

# Annotation Guidelines and User Manual for the annotation software ‘GATE TEAMWARE’

## The task

In this task, **you are asked to use an online (web) software, read 675 social media messages and say about each one whether it is true or untrue and if it is untrue - whether you think it may contain “disinformation”.** By “disinformation” we intend “**false or misleading content that is spread with an intention to deceive** or secure economic or political gain and which may cause public harm”, as defined in the European Commission's Communication on the European Democracy Action Plan<sup>1</sup>.

You have to state about each message whether it is true or untrue and whether it contains disinformation or not **on the basis of your professional experience and/or by using your professional methods to check statements.**

In the fields of Computational Linguistics and Natural Language Processing, this kind of task is called “annotation” or “annotating”. We will call it like this from now on.

You will be provided instructions on how to use the annotation software and how to annotate. Such instructions (called “Annotation Guidelines”) are provided in this document, as well as in the annotation software.

## Anonymity

**Your participation is anonymous, as are the decisions you make about each social media message. Only the organiser knows who the participants are, so act freely.**

## The time for making decisions is recorded

**Please note:** The annotation software automatically records the time you take to make your decisions. This is necessary and important for the research that we are conducting. As it is clear that you will not complete the annotation of all 675 social media messages at once, **please make sure to either close the browser window in which you are annotating or log out of**

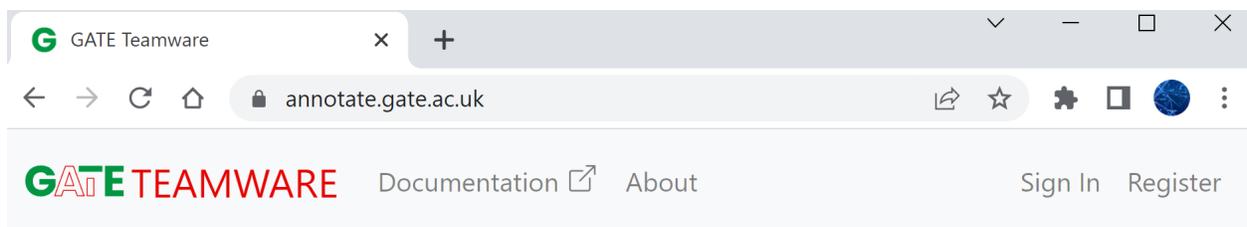
---

<sup>1</sup> <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020DC0790&from=EN>.

**the annotation software if you decide to interrupt your annotation.** This will ensure that the tasks you have completed up to that point are saved, the time recording is paused, and it will restart the next time you log in and press "**Annotate**". This will allow you to continue from the social media message at which you interrupted your annotation.

## Access to the software

The software is accessible online (you need Internet access). To use it, you need to open a browser on your desktop computer, laptop, or phone. We recommend using a device with a larger screen. To access the software, please copy or click on the following link: <https://annotate.gate.ac.uk>.



**Short instructions for working with the annotation software:**

## Logging into the software

Enter the username and password we have created specifically for you by clicking on "**Sign In**":

## Sign In

Username

Password

This is what the homepage looks like:



## How to start annotating

To see the social media messages which need to be annotated, you have to click on **“Annotate”**.

By **“annotating”** (and **“annotation”**) we mean adding a category or additional information to a given text. In this case, you need to indicate whether a social media message is true or untrue and, if it is untrue, **whether it possibly contains intentionally spread disinformation**. This is

done by selecting from the specified categories and possibly adding text in a text field. Shortly, you will see how this is done.

## Where are the Annotation Guidelines

After clicking on **Annotate**, you will see the Annotation Guidelines:

## Annotator guideline

### Следвайте следните правила при аотиране:

1. Вземете решение дали съобщението е вярно или не, или съдържа ли целенасочена дезинформация, или не, на базата на Вашия опит, обща представа за нещата, или проверете, ако съобщението съдържа линк.
  - Когато съобщението съдържа достатъчно информация и линк, игнорирайте линка.
  - Когато самото съобщение не дава достатъчно контекст, но има линк, моля отворете линка и проверете.
  - Ако се случи съобщението да не дава, според Вас достатъчен контекст и да не съдържа линк, проверете по друг начин, на базата на Вашия професионален опит, или поне в Google, като се стараете да взимате възможно най-информирани решения.

