Introduction:

With the advent of social media, and its increased accessibility to what used to be considered private information, there is a need to improve the security of account recovery mechanisms. Some examples of account recovery mechanisms are URLs sent through email, or PINs the user determined when creating the account, or Backup Codes – similar to PINs but longer and, the use of security questions. Common security questions include: “In what city were you born?”, “What is the name of your favorite pet?”, “What is your mother’s maiden name?”, “What high school did you attend?” As an example, consider you are a person that wants to access another users’ Facebook account but you don’t know the users’ account password. You could investigate that users’ account, identify who their mother is, then look for their parents’ last name and you could find the users’ maiden name. If you wanted to find the user’s high school, you could look through their pictures and look for graduation photos where you could see the high school’s name. Some users even put the high school where they attended in their biography section; the same goes with the city in which they were born. Clearly, using the information available in a social network can allow for an unauthorized person to deduce the information to answer an account’s security questions. It has been estimated that cybercrime will cost the world $10.5 trillion *annually* by 2025. In other words, that’s approximately $1,350 per person, every year. If security questions are improved, it could help to reduce that amount considerably.

This work aims to develop a software system capable of going through a user’s social media account and find what security questions can be answered from the information extracted. Firstly, the user’s account will be parsed and stored in a Knowledge Graph. This Knowledge Graph will be based on Ontologies that describe social and physical relationships. Then, the information stored in the Knowledge Graph will be searched to identify the user’s exposure of sensitive information that may be used to answer the account recovery security questions. Moreover, this system could potentially generate a new set of security questions that cannot be easily answered by using the information that the user has already posted.