AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

SPRAWOZDANIE Z INFORMATYCZNYCH NARZĘDZI PRACY GRUPOWEJ

SCRUM CHALLENGE - SPRINT PLANNING

Automatyczna klasyfikacja tekstów

Przygotowali: Zofia Trzaskalik Iga Zieja Jakub Kisiel Kamil Machoń Krzysztof Kaplyta

Produkt:

Tematem naszego projektu była "Automatyczna klasyfikacja (kategoryzacja) tekstów". Mieliśmy napisać program, który automatycznie, za pomocą wyszukanych słów kluczy, dopasowywał by nowy tekst do odpowiedniej kategorii, oraz porównując nowy tekst z innymi z danej kategorii uczył by się nowych słów-kluczy.

Oprócz tego, mieliśmy za zadanie wykonanie sprawozdania z wykonanej pracy przy tworzeniu powyżej opisanego programu.

Metodyka:

Pracowaliśmy grupowo wykorzystując metodykę Scrum. Jest to, poznana przez nas na zajęciach, metoda ustalająca sposób pracy. Polega na planowaniu sprintów, które są czasowo ograniczone (u nas czasem jednego tygodnia), a w nich wyznaczenie celów i zadań na ten okres. Każdy członek zespołu Srumowego ma wyznaczone zadania do wykonania, które także są określone w czasie. Scrum zakłada, że praca jest systematycznie sprawdzana, podczas codziennych spotkań (u nas w formie raportów) oraz podsumowywana na koniec każdego ze sprintów.

Narzędzia:

Używaliśmy narzędzi: GitHub, Trello, oraz dodatki do drugiego z nich: Plus for Trello i Scrum for Trello.

GitHub w naszej pracy służył głównie jako repozytorium, w którym mogliśmy przechowywać wszystkie stworzone pliki, potrzebne do wykonania zadania. Na jego podstawie można obecnie zobaczyć kto jakie pliki stworzył i zmodyfikował, oraz w jakim miejscu dokonał zmian. Każda operacja była wykonywana dzięki commitom zawierającym opis wykonanej operacji.

Trello natomiast pomogło nam w planowaniu pracy. Dla każdego sprintu tworzyliśmy nową tablicę, a w niej listę z celami, oraz zadaniami określającymi jak dany cel osiągnąć. W każdej tablicy także znajdowała się lista "Done", do której przenosiliśmy wykonane zadania. Każde z zadań zawierały: opis (jeśli był konieczny), przydział osoby odpowiedzialnej, ostateczny termin wykonania, oraz check-listę, jeśli zadanie dotyczyło paru osób i był potrzebny dodatkowy podział. Po za tym istniały etykiety, dzięki którym można było oznaczyć czy w zadaniu pojawił się jakiś problem oraz czy został on rozwiązany

Dzięki Plus for Trello mogliśmy swoją codzienną pracę raportować zaznaczając jaką część pracy z danego zadania wykonaliśmy i ile nam ono zajęło, oraz dodając komentarz na temat tego co w tym zadaniu udało nam się zrobić.

Dodatek Scrum for Trello pozwolił nam na dodawanie planowanego czasu, jaki poszczególne zadania powinny zająć. Poprzez takie działanie mogliśmy lepiej zaplanować co w danym sprincie uda nam się czasowo wykonać, a na co nie starczy nam już czasu.

OPISY SPRINTÓW:

Sprint pierwszy:

Cele:

- Wybór środowiska informatycznego
- Ustalenie sposobu magazynowania plików
- Ustalenie podstawowych 20 kategorii tekstu
- Stworzenie magazynu plików
- Stworzenie programu do wczytywania plików

Zadania:

- Wybór środowiska informatycznego:
 - porównanie środowisk pod względem wczytywania plików i porównywania ich

osoba odpowiedzialna: Jakub Kisiel, czas realizacji: do 12.05.2016, godz 23:55

ocena znajomości i łatwości użytkowania

osoba odpowiedzialna: Jakub Kisiel, czas realizacji: do 8.05.2016, godz 23:55

zrobienie instrukcji do najważniejszych funkcji, które zostaną użyte w programie

osoba odpowiedzialna: Jakub Kisiel, czas realizacji: do 9.05.2016, godz 23:55

wybór systemu operacyjnego

osoba odpowiedzialna: Zofia Trzaskalik, czas realizacji: do 9.05.2016, godz 23:55

- Ustalenie sposobu magazynowania plików:
 - sprawdzenie dostępnych opcji magazynowania plików

osoba odpowiedzialna: Iga Zieja, Krzysztof Kaplyta, czas realizacji: do 5.05.2016, godz 23:55

porównanie trudności obsługi

osoba odpowiedzialna: Krzysztof Kaplyta, czas realizacji: do 5.05.2016, godz 23:55

wybór jednego sposobu

osoba odpowiedzialna: Krzysztof Kaplyta, czas realizacji: do 6.05.2016, godz 23:55

