, using the SMT data is often redundant. The one potential advantage is future identification of understeer or oversteer that the driver does not report. Because of this, we did not pursue using SMT data to highlight important events, choosing instead to employ our NER model.

4 Sequence Visualization

Our audio summarization and event detection algorithms provide two important results: (1) an extractive text summary of driver–pit crew communication, and (2) a set of pit stop and balance events identified from the audio transcript with supporting context. Our final goal was to combine this information with SMT data (GPS position, speed, acceleration, throttle, brake, and steering position) and present it using an interactive visualization dashboard. The dashboard is designed as a post-race analysis tool for our NASCAR collaborators, allowing them to identify relationships and patterns and compare multiple drivers during a race.

The dashboard is built as a web application for ease of use since this supports access from any Internet-connected web browser. jQuery (OpenJS Foundation, 2023), Highcharts (Highcharts, Inc., 2023), and Javascript are used for UI widgets, visualizations, and dashboard control in the front-facing web application. Python and Flask (Bardhan, 2023) are used for backend control, communicating with Azure and a corresponding Data Lake storage repository, and Azure's speech-to-text system.

All visualization designs use a *data-feature mapping* and a *framework* to define how data is represented. Recall the two goals for our users are *summarization of driver–pit crew communication* and *identification of important events*. For each driver, the dashboard displays information using a modified Sankey diagram as its visual foundation. Vertical bars (nodes) in the Sankey plot represent important events. A driver's race is displayed as connected links separated by important event nodes. Summaries of driver–pit crew communication and overlaid as labels on the Sankey diagram, and through pop-up dialogs that can be displayed on demand by a viewer. Line segments between events are visualized to present driver performance and SMT data using the following data-feature mapping (Figures 6, 7).

- *time:* horizontal x-position
- *driver:* horizontal line
- *average lap time:* line thickness, thicker for slower
- *minimum corner speed:* red–blue color gradient, blue for faster, red for slower
- *average on-throttle time:* translucency, more transparent for longer on-throttle time
- *average understeer gradient:* color of vertical event axes, darker for larger understeer

The data-feature mapping was chosen based on the relative importance of each data attribute according to our NASCAR collaborators and our extensive knowledge of human visual

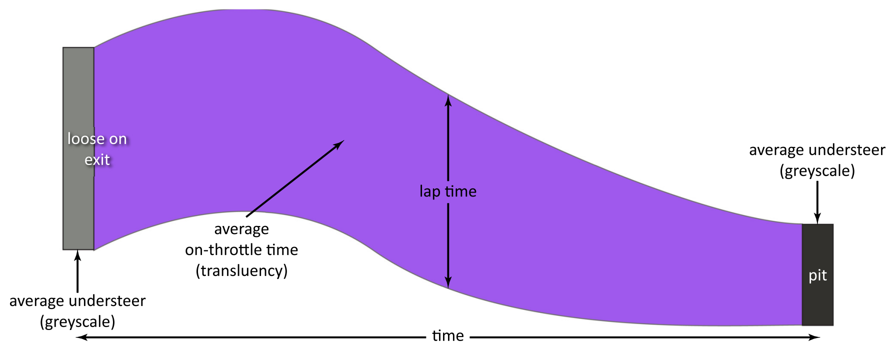


Figure 6: Visualization of a driver segment between loose on exit and pit stop events: important events → vertical bars; average understeer → bar greyscale; lap time → line thickness; average on-throttle time → translucency (more opaque for less on-throttle time); minimum corner speed → blue–red color gradient (blue for faster, red for slower).

perception and its ability to interpret basic visual features like color, texture, and motion, both in isolation and when combined in a single visualization (Healey and Enns, 1999; Huber and Healey, 2005; Healey and Enns, 2012; Healey and Sawant, 2012). Once perceptually optimal data-feature mappings are constructed, the resulting visualizations are modified based on end-user feedback to respect domain context and user preferences (*e.g.*, meteorological maps use color hue to represent temperature, so using a different visual feature to display temperature is avoided since it would confuse most viewers). The visualization was augmented by including contextual information from the summaries and important event detection. For example, vertical event nodes were labeled, and pop-up dialogs with summaries and telemetry data were displayed when users hovered over an event node (Figures 6, 7b).

