

Annotation guidelines

This document outlines the guidelines for a manual annotation task designed to evaluate the performance of a Large Language Model (LLM) on an automatic event detection task.

Task Context

Process Mining comprises a set of techniques and methodologies that enable the analysis of business processes based on data recorded in information systems. These techniques allow for the discovery, monitoring, and improvement of real processes through the extraction of knowledge from event logs available in information systems.

Task Objective

The automatic event detection task helps enrich log files useful for other tasks such as process discovery. Your annotation will help evaluate the methodology applied in this experiment with the aim of finding relevant (activity, date) pairs that can be included in the logs.

In other words, your contribution to manual annotation will serve to create a gold standard to verify the effectiveness of automatic algorithms in detecting significant events within business processes, correctly identifying the activities performed and their corresponding timestamps.

Annotation Procedure

This section provides details about the inputs, tasks, and tools you will use for annotation.

Input

Tenders Electronic Daily (TED) contains all the notices and calls for tender published in the Supplement to the Official Journal (of the European Union)¹.

You will be provided with the following fields for each of the 50 instances/caseid:

- **caseid**: unique identifier of the TED documents
- **complementary_info**: the portion of text from the TED documents containing useful activities to be included in the logs. This textual content is what's inserted into the prompt to request the recognition of activity-date pairs
- **name-of-activity**: corresponds to the activity, or event, detected by the LLM within the complementary information

1

https://single-market-economy.ec.europa.eu/single-market/public-procurement/digital-procurement/tenders-electronic-daily_en

- **date:** contains the date corresponding to the activity, or event, detected by the LLM within the complementary information

Task

Your task is to assess the correctness of the LLM output based on your expertise.

For each activity-date pair detected, you will fill in the following fields:

- **label:** for each instance, annotate with one of the following values:
 - Ok (the event/date is the correct one **)
 - No (the event/date is wrong)
- **note:** if you deem it necessary, you can add some notes about your annotation of the instance

Annotation Tool

You will be provided with a spreadsheet containing the case IDs, the corresponding texts, the pairs (activity, date) automatically detected, and two columns: one for labeling and one for writing eventual notes.

Important Considerations

- **Ambiguity:** If you encounter any ambiguity in an LLM's output or the evaluation criteria, please document it clearly in the provided space for comments.
- **Consistency:** Maintain consistency in your annotations throughout the task. If you are unsure about a specific case, refer back to these guidelines or consult with the research team.
- **Focus on the Task:** Your focus should be solely on evaluating the correctness of the LLM outputs based on the given criteria. Do not consider other factors such as the LLM's writing style or formatting.
- **Confidentiality:** The information provided for this annotation task is confidential and should not be shared with external parties.

By carefully following these guidelines, you will contribute valuable data to this research project, ultimately helping us evaluate our experiment. We thank you for your participation in this experiment.

** Italian type of activities to be considered "correct":

- Delibere,
- Provvedimenti,
- Decreti,
- Determinazioni,
- Atti,
- Direttive,
- Ordinanza

- Comunicati

Attention:

- do not consider "Delibera sindacale" and neither "Delibera + adjective"
- do not consider the award as having become effective ("L'aggiudicazione è divenuta efficace")
- do not consider the contract has been awarded ("L'appalto è stato aggiudicato")
- do not consider if it refers to a previous procedure, e.g. "Determina n. 113/2022 di presa d'atto procedura di gara deserta"
- do not consider protocol of understanding related to collaboration relationships ("Protocollo d'intesa relativo ai rapporti di collaborazione")
- do not consider protocol of Understanding ("Protocollo di Intesa")
- do not consider in "derogation of art. 51 of D. Lgs n. 50/2016", do not extract "D. Lgs"
- do not consider the executive date
- do not consider "the procedure has been revoked in self-protection" ("è stata revocata in autotutela la procedura")
- give greater weight to "DELIBERA DI AGGIUDICAZIONE" if it's in the text at the expense of other forms (e.g., delibera a contrarre)
- if there is "appalto è indetto" do not consider it
- do not extract the date if there is "Richieste di informazioni e/o chiarimenti"
- if it finds "determinazione dirigenziale DAD" or "determinazione dirigenziale DIM" do not consider DAD and DIM in the activity name
- ignore "Data di conclusione del contratto d'appalto"
- ignore an event if "certificato" (e.g., from the phrase: "Decreto di aggiudicazione efficace n. 12997 del 10/06/2022 certificato il 01/07/2022")
- delete duplicate events