Data Collection for ASSISTments 2.0's DRIVER-SEAT and QUICK-Comments

An Interactive Qualifying Project Submitted to the Faculty of Worcester Polytechnic Institute
For the Fulfillment of the Requirements for the Degree of Bachelor of Science

Submitted By:
Skyler Kim

On:
11 March 2019

Submitted To:
Professor Neil T. Heffernan, Advisor
Cristina Heffernan, Advisor

Abstract

With assignments being assigned daily, it is difficult for teachers to give each student meaningful feedback. The current ASSISTments 2.0 has plans for enhancing student feedback through their DRIVER-SEAT and QUICK-Comments system. New features were added to ASSISTments 2.0 to collect data on teacher comments for these planned features. The comment data to be collected consists of open response feedback and comments on a student's actions while completing an assignment. This data will be used to train the models needed to suggest meaningful and personalized feedback teachers could give their students.

<u>Acknowledgements</u>

Table of Contents

Table of Figures

Introduction

ASSISTments is a free online classroom management system created by Neil and Cristina Heffernan and hosted by Worcester Polytechnic Institute. ASSISTments aims to be a platform to help students receive feedback on the work they have done. At the same time, it gives feedback to the teachers by telling them the commonly incorrect problems, and makes suggestions as to what the teacher should spend more time on.

Two of the currently planned features for ASSISTments that recently received funding are DRIVER-SEAT and QUICK-Comments. Both build on the idea of feedback for the student by suggesting comments that teachers can give to students. DRIVER-SEAT focuses on feedback on the steps a user took to complete an assignment, while QUICK-Comments focuses on the feedback a teacher could give a student based on the content of their open response question submissions. Both of these features have plans to use deep learning models to identify different types of actions or responses and produce useful and meaningful feedback for the students. It is ideal if the models can be trained using student and teacher data already in the ASSISTments system.

The goal of this project is to create methods for data collection amongst current users of ASSISTments 2.0. This data will be the training set for creating the deep learning models needed to produce meaningful and personalized feedback on student actions. This goal will be achieved by enhancing how teachers give feedback to their students and expanding the content teachers can comment on.

Background

2.1 ASSISTments

ASSISTments is an online tool to assist teachers and students with assigned homework. Students use ASSISTments to complete homework assignments or in class assignments. If a student gets stuck along the way, the ASSISTments platform helps them with the topics they are confused about. Teachers use ASSISTments to quickly grade large amounts of homework and look at reports for both each individual student and for their class as a whole. These reports include average grades, common incorrect answers, and information on which students needed assistance (Burnett, Easy).

ASSISTments aims to give feedback to both teachers and students. When a student is solving an assignment, they get instant feedback for all the numeric text and multiple choice questions. If a student gets stuck on a problem, they can ask the ASSISTments tutor, or the web application used for completing assignments, for hints or help with a problem. ASSISTments is also feedback for teachers because it will organize the responses and grades of all students. This makes it easy for teachers to identify the problems that their students struggled on. Teachers can then take this information to reconstruct their classes to focus more on the subjects that students had trouble with. Without a computerized system, teachers would have to spend a large amount of time grading and keeping notes on the problems students got wrong (Burnett, ASSISTments).

The ASSISTments team is currently developing a newer version of ASSISTments called ASSISTments 2.0. This new platform integrates more with other classroom learning management systems to assist teachers and students with availability. The site currently interfaces with Google's Google Classroom, and has plans to integrate with Edmodo and Moodle in the future. To the students, ASSISTments 2.0 lives almost entirely inside the external learning management system (LMS) of choice. A student can access assignments and grades through different assignments that are assigned through their chosen LMS. These assignments are links that redirect to the ASSISTments Tutor, where students complete and answer questions just as they did in the original ASSISTments. There is never a need for a student to access the ASSISTments 2.0 website manually.

The teacher's experience is also changed by limiting the amount of time spent on the ASSISTments 2.0 website. Teachers assign problems to their LMS through the front page ASSISTments 2.0 site, but can then manage due dates and sending grades through their LMS of

choice. Clicking on the same assignment link that students use will redirect teachers to the report features of ASSISTments, which includes automatic grades for students and options for open response grading.

ASSISTments 2.0 is still being heavily developed and has plans for two large features: DRIVER-SEAT and QUICK-Comments. Even though ASSISTments 2.0 is still in development, teachers can still access and use the major parts of the website that make up the ASSISTments system.

