# Assignment 5

## The habitability of a planet

The habitability of a planet depends on key factors. Liquid water, regulated by distance from the star and atmospheric conditions, is fundamental. Atmospheric composition stabilizes temperatures and shields from radiation. Planet size and gravity affect atmosphere retention and water support. Orbital stability ensures a conducive climate. Geophysical factors like tectonics and magnetic fields regulate surface conditions. Together, these factors determine a planet's potential to sustain life.

We analyze exoplanet features from transit data (Kepler, TESS, CoRot, Hubble, Spitzer) to classify and rank planets based on Earth Similarity Index (ESI). In this assignment we will perform analysis with physical features.

## Question 1 (Model Training)

### Methodology

1. **Data Preprocessing**

   - Load the provided CSV dataset containing exoplanet features to train a machine learning model. Click here to download

2. **Train a Machine Learning Model**

   - Consider the following features to train your model:

     'P_MASS', 'P_RADIUS', 'P_PERIOD', 'P_SEMI_MAJOR_AXIS', 'P_ECCENTRICITY', 'P_INCLINATION', 'P_TEMP_SURF', 'P_TEMP_EQUIL', 'P_GRAVITY', 'P_DENSITY', 'S_TEMPERATURE', 'S_LUMINOSITY', 'S_RADIUS', 'S_MASS', 'S_HZ_OPT_MIN', 'S_HZ_OPT_MAX', 'S_HZ_CON_MIN', 'S_HZ_CON_MAX' versus 'P_ESI'.

     Train a regression model using RandomForestRegressor to predict the Earth Similarity Index (ESI) based on the exoplanet features provided. Check mean squared error and r2 score.

   - Also, plot the weights of each feature to analyze their contribution to depicting ESI.

     (**Interestng observation** - This will not necessarily reflect the exact contribution of each feature in scientific analysis, this will show what our model predicted; you may notice ambiguity with mass and gravity weight contribution if you try creating a new column for gravity calculating it using mass and radius, it's suggested to try and see it).

3. **Select Top 10 Planets with Highest ESI**

   - Once the model is trained and validated, use it to predict the ESI for exoplanets in both the Kepler and TESS datasets.
     Kepler data: Click here for Kepler data, TESS data: Click here for TESS data
   - Identify and list the top 10 exoplanets with the highest predicted ESI from each dataset.

# Question 2 (Research)

## Methodology

1. **Research and Documentation**

   - Select any 2 potentially habitable exoplanets from the lists of top 10 most earth like exoplanets identified in Question 1. ( Any 2 out of both lists, coordinate among yourselves and avoid repetition of exoplanets selected for research)
   - Research each exoplanet to document its characteristics, challenges that humans will face if we attempt to colonize them in future, and potential strategies to overcome these challenges. Finally prepare 1 page report for each.