# Annotation Guidelines

## Dataset information

We provide one dataset of tweets in Dutch from Brussels between 2015 to 2019.

The dataset will be available on the Tagtog annotation platform.

## Annotation tasks

1. Identify whether a tweet is traffic related or non-related

The first task is to identify whether a given tweet is traffic related or non-related.

**Label definitions**:

1. **Traffic related**: This type of tweets reports **non-recurring events** and **traffic flow conditions**.

**Non-recurring events** are the events that generate an abnormal increase in traffic demand or reduce transportation infrastructure capacity. The examples of non-recurring events include traffic crashes, disabled vehicles, highway maintenance, work zones, road closure, vehicle fire, traffic signal problems, special events, and abandoned vehicles.

**Traffic flow conditions** includes events such as daily rush hours, traffic congestion, traffic delays due to high traffic volume and jammed traffic. Also, any tweets that disseminate new traffic rules, traffic advisory, and any other information on transport infrastructures.

1. **Non-Traffic related**: Any tweet that does not fall into the Traffic related category should be labeled as non-traffic related.

You should manually annotate each tweet. By label definitions, if a tweet (you can evaluate the translation of the original tweet text) is identified as traffic related, you label the tweet as **true**, otherwise label **false**.

1. If a tweet is identified as traffic related, extract relevant information from it.

The second task is to get more information from the traffic related tweets. Slot filling problem aims to identify relative text spans that can answer predesigned slot filling questions.

Here are three potential slot filling questions for the traffic related tweets:

**What** happened during the reported traffic related tweets?

**Where** did the reported traffic related tweets happen?

**When** did the reported traffic related tweets happen?

What is the **CONSEQUENCE** of the reported traffic event?

Try to answer these four questions using the text spans from the tweet, highlight the identified text spans.