2.2 DRIVER-SEAT

Majority of math teachers assign daily homework that require the student to explain their reasoning. Even though the assignment requires the student write out an explanation, it is difficult for a teacher to find the time to go through every answer on every assignment for every student and give meaningful feedback (Heffernan, QUICK).

DRIVER-SEAT stands for Dialogue Reinforcement Infrastructure for Volitional Exploratory Research – Soliciting Effective Actions from Teachers. The aim of DRIVER-SEAT is to help generate and suggest the meaningful and personalized feedback a teacher wants to give their students, but cannot find the time to write out (Heffernan, QUICK). Current method for enhancing student performance only focus on enhancing the top student's abilities. The planned system for DRIVER-SEAT and its comment generation system will assist students of all knowledge bases (Duffy). Comment generation for DRIVER-SEAT will be done using technologies similar to Google's Smart Reply feature used in Google's Gmail (Heffernan, DRIVER).

Google's Smart Reply technology works on a deep learning model called Long Short-Term Memory (LSTM) (Kannan). LSTM is a form of Recursive Neural Network (RNN), meaning that different components of the input are passed recursively through the available models. The output of one model is passed in as one of the inputs to the next instance of the model. In the case of natural language processing, the output of one model is passed in along with another token in the sentence or phrase to be analyzed (Bottou).

The problem with RNNs is that RNN models are only good at finding the best possible token to make a sentence complete in a given instance. RNN models are not able to look at the context of the sentence or word in a larger paragraph. LSTM fixes this problem by introducing a

memory cell that has the ability to retain and forget certain pieces of information about the data being read. This memory component is carried across each recursive call in a RNN and helps the model maintain dependencies throughout the analyzation and generation new data (Greff).

Google's Smart Reply LSTM model aims to find a set of the best possible responses for a given original message. The model does this by scoring appropriate responses and finding the highest scoring response. The method for scoring responses is done by finding the conditional probability of a particular response being sent in response to the original message (Kannan).

DRIVER-SEAT will use a similar model to select an appropriate response to an action set. These responses will be chosen from a collection of trusted messages that teachers can give their students. The list of comments that will be available to teachers to quickly send their students will be built by selected math and science teachers. At the same time, the data generated by these teachers and their students will be used to test various deep learning models to determine how and when to suggest certain feedback (Heffernan, Whitehill). At the moment, there is no way to collect data on the feedback a teacher gives a student based on the actions they took to complete the assignment. In ASSISTments 2.0, there is no way for a teacher to see the actions a student took to complete an assignment.

2.3 QUICK-Comments

ASSISTments 2.0 can currently handle automatic scoring for multiple choice questions and short numerical answers, but open response questions must be graded by the teacher. The feedback a teacher gives includes a numerical grade along with a short comment on the student's response. With the large amount of open response questions a teacher must grade, it is difficult to write detailed and meaningful feedback to help each student improve (Heffernan, QUICK).

QUICK-Comments is a part of the DRIVER-SEAT infrastructure that focuses on suggesting comments for teachers to give on student open response questions. The system is planned to show multiple suggestions for potential comments to give students. Teachers will still be given the option to write their own comment if they still wish to write their own comment (Heffernan, QUICK).

QUICK-Comments will also use technologies in Google's Smart Reply to generate meaningful responses to student's open response submissions (Heffernan, QUICK). QUICK-Comments will involve more natural language processing than the action comments used in

DRIVER-SEAT since the system has to parse and understand what the student is saying. The system will have to understand if the student is saying the correct answer in addition to any incorrect answer that might be submitted. At the same time, the system has to determine an appropriate comment to leave the student in order to help push the student towards the correct answer or method. LSTM models will play a stronger role here since LSTM models can be used for analyzing meaning and appropriateness of natural language in addition to generating and selecting meaningful comments (Kannan).

The model for generating comments for QUICK-Comments will also be developed by the data generated by the teachers selected to develop DRIVER-SEAT. These teachers will be tasked with assessing the comments of other teachers in addition to providing their own (Heffernan, QUICK). Currently, Professor Heffernan is using his artificial intelligence classes as a means to test and create different models that would use the selected teacher's data in creating the QUICK-Comment system (Heffernan, Grants).

In ASSISTments 2.0, there are ways for teachers to give students feedback on open response problems through Essay Grading. However, teachers do not grade students anonymously. The QUICK-Comments system would ideally work solely on the content of a student's response, and training the models needed for QUICK-Comments on current data might introduce the bias that a teacher has when they know the student they are grading. In order to eliminate bias in comments and grades, a new system is needed to collect information to create the QUICK-Comments system.

