Web Application Privacy Manager

Eric de La CELLE
ead9@kent.ac.uk

MSc Computer Security

Project by Florian BAHEUX (fb229), Maxence POULAIN (mmgp2) and Eric de La CELLE (ead9)

Words count: 6300

Abstract

Nowadays, Internet and all its aspects are daily used by millions of people. The boom of the social networks has change how people communicate and share with each other. Our personal information are spread everywhere on the Internet and it can be really difficult for people to control and protect them.

The main problem of these leaks of data are caused by unaware people who put online everything of their life without thinking about the consequences before. Moreover, there is no real tool available for users to manage their privacy on all their favourite social networks (Twitter, Facebook...) or the Internet in general.

Our project is based on this problematic and our desire to provide a tool for anyone who wants to take control about their private life online. The outcome is not same as our expectations but it can surely be useful to people to show them how much of themselves are displayed.

Table of contents

I. Introduction

A. Context

This document is a report done for the Project Research module of the MSc of Computer Security at the University of Kent. The subject of the project is to realize a ***Privacy Manager*** to help people to control their personal information over the Internet. All the idea, design and implementation were the result of the work of Florian BAHEUX, Maxence POULAIN and me[1].

B. Problem the project tries to solve

The three of us wanted to use what we learned this year as the bases of our future idea. We wanted something which groups security, the news, Internet and the privacy. After multiple meeting to find something that matches our desires, we were coming out with the idea of a Privacy Manager.

The aim of the project is to create a web interface for people who want to improve their privacy over the Internet. This idea comes from what we learn from this year and from the actuality (for example: numerous cases of leak from Facebook which leads to dismissal…). We want to warn people about the vicious side of social networks.

We wanted to create an interface where all majors - interesting website in term of privacy - social networks and their own privacy settings are present. Moreover, we wanted to create a tool capable of searching online for the connected person - using his name and last name - on different search engines (such as Google, Bing, Yahoo…).

C. Incentive

This project was a good opportunity for all of us to learn more about technologies we have not already work with and our desire to learn more. It is

[1] Referred as 'we' in this document.

a project that follows the news (social networks and privacy seems to appear a lot nowadays and it happens to be a real problem sometime). After some reflection, we have chosen Python that is a language widely used and requested by many companies (it is a big plus when searching for a job to know about python and actual languages). Moreover, it was a good occasion to see how web API are built and used, see what is accessible from a third-party application and what we can do with those data.

D. Aim of the dissertation

In this document, I will first do a review about the privacy on the Internet with a focus on Twitter which is my main subject in the project. Then, I will present more about the project itself and our organization as a group to give detail about how we worked together on a same project but different parts. After this, I will show my research about all the different potential social networks I could work on and then present my part of the project. Following this, I compare the final product we did with our objectives at the beginning. From this point, I will do an analysis of what I have done and then talk a bit about the evolution perspectives for our web application before concluding.

II. Literature review

A. Privacy on the Internet

With the development of the social networks (Facebook, Twitter, MySpace, LinkedIn, Instagram, etc.) the privacy of the users has been compromised. The first role of these websites was to keep in touch with people who are far away from each other. Nowadays, social networks allow everyone to connect to other people just by searching their name. It is possible to get information about almost everyone (depending on how well they protect themselves).

In 2011, Bob Sullivan[2] said:

> *"If you feel like someone is watching you, you're right. If you're worried about this, you have plenty of company. If you're not doing anything about this anxiety, you're just like almost everyone else."*

This sentence expresses the malaise some people can feel because of how their privacy is exposed everywhere to everyone. We can ask ourselves: why? Why those websites ask us more and more personal information? Is it to improve what our friends see? Create new links with new people?

We can answer both yes and no:

- **Yes**: they used our information to improve our visibility and share always more with others. They want to create a page at your image: using what you like, what you dislike, what you are doing in your life to provide always more content you may like or be aware of.
- **No**: When your privacy is used, there is always an economic reason behind. If they use something, they will earn money from it. When you register in a social website, all of what you put online, becomes the property of the website because you have agreed their term of use. That means your data can be sold to anybody, mostly to do targeted advertising (a simple example is Facebook: if you like

[2] https://en.wikipedia.org/wiki/Bob_Sullivan_(journalist)

different pages about a same topic, ads on your page will change too).