- Ustalenie podstawowych 20 kategorii tekstu:
 - ustalenie przez każdego 10 kategorii

osoba odpowiedzialna: wszyscy (stworzona była check-lista), czas realizacji: do 6.05.2016, godz 23:55

odrzucenie powtarzających się kategorii

osoba odpowiedzialna: Iga Zieja, czas realizacji: do 10.05.2016, godz 23:00

wybór właściwych 20 kategorii

osoba odpowiedzialna: Zofia Trzaskalik, czas realizacji: do 10.05.2016, godz 23:55

- Stworzenie magazynu plików:
 - stworzenie podfolderów z kategoriami

osoba odpowiedzialna: Jakub Kisiel, czas realizacji: do 12.05.2016, godz 23:55

wybór fizycznego miejsca przechowywania plików

osoba odpowiedzialna: Jakub Kisiel, czas realizacji: do 8.05.2016, godz 23:55

- Stworzenie programu do wczytywania plików:
 - stworzenie pliku nagłówkowego

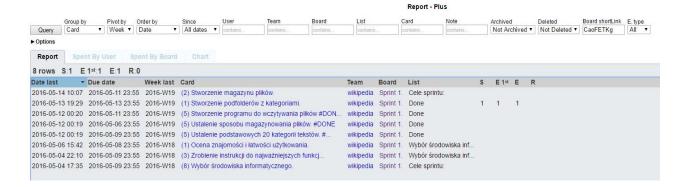
osoba odpowiedzialna: Kamil Machoń, czas realizacji: do 10.05.2016, godz 23:55

stworzenie pliku zawierającego funkcję do wczytywania plików

osoba odpowiedzialna: Kamil Machoń, czas realizacji: do 11.05.2016, godz 23:00

stworzenie pliku main

osoba odpowiedzialna: Kamil Machoń, czas realizacji: do 11.05.2016, godz 23:55



Sprint drugi:

Cele:

- Podanie słów kluczowych dla każdej kategorii
- Stworzenie funkcji szukającej podanego słowa w tekście
- Stworzenie funkcji szukającej każdego z podanych słów w tekście
- Stworzenie funkcji zamykającej plik

Zadania:

- Podanie słów kluczowych dla każdej kategorii:
 - Wymyślenie i przyporządkowanie do każdej kategorii słów-kluczy

Osoba odpowiedzialna: Krzysztof Kaplyta, Iga Zieja

Czas realizacji: 19.05.16

- Zapisanie słów kluczy w odpowiednim dokumencie w repozytorium

Osoba odpowiedzialna: Jakub Kisiel

Czas realizacji: 19.05.16

- Stworzenie funkcji szukającej podanego słowa w tekście:
 - Stworzenie deklaracji funkcji w pliku zawierającym deklaracje funkcji

Osoba odpowiedzialna: Zofia Trzaskalik

Czas realizacji: 17.05.16

- Dodanie możliwości przetestowanie funkcji w pliku main

Osoba odpowiedzialna: Zofia Trzaskalik

Czas realizacji: 17.05.16

- Stworzenie definicji funkcji w pliku zawierającym definicje funkcji

Osoba odpowiedzialna: Zofia Trzaskalik

Czas realizacji: 17.05.16

- Stworzenie funkcji szukającej każdego z podanych słów w tekście:
 - Przerobienie funkcji szukającej słowa w tekście, tak żeby szukała każdego słowa po kolei

Osoba odpowiedzialna: Zofia Trzaskalik, Kamil Machoń

Czas realizacji: 18.05.16

- Stworzenie tablicy na wpisywane słowa przez użytkownika

Osoba odpowiedzialna: Zofia Trzaskalik, Kamil Machoń

Czas realizacji: 18.05.16

Przetestowanie funkcji

Osoba odpowiedzialna: Zofia Trzaskalik, Kamil Machoń

Czas realizacji: 18.05.16

- Stworzenie funkcji zamykającej plik
 - Przetestowanie programu i poprawienie ew. błędów

Osoba odpowiedzialna: Jakub Kisiel

Czas realizacji: 19.05.16

Sprint trzeci:

Cele:

- Sprawdzenie, czy wszystkie otwarte pliki zostały zamknięte
- Stworzenie funkcji umożliwiającej zapis pliku w wybranym miejscu
- Stworzenie funkcji przydzielającej kategorie tekstowi
- Przetestowanie programu
- Stworzenie docelowej bazy danych

Zadania:

• Sprawdzenie, czy wszystkie otwarte pliki zostały zamknięte:

- Sprawdzenie czy w każdym miejscu w programie otwarte pliki zostały zamknięte
- Dodanie zamknięcia pliku w odpowiednich miejscach w programie

Osoba odpowiedzialna: Jakub Kisiel.

- Stworzenie funkcji umożliwiającej zapis pliku w wybranym miejscu:
 - Znalezienie artykułów / stron zawierających możliwe rozwiązania problemu
 - Wybranie najlepszego dla nas sposobu rozwiązania problemu
 - Utworzenie odpowiednich elementów kodu w programie

Osoba odpowiedzialna: Krzysztof Kaplyta.