Implementation

3.1 Student Actions Report

Creating the model needed to generate comments for DRIVER-SEAT requires two types of data: the types of actions they notice a student doing and the comment the teacher give for those actions. Groups of actions would go towards identifying the diagnosis a teacher could give a student, while the comments would go towards the possible responses to those groups of actions.

Showing the actions of a student is now available through the Student Actions Report page. This page is accessible by clicking a student's name on the Item Report page. The page shows a table of every action a particular user took while attempting the assignment. The columns of the table are the timestamp of the action, the type of action, any response that went along with the action, and a selectable actions column. The top of the page has options to show more information about each problem, or swap to another student's Student Actions Report.

The actions that are shown are categorized into three types of actions: response actions, assignment metadata, and assistance actions. Response actions are actions involving a prompt for a student's response, whether it be an open response problem or a numeric answer field. The response actions include the actions of Student Response, Student Submission, Work Submission, and Comment Submission. Both open response problem submissions and numeric answer field submission are saved as Student Response actions, and are the most common type of action to report. The other actions of this category do not appear as often as Student Response, but do have the ability to be shown in the Student Actions Report.

If the Student Response action corresponds to a numeric answer field, the row in the Student Action Report will be highlighted a certain color depending on the correctness of the answer. Correct answers are colored in green and incorrect answers are colored in red. If the Student Response action references an open response submission, then the row will be colored dark blue. The difference in coloring is meant to help teachers identify the open response questions that need their grading. This color scheme matches the exact color scheme used in the Student Actions Report in the original ASSISTments. Having consistency between the two versions gives the teachers a sense of familiarity as they transition from the original ASSISTments to ASSISTments 2.0.

+ 0 mins 1 secs	Continued to Next Problem		
+ 0 mins 0 secs	Started a Problem	What is the scale factor from polygon ABCD to polygon PQRS?	
+ 0 mins 6 secs	Answered Incorrectly	1	
+ 0 mins 2 secs	Answered Incorrectly	2	
+ 0 mins 2 secs	Asked for an Answer		
+ 0 mins 4 secs	Answered Correctly	3/2	
+ 0 mins 0 secs	Finished a Problem		
+ 0 mins 0 secs	Completed Assignment		

+ 0 mins 0 secs	Finished a Problem		
+ 0 mins 1 secs	Continued to Next Problem		
+ 0 mins 0 secs	Started a Problem	Use tracing paper to compare angle NOP with its corresponding angles in Copy 1 and Copy 2. What do you notice?	
+ 0 mins 3 secs	Submitted an Essay Answer	I noticed that the angles in Copy 2 are larger than those in Copy 1. This is probably because Copy 2 is stretched more than Copy 1.	
+ 0 mins 0 secs	Finished a Problem		
+ 0 mins 0 secs	Continued to Next Problem		

Figure 1: Action Coloring on Student Action Report

Assignment metadata are actions that involve the state of a problem a student is working on. These mainly focus on when a student starts and stops a particular problem, and can be paired with the timestamp to get an idea of how long a student spent on a particular problem. The assignment metadata actions are Problem Started, Problem Resumed, Problem Finished, Continued to Next Problem, and Assignment Finished. Problem Started and Problem Finished are used to describe the start and end of each problem, and Assignment Finished is used to signal when the student has completed the assignment.

The Problem Resumed action is used to describe the situation when a student reopens the ASSISTments tutor to continue solving a problem after they have already started the problem, but navigated away from the tutor. Showing this may be useful to teachers to see how long breaks away from an assignment can affect a student's grades. The thought behind logging this action was the case where a student answers questions incorrectly, resumes the assignment after

a long break, then finishes the assignment with high accuracy. In this situation, it is implied that the student realized they are not familiar with the content of the assignment, left to study the section more, then came back to complete the assignment. This is a prime action a teacher can comment on and ask students about their strategy for completing assignments.

Assistance actions are actions where students are either prompted with the correct answer or select options to assist them on a problem. The assistance actions are Hint Request, Scaffolding Request, Answer Request, Explanation Request, and I'm Stuck Button Click. The Hint Request, Answer Request, Explanation Request, and I'm Stuck Button Click actions have their row in the table colored in orange. This helps teachers quickly identify places where the student has problems. This color key is the same as the Student Actions Report in the original ASSISTments, and aims to help teachers in the same way the colors for Student Response action help teachers.