We have here the bases of how your privacy can be exposed (using the example of Facebook). I will now talk more about Twitter which was my part in the project and how our privacy can leak.

Some interesting data from different year shows that the tweets are mostly about the author more than a link to a music, video, article…

Honeycutt and Herring - 2008	41% tweets are information about the author himself
Naaman, Boase, & Lai - 2010	50% tweets are information about the author himself, other half is information about someone else or something else.

Privacy on Twitter[3]

Now, if we simply got the information about twitter itself, we can see there are 316 000 000 users monthly and 500 000 000 tweets sent per day[4]. With the previous data, we can estimate that nearly 250 000 000 are information about the author. And, if we continue by saying: there is 0.1% of those tweets which content sensitive information about the author or somebody he knows that represents 250 000 tweets per day, which is a huge amount when personal information are involved.

The next question is: what kind of information can leak? Firstly, we need to remember that tweets are public (readable for anyone who got a Twitter account) if you have not changed your privacy settings.

It appears there are three different kinds of leak that can happen because of tweets: ***the vacation tweets, the drunk tweets and the disease tweets***[5].

[3] http://www2.research.att.com/~bala/papers/ica10.pdf (see bibliography for more information)
[4] https://about.twitter.com/company
[5] https://www.cs.indiana.edu/~kapadia/papers/loosetweets-wpes11.pdf (see bibliography for more information)

• **Vacation Tweets**: they are sensitive only when they inform about where and when the people are going because it can expose them to burglars. Moreover, if the tweet comes from a mobile device, information about where it has been sent can appear on the tweet. So, if someone says when and where he is going with the location, it is easy to take advantage of these data.

• **Drunk Tweets:** in this case, tweets are used to express an emotional status or something you will not tell if you were sober (ex: sexuality, confessions, bodily harm, illegal activities…). Those messages can harm heavily the author and his relatives.

• **Disease Tweets:** over a nine month period (January to September 2010), there is more than 24 000 tweets referring to serious illness such as cancer, HIV, diabetes, etc. Usually, when someone refers to such condition, the concerned person will be in an unstable state of mind which can be an opportunity to take advantage of them which extremely dangerous.

After seeing those data, I think we can say that social networks - and principally Twitter in our case - can be a real weapon for malicious people. This is why we wanted to build an application capable of helping people to understand more the risks.

III. Aims

A. Original aims

1. Description of the objectives

The main goal of this project is to create a web interface where we could manage our privacy over the Internet. To do so, we wanted to gather all the information we could on the user (using search engines) and help them taking control over them. Moreover, we wanted to provide them with an easy way to edit their privacy settings on their favourite social networks.

We wanted to use several search engines to improve the accuracy of the results by comparing if all references appear on different website. The first hundred results are gathered and listed to inform the user about what may be him. This part will help the user to find information about himself and take control over it.

Then, we were thinking about what most of the people use today. Obviously, the social networks were the first and best motive for use and our project about privacy. After a lot of researches about different ones (Facebook, Twitter, Instagram, Pinterest, LinkedIn…) and what are exactly their purpose. We decided to add a privacy panel for Facebook and Twitter on our web application: they are the most commonly used and a lot of mistakes happen because users do not take care of their privacy. We wanted to show them and let them modify properly their settings.

Based on those objectives, we wanted to provide the best and easiest ***Privacy Manager*** for the users.

2. Critic on the objectives

Even when you know what you want to do, sometimes it appears you cannot do it because of what you need to have access to. In our case, the API provided by the different social networks were too secure about the privacy and security settings of the users. We were forced to modify the initial goals to fit the needs.

One of our biggest risk (written in our risk assessment) happened which was a big issue to deal with: we could not change (or sometimes get) the privacy settings. But even with this problem, our main was the same: we wanted to provide to the largest number a tool to help them. Our ambitions were just too high, we set them to a lower level where we would be able to work.