2. Решението дали съобщението е вярно или не **трябва да се отнася за датата, на която е било публикувано съобщението.**

Например:

**Дата:** 14.10.2020 г.

**Съобщение:** "Страх от COVID-19 отмени посрещането на Нова година в центъра на Виена"

Съобщението се отнася за посрещането на **Новата 2021 година във Виена**. Изберете дали твърдението е вярно или не като си представите/проверите дали е вярно за съответния момент, а не за други нови години.

3. Игнорирайте части на съобщенията, съдържащи бъдеще време (несъществуващ момент). Аотирайте само настоящ или минал момент.
4. Игнорирайте евентуални правописни грешки в съобщението.
5. Ако съобщението съдържа повече от едно твърдение, и част от твърденията са верни, а част - не, изберете "Частично". Ако имате време, моля копирайте в "Коментар" **частта от съобщението която е вярна**.

Например:

**Дата:** 7.10.2022 г.

**Съобщение:** "ГЕРБ-СДС получиша най-много гласове и 65 мандата."

**Частично**

Коментар: **вярно е, че "ГЕРБ-СДС получиша най-много гласове"**

6. Ако съвсем нямате никаква представа, дали цялото или част от съобщението е вярно, може да изберете "не знам". Моля, използвайте "Не знам" много рядко! Ако има повече от 15 "не знам", цялото Ви аотиране няма да бъде прието. Ако се наложи, направете проверка.
7. Когато съобщението цитира някого и:

1. **Това е конкретен човек с име** (Напр. "Бойко Борисов"), **или е неуточнен човек или хора**, като например: "Невролог: ..", "Наши учени..", "Австралийски учени", "Бивш министър..", **но можете да се сетите за кого става дума** (напр. това е известен човек/изказване):

Целта е да кажете, **дали е истина, че този човек е казал това на съответната дата**, или е целенасочена дезинформация, че го е казал, например:

**Дата:** 22.10.2021 г.

**Съобщение:**  Проф. Иван Чалъков: "Няма да се #ВАКСИНИРАМ, защото ваксина за 5 месеца не се прави!"

**Истина ли е, че Проф. Иван Чалъков е казал това? Целенасочена дезинформация ли е, че го е казал?**

**Внимание:** Ако човекът наистина е казал това твърдение, но самото твърдение на този човек представлява невярна информация или целенасочена дезинформация, тогава на първия въпрос (дали е вярно или не), кликнете на "Частично", и копирайте в първия Коментар цялото съобщение. На втория въпрос (дали съдържа дезинформация), кликнете на "Частично" и копирайте самото твърдение. Например:

**Дата:** 10.04.2022 г.

**Съобщение:** "Корнелия Нинова: Не изнасяме оръжие за Украйна."

Корнелия Нинова наистина е казала това твърдение, тъй че то съдържа частична истина. Обаче, както се разбра после, тя е излъгала, т.е. самото нейно твърдение съдържа дезинформация.

Въпрос 1 (вярно или не): Частично. Коментар 1 (коя част е вярна): "Корнелия Нинова: Не изнасяме оръжие за Украйна."

Въпрос 2 (съдържа ли дезинформация?): Частично. Коментар 2 (коя част е дезинформация): Не изнасяме оръжие за Украйна

2. **Това е неуточнен човек или хора, за които не знаете кои са:** целта е да кажете **дали казаното в цитата е вярно**. Например:

**Дата:** 29.11.2020 г.

**Съобщение:** "Невролог: Загубата на мирис е първи симптом при COVID-19, главоболието е остро"

Не знаете кой е неврологът. Истина ли е, че "Загубата на мирис е първи симптом при COVID-19, и че главоболието е остро", и ако не - целенасочена дезинформация ли е това?

8. Когато съобщението съдържа твърдение, трябва да се аотира по отношение на казаното в това твърдение, например:

**Дата:** 21.06.2022 г.

**Съобщение:**  Кабинтът "Петков" е престъпен и незабавно трябва да подаде оставка. !! НЕЗАБАВНО!

Игнорирайте правописната грешка "Кабинтът", тя идва от оригиналното съобщение.