A Scaffolding Request is when a student asks for assistance on a problem and is prompted with more questions. These sets of questions are known as scaffolding questions, and are smaller questions that guide the student towards a correct answer for the actual problem. Actions involving scaffolding problems are listed as scaffolding problems and answers in order to prevent confusion between normal problems and scaffolding problems.

By clicking on the cells in the Action column, a teacher can highlight the actions that spark their interest. When the teacher stops highlighting, a new box appears to the right of the table to allow teachers to write and submit comments. The box will appear next to the last action that the teacher has highlighted in order to draw the attention of the teacher and prevent them from having to search the Student Actions Report for the new prompt. These comments are sent to the server along with the highlighted actions for logging.
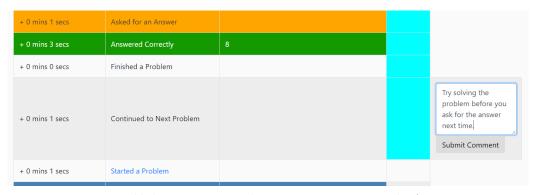

Figure 2: Highlighting Comments and Writing a Comment on a Student's Actions.

In order to simplify the amount of features added, the Student Actions Page currently has its action highlighting and comment posting functionality removed. It was decided that the ASSISTments team wishes to look at the results of QUICK-Comments before looking at the actions component of DRIVER-SEAT. As a result, the Student Actions Report has the selectable actions column replaced with an Essay Grading column. This column allows teachers to enter a grade and a comment for any open response problem. This functions exactly like Essay Grading, only it allows teachers to grade open response problems by student instead of by question. If there are no open response problems in the problem set, the Essay Grading column does not appear.
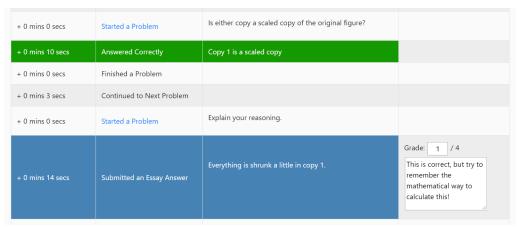
+ 0 mins 0 secs	Started a Problem	Is either copy a scaled copy of the original figure?	
+ 0 mins 10 secs	Answered Correctly	Copy 1 is a scaled copy	
+ 0 mins 0 secs	Finished a Problem		
+ 0 mins 3 secs	Continued to Next Problem		
+ 0 mins 0 secs	Started a Problem	Explain your reasoning.	
+ 0 mins 14 secs	Submitted an Essay Answer	Everything is shrunk a little in copy 1.	Grade: 1 / 4 This is correct, but try to remember the mathematical way to calculate this!

Figure 3: Essay Grading Integrated into Student Actions Report.

3.2 QUICK-Comments Essay Grading

The suggestions generated in QUICK-Comments is similar to DRIVER-SEAT, but only focuses on essay response questions. Training models for QUICK-Comments requires a sample of grades and comments for a given problem. It was decided to make use of the already existing Essay Grading part of ASSISTments as an example of the data to collect, but to incorporate new methods of open response grading to ensure feedback only looks at the content of the response.

The resulting application is a webpage displaying a single open response answer. This page is accessible to teachers by clicking the already existing Essay Grading link found on the Item Report page. The teacher can only see one student's response at a time. Limiting teachers to one student at a time helps the teacher concentrate on the current student. The student who wrote the open response solution is anonymized at first, but becomes visible once the teacher has

submitted a comment. The teacher can update their comment after they see the student name if a correction is needed.
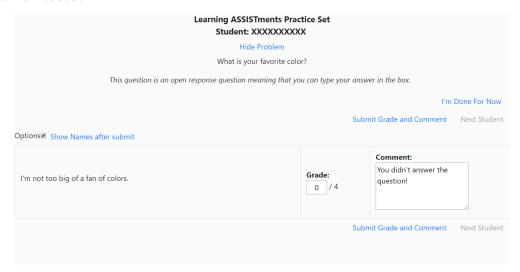