IV. Organisation

A. Working as a group

The Project Research allows us to build something from scratch, alone or in group. We were thinking it could be a good opportunity to improve our management skills. Of course, working as a group adds a lot of advantages, but increase the mistakes that could happen.

1. Team management

We were a small group of three people, but we needed to set some simple rules of work. We decided to fix days and hours of work using a calendar. Communicating was a priority to keep everyone aware of what have been done, where were the issues to quickly think together about it and lose as little time as possible. At the end of each day of work, we talked about the issues encountered that have not been solved to find a solution. We created some milestones to help our management.

2. Task splitting

We were a small group so it was easy to split the work. Knowing the number of social networks nowadays, we decided to split this part in two. We selected different website which could be interesting to study. We knew for sure that we would be able to create something with Facebook and Twitter, but less with other such as Instagram, Pinterest, LinkedIn, etc. So each one of us (Maxence and me) take one to be sure to have something to produce. The second part was the usage of search engines to see how far you are present on the Internet and what kind of data (videos, photos, others) or information (phone number, email address, address…) and try to control them. For better accuracy it is needed to use multiple search engines (Google, Yahoo, Bing, etc.) to compare the different results and their visibility online.

The more we were advancing on the project, the more we saw the issues we will have in the near future. We needed to do some meeting to discuss about

what was impossible to achieve and how to modify the project itself to fit our expectations.

3. Technologies

As mentioned in the Introduction, we used Python to realize this project. I will talk a bit more about what we used to develop our web application.

- Python 2.7[6]

Python is interpreted, object-oriented and high-level programming language. We decided to use this language because it is one we do not know yet and one of the most popular nowadays. It is an easy language to learn and a good opportunity for us to improve our knowledge.

- Django 1.8[7]

Django is a free open web framework written in Python. It is composed of a set of tools used to develop our website in a fast and easy way. It used the M-V-C (Model - View - Controller) architecture, which is useful to implement our work in the best possible manner.

- Scrappy 1.0[8]

Scrappy is a web crawling module in Python used by Florian to gather information about the user of our ***Privacy Manager*** on different search engines.

B. Milestones

To help us in our work, we decided to set some milestones which have been useful to keep a proper work pace.

[6] https://www.python.org/
[7] https://www.djangoproject.com/
[8] http://scrapy.org/

Event	Deadline
Beginning of the project & researches	06/29/2015
End of the researches & starting point of API and data crawling	07/05/2015
End of API & data crawling Starting point of Individual report and the front part	07/30/2015
End of the front part Our Personal Deadline	08/13/2015
Bugs fix + risk management overtime End of the project	08/27/2015
End of the individual writing	09/15/2015

V. Implementation phase

A. Research

We needed to do a lot of researches before starting the project itself because of the number of social networks nowadays. We have to see, for each selected website what his goal is, what his privacy settings are, will it be relevant to use it in our application? Then, we have to search for connection, which is usually the same for the different social networks: OAuth.

1. Social networks

In this part, I will present the different social networks on which I have done research for the needed of the project.

LinkedIn[9]

LinkedIn is a social network used by both companies and people to find jobs easily and create a network of contacts. The goal when you register on this website is to improve your visibility on the Internet and be able to quickly find the job that corresponds to you.

After seeing the documentation of the API[10], it appears that you came get all the information displayed on the user profile. To do so, you need to ask the permission (for example, for the user profile, you need the `r_basicprofile` member permission). All data seem to be protect and you need to ask permissions to have access on it and so you need to justify yourself to LinkedIn.

This kind of social networks is used to propagate your data to as many people as possible to enlarge your connections and opportunities. It is not our domain, we want to restrain access of the personal data's user to let only what he wants to display and no more. If we applied the same way of thinking we have for other websites, it will be better to hide addresses, mobile numbers, location,

[9] https://www.linkedin.com/
[10] https://developer.linkedin.com/docs/rest-api

etc… which will result in a useless profile. We decided to put aside this website for the moment.