- Stworzenie funkcji przydzielającej kategorie tekstowi:
 - Stworzenie pętli w programie, która dla każdej kategorii, będzie sprawdzać zgodność słów kluczy z tekstem
 - Stworzenie 5 liczników (odpowiadającym ilości słów kluczy w tekście) do porównywania
 - Porównanie liczników
 - Zwrot numeru kategorii

Osoba odpowiedzialna: Zofia Trzaskalik oraz Kamil Machoń.

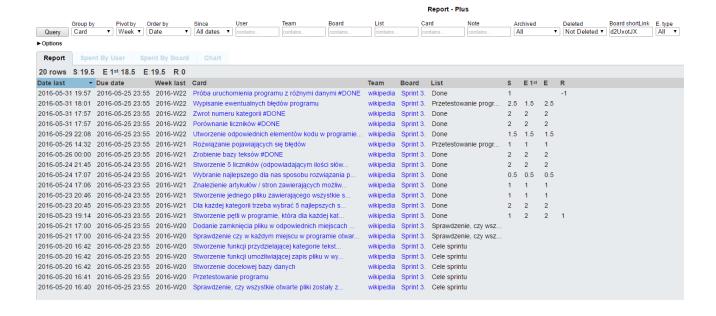
- Przetestowanie programu:
 - Próba uruchomienia programu z różnymi danymi
 - Wypisanie ewentualnych błędów programu
 - Rozwiązanie pojawiających się błędów

Osoba odpowiedzialna: Zofia Trzaskalik oraz Kamil Machoń.

- Stworzenie docelowej bazy danych:
 - Dla każdej kategorii trzeba wybrać 5 najlepszych słów kluczy z podanych
 - Stworzenie jednego pliku zawierającego wszystkie słowa klucze ze wszystkich kategorii
 - Zrobienie bazy teksów

Osoba odpowiedzialna: Iga Zieja.

Wszystkie zadania zostały wykonane w określonym terminie. Poniżej przedstawiamy tabelę podsumowującą cały sprint.



Sprint czwarty:

Cele:

Głównym celem sprintu czwartego było podsumowanie dotychczasowej pracy zrealizowanej przez cały zespół.

W sprincie czwartym wygospodarowaliśmy sobie czas na sfinalizowanie niedokończonych zadań z poprzedniego sprintu.

Zadania:

Wykonanie sprawozdania z pierwszego sprintu

Osoba odpowiedzialna: Zofia Trzaskalik Czas realizacji: do 1.06.2016 do 00:00

Podsumowanie sprintu drugiego

Osoba odpowiedzialna Kamil Machoń. Czas realizacji: do 1.06.2016 do 00:00

Przegląd sprintu drugiego

Osoba odpowiedzialna: Krzysztof Kaplyta. *Czas realizacji*: do 1.06.2016 do 00:00

- Zrelacjonowanie wstępnych założeń całego projektu.

Osoba odpowiedzialna: Zofia Trzaskalik.

Za zadanie miała opisanie produktu, który mamy dostarczyć; metodykę działania i narzędzia których użyliśmy do zrealizowania projektu.

Czas realizacji: do 1.06.2016 do 00:00

- Przygotowanie dokumentacji naszych osiągnięć formie papierowej

Osoba odpowiedzialna: Iga Zieja. Czas realizacji: do 2.06.2016 do 14:00

 Każdy z uczestników, miał za zadanie opisać swój wkład w poszczególne sprinty, raportowanie oraz aktywność udokumentowaną narzędziami. W tym sprawozdaniu każdy miał podać procentowo, wkład własny w każdy ze sprintów.

Czas realizacji: do 1.06.2016 do 00:00

- Do celów tego sprintu należało również przygotowanie prezentacji naszego projektu.

Czas realizacji: do 2.06.2016 do 14:00

Sprint czwarty, jako ten najkrótszy opierał się głównie na zebraniu informacji ze wcześniejszych sprintów i zintegrowaniu ich.

Osobą odpowiedzialną za zrobienie sprawozdania z czwartego sprintu był: Jakub Kisiel

