﻿

# Dataset explanation

**Excel document:** findm-annotations-of-images

**Location:** Proj-FINDM > Data > Raw > image-search > pfizer\_20210223\_dataset

We have the following columns:

* ﻿**Phone Number Present:**

Yes/No

* **Medicine:**

Yes/No

* ﻿**Other English PII present:**

Yes/No

* ﻿**Foreign English PII possibly present:**

Yes/No

* ﻿**Logo:**

Yes/No

* **﻿ Phone Number:**

If present the number string as it appears in the image/ NA if not present

* ﻿ **Whatsapp:**

Yes/NA

* ﻿ **Tencent QQ:**

Yes/NA

* ﻿ **DingTalk:**

<username> /NA

* ﻿ **Line In:**

<username> /NA

* **﻿ Skype:**

<username> /NA

* ﻿ **Wickr:**

<username> /NA

* ﻿ **Threema:**

<username> /NA

* ﻿ **Wechat:**

<phone number> /NA

* ﻿ **Email:**

<email> /NA

* ﻿ **Webxloo:**

<username> /NA

**Content:**

**NA:** It is not present on image

**<username>:** It was present on image and a string representing the username was also present

**<email>:** It was present on image

**Whatsapp Yes:** The logo or the word “Whatsapp” were present on image

**Medicine Yes:** Images that have medicines in them

**Other English Pii present Yes:** Images that have usernames or references to other social media and messaging platforms other than whatsapp

**Foreign English possibly present Yes:** Images that have possibly personal identifiable information, we need to analyze the text and translate it to verify that there is contact information on them.

**Explanation of social media and messaging platforms (From Wikipedia and the web)**

|  |  |  |
| --- | --- | --- |
| **Messaging Platform** | **Explanation** | **Country it was developed** |
| Tencent QQ | Also known as QQ, is an instant messaging software service and web portal developed by the Chinese tech giant Tencent. QQ offers services that provide online social games, music, shopping, microblogging, movies, and group and voice chat software. | China |
| Telegram | Telegram is a freeware, cross-platform, cloud-based instant messaging software and application service. The service also provides end-to-end encrypted video calling, VoIP, file sharing and several other features | United Kingdom / United Arad Emirates |
| DingTalk | Enterprise communication and collaboration platform developed by Alibaba Group | China |
| Line In | Freeware app for instant communications on electronic devices such as smartphones, tablet computers, and personal computers | Japan |
| Wickr | Instant messaging apps allow users to exchange end-to-end encrypted and content-expiring messages, including photos, videos, and file attachments | United States |
| Threema | Paid open-source end-to-end encrypted instant messaging application for iOS and Android | Switzerland |
| Wechat | Chinese multi-purpose messaging, social media and mobile payment app developed by Tencent. | China |
| Webxloo | Integrated communication platform built for teams | United States |