Instagram[11]

Instagram is used to share your life online, usually for your friends and sometimes for everyone (depending on your activity). Using photos, videos and filters to express an emotion or show something to people. We were thinking of this social network because of the numbers of images that are posted online every day (70 000 000 photos per day[12]).

The Instagram API[13] provides methods to get user feed, relationships, location, media… and way more, but not the settings of the user about the privacy of the account. It appears there are two different types of account depending of the parameters the user sets himself on ***Instagram***: public (anyone can see what the user posts online) and private (only the authorized people by the user can have access to his data).

If I wanted to do something really useful, it would have been more relevant to do a project only on this particular social network. I will now explain why I think that way: Instagram is about - principally - photos, some text accompanying them (a title or a description) and tags (using the char "#"). What would have been the best to do for this website is an algorithm to detect particular sensitive things on photos (blood, drugs, alcohol, tears, etc.) which have been selected by searching inside the user's data (and maybe the user data's friend using specific tags to be more accurate) and warn him about displaying such content online. It would have been too much for us to realize something that big for what we intended to do.

[11] https://instagram.com/
[12] https://instagram.com/press/
[13] https://instagram.com/developer/?hl=en

Pinterest[14]

Pinterest provides you topics you will like by giving some keywords about things you appreciate (Stars Wars, Cats, Movies, Cakes…) and then create a wall of pints for you. The goal is to select what you want to keep: something important or which you really like, to pin it on your board.

There is not so much personal information displayed about the user on Pinterest[15], only the name and last name are needed for an account. Moreover, you do not have to post any kind of information yourself because they propose to you content you may like and if you want more you just have to indicate it. Here, the privacy is not taking from the outside, but more from the inside because the website can use the keywords you have used on search engines (ex: Google), use visited websites or information giving by their partners. They can learn a lot from a person's researches and it can be harmful for the user if they keep those data. In addition, there is no API calls to get those settings (even in read-only mode) and we cannot change the privacy of a board (***"secret"*** is used for personal gallery and authorized people).

The privacy management I could have done with this social network is a bit different from what we were thinking at the beginning: we wanted to protect personal information from other users, ***protecting from the outside***. Here, it is more a protection work ***from the inside*** although it could be interesting to add this part in the future.

Twitter[16]

Twitter is one of the most used social network around the world with 300 000 000 active users monthly. The idea is to ***tweet*** something, a ***tweet*** is a message which maximum length is 140 characters. Everything can be posted online using this: a feeling, a photo, a video, an article…

This time, the API[17] provides us some interesting things concerning the privacy: the major part of the settings are available to receive. Moreover ***it seems***

[14] https://www.pinterest.com/
[15] https://developers.pinterest.com/docs/api/overview/
[16] https://twitter.com/
[17] https://dev.twitter.com/rest/public

possible to edit them (there are a POST and a GET requests available). There is the social network with the most available settings to work with and so the most interesting one. I needed to register myself as a developer on the dev part of Twitter to get the different tokens I will need to connect the application with a user and use the API. To be able to exchange with Twitter, it is necessary to use an OAuth connection.

2. OAuth[18]

What is it?

OAuth is an authorization framework used to obtain limited access to a user account (Facebook, Twitter, LinkedIn, etc… almost all social networks which authorize third part application use an OAuth authentication to gain access to their services). By using this protocol, it allows a user to approve one application interacting with another without giving his password to the third part application. Each website which uses the OAuth framework provides the different tokens needed by a third part application to log in:

- The Consumer Key (API Key)
- The Consumer Secret (API Secret)
- The Access Token and the Secret Access Token

How it works?[19]

OAuth works in steps. There are six different steps that need to be done before an application can truly communicate with another.

Step 1: The user shows his intention to connect to another application with his account (his Google+ account, Twitter account, Facebook account, etc…).

Step 2: The consumer (the third part application, here our Privacy Manager) asks for the permissions to the provider (the type of account used) and gets the different tokens access.

[18] http://oauth.net/
[19] http://blog.varonis.com/introduction-to-oauth/

Step 3: The consumer gives to the user the access token and then redirect him to the service provider. The user will have to connect to his account to authorize the consumer to gain access over some part of his account. Before login, the different needed permissions are showed so the user will know what the application wants to access.

