

From: "Jan Hajic" <hajic@ufal.mff.cuni.cz>
To: "Ufal-I mail conference, local DL" <ufal-l@ufal.mff.cuni.cz>
Sent: Friday, 30 December, 2022 13:44:11
Subject: PF 2023

Dear all:

As is now becoming a 'tradition', Bára sent a nice overview of ÚFAL's activities in the past year, and I am now writing here mainly about LINDAT/CLARIAH-CZ, which is the biggest part of our institute, and related quite a lot to many ÚFAL activities. My apologies for a rather long letter, but at least it is all in one place :-).

First of all, I would like to thank everyone, especially those working for LINDAT/CLARIAH-CZ and to LINDAT's many contributors, for another excellent year of work. This year, I would like to specifically thank first and foremost to the core IT/technical team of LINDAT, who – despite being relatively small team – managed to keep LINDAT running, to add new services (including the Ukrainian translation with the new UI), add new collections from very diverse sources, such as the National Film Archive and several others we have not catered for before, and initiate a big overhaul of the repository (upgrade to DSpace 7), which has just been delivered and will soon be tested and eventually deployed. Big thanks go also to those involved in the EOSC CZ program, where – thanks to a very active involvement of the LINDAT/CLARIAH-CZ core IT team - we eye some additional funding (from the structural funds) starting 2024 and then, hopefully, also in 2025. The same holds for the cooperation with the Charles University central library, which will deploy our repository solutions for university's open data repository, and which also helped to secure some additional funding already for 2023. Of course, there have been many other additions to the LINDAT/CLARIAH-CZ data collection and to its services – including two new UD releases, as always, or the Ukrainian MT. And, not to forget, work continues full speed on additional corpus resources, to be released in 2023, such as a new version of PDT and our lexical resources. LINDAT/CLARIAH-CZ has been involved in the organization of a few events – from the CLS summer school through shared tasks to the EBDVF (BDVA) event and the CLARIN Annual Conference: my big thanks to the main LINDAT and ÚFAL organizers and to all who participated and helped to make these events very successful. Our big thanks go also to the UFAL IT team for installing the new set of powerful computing facilities coming from the LINDAT/CLARIAH-CZ structural funds grant – I hope that everybody enjoys the now much more capable GPU cluster.

LINDAT/CLARIAH-CZ and several team members have been involved in EU and national projects that have all successfully finished in 2022 (see also Bára's overview for a full list). I would like to thank the PIs, Co-PIs, work package leaders and all team members in these projects for their effort, which contributed greatly to the success of these projects. Just as an example, I should mention the now closed ELG project, where our FSTP team managed the organization of the external partner selection and funding in such a way that the final review stated that "The open calls have been professionally organized and they got massive success [...]". Excellent work!

As you know from our department meetings, it has been very difficult year in terms of securing future funding. While in 2022 we have not been affected by any cuts (except for projects that have finished), for 2023 the situation is worse. Most importantly, LINDAT/CLARIAH-CZ has been cut 25% of the requested amount (which has been only slightly higher than we had in 2022) as a whole. Needless to say, it could have been worse: ÚČNK and some other infrastructures has been cut even more. Thank to budget negotiations with our partners (and thanks to those who have been successful in getting new grants, see below) we have been able to secure budget not substantially lower than in 2022. However, this does not take into account inflation – so the budget is not sufficient to cover that in full, of course.

While it might not be obvious, it was a big help (and relief) for LINDAT/CLARIAH-CZ that new projects at ÚFAL have been secured which either already started in 2022 or are starting early 2023, as Bára has listed, too: five Horizon 2020/HE projects (including one ERC grant), three GAČR grants, two NAKI grants and one MŠMT-funded one and one PRIMUS (by Charles University) – well done! We at LINDAT/CLARIAH-CZ hope to cooperate closely with these projects regarding their use of data and especially their deposit needs for their results.