Figure 7 compares a single-driver visualization versus a multiple-driver visualization. Common insights include the important events during the race for a driver, performance over the race, and perhaps most importantly, how performance changed if interventions to the car were performed during pit stops. The multiple-driver visualization adds the ability to compare cars, for example, to search for common or unique events, to see when the cars pit together versus when they pit separately, and how performance for the cars changes between interventions (*e.g.*, does one car get better and another get worse, suggesting a good versus a bad pit stop strategy?)

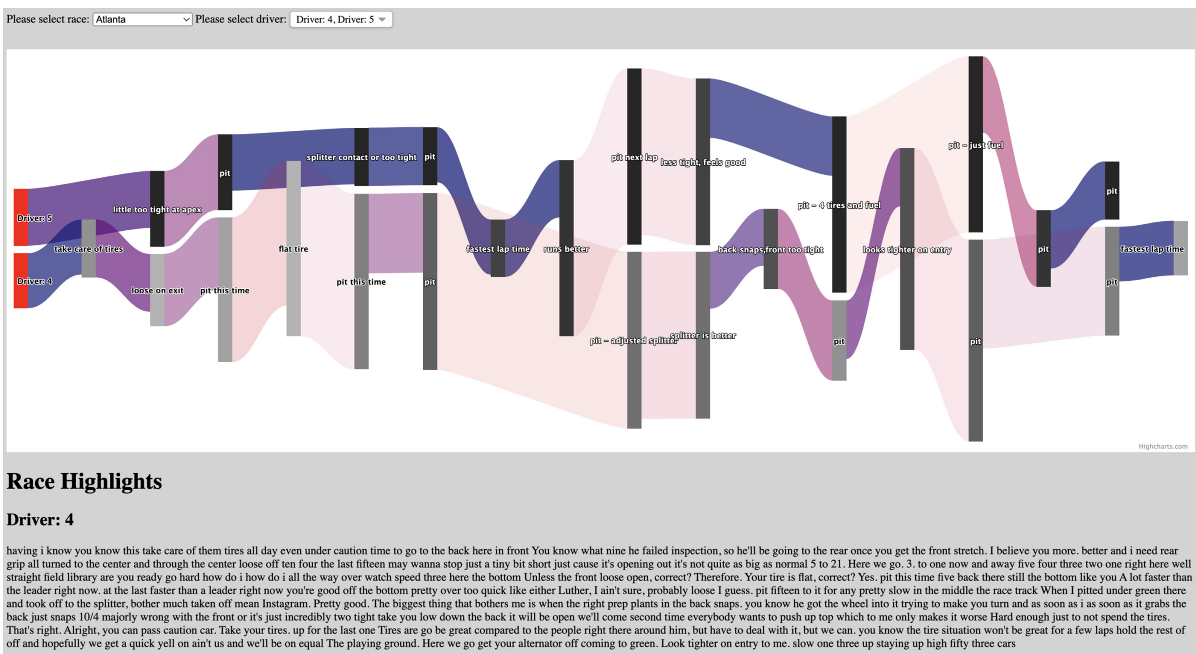
The current multi-driver overlap between different Sankey plots is being removed to avoid translucency–color ambiguity. Limited vertical space drove this decision, but our users noted they would only compare two or three drivers at any one time, eliminating this concern.

4.1 Evaluation

We demonstrated our dashboard to our NASCAR collaborators—a motorsport simulation manager, a crew chief, and a software engineering manager—then allowed them to use the tool to explore different races and drivers. The collaborators are engineers currently involved with NASCAR teams. All were male, with ages ranging from 25–44. One had a Ph.D. and two had Bachelor's degrees. One had 11–15 years of experience with NASCAR, and two had 16–20 years of experience.

Participants P_1, P_2, and P_3 were allowed to use the tool for as long as they liked. Next, we asked them to complete a questionnaire to collect their feedback (Tbl. 5). The questionnaire

(a)



(b)

Figure 7: Sankey flow visualizations: (a) a single driver 4 visualized with 14 important events during the race, highlighting important events and car performance indicating the effect of interventions between events (*e.g.*, lap time by height, corner speeds by color, and so on); (b) two drivers visualized in the dashboard, showing race and driver selection dropdowns and driver–pit crew transcript summaries, highlighting the difference in car performance, when the cars pitted together or individually, common versus unique issues with the cars, and so on.

was designed to assess the dashboard's design and its usefulness during daily tasks. Because of our small sample size, we consider our results anecdotal. More volunteers would be preferred, but not surprisingly, access to active NASCAR participants is difficult to secure.

