Report on the outcomes of a Short-Term Scientific Mission[1]

Action number: **CA19102**

Grantee name: Livio Robaldo – School of Law, Swansea University (UK)

Details of the STSM

Title: Fine-tuning GPT-3 for processing legal documents

Start and end date: 18/09/2023 - 22/09/2023

Host institution senior affiliated: Davide Liga, prof. Leon van der Torre

Description of the work carried out during the STSM

Description of the activities carried out during the STSM. Any deviations from the initial working plan shall also be described in this section.

This STMS took place just after Davide Liga and I received the notification that our paper in [1] was accepted for publication on Elsevier's journal Computer Law & Security review[2] (on the 4th September 2023). This paper represents the first step of our long-term research plan: combining the expertise of D. Liga in GPT with my expertise in symbolic-based LegalTech technologies, e.g., LegalDocML[3] and LegalRuleML[4]. [1] uses the LegalDocML and LegalRuleML annotations that I developed in [2] for fine-tuning GPT-3 fit to classify legal rules (obligations, permissions, and constitutive rules) in the GDPR.

In light of this, D. Liga and I spent most of the STMS to discuss future works of this paper. This include both new publications that extend the work in [1] and project proposals.

Specifically, we brainstormed new research ideas to design a new GTP-based pretrained language model (PLM) able to not only classify the legal rules, but also recognize them in text. We concluded that the LegalDocML structure could be very useful to this end, as the textual span of the legal rules are often obtained by the same structural elements (e.g., the introductory text of a paragraph) with the associated sub-elements (e.g., the points of the paragraphs). LegalDocML encloses these (sub-)elements within specific XML tags; thus, it facilitates the extraction/composition of the legal rules. Doing so in free text would be more difficult, as the PLM would be also in charge of recognizing the structure of the legal document, which would introduce a percentage of error in the final results.

In addition, we agreed that in order to facilitate the construction of fine-grained semantic representations, e.g., the LegalRuleML rules presented in [2], on which it would be possible to execute automated

[1] This report is submitted by the grantee to the Action MC for approval and for claiming payment of the awarded grant. The Grant Awarding Coordinator coordinates the evaluation of this report on behalf of the Action MC and instructs the GH for payment of the Grant.
[2] https://www.sciencedirect.com/science/article/pii/S0267364923000742
[3] https://www.oasis-open.org/committees/legaldocml
[4] https://www.oasis-open.org/committees/legalruleml

compliance checking, the extracted legal rules should be encoded in the form "it is MOD ACTION when CONDITION", where "MOD" refers to one possible deontic modality (obligation, permission, etc. see [3]). For a simple example, consider this excerpt of text taken from the legal obligations in force in the UK for drivers/riders, available online at [4]:

"Before you drive or ride you must have the correct driving licence"

This text conveys a legal obligation, but extracting the text as it is will not facilitate the construction of a corresponding logical rule that would represent its meaning and on which we could automatically infer compliance or non-compliance of the obligation. Therefore, the PLM should not only extract this obligation, but also *paraphrase* it in the form above:

"it is obligatory to have the correct driving licence when you drive or ride"

In light of this, we plan to develop a new PLM able to both extract legal rules from LegalDocML and convert them in the form above. The first step to achieve this goal is to create a parallel corpus on which fine-tuning the PLM. We are currently looking for small grants to which applying in order to obtain the resources needed to create this corpus (see next section).

Description of the STSM main achievements and planned follow-up activities

Description and assessment of whether the STSM achieved its planned goals and expected outcomes, including specific contribution to Action objective and deliverables, or publications resulting from the STSM. Agreed plans for future follow-up collaborations shall also be described in this section.

The STSM fully achieved the "Expected outputs and contribution to the Action MoU objectives and deliverables" stated in the STSM application: (1) starting new joint research and (2) brainstorming ideas to prepare research grant applications.

Concerning (1), D. Liga and I started a new set of experiments aiming at improving the accuracy of the PLM presented in [1]. We have also started writing a new paper that we will try to submit to LREC-COLING conference (https://lrec-coling-2024.org). Alternatively, we will submit it to the next editions of Jurix, Jurisin, or ICAIL. In addition, we started examining similar tasks in LegalBench[5], a new collaborative effort to build and distribute open access benchmark for legal reasoning through large language models. The benchmark currently consists of 162 tasks gathered from 40 contributors, and the research centre of Mishcon De Reya (MDR) has now decided to use in place of the RAIN dataset, which is no longer available. An ambition of D. Liga and I is to contribute to the LegalBench initiative thus creating new tasks based on our future research.

Concerning (2), as explained above we are now looking for resources to create a parallel corpus of extracted & paraphrased legal obligations, as explained in the previous section. To do so, we need to hire annotators, whim I plan to recruit through student internships for the students at the School of Law of Swansea University, in which I am affiliated. Therefore, I am now started writing grant applications, in which D. Liga acts as external collaborator, for gathering the resources to support the student internships.

Currently available calls to which I will apply are:

- Swansea University Impact Acceleration Account[6].
- British Academy/Leverhulme Small Research Grants[7].
- Swansea University Paid Internship[8].

[5] See https://www.linkedin.com/feed/update/urn:li:activity:7102731645207621632/; https://hazyresearch.stanford.edu/legalbench
[6] https://www.swansea.ac.uk/research/research-impact/impact/impact-acceleration-account
[7] https://www.thebritishacademy.ac.uk/funding/ba-leverhulme-small-research-grants
[8] https://www.swansea.ac.uk/som/student-opportunities/careers

Once we will manage to obtain the resources to create a small initial corpus, we will fine-tune our PLM on it to assess whether it is able to perform the same task. We are rather confident it will because GPT can be trained by specifying prompts, and it seems rather intuitive to define those corresponding to the MOD, ACTION, and CONDITION variables of the form above.

Once we will obtain results, we will publish them in conference/workshop papers, and we will possibly contact the administrator of LegalBench to create a new task therein.

The creation of this corpus is preparatory for a future application to a joint research project to the EPSRC-FNR bilateral call[9]. In this project, we will propose to develop a novel GPT-based PLM not only able to automatically reproduce the annotations in the corpus but also to evolve them until the fine-grained annotations in [2].

This call is a UK call intended to support research collaborations between research centres in the UK and in Luxembourg. The UK is the single country in the world in which it is possible to freely download all legislation in LegalDocML, thus we believe our research ideas to be rather attractive for UK research calls.

References

[1] Liga, D. and Robaldo, L.: *Fine-tuning GPT-3 for legal rule classification*. Computer Law & Security Review, to appear. Available at https://www.sciencedirect.com/science/article/pii/S0267364923000742

[2] Robaldo et al.: *Formalizing GDPR provisions in reified I/O logic: the DAPRECO knowledge base*, The Journal of Logic, Language, and Information, Vol 29. 2020.

[3] https://plato.stanford.edu/entries/logic-deontic/#TradScheModaAnal

[4] https://www.gov.uk/legal-obligations-drivers-riders

[9] https://www.ukri.org/about-us/epsrc/relationships/international-agreements/lead-agency-agreement-with-luxembourg