Figure 4: Driver Seat Essay Grading

The application is designed to have the teacher grade each response one by one. The order of the responses is randomized every time to help discourage teachers from recognizing patterns in the responses and attribute them to certain students. After grading all their students, the teacher will be sent to the already existing Essay Grading page. If a teacher were to return to the Item Report and click on the Essay Grading link again, they will be redirect to the original Essay Grading they are familiar with. Every grade and comment handed out by the teacher inside the new application will be recorded for the student they correspond to.
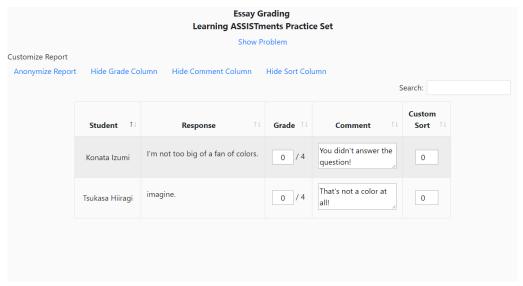

Figure 5: Essay Grading Updates After Completing Driver Seat Essay Grading.

It is important to keep students anonymized during the feedback process because it forces teachers to grade only the content of the open response answer. Having a teacher's opinion of a student mixed in with their feedback on the content of the response will pollute the training data. The resulting model needs to characterize the response a student gives, not characterize the student themselves. Eliminating the name and personality of their students will help the teachers only focus on the content of the open response answer and will hopefully reduce noise in the model to be constructed.

Each comment made by a teacher is saved in the ASSISTments 2.0 database under the DriverSeatEssayGradingLog table. The data saved along with the comment includes information on the student's response (also known as the ProblemLog), the external ID of the teacher, a flag for whether a comment was made anonymously, and a flag if the graded student belongs to the grading teacher. There are no reference keys for this table because it is meant to be temporary. The table will only live on the ASSISTments 2.0 database during the data gathering and model training. The current plan is to remove the table after a stable model is created.

The option to see student names is still available after grading in case the teacher must provide special accommodations to the student. It is possible the teacher sees the response the student gave in a new light after knowing which student made the response. For example, a teacher might want to change their harsh comment on a wrong answer if they find out the student behind the response has special needs. Both the updated comment and score and the original comment and score are saved on the DriverSeatEssayGradingLog table. The original comment is flagged as a comment and grade that were made while the teacher had no knowledge of the student.

The application has additional responses shuffled in to the answers the teacher needs to grade. These responses belong to students that are not in the teacher's class, but have a record of a response for the problem being graded. These students are not revealed to the teachers after submitting a response for privacy reasons. The current amount of extra responses to a teacher's workload is 10% of what the teacher currently has to grade. A response of a student will only appear for another teacher if that student already received feedback on it. Only previously graded assignments are visible to other teachers to force responses with feedback on them to receive more feedback.

Making teachers grade other teacher's student's responses allows for a comparison to be made on how feedback differs per user. The comparison is needed to remove the bias in a teacher towards specific types of responses. Having more than one teacher assess a response allows the feedback to be compared against itself. The similarities in the feedback for the same problem are an indicator for information that is objective to the content of the response. Differences in feedback may indicate the teacher putting their own bias and opinion on an answer. The differences can also be used to look at the writing style and levels of detail in the feedback.

A teacher's grade and comment on another teacher's student is not recorded in the student's gradebook, but is still saved in the DriverSeatEssayGradingLog table. This means that another teacher does not have influence on the grades of students that do not belong to them, but the feedback is still saved in a way that lets analysts look at the differences in feedback.

3.3 DRIVER-SEAT Teachers

The new Essay Grading application mentioned previously is currently only available to selected teachers. It was decided to only have the application open to these select teachers in order to keep the data slightly more controlled. Most teachers don't give very useful feedback to students because of the time it takes to construct each comment. This is the reason why DRIVER-SEAT is being created. Having a lot of meaningless and shallow comments from a lot of teachers will prevent useful messages from being produced by the resulting model. By selecting a small set of teachers and instructing them to write thorough comments, the model will not contain as much noise for bad comments. These teachers are also being rewarded for writing their comments, so there is extra incentive for writing well thought out feedback.

ASSISTments 2.0 knows which accounts are labeled DRIVER-SEAT teachers through an Attribute that is linked to all DRIVER-SEAT teacher external reference ID. An Attribute is a custom tag in the ASSISTments 2.0 database that serves as a metadata recorder. A new Attribute called driverSeatTeacher was added, and a procedure was written to tag selected accounts with this attribute.

