6 Jun 2023

# Crowdsourcing and AI

— UNIVERSITÉ DE GENÈVE —

---

## Final Project:

# Where2Go.CH

---


*Authors:*

- Patrick ARLT
- Jose Daniel AGUILERA
- Mohsen HASSAN NAEINI


Github repo*:* https://github.com/Mohsen-HN/Where2Go


*Teaching group:*

Prof. François Grey,
Amudha Ravi Shankar


19th of June, 2023

UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES
Département d'informatique

# TABLE OF CONTENTS:

# 1. Short Summary

Traveling and tourism are getting more and more attention lately. People are traveling a lot and with the uprising of social media, they are often sharing their trip experiences and pictures. This allows other users to discover maybe new places and sometimes even motivates them to plan a trip to those locations. Unfortunately, to do that, we have to know where those pictures are located.

On Instagram, you can add location tags and captions to your posts. Usually, people use location tags to indicate where those pictures were taken, but sometimes they use other ones such as fun ones, or totally wrong location. They even sometimes use multiple location hashtags to attract more people to see their posts. Thus the location that has been tagged isn't always correct and might fool people. This is where our idea came from: being able to validate the geolocalisation of media posted on social media.

**Key words:** geolocalisation, social media, traveling, tourism, Switzerland, crowdsourcing and AI.

# 2. Introduction

## 2.1. Choice of the project

After our group was constituted, we were assigned the topic: "Geolocalisation on social media data". The aim is to be able to geolocate data taken from social media. We then had to find a subject on which we could work. The first one we found was about geolocating images from Twitter and Instagram of well-known football stadiums in Europe. Unfortunately, this idea was out of scope as it was more about the classification of which stadium it is rather than a real geolocalisation problem. The second idea we came up with was related to tourism activities. Nowadays, with the increase in social media usage and better recommendation algorithms, people who like traveling often see the content of beautiful places to travel to. Unfortunately, users that post those don't always include the location, location tags, or hashtags. Some users even post a picture of a place and label it with plenty of hashtags to attract people to see their post, but only some indicate the correct location.

This is where our idea came from: we would like to build a tool that can geolocate, or at least geolocate the city where the picture has been taken so that people can visit that place. At first, we thought about doing this worldwide. Unfortunately, this would need us to gather too much data. Then we thought about using well-known monuments, but this solution might be a bit too easy. Finally, we decided to restrain ourselves to Switzerland's big cities such as:

- Geneva,

- Lausanne,
- Zurich,
- Basel,
- Bern,
- Lucerne,
- Lugano,
- Zermatt

This restriction seemed more simple for the task we would like to achieve. Moreover, as we are located in Switzerland, this will be easier to motivate Swiss people to participate and this will also provide us with better results, as Swiss people generally know well those cities and might not know places from all around the world or even all around Europe.

# 2.2. Related work

Regarding the usage of the crowd and the collaboration of people to geolocate places, buildings, monuments, etc there exists some online projects that are similar to ours. Still, there are some differences: for instance, they demonstrate the utilization of user participation and validation to enhance geolocation information, maps, or location-based services. While our project focuses specifically on validating geolocation data associated with media shared on social media platforms.

1. GeoGuessr: GeoGuessr is a popular online game that uses Google Street View images to challenge players to guess the location of various places around the world. It relies on visual cues and geographical knowledge to validate the accuracy of players' guesses.
   https://www.geoguessr.com/es

2. OpenStreetMap (OSM): OpenStreetMap is a collaborative mapping project where users contribute to creating a free and editable map of the world. It involves a community-driven approach to validate and update location data, including points of interest, roads, and landmarks.
   https://www.openstreetmap.org/#map=8/46.825/8.224

3. Mapillary: Mapillary is a platform that uses crowdsourced street-level imagery to create a global visual representation of the world. Users contribute by capturing images using their smartphones or other devices, helping to verify and update geolocation information.
   https://www.mapillary.com/?locale=es_ES

4. Google Local Guides: Google Local Guides is a community-driven program where individuals contribute information and review about places they visit. They validate and enhance location data by sharing details, ratings, photos, and reviews of various establishments and attractions.
   https://maps.google.com/localguides/

5. Humanitarian OpenStreetMap Team (HOT): HOT is a project that harnesses the power of volunteers to map areas affected by humanitarian crises or natural disasters. It involves validating and updating geolocation data to assist humanitarian organizations in their relief efforts.
   https://www.hotosm.org

# 2.3. State of the art

### ❖ Estimating Geographic Information from a Single Image: [1]

This paper is of great importance as it addresses the challenging task of estimating geographic information directly from a single image. It is related to our project since we take data from an image but with a different approach. The way in which they achieve geolocation is by developing innovative algorithms and techniques, which enables the extraction of valuable geospatial metadata, such as geolocation and orientation, from images.

