Opis problemu: Mamy grę komputerową RPG w fazie wczesnego dostępu. Chcemy zbierać dane dotyczące rozgrywki graczy, takie jak: ruch postaci, interakcje z otoczeniem, odbyte walki, postępy. Dane należy następnie przeanalizować, aby lepiej zrozumieć zachowania graczy i dokonać predykcji trendów. Celem jest wyeliminowanie z gry problemów technicznych i dostosowanie parametrów rozgrywki, takich jak ekonomia w grze, stopień trudności przeciwników itd.

Stosowana chmura: **Azure**

Wybrane narzędzia:

- **Azure Event Hubs** – usługa pozyskiwania danych w czasie rzeczywistym. Umożliwia przesyłanie strumieniowe milionów zdarzeń na sekundę. Pozwala na prowadzenie analizy w czasie rzeczywistym i reagowanie na zdarzenia w momencie ich wystąpienia. Jest warstwą pozyskiwania zdarzeń z dowolnego źródła do ich przesyłania -  integruje się z usługami danych i analiz na platformie Azure.

- **Azure Stream Analytics** – usługa do analizy dużej ilości danych w czasie rzeczywistym. Korzysta ze składni języka SQL i może być rozszerzony za pomocą niestandardowego kodu języka JavaScript lub C#. Dane wyjściowe usługi można skierować do wielu systemów magazynowych.

**-Azure Data Lake Storage** – oparte na chmurze rozwiązanie typu data lake. Usługa Data Lake to pojedyncze, scentralizowane repozytorium w którym można przechowywać wszystkie dane, zarówno ustrukturyzowane jak i w formacie nieprzetworzonym. Zestaw funkcji do analizy big data opartych na Azure Blob Storage.

**-Azure Data Factory** – usługa przekształcania i integracji danych, oraz organizacji przenoszenia ich na dużą skalę. Umożliwia konstruowanie procesów ETL oraz ELT. Przeznaczona do użycia na nieprzetworzonych, niezorganizowanych danych big data na różnych systemach magazynowania, w celu doprowadzenia ich do stanu umożliwiającego analizę.

**- Azure Databricks** – usługa udostępnia najnowsze wersje Apache Spark. Umożliwia łatwą konfiguracje środowiska i automatyczne skalowanie. Usługa Azure Databricks obsługuje języki Python, Scala, R, Java i SQL, a także struktury i biblioteki nauki o danych, takie jak TensorFlow, PyTorch i scikit-learn.

**- Azure Machine Learning** – usługa do tworzenia i wdrażania modeli uczenia maszynowego oraz zarządzania nimi. Korzysta z zaawansowanej infrastruktury sztucznej inteligencji. Przyspiesza cykl życia projektu uczenia maszynowego.

Opis Architektury:

Zbieranie danych: Usługa Azure Event Hubs będzie służyć do przesyłana strumieni z gry komputerowej. Dane będą mogły być zbierane w czasie rzeczywistym. W celu przefiltrowania i przetworzenia danych można zastosować usługę Azure Stream Analytics, która dane przekaże do magazynu Data Lake.

Przechowywanie danych: Dane przechowywane będą w usłudze Azure Data Lake Storage. Do zarządzania przepływem danych i ładowania ich do odpowiednich magazynów można zastosować usługę Azure Data Factory.

Przetwarzanie i analiza danych: Kiedy już dane znajdują się w magazynie, możemy przejść do ich analizy. Do wykonywania operacji na tak dużym zbiorze danych przydatna będzie usługa Azure Databricks. Można będzie tam dokonać analizy danych w celu identyfikacji wzorców i zjawisk. Azure Machine Learning będzie służyć do trenowania i zarządzania modelem utworzonym na podstawie zebranych danych z gry.

Wykorzystanie danych i modelu: Model może pomóc zrozumieć, w jaki sposób gracze wykorzystują zasoby w grze, takie jak: waluta, umiejętności, przedmioty. Model może przewidywać trendy w użyciu tych zasobów i pomóc w zarządzaniu ich dostępnością. Przez zrozumienie zachowań i preferencji graczy możliwa będzie optymalizacja różnych elementów w grze; umiejscowienie ofert czy zadań, odpowiednie ceny przedmiotów, zbalansowane nagrody, ilość przeciwników itd. Model może być użyty do testowania nowych rozwiązań, takie jak: wprowadzenie nowego przedmiotu, poziomu, zmiana ceny.

Model predykcyjny może też zostać zaimplementowany jako API i zintegrowany z samą grą. Wtedy może służyć do dopasowania poziomu trudności indywidualnie pod danego gracza, może też sugerować indywidualnie kolejność wykonania zadań, pomóc określać graczowi cele.

Można również określić popularność danych funkcji lub elementów gry, co może być użyteczne przy planowaniu aktualizacji ( jeśli coś jest dużo używane przez graczy, oznacza, że im się podoba i można te elementy rozwijać, mało używane funkcje być może są nie zbalansowane i należy się im pochylić i naprawić ).