Questions Q1 and Q2 asked how extracted events compared to what was seen during a race, and Q3 checked for overlap between audio and telemetry data. Results showed a strong correlation between extracted events and in-race observations, as well as a low overlap between

Table 5: Visualization dashboard evaluation questionnaire.

		Poor				**Good**
Q1	Based on experience, how do our race summary results compare to what you observe during the race?	1	2	3	4	5
Q2	Based on experience, how do our detected events compare to what you observe during the race?	1	2	3	4	5
Q3	Based on experience, what is the overlap between information contained in the audio and telemetry data?	1	2	3	4	5
		Low				**High**
Q4	In your daily work, how would you evaluate the importance of each of the extract pit stop information?	1	2	3	4	5
Q5	In your daily work, how would you evaluate the importance of each of the extracted balance information?	1	2	3	4	5
Q6	In your daily work, how important would it be to have access to the information presented by the dashboard?	1	2	3	4	5
		Low				**High**
Q7	How reliable do you consider audio data from races in general?	1	2	3	4	5
Q8	How reliable do you consider telemetry data from races in general?	1	2	3	4	5
Q9	What information does our tool provide you, in addition to the current tools that you already use?					
Q10	What additional information would you like to see in the tool?					
Q11	Would you use the tool in your role?	Yes Maybe No				
Q12	If you answered yes in the previous questions, please explain how you would use this tool in your job.					
Q13	What would you change in the tool?					

audio and telemetry data (Tbl. 6a). This confirms that audio data may provide additional useful information and that our system is analyzing it as our collaborators expected.

Questions Q4–Q6 ask about the importance of pit stop information, balance information, and access to the dashboard, respectively (Tbl 6b). The two lower scores for Q4 and Q5 represent differences based on participant: the crew chief is more interested in pit stops than either of the engineers (Q4), and one engineer and the crew chief are more interested in balance issues than the second engineer (Q5). Despite these responsibility-driven differences, all three collaborators agreed access to our dashboard would be important and useful (Q6). All the participants also agreed that, overall, both the audio and telemetry data are reliable (Q7–Q8, Tbl. 6c).

The remaining questions allowed for freeform responses. Q9–Q10 asked about information the dashboard provides that is currently unavailable and what additional information our collaborators would like to see included.

- Q9: New information available through the dashboard.
  - P_1: Mentions of balance and balance change through the race for the competitor vehicles

Table 6: Visualization dashboard evaluation questionnaire results.

Question	$\mathbf{P}_1$	$\mathbf{P}_2$	$\mathbf{P}_3$	μ (σ^2)
Q1: Do summary results and race observations correspond?	5	5	4	4.76 (0.58)
Q2: Do detected events and race observations correspond?	5	4	4	4.33 (0.58)
Q3: Overlap between telemetry and audio data?	1	1	2	1.33 (0.58)

1: poor . . . 5: good

(a)

Q4: Importance of extracted pit stop information?	2	4	2	2.67 (1.15)
Q5: Importance of extracted balance information?	4	4	2	3.33 (1.15)
Q6: Importance of dashboard?	5	5	3	4.33 (1.15)

1: low . . . 5: high

(b)

Q7: Reliability of audio data?	4	3	4	3.67 (0.58)
Q8: Reliability of telemetry data?	4	4	3	3.67 (0.58)

1: low . . . 5: high

(c)

  - P_2: Quick race summary, quick access to important race events, and ability to see track change affecting the entire field
  - P_3: Quick race review, use to pick out pit laps + green laps, use to correlate other data sources.

- Q10: Information you would like in the dashboard.
  - P_1: More balance type metrics correlated to driver feedback
  - P_2: Multiple race overlay
  - P_3: Vehicle proximity to pick out clean lanes

Multiple participants noted the ability to monitor competitors or overall track conditions. Monitoring competitors and track conditions makes sense since information about the team's driver is readily available through existing communication, telemetry data, and proprietary information collected by the team during the race. Information requested included the correlation between audio and telemetry data, multi-race visualizations, and information about competitor locations on-track.

Finally, questions Q11–Q13 asked whether our collaborators would use the tool and, if so, how they would integrate it into their current workflow. All three participants said they would use the tool. Integration focused on race preparation, monitoring track changes, data correlation, and post-race analysis to improve existing simulation tools. They also suggested custom plots and real-time updates to the dashboard during a race.