Step 4: If the user agrees about what the consumer wants access to, he logs on to his favourite social network account by giving back the request token the consumer gives him. After the connection is done, the user is redirected back to the application.

Step 5: Now the token given by the consumer is certificated by the user and the provider. The consumer needs to exchange the token for a pair of access token which will be used to communicate with the provider.

Step 6: This is the "***final step***", at this moment, the consumer can ask the provider about anything the user granted it access to just by making requests and giving the access token.

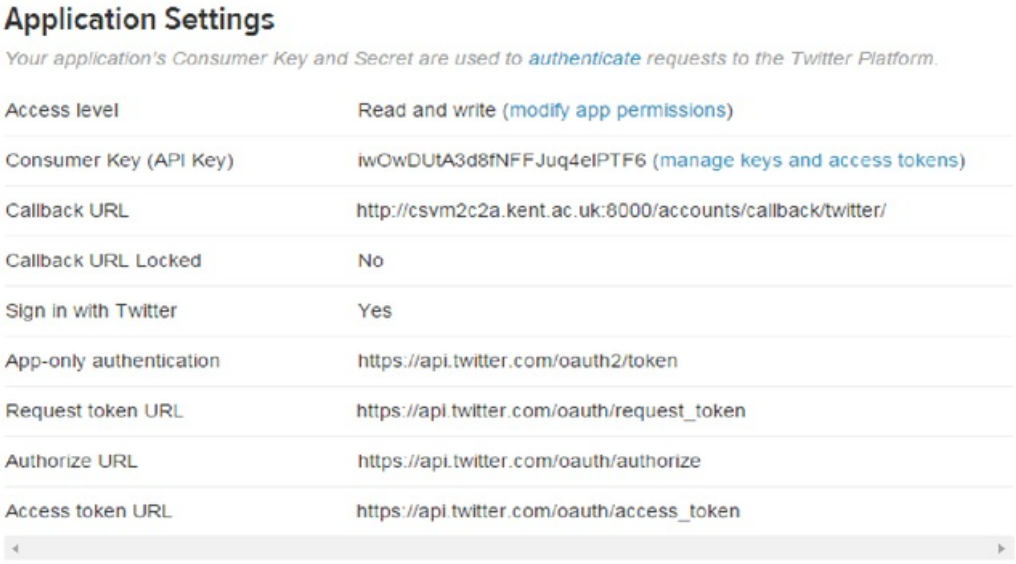
B. Development key-points

This part will be used to present the different steps of the development for the Twitter part of our Privacy Manager.

1. Connection with OAuth

Requirements

There were some requirements needed by Twitter to be granted access on it. The first one was to register the application on the dev part. I provided them some minor information about it and then I had all the links and tokens to create the OAuth connection.



Application Settings

Your application's Consumer Key and Secret are used to authenticate requests to the Twitter Platform.

Access level	Read and write (modify app permissions)
Consumer Key (API Key)	iwOwDUtA3d8fNFFJuq4elPTF6 (manage keys and access tokens)
Callback URL	http://csvm2c2a.kent.ac.uk:8000/accounts/callback/twitter/
Callback URL Locked	No
Sign in with Twitter	Yes
App-only authentication	https://api.twitter.com/oauth2/token
Request token URL	https://api.twitter.com/oauth/request_token
Authorize URL	https://api.twitter.com/oauth/authorize
Access token URL	https://api.twitter.com/oauth/access_token

The important things are:

- Access Level: be able to get (read) and send (write) data to Twitter, for what I intended to do: I wanted to get the privacy settings and then edit them from out web interface.
- Consumer Key: the key used by the consumer to link itself with Twitter by the intermediate of the user and OAuth.
- Callback URL: the link where Twitter will redirect the different token after the user has log on with his account.

These were the requirements needed to do before the real connection.