The method they use combines visual content analysis, image retrieval, scene matching, and spatial consistency scoring, the method aims to estimate the geographic location of a single image, even when no explicit geolocation metadata is available.

### ❖ Geolocating Images with Crowdsourcing and Diagramming: [2]

This paper's significance lies in its exploration of the potential of crowdsourcing and diagramming techniques for geolocating images. By leveraging the collective intelligence of a crowd and utilizing diagramming tools, it offers an alternative approach to address the challenge of accurately geolocating images. This work not only demonstrates the effectiveness of combining human input with automated methods but also highlights the importance of collaborative efforts in enhancing geolocation accuracy and expanding the scope of geospatial data analysis.

For the crowd part they use the same logic as us, asking people to identify the image, the possible answers are “yes/maybe or no”.

### ❖ Camera Geolocation From Mountain Images: [3]

We found this paper on camera geolocation from mountain images tackles the specific challenge of estimating camera positions from mountainous terrain images, which often lack obvious visual cues for geolocation. By developing advanced algorithms and leveraging features unique to mountain landscapes, this research enables precise camera geolocation, facilitating applications such as landscape analysis, environmental monitoring, and tourism planning in mountainous regions.

The approach involves identifying distinctive features in the images, such as peaks and ridge-lines, which serve as reference points. These features are then extracted and matched across multiple images using computer vision algorithms.

- **Large Scale Visual Geo-Localization of Images in Mountainous Terrain:** [4]

This paper presents a promising approach for our work, as it focuses on recognising pictures of landscapes with mountains, which aligns with our project's objective in Switzerland. The authors specifically target mountainous terrain and propose a method that utilises the visible horizon or skyline in mountainous images to estimate the camera's 2D position and viewing direction. Their approach involves constructing a robust visual database of contour words, employing a sky segmentation technique to handle potential haze, and implementing an inverted file system for efficient search and verification. By leveraging these techniques, the method demonstrates potential for accurate recognition and geolocation of mountainous landscape images.

# 3. Methodology

## 3.1. Creating the dataset

As said in the introduction, the goal of this project is to geolocate data extracted from social media. At first, we thought about using the tool called *Visualcit* (http://visualcit.polimi.it:7778/) which can be used to automatically build a dataset based on images taken from Twitter, Reddit, and Flickr given certain keywords and filters.

We thus wanted to extract a dataset from Twitter using this tool. Unfortunately, this solution became paying for Twitter but not for the 2 other sources. Anyway, these two socials are not well known and much less used compared to Twitter or Instagram for instance. Also, people tend to post more often pictures of travels on Instagram compared to Twitter as this social is designed to share photos only. Thus we decided to use Instagram as our source of images to build our dataset.

We then had to decide how we would extract data from Instagram. We came up with the following solutions:

1. Ask the crowd to provide them to us
2. Find a pre-build dataset
3. Find a tool that allows us to extract them for us
4. Write a script that does that for us given some location tags or hashtags
5. Build the dataset by hand

The first solution didn't motivate us, as we would like to involve the crowd in later steps of the project.

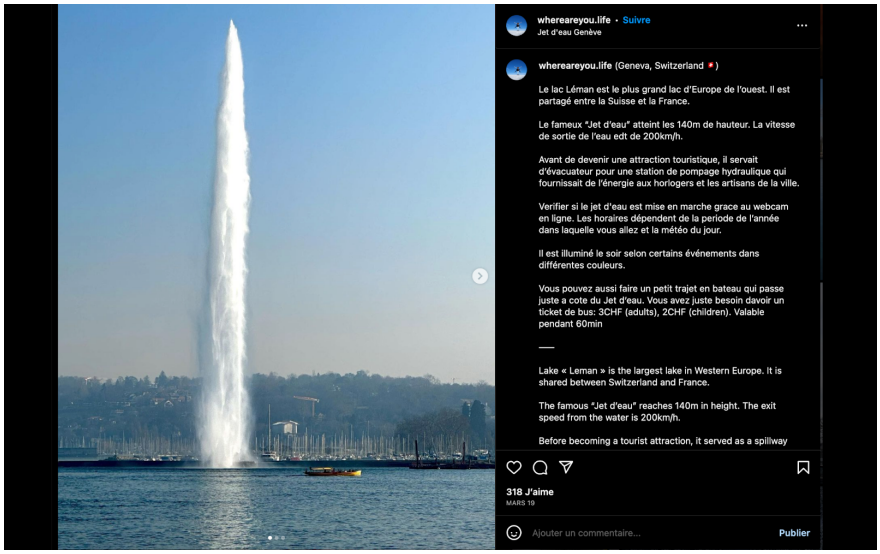
The second one seemed interesting. After some research, we find 3 pre-built datasets. Still, unfortunately, neither of them suited us: either they did not contain images, either they didn't include location tags or either they didn't include hashtags.