Our conclusions from the questionnaire were: (1) the dashboard offers important information currently not easily accessible to our collaborators; (2) audio transcripts offer an important opportunity to monitor competitors; and (3) the dashboard would be used as an addition to the current suite of analysis tools.

5 Conclusions and Future Work

We propose transcribing driver–pit crew audio captured during NASCAR events to: (1) summarize a driver's race based on events important to the team; (2) extract important events and supporting context from the transcripts; and (3) combine the summaries, important events, and race telemetry data into a visualization dashboard. The dashboard allows post-race analysis of audio and telemetry data for one or more drivers during a user-selected race.

Audio recordings are converted to text and summarized using an unsupervised text analytics approach. Important events are extracted from the transcript through named entity recognition and relationship extraction. The resulting data is merged with SportsMEDIA Telemetry data and visualized using a modified Sankey diagram integrated into a web-based interactive dashboard. Anecdotal feedback from three NASCAR collaborators confirmed the dashboard provides novel and useful information and would be used as part of the team's analytic workflow.

5.1 Limitations

While NER models are robust and can detect entities based on context, any entity not seen during training will not be identified. Moreover, if the vocabulary used during the race changes significantly, the NER models would need to be re-trained.

Our methodology provides insights not previously available to our collaborators, but it is limited to the contents of the audio recording. If an event is not discussed, it will not be available in either the race summary of the set or important events.

Finally, the system is currently designed as a post-race analysis tool. Important event identification could be implemented in real-time, but summaries, by their nature, require a race to conclude before they can be generated. Partial "summaries to date" may be possible by analyzing driver–pit crew communication throughout the race, but further investigation is needed to determine if this would produce useful information.

5.2 Future Work

Our investigation of NASCAR audio summarization highlighted numerous avenues for future work. Promising items are listed here.

1. Moving from post-analysis to real-time by creating an Azure pipeline that links real-time transcription with stream processing using Azure Databricks and PowerBI.
2. Implementing a way to suggest the type of pit stop a driver should perform.
3. Constructing predictive machine learning models for important events. Examples include caution flags, pit stop types, and green flag strategies.
4. Creating an ontology of pit stop and balance terms to facilitate future event detection.
5. Including custom plotting capabilities in the dashboard.
6. Including a multi-race visualization capability in the dashboard.
7. Correlating a driver's verbal indication of a car's balance to the SMT balance data to build driver-specific metrics for understeer and oversteer.
8. Performing sentiment analysis on driver–pit crew text (Healey et al., 2021) to augment the text and possibly weight the importance of different events based on emotional dimensions like activation (Russell et al., 1989).

Another area of future work involves topic clusters. LDA, and topic models in general, tend to generate topics that are heavily influenced by term frequency or that are difficult to interpret based on their term weights. One potential solution is guided LDA, where a small collection of seed word sets "guide" the algorithm towards topics (Jagadeesh et al., 2012). We would allow our NASCAR collaborators to choose seed term sets to form topics based on those terms. Another suggestion is BERTopic, where BERT document embeddings cluster documents and apply TF-IDF to generate topic representations (Kenton et al., 2019; Grootendorst, 2022). BERTopic produces semantically coherent and possibly more relevant topics for use in our system.

Supplementary Material

Python code for: (1) processing the transcribed driver–pit crew text, and (2) generating a web-based visualization of important events for a user-selected race and one or more drivers have been uploaded to the GitHub repository https://github.com/cghealey/JDS. Instructions on how to run the code are shown in the `README.md` file.

Acknowledgement

The authors would like to thank our NASCAR collaborators from Richard Childress Racing and General Motors, who dedicated valuable time and effort to explain NASCAR-specific concepts, evaluate our work, and make suggestions for improvements to the analytics and visualizations.

References

Aggarwal C (2018). *Machine Learning for Text*. Springer, New York, NY.

Amazon (2022). Amazon Transcribe. Accessed: 03-Feb-2023.

Bardhan S (2023). Deploying a Flask web app on Microsoft Azure. https://medium.datadriveninvestor.com/deploying-flask-web-app-on-microsoft-azure-89cea17e9114. Accessed: 03-Feb-2023.

Bernstein S (1927). Sur l'extension du théoréme limite du calcul des probabilités aux sommes de quantités dépendantes. *Mathematische Annalen*, 97: 1–59. https://doi.org/10.1007/BF01447859

Blei DM, Ng AY, Jordan MI (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3: 993–1022.

