# TM18004 - Privacy and Security on Online Social Media
# Assignment 2

**Instructions :**

- Languages allowed : Python/Java
- You are free to use API documentation but if referring to any other sources, please cite them
- Please write your own code. All codes will be tested for Plagiarism and if found, institute policy for plagiarism will be followed.
- Document your code properly.
- You can use any database for storing the data but it will be tested at the time of demo.
- Write all the analysis along with graphs, charts etc in analysis.pdf
- Make a readme.txt file with the instructions how to run the code. All libraries, sources etc used should be properly mentioned in it.
- Do the Assignment individually
- Zip all your code files along with analysis and readme file in RollNo_Assignment1.zip format. Example 201402230_Assignment2.zip

**Ques 1** Using Facebook Graph API, given an access token : (50 marks)

a) Collect the list of friends of user
   i) Make a pie chart of the locations and % of friends belonging to it
   ii) Make a pie chart of languages known by friends
   iii) Make a time series graph in which x axis is the month and y axis is the number of friends having their birthday in x month

b) Collect the posts on the timeline of user.
   i) Analyze the count of reactions and the frequency of words present in the post and draw some inferences.

c) Given the access token of User A, can you collect the above information for User B? If yes, how?

**Ques 2** Consider any of the IIIT Hyderabad group that you are a member of and build up a social graph where nodes are the members and edges denotes the connections. Write a code to collect the data and visualize the graph using Gephi tool ( 20 marks). Feel free to be creative in terms of what graphs to plot from the data. At least do 4 plots.

**Ques 3** Collect data for two verified city police Twitter handles (one from delhi / mumbai / bangalore police with in India, and one from outside India like Boston / Metropolitan / Baltimore in the US / UK) Collect data for 1 Sept 2018 - 5 Oct 2018. (5 marks)

**Ques 4.** Compare the difference in behaviour and activity of both the handles (collected in Ques 3) in terms of : (25 marks)

1. Frequency of tweets (Activeness on twitter)
2. User response to their tweets (Number of likes, retweets and replies). Create one score using all these and mention the function you used to create the score.
3. Major concerns the city police is addressing in the two pages
4. How the tweet contents affect their popularity in terms of only text, images, videos, i.e compare the different media content
5. Any other interesting observations