

# UniDive MC meeting 2

22 March 2023, 9:00-10:45, online

## Minutes

1. [AGATA SAVARY: 5 min.] Welcome to participants, verification of the quorum and adoption of agenda
  - a. The quorum is 2/3 of COST Full and Cooperating Members
  - b. [Attendance list](#) is checked and the quorum is confirmed: 31/36 countries represented
  - c. Reminding the MC vote rules on e-cost:
    - i. If two representatives of the same country cast different votes (approved vs. objected), the vote for this country is considered void
    - ii. If possible/appropriate, discuss with your co-representative to agree on a common vote
    - iii. Vote result visibility: every MC member can see the votes of all countries and of all MC members
  - d. Giuseppe Lugano (our COST Scientific Officer) could not join but sends his best greetings to the Action MC and remains available to address any doubts or questions that emerged during the meeting
2. Information to the MC
  - a. [AGATA SAVARY: 5 min] Recap of the [minutes](#) of the MC meeting on 23 Sep 2022, e-votes and matters arising since the last meeting
    - i. Some items from the brainstorming sheets from Brussels were included in the WG discussions in Paris-Saclay
    - ii. The Work and Budget Plan voted in Brussels is being implemented, with 40,000 EUR more in the budget - see below
    - iii. Part of the tasks defined in Brussels have been partly accomplished: criteria for WG membership (see below), STSM calls, [1st general meeting at Paris-Saclay](#):
      1. 2 days; 93 participants;
      2. [Introduction to UniDive](#); 2 plenary talks about the origins of UniDive
      3. 38 abstracts presented as posters (66 submissions, 4 withdrawn)
      4. Extended Core Group meeting
      5. 8 WG sessions (2 per WG)
    - iv. No questions or comments from the MC
  - b. [AGATA SAVARY: 5 min.] Core Group: report from the Core Group, including delegated decisions.
    - i. We had 6 (Extended) Core Group meetings so far. All past and upcoming [CG meetings](#) are documented at the Action's website.
    - ii. The MC delegated the following powers to the CG:

1. Approving/refusing WG applications - see below
2. Selecting event participants who will be reimbursed - see below
3. Adjusting the location and dates of activities (in close consultation with the Local Organizer) - see below
- iii. Including the GH Manager (Beata Wójtowicz) in the Core Group (**VOTE**)
  1. Beata Wójtowicz presents herself
  2. No questions or comments from MC
  3. Vote result: 0 against, 0 abstentions, all in favor
- iv. **Decision 1:** Beata Wójtowicz joined the Core Group
3. Action Membership:
  - a. [AGATA SAVARY: 2 min.] New COST Members represented in the MC
    - i. 35 countries were represented at the MC in September 2022
    - ii. Armenia nominated 1 new member: Anna Danielyan (Anna presents herself)
    - iii. The total number of countries of the MC is now 36 (out of all 43 COST, Cooperating and Partner Members - see the [list](#))
    - iv. 23 of those (64%) are Inclusiveness Target Countries
  - b. [DAN ZEMAN: 5 min.] Action Participation: WG membership and applications
    - i. Rules and processes for WG membership
      1. We try to be very inclusive. Almost everyone who applies for WG membership is approved. No one has been rejected but a few people have been “put on hold” if their profiles seemed to be very vague and/or unrelated to what we do. One person got back to us and was subsequently admitted.
      2. Technically: Only Agata gets notifications about new applications and only Agata and Dan can approve them. Most are approved by Dan when he learns about them. Then Olesea is informed by Dan to add these people to the corresponding mailing lists. Once approved, people can ask either of us to have their WG assignment changed.
    - ii. Current results (see also slides 13-14 of the [Paris-Saclay opening](#))
      1. 236 applications so far (+ 4 MC members who have not applied for WG membership)
      2. 216 approved, 20 in the queue (12 recent applications to be considered; 8 older “irrelevant” applications on hold)
      3. Approved by working group:
        - a. WG1: 141
        - b. WG2: 113
        - c. WG3: 124
        - d. WG4: 91
        - e. all four WGs: 24
      4. Approved by other criteria:
        - a. Young researchers: 72
        - b. From ITC countries: 126