Thanks to those research, we put our hands on a Python library called Instaloader which allows its user to extract Instagram images and videos along with their captions and other metadata based on specific criteria. It will enable downloading Instagram images using:

- an account name,
- a location tag,
- a hashtag,
- etc, …

To create our first noisy dataset, we created a Python script that uses this library and which allowed us to gather images of each city listed above, according to location tags but also traveling hashtags such as “genevatravels”. We call it noisy because there are a lot of pictures that contain commercial advertising or totally wrong content.

The dataset’s images are as shown in the example bellow:




# 3.2. Validate the data using the crowd

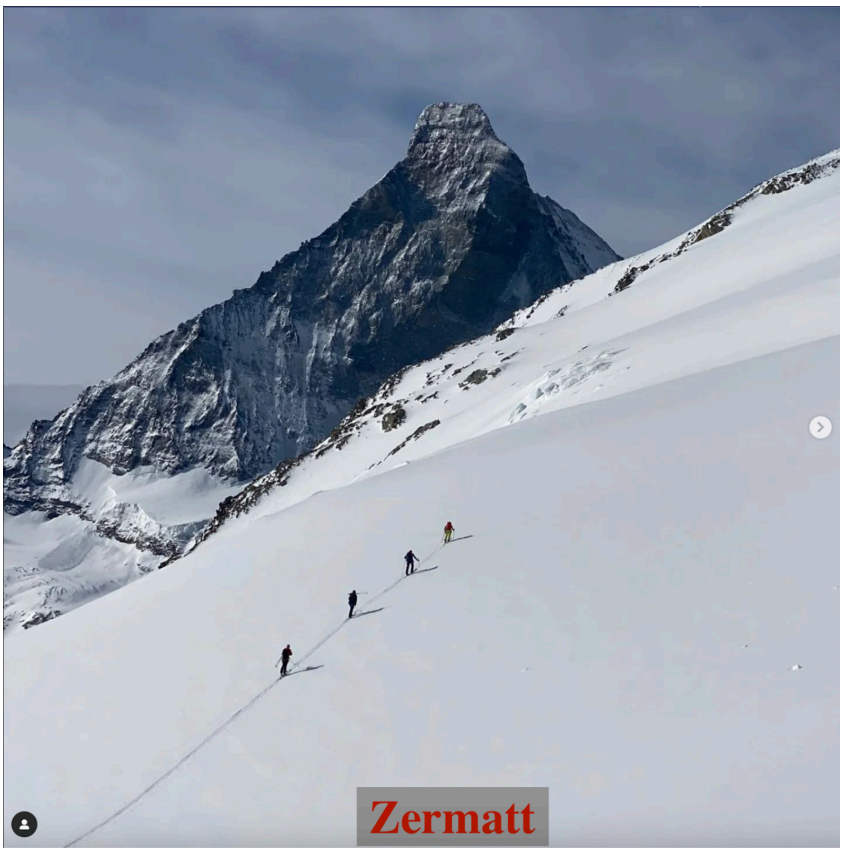
Now that we have our clean dataset, we had to look for a tool we could use to ask the crowd to validate the location of the pictures taken. We directly thought about CS Builder website (https://lab.citizenscience.ch/) that allows us to create tasks regarding images.

We first tested CS Builder by creating a test project with only one simple task: ask the crowd if “the image is indeed located in Geneva”, as we were only using the images of Geneva for the test.

For the final version, as the name of the city figures in the question, we will need to customise each question according to the name of the city in which the picture as been taken. And as we will upload all the images from all the cities of the dataset on CS Builder, then it will be shuffled. We thus figured out it would be complicated to do that customisation of the questions.

We thought about creating multiple CS Builder projects, according to each city, and uploading only the pictures related to this city on the project. Unfortunately, this would be really inefficient and redundant.