Истина ли е, че на съответната дата кабинетът "Петков" е бил престъпен? Целенасочена дезинформация ли е твърдението, че на съответната дата кабинетът "Петков" е бил престъпен?

After reading them, you can click on “Hide”, so that they get partially hidden. They will be available during the whole annotation process and you can always click on “+Show” so that they are fully displayed.

Now click on the blue button that says: “**Start Annotating**”:

## Annotate a document



### Starting annotation

You're about to start annotating documents. Please make sure you've read the annotator guideline to make sure you know what's expected from your annotations.

Start annotating

## How to annotate social media messages

**Now you should see the first social media message** for which you need to say whether you believe it is true or not (untrue), and if it is “untrue”, whether it contains intentionally spread disinformation (i.e., whether the person who wrote the message intentionally lied and spread false information) or not (for example, the person who wrote the message simply shares information they saw somewhere, which they believe is true and do not know that it is untrue).

You should be able to see this:

## Annotate a document



**Дата:**  
12/14/2021

**Съобщение:**  
Трус със сила 7,3 по Рихтер удари Индонезия и припомни за трагедията от коледата на 2004  
г.<https://indianexpress.com/article/world/indonesia-earthquake-tsunami-warning-live-updates-7671730/>

Съобщението съдържа ли истина или не?

Да  Не  Частично  Не знам

**Коментар:**

Съобщението съдържа ли целенасочена дезинформация, или не?

Да  Не  Частично  Не знам

**Коментар:**

Each social media message includes the date on which it was published/posted on social media. The date is provided because a statement may have been (or considered) untrue at a given time but later turned out to be true or the opposite. **The decision whether the social media message is true or untrue should be made by considering whether it was true or not at the date of its publication.**

Read the date, the message, and after that, following the guidelines in "+Show", select whether the message is true or untrue by clicking on the corresponding circle. The message may be partially true, then click on "Partially" and **copy (it is faster) or rewrite manually the part of the message which is true in the first "Comment:"**.

After that, try to give your competent opinion on **whether at least part of the social media message could have been written with the purpose to deceive those who will read it. If you have selected "Yes" (contains disinformation) or "Partially", copy or rewrite manually the part of the message that you believe contains intentional disinformation into the second "Comment:" field.**

An example of such a message is: "Kornelia Ninova: We are not sending weapons to Ukraine." According to the guidelines, Kornelia Ninova did really say this, so the whole message is true, but her statement is a lie, i.e. disinformation. In this case, you have to click "Yes" on the first question, and copy the entire message into the first "Comment:" field: "Kornelia Ninova: We are not sending weapons to Ukraine." On the second question, click "Partially" and copy into the second "Comment" field only the part "Not sending weapons to Ukraine".

When you are sure that you have finished annotating this social media message, click on **"Submit"**. If you want to make a correction, such as removing all category choices, click on

"Clear" (**Attention: This will delete all your category choices and any text you have copied or written for the currently annotated social media message!**).

Please note that **you cannot go back to previous social media messages** for which you have already clicked on "Submit", so be sure that your decisions are final before clicking on "Submit".

After clicking on "**Submit**", you will see the next social media message that you need to annotate.

## How to understand how many messages are left to be annotated

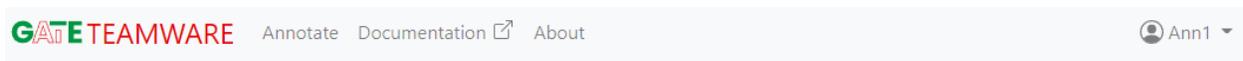
Above each social media message, on the top right, you can see the following three fields:



The first one shows the internal number of the social media message you are currently annotating (in the field in **grey**). This number is not important to you.

The next field shows the number of messages which have already been annotated (in the field in green). Finally, in the yellow field, you can see the number of messages that still need to be annotated. Since you know that you need to annotate 675 messages, this number will be decreasing and will indicate the remaining messages.

When you are done annotating all social media messages assigned to you, you will see this:



### Thank you for participating!

Thank you, all of your annotation tasks for the project have been completed!

It's now possible to participate in annotating for other projects. Inform a project manager of your username if you wish to do so.

**If you have any questions, please write to me or call me.**

**Thank you very much for your help :)**