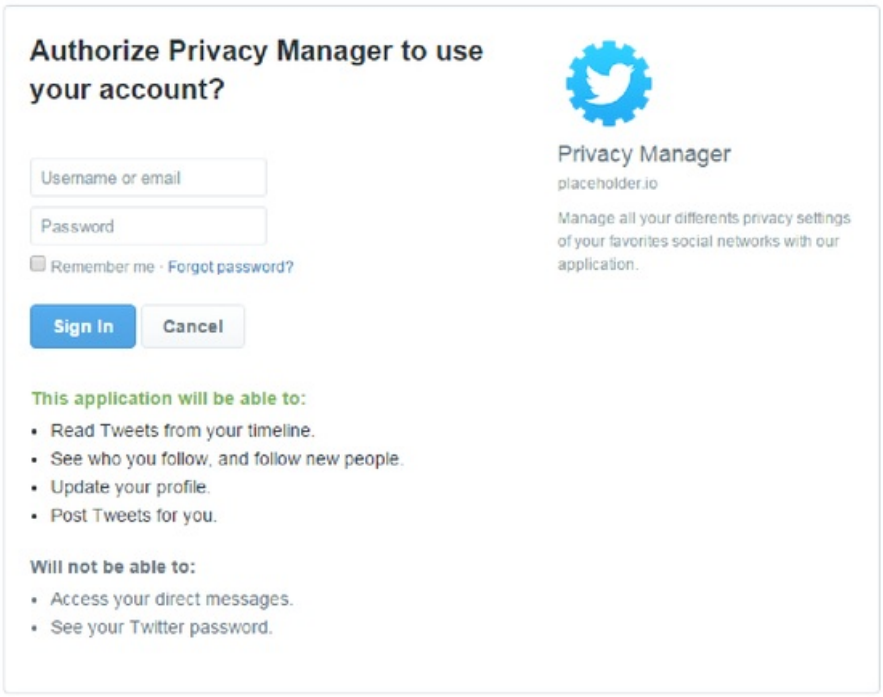
Connection

After all the access from Twitter, I needed to integrate the OAuth connection. I tried different Django module such as Twython Django[20], Django Twitter OAuth[21] and Django All Access[22]. I discussed with Maxence about this part because he was working on the Facebook part. We agreed to choose *Twitter All Access* because it was the only to propose different social network

[20] https://github.com/ryanmcgrath/twython-django
[21] https://github.com/henriklied/django-twitter-oauth
[22] https://github.com/mlavin/django-all-access

connections. Our application will be lightened by using only one module rather than two. I will not explain more about OAuth knowing that I already explain what it is and how it works previously.

Results



The user can now use his Twitter account to log in our *Privacy Manager*. This allowed me to interact with the API and get information about the connected user.

2. Using the API

Class

Firstly I needed to build a class to do the API calls and the basic verification. Before sending the request, I verified if the user is connected with his Twitter account on the Privacy Manager. Then, I get the access token needed by each call done. Each call is done the same way and need parameters:

- The method used GET (getting the data from Twitter to our web application) or POST (posting the new data set by the user to Twitter) for the call.

- The URL for the API call, if some options are needed (more or less data inside the JSON response) they can be added at the end of the URL.
- The access token, the one certificated by the user and the provider when log in the application with the Twitter account.

After the class was created, I started to get the data from my personal account to use them and see what I get and what I could do with them.

Fetching data

After using the API call written in the class before, I obtained those JSON data for the `account/settings` of my Twitter account. I will now describe what the important elements are inside and what they are related to.


```
{
    "allow_dm_groups_from": "following",
    "use_cookie_personalization": true,
    "language": "en-gb",
    "display_sensitive_media": false,
    "time_zone": {
        "utc_offset": -25200,
        "name": "Pacific Time (US & Canada)",
        "tzinfo_name": "America/Los_Angeles"
    },
    "discoverable_by_email": false,
    "always_use_https": true,
    "protected": true,
    "discoverable_by_mobile_phone": false,
    "allow_contributor_request": "following",
    "geo_enabled": false,
    "allow_dms_from": "following",
    "sleep_time": {
        "start_time": 0,
        "enabled": true,
        "end_time": 8
    },
    "screen_name": "xxxxx"
}
```