The solution we came up with was to create only one CS Builder project but change the image we would give as input: add in the image, the supposed location tag like so:




This allows us to ask a single question to the crowd:

“Is the city name at the bottom of the image the correct location of the photo?”

Then, this question will have the following options to answer with:

- Yes, it is the correct city
- No
- Not sure (I don't know)

We also thought about adding two other questions for each image on the CS Builder project:

- if someone selects "No" for the first question, then a new question appears and ask if the person knows where this image has been pictured.
- Then, whatever has been answered to the first question, we ask people if they know the exact location of the picture. To do so, people have access to a map on which they can indicate the precise coordinates of the picture's location.

Thus, by using the crowd to validate our data, we will be able to create a clean dataset of images, with their correct location. We will then be able to use this dataset to train and test AI models.

# 3.3. IA Crowd part

Once the CS Builder was created and ready for contribution, we had to create an AI crowd project. AICrowd is a collaborative platform that allows users to propose project ideas and use this collaborative hub to ask for help from the crowd to generate an AI solution to their problem. Thus this website regroups a lot of artificial intelligence and machine learning passionate people and provides challenges for them.

In our case, we first created an AICrowd project on their platform. Unfortunately, the project had to be validated which is a step that might take a lot of time. Also, as we simply need to create a temporary project, It might not get published easily.

That is why in the meantime, we created a GitHub page which is basically a description of our AI Crowd project and which regroups all the information we had included in the IACrowd project itself. You will find more information about the AI Crowd project in section 4.2. AI Crowd deployment.

# 4. Results

## 4.1. Results from CS builder

For the validation of our data it took us about 2 weeks two find five participants who agreed to contribute to the project. At first we had high hopes to get the results we need but it took us more time and two more person to contribute. The reason why we needed five people was because we set the redundancy of each task to five. We encountered a small issue, season for why we had to search for two addition contributors, where 2 of the five users had not completely do all the 144 tasks and did partially the work. Therefore, two additional people were needed to complete validation of all 144 tasks and for us to have full statistics.

The results of all tasks can be exported as a CSV fill. Inside this file we can find the result for each task from each user. For each row, meaning each task, we have the results of the 3 questions that we have mentioned in *section 3.2*. For the first question, *"Is the city name at the bottom of the image the correct location of the photo?"*, approximately 71% of the task/images were answered with *"Yes, it is the correct city"*, 21% were answered with *"Not sure"* and only 8% were answered with *"No"*. If the answer to the first question was *"No"*, the user gets the optional question where they can type what they might think the correct city where the image was taken from is. Out of the 8% answered with *"No"*, ⅔ of them were followed up with the correct city name. And ⅓ where left empty because the user might know the images is not from that labeled city but at the same time not know where the correct city might be.

For the third question we ask an optional question where users can point where they might think the exact location where the image was taken. Not that many tasks had this question answered. One interesting fact that was observed was that we can observe some user were more skilled at exactly geolocating on the map the location of the image while some other user barely answered this questions for most tasks.

We also feel that the top 3 cities that had the a lot of certainty of their location, both for first question and the map question, were Lucerne, Geneva and Bern, and then followed by Basel. This is slightly unexpected result since Zurich is the largest city in Switzerland but yet many of our contributors were not sure of their answers for this city. This issue can either be from the contributor lack of knowledge for this city be or maybe even an issue with taggings from social media because a lot of user might mention Zurich when they want to tag Switzerland in general and therefore create a more confusing/false dataset for this city. This is one of the reason why we use the crowd to remove these false data and validate the correct ones.

# 4.2. AI Crowd deployment

As said earlier, AI Crowd is a collaborative platform that allows users to propose project ideas and use this collaborative hub to ask for help from the crowd to generate an AI solution to their problem. This allows using the diversity of knowledge among machine learning experts and passionate from all around the world. This allows us to gather a large number of solutions with different approaches, methods, and ideas on how to resolve this problem.

In our case, we need to create a machine-learning model that is able to predict the name of the Swiss city in which an image (extracted from Instagram) has been taken. Participants are encouraged to explore state-of-the-art deep learning models, such as advanced CNN architectures, attention mechanisms, and transfer learning, to enhance geolocalization accuracy.

In the scope of this project, we will simply create a mock AI Crowd project. If we wanted to really create an AI Crowd project, we would create a much more dense dataset that we will split into two subsets: one training set which will be shared with all the participants so they can train their network on those data, and another on which will be our private test set on which we will evaluate all the propositions made. As this second dataset will be private to us, the model proposed will not know those images and thus, we will be able to evaluate and rank the best models in terms of accuracy.