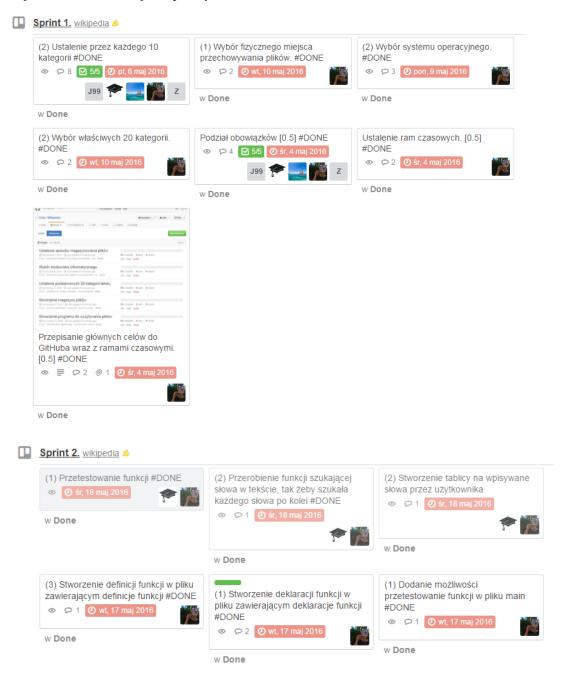
Czas realizacji: do 2.06.2016 do 14:30

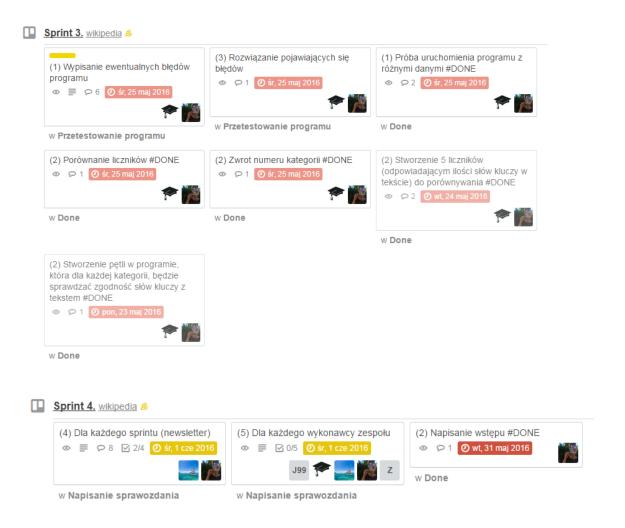
WKŁAD W PRACĘ POSZCZEGÓLNYCH CZŁONKÓW ZESPOŁU

Zofia Trzaskalik

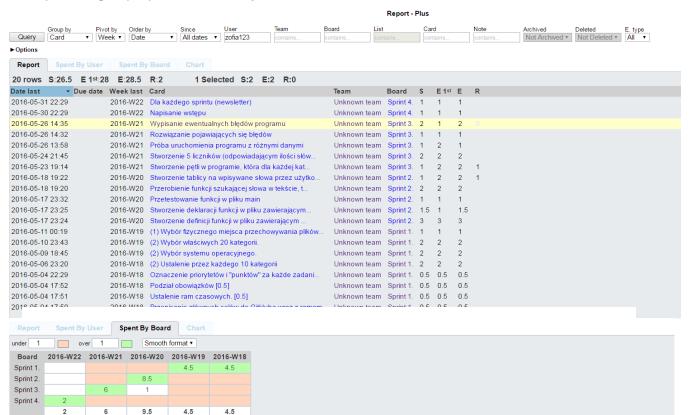
Poza działaniami związanymi z wykonywaniem zadań, moja aktywność polegała na: tworzeniu tablic na kolejne Sprinty, w dużej mierze wymyślaniu celów i zadań, które powinien zespół wykonać. W każdym ze Sprintów zajmowałam się także określeniem ram czasowych. Jako Scrum Master w pierwszym Sprincie dbałam o to, aby każdy członek zespołu wiedział jakie zadanie ma przydzielone, oraz wykonał go w terminie. Sprawdzałam także poprawność wykonanych zadań, oraz czy dokumentacja wykonania zadania znajduje się w repozytorium.

Moje zadania na kolejne Sprinty:

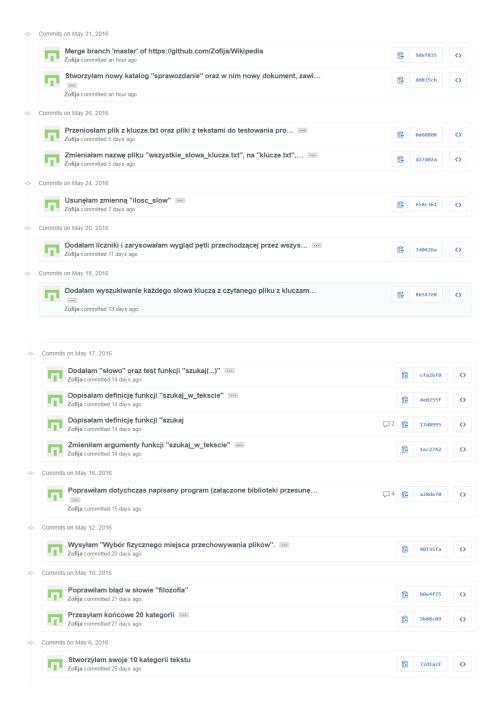




Wszystkie raporty wykonane w Plus for Trello:



Commity wykonane na GitHubie:



Wykres określający ilość wykonanych commitów w czasie:

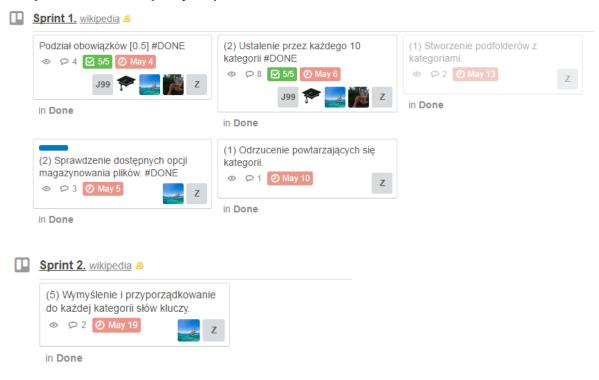


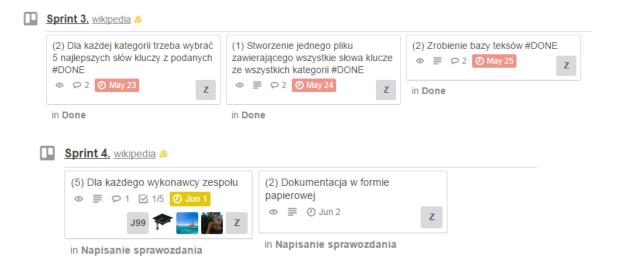
Oceniam mój wkład w pracę zespołu na 35%

Iga Zieja

Moja aktywność podczas pracy nad projektem polegała na przypisywaniu się do zadań i realizowaniu ich w określonych terminach. Moje zadania skupiały się w większości na pracy z kategoriami, wyszukiwaniu do nich słów kluczy i tym podobne.

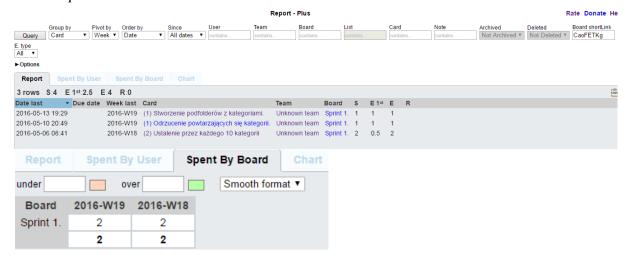
Moje zadania na kolejne sprinty:





Raporty wykonane w Plus for Trello:

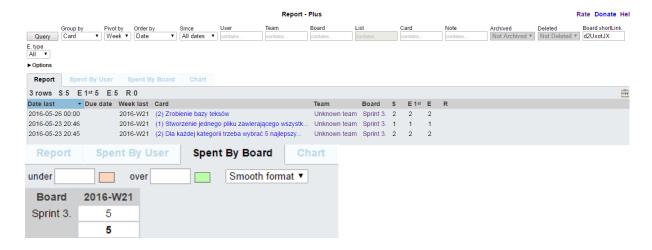
• Sprint 1



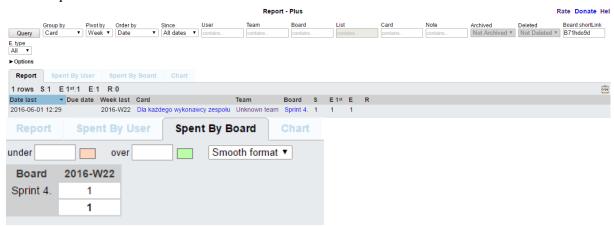
• Sprint 2



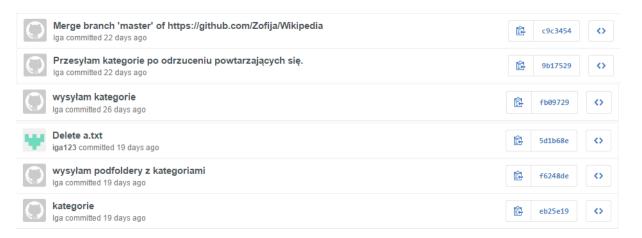
• Sprint 3

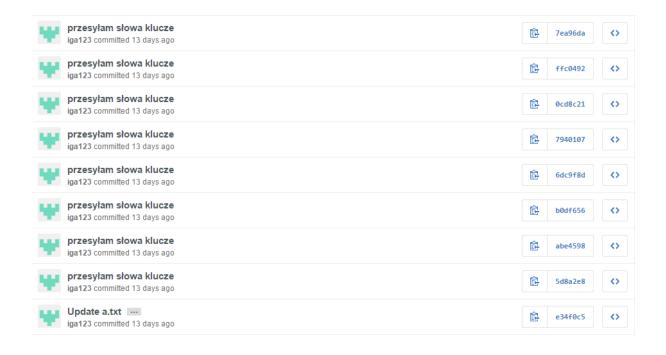


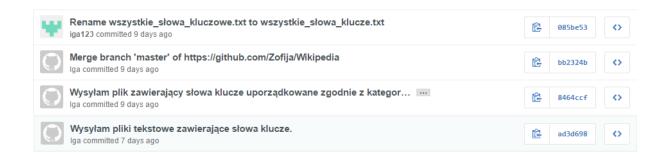
• Sprint 4



Commity wykonane na GitHubie:







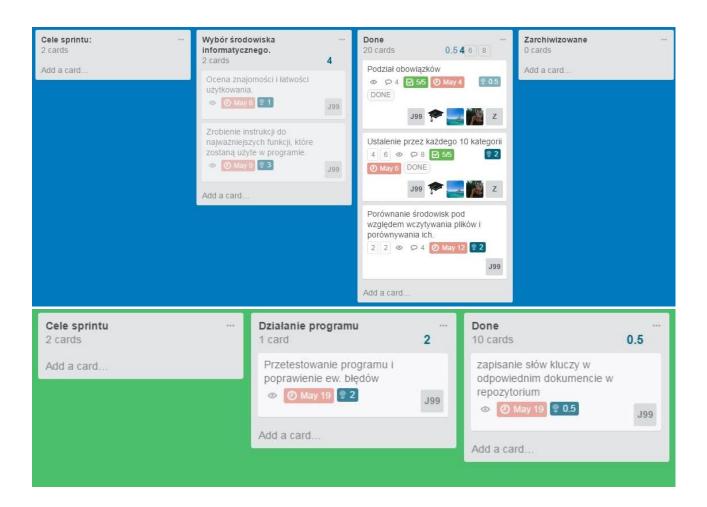
Wykres określający ilość wykonanych commitów w czasie:



Oceniam mój wkład w pracę zespołu na 15%

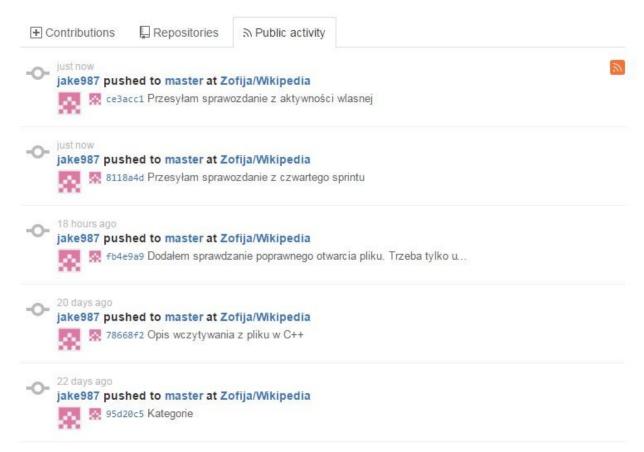
Jakub Kisiel

Moja aktywność podczas tworzenia sprintów opierała się na ewentualnym dodawaniu zadań, które zostały pominięte lub zbyt ogólnie przedstawione. Moje zadania opierały się głównie na aspektach teoretycznych takich jak opis języka programowania, ale również kontroli i poprawy potencjalnych błędów.







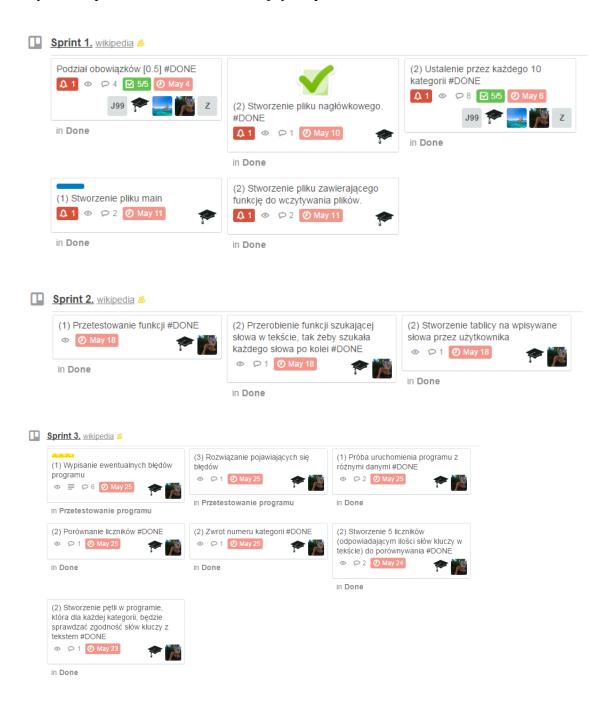


Podsumowując, moja zaangażowanie w pracę zespołu było zmienne. Swoją aktywność określam na 18,5%.