In and for 2023 (and beyond), we will now be working on a new 4-year contract with the MŠMT for LINDAT/CLARIAH-CZ continuation (the government have finally approved all the infrastructures eligible for future funding on Dec 14, 2022). We are welcoming a new set of partners from the EHRI CZ consortium – we will now be 17 partners from 15 institutions. However, the guaranteed level of funding for 2024-2026 is substantially lower than that for 2023 (about 10M less, for each year), actually decreasing towards the end of the four-year period. As mentioned above, we hope to get something from the EOSC CZ structural funds, but it might not cover all of the decrease. We thus urge you to consider applying for new grants in 2023 (to start in 2024) as much as possible – from ERC through “standard” Cluster 4 and 2 Horizon Europe grants (please note the Natural Language Understanding call I have already emailed about, with March 2023 deadline), to the usual national ones like GAČR, TAČR, NAKI and MŠMT (including OP/structural funds) ones. Please note that our contacts through LINDAT/CLARIAH-CZ (especially, our partner institutions directly involved in it, as well as the EU grants with direct LINDAT/CLARIAH-CZ involvement) might be useful to form relevant consortia for applying for both international as well as national grants (as some of you actually already did, and succeeded at that). Also, please consider the possibilities offered by Charles University, such as the student-oriented GAUK grants. Every million counts, as I often use to say :). I should not forget to mention industrial and public administration cooperation: even though the current level of income from such cooperation is minimal, you never know – we welcome any activity going in that direction, and can contribute to it both technically at LINDAT/CLARIAH-CZ as well as legally through our well-established cooperation with the Charles University’s CUIP a.s. technology transfer company.

Regrettably, at least from a certain perspective, we also had to introduce some new policies regarding LINDAT/CLARIAH-CZ personnel expenses, the PBML journal and the SCTL book series. Within LINDAT/CLARIAH-CZ, not only for financial reasons but also for the expected new style of future evaluations carried out by the Government Office and the MŠMT, we can only fund those teams the result of which will bring more users and their results (publications, adaptations of our data etc.), as opposed to “just” language resources exposed in our repository. This has already affected some of you, but for the continued operation of

LINDAT/CLARIAH-CZ, we do not have another option. Please take it more of a future opportunity rather than as a restriction; I will be happy to explain more personally as we head towards 2023 and especially as we will start preparing for the next period of LINDAT, i.e., beyond 2026. For the SCTL book series, we will only publish books in exceptional cases where no other (prestigious or more appropriate, especially foreign) publisher is not available; please take this into account when planning your next grant. And finally, as you might know, we are – after more than 60 years and 120 issues – discontinuing the printed version of PBML, going to electronic edition only.

Even as some might perhaps see “PR” as a distracting activity, I believe it brings us some visibility and perhaps new students, collaborations and invitations to national and international projects. If you know about anything related to LINDAT/CLARIAH-CZ that is worth to show to the world, let me know and I will be happy to tweet through @LindatClariahCZ. And if you are on twitter, please follow and retweet whatever you like!

For equipment, while we still have some funding at the EXPRO grants in 2023, we will be again applying for the OP (structural) funds aimed directly at research infrastructures, once it is announced. It should then be available sometime in 2024-2026. So our computing cluster will hopefully stay and possibly expand in the future. In the meantime, we are also building ties with the other Czech e-infrastructure institutions (such as IT4Innovations and CESNET, mainly through the new EU HPLT project), which might enable us to use their and their EU partners’ computing capacities to use for our large students’ or any other compute-intensive research projects.

While the future, especially starting 2024, is quite uncertain and pretty much depends on external conditions (the war, energy prices and supplies instability, inflation, and more), over which we have little influence, we at LINDAT/CLARIAH-CZ seem to be now able to go through 2023 with some certainty, and to continue to support your and your students’ excellent research and teaching, as we have been doing so far, at least at the current funding level.

To finish in an optimistic tone, we at LINDAT/CLARIAH-CZ are looking forward to 2023, to cooperation with all of you and with our external partners.

I wish you, your families, colleagues and friends all the best for 2023!

Jan Hajič

PI and coordinator, LINDAT/CLARIAH-CZ