Also, note that the dataset will be the one validated by the crowd on our CS Builder project.

Of course, we will ask participants to share their trained model, their code, and any necessary dependencies or instructions for reproducibility. Additionally, participants are encouraged to submit a brief description of their approach, highlighting the key techniques and methodologies used in their AI model. This provides valuable insights into their submission and encourages knowledge sharing among participants.

For the best solution proposed, we will provide some prizes: the first place will be rewarded with a three-day getaway to their chosen city in Switzerland, all expenses covered. For the other participant being between the second and fifth position, they will be rewarded with cash rewards, recognition on the AIcrowd platform, and potential collaboration opportunities with our project team.

As said earlier, in the scope of this project, we will simply create a GitHub page that will contain the whole description of the project. It will contain all the information we would have to include in a real AI Crowd project. You will find the GitHub page on this link: https://mohsen-hn.github.io/Where2Go/

# 5. Summary and Outlook

In this project, we explore ways to use the crowd to solve a problem: geolocate images uploaded on social media.

The first step was to find a subject. We came up with the idea of geolocating tourism pictures from Instagram to enhance and foster touristic activities and visits. We then had to figure out where the crowd would be useful for us.

Social media like Instagram allows users to include location tags and hashtags in their posts. Unfortunately, users don't always indicate the correct location and even sometimes don't indicate the location at all. So first, to build a clean dataset that could be used for an AI task, we will ask the crowd to validate the location of some images extracted from Instagram. Thus they have to confirm the location tag used by the user so that we are sure about the location where it has been taken. To do so we used the CS Builder website that specifically allows us to do this validation task and share it with the crowd so they can contribute and help us build this dataset.

Then, we created a mock AI Crowd project where we ask the crowd to come up with innovative IA models which, given the validated dataset, will be able to predict the location of the picture. We will provide the crowd with a restrained part of the dataset such that they can train their model on it and then we will evaluate all the solutions proposed on a test set that will be kept private and unknown to any model.

As the AI Crowd part is fictive in this project, an outlook for this work would be to first extract a larger number of images such that the dataset is large enough to be able to train correctly a model on it.

Another improvement would be then to extend the dataset to neighbouring countries like France, Germany, Italy, etc, and little by little extend the number of countries and cities known by the model.

Finally, we would try to gather as many people as possible to help us develop this neural network model, by first asking our classmates to help us. Then we could organise university events to gather people to help us, in the same way as a hackathon. Finally, If the project is becoming really concrete, we would ask the university if we could have some kind of subvention so we can propose a cash price for the winner of the AI Crowd project.

Also, once we obtain an accurate neural network model that performs accurate predictions, we could for example create a google chrome extension that would allow users to directly predict where an image has been taken without using anything else than google chrome.

6 Jun 2023

# Who did what in the team:

| Photos | Names | Roles and distributions: |
|---|---|---|
|  | Mohsen HASSAN NAEINI | Role: Organiser<br><br>Distribution: Report, Data Extraction, Creating CS-builder (mainly by Mohsen), CS-builder contribution management, Github pages, presentation. |
|  | Patrick ARLT | Role: Documenter<br><br>Distribution: Report (mainly by Patrick), Data Extraction, Creating CS-builder, ai-crowd project deployment on AI-Crowd, Github pages, SDGInnProgress documentation. |
|  | Jose Daniel AGUILERA | Role: Presenter<br><br>Distribution: Report, Data Extraction, stat-of-the-art and research for related works, presentation. |

# References

[1] HAYS, James et EFROS, Alexei A. IM2GPS: estimating geographic information from a single image. In : *2008 ieee conference on computer vision and pattern recognition*. IEEE, 2008. p. 1-8.

[2] KOHLER, Rachel, PURVIANCE, John, et LUTHER, Kurt. Geolocating images with crowdsourcing and diagramming. In : *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI 2018)*. 2018.

[3] CHEN, Yi, QIAN, Gang, GUNDA, Kiran, et al. Camera geolocation from mountain images. In : 2015 18th International Conference on Information Fusion (Fusion). IEEE, 2015. p. 1587-1596.

[4] BAATZ, Georges, SAURER, Olivier, KÖSER, Kevin, et al. Large scale visual geo-localization of images in mountainous terrain. In : Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part II 12. Springer Berlin Heidelberg, 2012. p. 517-530.