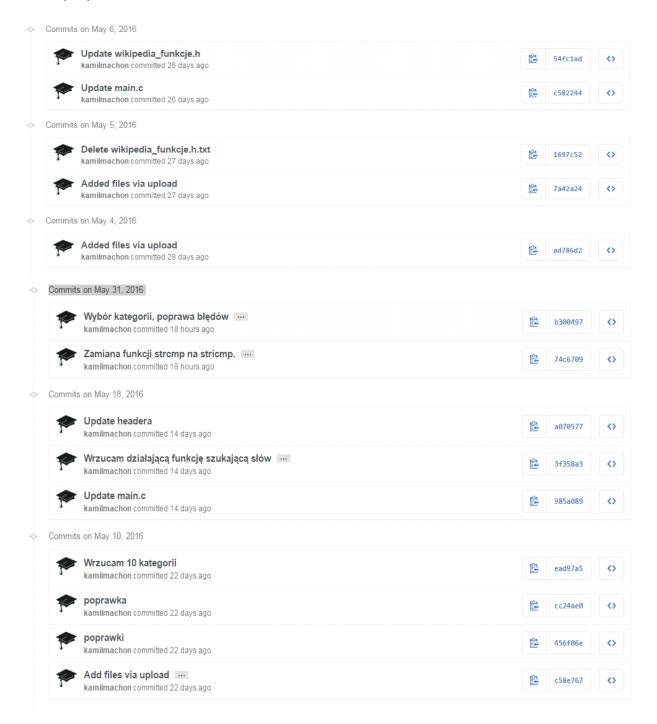
Kamil Machoń

Moim zadaniem w przedstawianym projekcie było przede wszystkim napisanie kodu wykonywalnego. Pisałem program, który był celem poniższego projektu, poprawiałem zauważone przezemnie błędy, starałem się zwiększyć czytelność kodu, oraz znaleźć najbardziej optymalny algorytm dla programu. Podczas drugiego sprintu przejąłem role Scrum Mastera i dbałem o to, by poszczególne zadanie wykonane były poprawnie i w terminie.

Wykonane przezemnie zadania w kolejnych sprintach:



Commity wykonane na GitHubie:



Wykres aktywności:



Aktywność udokumentowana dzięki Plus for Trello



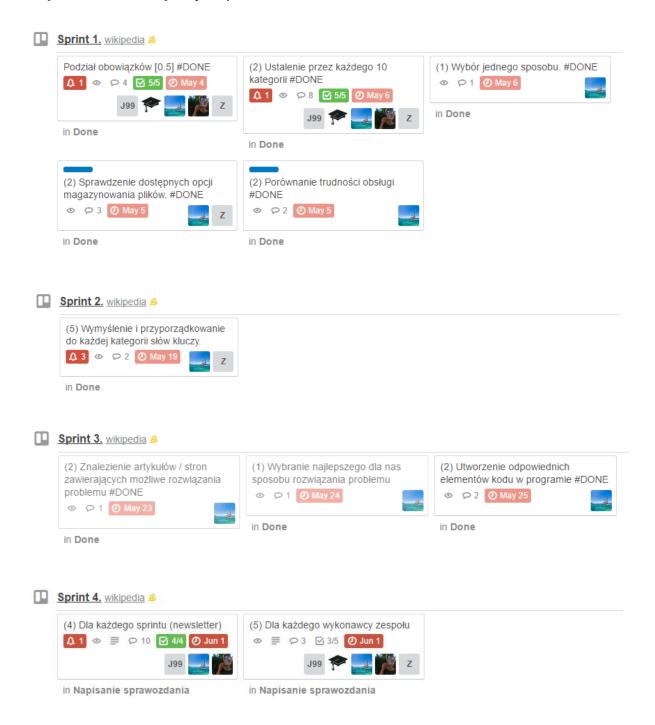
| 4 rows S:5.5 E 1st:5.5 | | E:5.5 R:0 | | | | | | | |
|------------------------|----------|-----------|--|--------------|-----------|-----|-------|-----|---|
| Date last | Due date | Week last | Card | Team | Board | S | E 1st | E | R |
| 2016-05-31 19:5 | 7 | 2016-W22 | Próba uruchomienia programu z różnymi danymi | Unknown team | Sprint 3. | 1 | 1 | 1 | |
| 2016-05-31 18:0 | 1 | 2016-W22 | Wypisanie ewentualnych błędów programu | Unknown team | Sprint 3. | 0.5 | 0.5 | 0.5 | |
| 2016-05-31 17:5 | 7 | 2016-W22 | Zwrot numeru kategorii #DONE | Unknown team | Sprint 3. | 2 | 2 | 2 | |
| 2016-05-31 17:5 | 7 | 2016-W22 | Porównanie liczników #DONE | Unknown team | Sprint 3. | 2 | 2 | 2 | |

Swój wkład oceniam na 20%

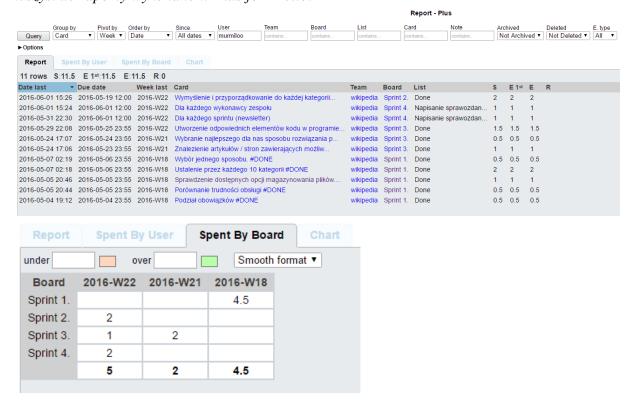
Krzysztof Kaplyta:

Moja działalność związana z projektem opierała się na aspektach teoretycznych. Brałem również udział w tworzeniu programu.