- iii. Agata Savary's comments:
      - 1. New applications come every week, but AS will inform DZ more often about them
      - 2. MC representative could communicate with participants from their country and build national level networking
- 4. [NURIT MELNIK: 10 min.] Grant Awarding by the Action - update on granted STSMs
  - a. Two calls: guests and hosts (ongoing)
  - b. 2 host applications
  - c. 13 guest applications (+1 for 10/2023)
  - d. 4 applications due to hosting proposals (2 applications for each of 2 hosts)
  - e. Diverse in terms of content, target languages, applicants' career stage and STSM duration
  - f. The total money requested was **32,610 euro**, whereas our total budget for the year is **13,600**.
  - g. 7 applications were accepted (**total: 14,150 euro**, thanks to extra budget)
    - i. From: Israel, Spain, Netherlands (2), Ireland (2), Finland
    - ii. To: Belgium, France, UK (2), Italy (2), Netherlands
  - h. 4 on stand-by, depending on budget left
  - i. 2 rejected (double applications)
  - j. Short discussion about the topics of STSMs
    - i. [Jolanta Kovalevskaitė] Can we know the topics of the STSM applications?
    - ii. [Atul Kumar Ojha] They are visible on e-cost [here](#), but only for the Core Group members.
  - iii. 7 topics recopied from the applications:
    - 1. Transitions all the Way Down: From Characters to Full Document Annotation in One System
    - 2. Designing "Paidiom": a text-preprocessing algorithm to improve the neural machine translation of multiword expressions
    - 3. Enhancing the NCTTI dataset with Romanian nominal compounds
    - 4. Laying the foundations to create Universal Dependencies treebanks for low-resourced Dravidian languages
    - 5. Universality, diversity and idiosyncrasy in language technology
    - 6. Syntactic and Lexical Variation in Low Saxon
    - 7. Analysis of Non-compositionality Handling in Transformer-based Models
- 5. [BEATA WÓJTOWICZ: 10 min.] Budget status: summary from the Grant Holder.
  - a. For Paris meeting an average reimbursement rate is 831 EUR (based on 47 submitted claims) against the estimated 980 EUR
  - b. Additional budget: 40,000 EUR from COST left-overs (we can spend additional 35,000 EUR)
  - c. Work and Budget Plan amendment process ongoing
  - d. Split of the additional budget
    - i. Reimbursing 10 persons more at Paris-Saclay

- ii. Reimbursing 16 people more at the Istanbul meeting
  - iii. Increasing the STSM budget
  - iv. Increasing the bank fees by 200 EUR
- e. Due to technical issues on e-cost, editing WBP for Paris-Saclay and OERSA is no longer possible. Therefore, these 2 amounts are assigned to STSMs. The actual STSM budget will have to be adjusted after the Paris-Saclay reimbursements.
- f. Agata Savary shares the current WBP and the draft of the new WBP and compares the differences
- g. E-cost vote on the WBP amendment is ongoing
- 6. Follow up and discussion
  - a. [AGATA SAVARY: 5 min.] Implementation of the COST Excellence and Inclusiveness Policy
    - i. Policy for selecting the 1st general meeting participants
      1. Abstract presenters + ECG members + the GH Manager
      2. Local participants: LISN, Paris, close to Paris, with no accommodation, only short/middle distance costs
      3. People with leading roles in the communities and seminal works
      4. The others chosen so as to balance the percentage of YRI, countries, gender and WG membership
      5. For details - see the [minutes](#) from the CG meeting 4
    - ii. Statistics from STSMs - among the 7 accepted STSM applications,
      1. 3 women (43%)
      2. Only non-inclusiveness target countries benefit from STSMs (as senders or receivers); applications from ITCs should be encouraged; see also page 21 of the Paris-Saclay [opening slides](#)
    - iii. No questions or comments from the MC
  - b. Progress of each working group
    - i. [CARLOS RAMISCH: 5 min.] WG1 - [Details](#) and [slides](#)
      1. **Session 9 - Breaking the ice** - annotation "speed dating" - 10-15 people presented their annotation environments. Various topics: annotation interfaces, data formats, genre-specific challenges...
      2. **Session 11 - Brainstorming** - 2 parts, "what" can we collaborate on, and "how" to make progress. Discussion in subgroups:
        - a. *Linguistic typology and multilingual corpus annotation*
          - i. What: Mutual benefits: Unification of terminology in guidelines, Cross-lingual concepts,...
          - ii. How: monthly online meetings, shared documents, fewer manageable problems to address
        - b. *Extensions to MWE annotation guidelines*
          - i. What: Address new MWE categories (priority on nominals). Challenges: lemmatisation, tokenisation, compounding, rich morphology
          - ii. How: Shared documents, webinars, invited talks