Future Outlook

Currently, after a teacher gives feedback on an open response problem, the student's grade is updated within ASSISTments 2.0. This means that if the student were to ever go back and look at the assignment in Google Classroom again, they would see a grade for an open response problem along with a comment. With the Essay Grading functionality being added to both new pages, it was realized that there is no way for a student to know when their open response problems have been graded. This is especially problematic for problem sets where all the questions are open response problems, which is the target audience of the QUICK-Comment feature.

This problem can be solved with a notification system to alert users when they have new feedback. The potential systems looked at were an internal notification system and making use of notifications in Google Classroom. The internal notification system would alert the student if they had any grade or feedback updates since the last time they had visited the site. There would then be the option to take the student to view these updates. There is also the google classroom option, where students could receive messages or notifications through Google Classroom about their grade changing. These changes would then be handled by Google Classroom, and the student would be sent emails and notifications through the Google account infrastructure.

Both the internal notification system and making use of Google Classroom have major drawbacks that prevented them from being fully designed and implemented. The internal notification system would be ineffective because the student rarely needs to access the ASSISTments 2.0 website. Students only need to access the ASSISTments 2.0 Tutor to complete assignments. It would be disorientating if a student was prompted to look at their last homework assignment every time they tried to complete a new one.

Google Classroom notifications also has major drawbacks because of the clutter it creates. The two ways to create a personalized message from the teacher to the student through the Google Classroom API is by creating a "class material" and sending it to an individual student, or by creating a "class announcement" and sending it to an individual. Both cause a message to appear in the student's classroom stream and properly notify the student of a grade. However, every announcement and material added by the teacher would also show up in the teacher's stream. The teacher would have their stream quickly cluttered by every feedback message sent to every student in their class. This is not ideal for the teacher as it makes it

impossible for the teacher to find anything in their stream, and can hinder their workflow. The ideal messaging system through Google Classroom would be to add a private comment to the assignment itself, which is possible from the teacher account in Google Classroom. The Google Classroom API does not allow for these private comments on assignments to be made yet, so using Google Classroom to send notifications is currently on the backlog.

As QUICK-Comments and DRIVER-SEAT start to develop and models start to be trained, it will become increasingly more important for a notification system. Both new features focus on generating feedback for students, and a lack of a notification system will make it impossible for them to receive the feedback from the teacher.

References

Bottou, Léon. "From machine learning to machine reasoning." Machine learning 94.2 (2014), 13
    February 2013, doi: 10.1007/s10994-013-5335-x.

Burnett, Andrew. "ASSISTments Overview." ASSISTments, WPI, ASSISTments, 2016,
    www.aboutus.assistments.org/overview.php.

Burnett, Andrew. "Easy Ways To Get Started." ASSISTments, WPI, ASSISTments, 2016,
    www.aboutus.assistments.org/easy-ways-to-get-started.php.

Duffy, Alison. "NSF Funds Effort to Enhance a Mathematics Online Tutoring System Developed
    at WPI to Make It Easier for Teachers to Provide Feedback to Students." WPI, WPI,
    ASSISTments, 3 Oct. 2018, www.wpi.edu/news/nsf-funds-effort-enhance-mathematics-
    online-tutoring-system-developed-wpi-make-it-easier.

Greff, Klaus, et al. "LSTM: A Search Space Odyssey." IEEE Transactions on Neural Networks
    and Learning Systems, vol. 28, no. 10, 4 Oct. 2017, pp. 2222–2232.,
    doi:10.1109/tnnls.2016.2582924.

Heffernan, Neil, and Jacob Whitehill. "Putting Teachers in the Driver's Seat: Using Machine
    Learning to Personalize Interactions with Students (DRIVER-SEAT)." Research.gov,
    NSF Online Grant Management, 18 Aug. 2018, grant:1822830.

Heffernan, Neil. "DRIVER-SEAT." Neil Heffernan, Neil Heffernan, 2019,
    www.neilheffernan.net/projects/funded-projects/drivers-seat.

Heffernan, Neil. "Grants." Neil Heffernan, Neil Heffernan, 2019,
    www.neilheffernan.net/bio/grants.

Heffernan, Neil. "QUICK-Comments." Neil Heffernan, Neil Heffernan, 2019,
    www.neilheffernan.net/projects/funded-projects/quickcomments.

Kannan, Anjuli, et al. "Smart Reply: Automated Response Suggestion for Email." Proceedings
    of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data
    Mining - KDD '16, 15 June 2016, doi:10.1145/2939672.2939801.