Bokhove C, Downey C (2018). Automated generation of 'good enough' transcripts as a first step to transcription of audio-recorded data. *Methodological Innovations*, 11(2). https://doi.org/10.1177/2059799118790743

Brinch S (2019). Charles-Joseph Minard's map of Napoleon's flawed Russian campaign: An ever-current classic. Accessed: 03-Feb-2023.

Chaudhuri S, Das G, Srivastava U (2004). Effective use of block-level sampling in statistics estimation. In: *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data (SIGMOD 2004)* (Weikum, G, König, C, Deßlock, S, eds.), 287–298. Paris, France. 10.1145/1007568.1007602

Chen W, Lao T, Xia J, Huang X, Zhu B, Hu W, et al. (2016). GameFlow: Narrative visualization of NBA basketball games. *IEEE Transactions on Multimedia*, 18(11): 2247–2256. https://doi.org/10.1109/TMM.2016.2614221

Coupland D (1995). *Microserfs*. HarperCollins, New York, NY.

Davies DL, Bouldin DW (1987). A cluster separation metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2): 224–227.

Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6): 391–407. https://doi.org/10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9

Explosion Inc (2023a). prodigy. https://prodi.gy. Accessed: 06-Feb-2023.

Explosion Inc (2023b). spaCy. https://spacy.io. Accessed: 06-Feb-2023.

Fu Y, Stasko J (2022). Supporting data-driven basketball journalism through interactive visualization. In: *ACM CHI Conference on Human Factors in Computing Systems (CHI 2022)* (Appert, C, Shamma, DA, eds.), volume 598, 1–17. New Orleans, LA.

Google (2022). Google Cloud speech to text. Accessed: 03-Feb-2023.

Grootendorst M (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure.

Gupta V, Lehal G (2010). A survey of text summarization extractive techniques. *Journal of Emerging Technologies in Web Intelligence*, 2(3): 60–76.

Healey CG, Dinakaran G, Padia K, Nie S, Benson JR, Caira D, et al. (2021). Visual analytics of text conversation sentiment and semantics. *Computer Graphics Forum*, 40(6): 484–499. https://doi.org/10.1111/cgf.14391

Healey CG, Enns JT (1999). Large datasets at a glance: Combining textures and colors in scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, 5(2): 145–167. https://doi.org/10.1109/2945.773807

Healey CG, Enns JT (2012). Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics*, 18(7): 1170–1188. https://doi.org/10.1109/TVCG.2011.127

Healey CG, Sawant AP (2012). On the limits of resolution and visual angle in visualization. *ACM Transactions on Applied Perception*, 9(4): 20:1–20:21. https://doi.org/10.1145/2355598.2355603

Heilmeier A, Graf M, Lienkamp M (2018). A race simulation for strategy decisions in circuit motorsports. In: *21st International Conference on Intelligent Transportation Systems (ITSC 2018)* (Zhang, W-B, Bayen, AM, Sánchez Median, JJ, Barth, MJ, eds.), 2986–2993. Maui, HI.

Hernandez K (2019). Live from Daytona 500: SMT digitizes 62-year-old race with broadcast, team data-tracking. Accessed: 03-Feb-2023.

Highcharts, Inc (2023). Highcharts. https://www.highcarts.com. Accessed: 03-Feb-2023.

Hori C, Furui S (2003). A new approach to automatic speech summarization. *IEEE Transactions on Multimedia*, 5(3): 368–378. https://doi.org/10.1109/TMM.2003.813274

Huber D, Healey CG (2005). Visualizing data with motion. In: *Proceedings IEEE Visualization Conference (Vis '05)* (Sliva, C, Gröller, E, Rushmeier, H, eds.), 527–534. Minneapolis, MN.

IBM (2022). IBM Watson speech to text. Accessed: 03-Feb-2023.

Jagadeesh J, Daumé H III, Udupa R (2012). Incorporating lexical priors into topic models. In: *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2012)* (Lapata, M, Màrquez, L, eds.), 204–213. Avignon, France.

Johnson BE (2011). The speed and accuracy of voice recognition software-assisted transcription versus the listen-and-type method: A research note. *Quantitative Research*, 11(1): 91–97.

Kenton JD, Chang MW, Lee KT (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2019)*, 4171–4186. Minneapolis, MN.