JSON data get from `account/settings`[23]

[23] https://dev.twitter.com/rest/reference/get/account/settings

Parameter	Related to
discoverable_by_email	Allow people to search for you using your email address.
protected	Allow only people who you approve to see your tweets.
discoverable_by_mobile_phone	Allow people to search for you using your mobile number.
allow_contributor_request	Depending your settings, you can be added to groups or not.
geo_enabled	When you tweet something, post where you were writing this tweet or not.
allow_dms_from	Activate / Deactivate the possibility that anyone can sent you a message.
sleep_time	Set the hours when you do now want to receive notifications on your mobile device.

Explanation of the parameters

Conclusion

I received all the parameters linked with the privacy settings (but none of the security settings were available from the API), I could try to edit them to see if what I intended to do was possible. It appeared the data were read-only because when I tried to change them from our web application I only obtained a *403 - Forbidden* error. I needed to find an alternative from the original goal.

3. Finding an alternative

Full Settings alternative

The first would have been to stop using the API and get all the parameter that were available on Twitter. Privacy as well as security would have appeared on our application but, without the values set by the user. It will just present to the user the different possibilities he has and explained in which way he can

protect himself. The OAuth connection is not needed anymore because there is not point of having some parameters with the actual settings of the user and the rest with random value (all the parameters which are not available from the API). The goal would be to recreate all the settings part of Twitter.

API alternative

The other solution is to continue with what I got from the API calls and create something to display each parameters with their value and explain in what way they are involved in the user's privacy. I will use only what I can fetch from the API and display them on our application. In this way, the user will be able to see what a third part application can gain access on and what it is possible to do.

Conclusion

In the end, I chose to keep what I had done so far because I thought it would be more relevant to show to the users the power they are allowing to a third application and the possible risks. Moreover, a help will be displayed for each setting to give information to the use.

4. Front end

The front part was quickly done thanks to Django and its tools (I only used HTML / CSS and very little JavaScript to do it). It is a simple page used to display the bases of the user's profile and his privacy settings. Each one of them appears with its own value (even if changing them will do nothing because the API does not allow it) to advice the user and a tooltip is provided to let people know what it is used for. A short explanation of why I have done this part and why I thought it was important is provided in the other page on the website.

The architecture of the main page of my part is simply structured:

- Before the connection: the user will see a simple link to connect his Twitter account to our application (a verification of connection is done obviously). He is sent to twitter to allow our Privacy Manager to be grant access on is account then redirected to the home page.

- After the connection: the same verification is done but this time, it is validated. Then, the content changes to let the user see his profile (basic) and all accessible parameters concerning his privacy and their own tooltip.

VI. Final product

A. Features

Our final version of the project provides the following functionalities:

- OAuth connection with Twitter
  - Rapid summary of the user's profile
  - The different privacy settings available
  - Tooltips about the purpose of each setting
- OAuth connection with Facebook
  - Data about the different media and their privacy
    - Posts
    - Videos
    - Albums
- A searching tool using your name and last name to find any relevant information about yourself over different search engine (Google, Yahoo).

B. Differences from primary objectives

At the beginning of the project, one of major objective was to provide a web interface where people could change all their privacy settings. A central station for all of their different social networks and their associated parameters. Unfortunately, the API of Facebook and Twitter were not providing any request to modify the settings from a third party. It was both a good and bad news for us because it showed us that social networks are aware of the potential leaks of their users' data from other application. But in the other hand, it forced us to revise our expectations.

C. Targeted users

We want to sensitize the largest number of people to let them improve their knowledge about their online privacy. Nowadays, the danger is yourself because no reflection are done before putting anything online. It is a major issue in our society that can cause serious repercussion to anybody who are not careful. However we can divide the targets as two different groups:

- The *young people* (10-18yo) who are born with Internet and know it from the earliest age. They use it every day without paying attention of possible consequences of their actions. Some work needs to be done to educate them.
- The *parents* who are not so bound with the technology and Internet in general. They need to learn to protect their child from ignorance and possible prejudices them can suffer.