Moje zadania na kolejne sprinty:

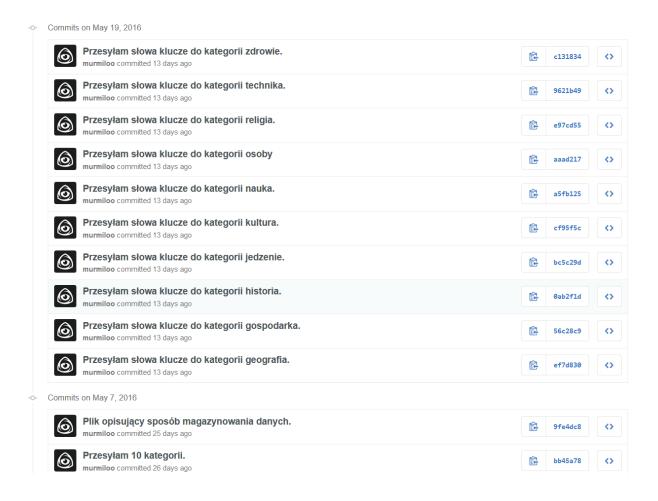


Wszystkie raporty wykonane w Plus for Trello:



Commity wykonane na GitHubie:





Wykres określający ilość wykonanych commitów w czasie:



Swój wkład w pracę zespołu oceniam na około 20%.

Podsumowanie

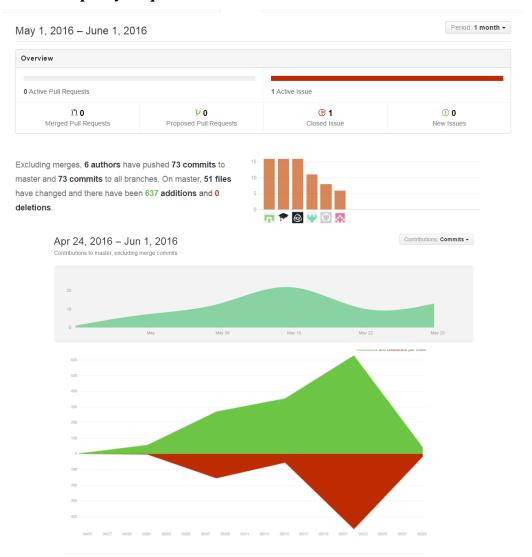
Podczas trwania całego projektu udało nam się wykonać niemalże wszystkie zadania, oraz osiągnąć prawie wszystkie zamierzone cele. Jednak przede wszystkim każdy z nas nauczył się jak powinna wyglądać dobrze zorganizowana praca w grupie, jak ważne jest dzielenie się obowiązkami, komunikowanie o tym, co się udało lub gdzie występuje jakiś problem, wyznaczanie drobnych, ale konkretnych zadań, żeby łatwiej było osiągnąć planowany cel.

Osiągnięcia

W wyżej opisanych Sprintach pokazaliśmy, co robiliśmy podczas minionego miesiąca pracy. Obecnie nasz produkt zawiera: program złożony z 3 plików (pliku nagłówkowego, pliku zawierającego funkcje używane w programie oraz z pliku main), katalogi z nazwami odnoszącymi się do poszczególnych, wymyślonych przez nas, kategorii tekstów, plik zawierający słowa klucze (dla każdej kategorii po 5 słów), przykładowe teksty, do testowania programu, oraz sprawozdanie z pracy.

Program kompiluje się bezbłędnie oraz nie zawiera żadnych błędów. Jest zdolny do wczytania z pliku podanego tekstu, następnie do porónania go z listą słów kluczy. W wyniku programu uzyskamy informację o tym, które ze słów kluczy zostały użyte w tekście oraz ile razy. Dzięki temu będzie można podjąć decyzję co do przynależności teksu do kategorii.

Podsumowanie pracy zespołu:



Plany

Jak wynika z powyższego opisu, nie osiągnęliśmy jeszcze górnej poprzeczki. Gdybyśmy kontunuowali nasze działania, z pewnością kolejnym celem stało by się szukanie nie tylko słów kluczy, ale także ich odmian, po przydzieleniu tekstu do danej kategorii porównanie go z pozostałymi tekstami oraz szukanie nowych powtarzających się słów. Trzeba byłoby też dopracować system przydzielania kategorii, aby mógł bezbłędnie określać gdzie dany tekst powinien zostać przydzielony, oraz który w momencie nie znalezienia odpowiedniej kategorii stworzy nową.