- c. *Extensions to morphosyntactic annotation guidelines*
      - i. What: 'error mining' might be useful. Improve UD guide: morphosyntax vs. semantics, notion of word
      - ii. How: Automated data evaluations, improvements, joint meetings of treebank creators
    - d. *Convergence between PARSEME and UD*
      - i. What: Challenging: not many people know both. Sharing best practices: UD-inspired validation script, PARSEME-inspired decision trees
      - ii. How: Regular meetings of guidelines developers, Support for newcomers
    - e. *Sharing tools, formats, and infrastructure*
      - i. What: Go beyond framework-centric inventories. Use existing centralized platforms (e.g. CLARIN).
      - ii. How: platform for messaging (e.g. Slack), platform for data sharing without storage limitations
  - 3. **Next steps**: Heterogeneous interests: how to benefit from the richness of backgrounds to make progress? More structured online survey on interests and priorities
  - 4. Chat discussion about the planned tasks/homework in WG1 [Jolanta Kovalevskaitė and Carlos Ramisch]
- ii. [VERGINICA MITITELU: 5 min.] WG2 - three parallel tasks with associated forces, each led by 2 co-leaders; see [report](#) from the two sessions on 17th March; coming this week: doodle for establishing the time for the monthly meetings of WG2
  - 1. Chat comment on increasing interactions between WGs and organizing combined WGs meetings [Jolanta Kovalevskaitė]
- iii. [JOAKIM NIVRE: 5 min.] WG3: Two sessions at first general meeting – see [report](#). Two main activities/deliverables: 1. Documentation of multilingual and cross-lingual NLP tools. 2. Multilingual evaluation campaigns. Focus of next meeting in Istanbul will be documentation of multilingual and cross-lingual tools. Subgroups will be set up for preparatory work on two aspects: 1. Taxonomy, 2. Infrastructure.
  - 1. Chat comment supporting the evaluation campaign of the tools [Jolanta Kovalevskaite]
- iv. [MARIE-CATHERINE DE MARNEFFE: 5 min.] WG4 - two sessions ([slides and summary of the discussions](#))
  - 1. Quantifying diversity: Small group discussion to break the ice: "What is your experience with quantifying diversity?", then focus on how do we define diversity exactly – this still needs to be refined
  - 2. Promoting diversity: "How would we choose 100 languages to be represented in a shared task?" Small group discussion and overall discussion.

The slides above give a summary of the discussions. Next steps:

- Need to decide on a way to share info within and across WGs since there is a lot of transversality
  - Create sub-tasks groups
  - v. All the above linked documents are now linked from the [WG pages](#) on the Wiki website
7. [OLESEA CAFTANATOV: 5 min.] Communication: Science Communication Plan, communication survey, expectations survey
- Communication Survey - to be sent on March 23, 2023, feedback needed within 1 week
  - Regarding Science Communication Plan - Each Action MC shall adopt a Science Communication Plan including a communication, dissemination, and valorisation strategy, as well as a plan to implement this strategy. The Science Communication Plan shall reflect the MoU in particular connecting to the aims and objectives of the Action. It is recommended that the Science Communication Plan is approved by the Management Committee **not later than 6 months** after the start date of the Action. It is recommended that the Science Communication Plan, including progress on implementation, is discussed on a yearly basis by the Action MC and reviewed or amended where necessary.
  - We call for 2 volunteers to contribute to the Science Communication Plan
  - Logo: a contest will be launched soon (it is being drafted)
  - Videos from the Paris-Saclay meeting are upcoming
  - See also page 20 of the Paris-Saclay [opening slides](#)
  - No comments from the MC
8. [ALINA WRÓBLEWSKA, AGATA SAVARY: 5 min.] Progress on MoU Objectives, WG tasks, deliverables, and Goals for the current GP
- a. Establishing internal and external communication means
    - i. Wiki website, mailing lists, upcoming communication survey
  - b. Establishing working relationships and structuring the community around the Working Groups
  - c. Understanding the state of the art for the work program of each WG
  - d. Planning the activity of each Working Group
    - i. For the 3 above items, see the WG reports above
    - ii. The work program might need updates due to a very fast evolution of the NLP domain; discussions started at Paris-Saclay
  - e. Extending the Action to new potential countries, languages and dialects
    - i. New country on board (Armenia)
    - ii. COST/Partner Member countries not yet represented in the UniDive MC: Austria, Cyprus, Denmark, Iceland, Luxembourg, Montenegro, South Africa
    - iii. A question about the language representation is planned in the communication survey
  - f. Facilitating and coordinating the development of language resources and tools for new under-resourced languages
    - i. The webinar in June will be dedicated to UD/PASREEM newcomers

- g. No questions or comments from the MC
- 9. Planning
  - a. Upcoming activities for Grant Period 1
    - i. [AGATA SAVARY: 1 min.] Online Webinar - training for newcomers to Universal Dependencies and PARSEME
    - ii. [GULSEN ERYIGIT or AGATA SAVARY: 5 min.] WG3 meeting in Istanbul
      - 1. Venue: ITU Maçka Hotel - a small hotel within one of the ITU campuses; located at the very city center
      - 2. The accommodation there is relatively inexpensive but pre-booking might not be accepted and some participants might have to take other hotels in the downtown and these have prices mostly above 100 EUR.
      - 3. It would be safer **not to lower** the daily allowance.
  - b. [AGATA SAVARY: 5 min.] Draft plans for Grant Period 2
    - i. 1 general meeting
    - ii. 1 training school
    - iii. Either another general meeting (preferred, to increase interactions) or 1-2 WG meetings, depending on the budget
    - iv. Discussion: what would be preferable - delay the discussion
    - v. [Jolanta Kovalevskaitė] Recall that we planned on online WG just before or just after Istanbul
- 10. AOB
  - a. For better visibility, please, link the UniDive Wiki website (<https://unidive.lisn.upsaclay.fr/>) from your personal websites
  - b. Strategies for getting visas from non-EU countries
    - i. Anna Danielyan from Armenia did not receive an appointment for a visa early enough and had to cancel her travel to Paris-Saclay
    - ii. This is a real problem for inclusiveness in UniDive.
    - iii. Next meetings will be organized more in advance to allow for long visa procedures
    - iv. Potential invitees might ask for invitations from the meeting organizers early enough, but should keep in mind that these do not yet guarantee reimbursement (an official e-cost invitation is needed).
- 11. Summary of MC decisions'
  - a. Due to some persons having arrived too late for the MC vote in item 2-b-iii, the vote is repeated for these persons. **Vote results:** 0 against, 0 abstaining, all in favor
- 12. Closing
  - a. The next meeting will take place online in October (planned for **18 Oct**, to confirm), with the WBP for year 2 on the agenda.

## TODO

- [ALL] Think about coordinating national groups including UniDive WG members

- [ALL] Answer the communication survey
- [ALL] Consider volunteering to help develop the Science Communication Plan (2 volunteers needed)
- [ALL] Link the UniDive Wiki website (<https://unidive.lisn.upsaclay.fr/>) from your personal websites
- [AGATA] Publish and advertise these MC minutes
- [OLESEA] Add Beata to the CG mailing list
- [AGATA] Notify Dan more often about WG applications
- [AGATA] Give the WG (co-)leaders access to the Wiki website.
- [ECG] Discuss online meetings for WGs 1, 2 and 4 around the date for Istanbul (to increase momentum and interactions)
- [AGATA] For the next meetings, check if generating an attendance list from a Zoom meeting is possible with the CNRS account.
- [AGATA] Confirm the date for the next MC meeting in October