Moreover, it is a good thing for both young and old people to be aware of the danger Internet can represent for anyone.

VII. Analysis and discussion

A. What could have been done better

At the beginning of the project, we were thinking of different things we would do like using standards when coding in Python or doing unit tests. In the end, we code like we learned at our school, so with our standards without using the PEP 8[24] or PEP 20[25].

We could have used a versioning system like Git or SVN to protect our work and avoid losses. However, the usage of Django was truly useful when we had to merge our different parts: we just have to put our part in the appropriate folder on the project, adding the path of each part inside the launching file and set the URLs. It would have been crucial to have one of this versioning tool if we were not doing any backup ourselves but, that was not the case.

B. Critical review

Our primary goal was to create a useful application where we could set all our social networks, basically a web interface for all the privacy and security settings available. This aim was maybe too ambitious or the result of our lack of knowledge on the subject.

In the end, our project was not what we wanted to do: the platform of control we wanted to create was transformed in a simple tool to help people to understand on the Internet about what they are doing and the consequences. Nevertheless, we learned a lot about social networks, API and the protection of their users' data.

[24] https://www.python.org/dev/peps/pep-0008/
[25] https://www.python.org/dev/peps/pep-0020/

VIII. Future of the project

A. Improving existing features

The project can be improved in many ways just by making a friendlier user interface with explanation about what we want to show and what are displaying on board. We principally used the application ourselves, so we knew how it works and how to use it, but the potential user do not.

In addition to the ergonomic and the usage, a good way of improvement could be to explore in-depth what we could display using the different API / other search engines.

B. New features

A lot of improvements can be done, there is a not limited list

- Improving the searching tool by adding the possibility to search with an image too. This feature can be achieved by using the tool of Google. This feature could improve a lot the performances and relevance of the data find. In the other hand, we do not know if it is safe for the user's data because Google could keep them. Researches about the working of it are needed.
- Adding the full settings option on the Twitter part in addition to the existence one to provide way more information to users. The first one will show what you can grant to a third-party application and what they can do with it. The second one will provide a full and clear help to improve your privacy protection.
- Adding new social networks to the panel like the ones I present in the **V.A.1** section of this dissertation. Some of them will need a lot of time to do something useful (ex: Instagram).
- Doing researches about all existing social networks (even if they do not provide any API) and list all of them, saying what their purpose is, and what kind of information can be exposed.

- Doing a ***prevention*** part: we could present articles, videos or images of people who have been victim of their own ignorance to sensitize people.
- A totally new section can be added to the project which is another part of prevention, but this time about all the methods people can use to swindle others (Fake identity, account stealing…).
- A simple page to present the importance of a good password on social networks (and any other type of website) and how to create a resistant one.

This project can be improved in many ways, the only limitation is what we are thinking at the instant because of the constant evolution of Internet.

IX. Conclusion

The realization of this project was for me an excellent opportunity in term of experience and learning. Moreover, I think working in group is always better to share ideas and opinions and it is closest thing of work in a company.

I was able to learn about a language (Python) I only knew from what I read about. The possibility of using Django in addition to Python was a major plus because of its notoriety and popularity nowadays: python is a powerful script language and with Django it becomes even more powerful (in the field of web development). Furthermore, I discover how API are built and how to use them and on top of that how OAuth works which could be really useful depending on what I will work in the future.

With the improvements I gave previously and mainly the new features, I think this project can become a good tool for anyone.

X. Appendices

Bibliography

• Lee Humphreys, Phillipa Gill, Balachander Krishnamurthy. 2010. *How much is too much? Privacy issues on Twitter*. [ONLINE] Available at: http://www2.research.att.com/~bala/papers/ica10.pdf. [Accessed 28 August 15].

• Huina Mao, Xin Shuai, Apu Kapadia. 2011. *Loose Tweets: An Analysis of Privacy Leaks on Twitter*. [ONLINE] Available at: https://www.cs.indiana.edu/~kapadia/papers/loosetweets-wpes11.pdf. [Accessed 28 August 15].