

ECF 2-juin-23

Objectif :

-
- **Organisation projet agile**

- Product Owner

Le Product Owner est le représentant du client, il :

- Définit les fonctionnalités du produit.
- Décide des dates de livraison et de leur contenu.
- Est responsable de la rentabilité du produit (ROI).
- Prioriser les fonctionnalités en fonction de la valeur métier.
- Ajuste les fonctionnalités et leur priorité avant chaque planification d'itération.
- Accepte ou rejette les fonctionnalités réalisées.
- Anime la réunion de planification de sprint.

Le rôle du Product Owner est assuré par un membre de l'équipe métier (MOA) et au besoin une assistance (AMOA)

- Equipe d'assistance au Product Owner.

Elle assiste le Product Owner et à ce titre peut intervenir dans le cadre des activités associées aux responsabilités de ce dernier. Plus concrètement, elle :

- Alimente et maintient le Product Backlog, et priorise les éléments de ce dernier.
- Ajuste les fonctionnalités prioritaires du Product Backlog (candidates au périmètre du prochain sprint) et s'assure que les pré-requis aux développements associés seront disponibles en temps voulu (exemples : décisions métier, jeux de données du SI amont,...).
- Rédige les User Stories associées aux fonctionnalités prioritaires et dessine au besoin des maquettes d'écran
- Répond aux questions soulevées par l'équipe de développement en cours de Sprint et complète au besoin les User Stories associées.
- Vérifie en cours de Sprint la bonne couverture du besoin des fonctionnalités terminées en collaboration avec l'équipe de développement.
- Rédige les plans de tests.
- Participe à la réunion de revue de sprint au cours de laquelle, elle aide le Product Owner à accepter ou rejeter les fonctionnalités présentées.
- Testé avant mise en production la conformité du produit dans son ensemble.

.Scrum Master

Le Scrum Master appartient à l'équipe de développement (MOE), il :

- S'assure que l'équipe est pleinement opérationnelle et productive.
- Établit une collaboration étroite entre l'ensemble des rôles et fonctions
- Supprime les obstacles rencontrés par l'équipe de développement.
- Protège l'équipe des interférences extérieures.
- Assure le suivi du processus.

.Équipe de développement

L'équipe de développement :

- Réalise les fonctionnalités du produit.
- Présente au Product Owner les résultats de son travail sous forme de démonstrations.
- Maintient à jour les spécifications détaillées du produit.
- Package et livre le produit.

Préparation:

Le Product Owner doit s'assurer que le Product Backlog (Liste des exigences attendues) est correctement ordonnancé en fonction de la valeur métier et du coût de chaque exigence. Pour les exigences dépourvues de coût, il peut solliciter l'équipe de développement afin de procéder à des séances d'estimation. Il participe à ces séances - assisté par son équipe - en répondant aux questions de l'équipe de développement. Il doit également s'assurer que le besoin associé aux exigences prioritaires du Product Backlog (candidates au périmètre du prochain Sprint) est suffisamment mûr (décisions métiers structurantes prises, besoin formalisé sous forme de User Stories,...)

- Rappel des ETL.
- Rappel de PySpark databricks.

Azure Databricks est un service d'analyse du Big Data rapide, simple et collaboratif basé sur Apache Spark, conçu pour la science des données

voici un exemple de code que nous pouvons taper sur azure Databricks:

créer un DataFrame avec Python

```
import pandas as pd

data = [[1, "Elia"], [2, "Teo"], [3, "Fang"]]

pdf = pd.DataFrame(data, columns=["id", "name"])

df1 = spark.createDataFrame(pdf)
df2 = spark.createDataFrame(data, schema="id LONG, name STRING")
```

Lire une table dans un DataFrame

```
spark.read.table("<catalog_name>.<schema_name>.<table_name>")
```

Charger des données dans un DataFrame à partir de fichiers

```
df = (spark.read
      .format("csv")
      .option("header", "true")
      .option("inferSchema", "true")
      .load("/databricks-datasets/samples/population-vs-price/data_geo.csv")
)
```

Combiner des Data Frames avec jointure et union

```
joined_df = df1.join(df2, how="inner", on="id")
unioned_df = df1.union(df2)
```

Filtrer des lignes dans un DataFrame

```
filtered_df = df.filter("id > 1")
filtered_df = df.where("id > 1")
```

Sélection de colonnes à partir d'un DataFrame

```
select_df = df.select("id", "name")
```

```
subset_df = df.filter("id > 1").select("name")
```

Afficher le DataFrame

```
display(df)
```

- Rappel des Machine learning , deep learning.

Le machine learning est une forme d'intelligence artificielle qui permet aux systèmes d'apprendre automatiquement et de s'améliorer à partir de l'expérience sans être explicitement programmés

Le Deep learning est un sous-ensemble de l'apprentissage automatique qui utilise des réseaux neuronaux artificiels pour modéliser et résoudre des problèmes complexes.

Sujet : API, Web Scraping : Vidéos

Première partie :

Pour un projet vu en formation expliquer :

- le workflow: (flux des données)

Étape 1 : Inspecter votre source de données

Étape 2 : Extraire le contenu HTML d'une page

Étape 3: Analyser le code HTML avec une belle soupe

- les outils utilisés: (framework , libraries)

```
import requests
from bs4 import BeautifulSoup
page = requests.get(URL)

page = requests.get(URL)
soup = BeautifulSoup(page.content, "html.parser")
```

- Le dataset: (API , fichiers , scrapping)

Analyser des films

http://www.imdb.com/search/title?release_date=2017&sort=num_votes,desc&page=1

```
url
='http://www.imdb.com/search/title?release_date=2020&sort=num_votes,desc&page=1'

#permet de charger l'url

response = get(url)

#permet d'importer la bibliothèque BeautifulSoup qui va nous permettre
d'extraire les données d'une page web

from bs4 import BeautifulSoup

html_soup = BeautifulSoup(response.text, 'html.parser')
```

```

#importation de la bibliothèque numpy et pandas qui sont de puissante

import numpy as np

import pandas as pd

#matplotlib permet la visualisation de donnée

import matplotlib as plt

from requests import get

from time import sleep

from random import randint

from warnings import warn

from IPython.display import clear_output

```

- L'outil utilisé pour la visualisation ainsi que les alternatives éventuelles.
Possibilité de partager le travail.

microsoft-teams pour un travail collaboratif

zoom

Jira

Liste de sujet :

Nom prénom	Projet	Thème	Remarque
Vianney	Vidéos	API, Web scraping	