Kryściński W, Paulus R, Xiong C, Socher R (2018). Improving abstraction in text summarization. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language*

Processing (Ciang, D, Hockenmaier, M, Tsujii, J, eds.), 1808–1817. Brussels, Belgium.
McKeown K, Hirschberg J, Galley M, Maskey S (2005). From text to speech summarization. In: *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)* (Barner, K, Pesquet, J-C, eds.), v–997, volume 5. Philadelphia, PA.
Microsoft (2022). Microsoft speech to text. Accessed: 03-Feb-2023.
Moratanch N, Chitrakala S (2017). A survey on extractive text summarization. In: *International Conference on Computer, Communication and Signal Processing (ICCCSP 2017)* (Srinivasan, R, Shahina, A, Vasuki, P, Malathy, EM, AruKumar, V, Sofia Jenifer, J, Pavithra, LK, Geetha, K, eds.), 1–6. Tamalnadu, India.
OpenAI (2021). GPT-3 powers the next generation of apps. Accessed: 03-Feb-2023.
OpenAI (2022). ChatGPT: Optimizing language models for dialogue. Accessed: 03-Feb-2023.
OpenJS Foundation (2023). jQuery. https://jquery.com. Accessed: 03-Feb-2023.
Padia K, Bandara L, Healey CG (2019). A system for generating storyline visualizations using hierarchical task network planning. *Computers & Graphics*, 78: 64–75. https://doi.org/10.1016/j.cag.2018.11.004
Payne M (2021). State of the art GPT-3 summarizer for any size document or format. Accessed: 03-Feb-2023.
Perin C, Vuillemot R, Fekete JD (2013). SoccerStories: A kick-off for visual soccer analysis. *Computer Graphics*, 19(12): 2506–2515.
Perin C, Vuillemot R, Stolper CD, Stasko JT, Wood J, Carpendale ST (2018). State of the art of sports data visualization. *Computer Graphics Forum*, 37: 663–686. https://doi.org/10.1111/cgf.13447
Pileggi H, Stolper CD, Boyle JM, Stasko JT (2012). Snapshot: Visualization to propel ice hockey analysis. *IEEE Transactions on Visualization and Computer Graphics*, 18(12): 2819–2828. https://doi.org/10.1109/TVCG.2012.263
Rekabdar B, Mousas C, Gupta B (2019). Generative adversarial network with policy gradient for text summarization. In: *13th IEEE International Conference on Semantic Computing (ICSC 2019)* (Bansal, S, Bloodgood, M, Persia, F, eds.), 204–207. Brisbane, Australia.
Rousseeuw PJ (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Computational & Applied Mathematics*, 20: 53–65. https://doi.org/10.1016/0377-0427(87)90125-7
Russell JA, Lewick M, Niit T (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology*, 57(5): 848–856. https://doi.org/10.1037/0022-3514.57.5.848
Sankey HR (1902). The thermal efficiency of steam engines. Accessed: 06-Feb-2023.
Spaȓch Jones K (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1): 1–11. https://doi.org/10.1108/eb026525
Stoll M, Krüger R, Ertl T, Bruhn A (2013). Racecar tracking and its visualization using sparse data. In: *1st IEEE Workshop on Sports Data Visualization* (Basole, R, Clarkson, E, Cox, A, Healey, CG, Stasko, J, Stolper, C, eds.), 1–6. Atlanta, GA.
Tharoor VV, Dhanya NM (2022). Performance of Indian cricket team in test cricket: A comprehensive data science analysis. In: *International Conference on Electronic Systems and Intelligent Computing (ICESIC 2022)* (Kavitha, M, Rajalakshmi, R, eds.), 128–133. Chennai, India.
Wienrich C, Reitelbach C, Carolus A (2021). The trustworthiness of voice assistants in the context of healthcare investigating the effect of perceived expertise on the trustworthiness

of voice assistants, providers, data receivers, and automatic speech recognition. *Frontiers of Computer Science*, 3: 1–12. 685250

Xu H, Cao Y, Ruipeng J, Liu Y, Tan J (2018). Sequence generative adversarial networks for long text summarization. In: *30th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2018)* (Alamaniotis, M, ed.), 242–248. Volos, Greece.

ZoomInfo (2023). ZoomInfo. Accessed: 16-Jun-2023.

Zuang H, Zhang W (2019). Generating semantically similar and human-readable summaries with generative adversarial networks. *IEEE Access*, 7: 169426–16943. https://doi.org/10.1109/ACCESS.2019.2955087