

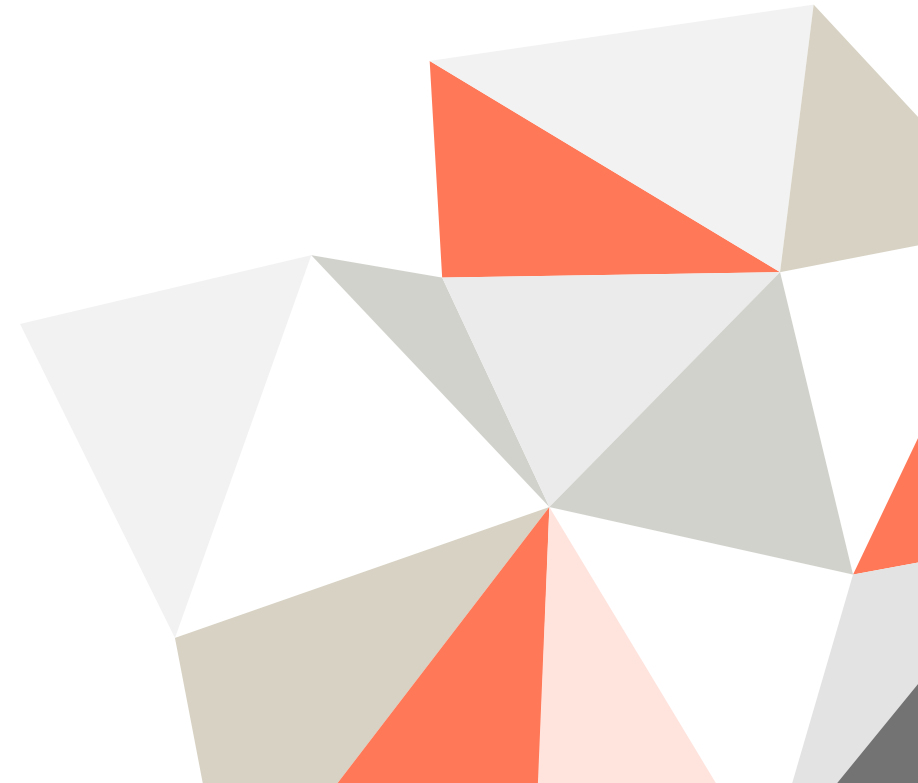
WELCOME CA21167

Universality, diversity and idiosyncrasy in
language technology (UniDive)

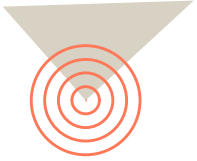


Your hosts today:

- Giuseppe Lugano (Science Officer)
- Andrea Tortajada (Administrative Officer)
- Agata Savary (Action Chair Candidate)



CA21167 - Information



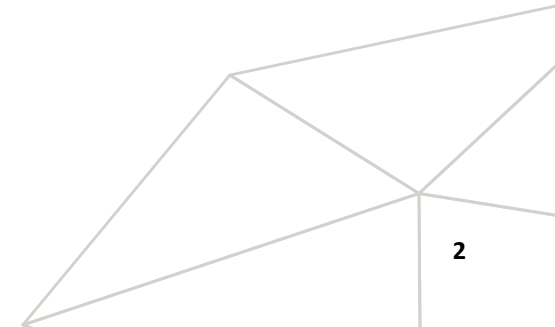
CSO approval: **27/05/2022**

Start Date: **23/09/2022**

End Date: **22/09/2026 (4 years)**

MC Chair: to be elected (today!)

MC Vice Chair: to be elected (today!)



Evacuation Instructions

In case of fire alarm*

- **Stop** all activity
- Try to **stay calm**, do not run
- Follow **COST staff** and the **evacuation signals**
- Emergency exits are located after the bathrooms at the entrance of the office
- **Do not use the elevators!**
- Join the meeting point, **outside** of the building

***Fire alarm test: every 1st Tuesday of the month at 2pm!**

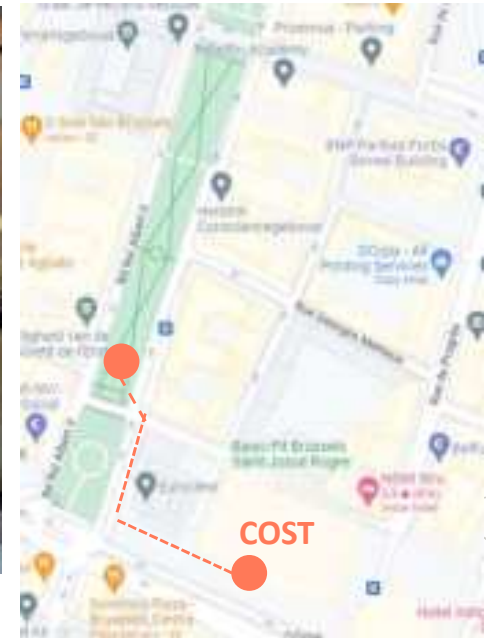




Evacuation Instructions

You will arrive at the Manhattan reception (ground floor)

- Exit the building and go to the right (Proxy Shop and Vlerick School)
- Then turn right of the Manhattan Center and cross the street to the meeting point on Bd Albert II

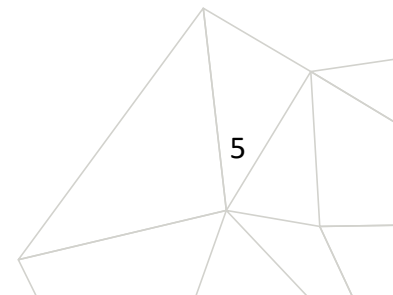


Evacuation Instructions



People with reduced mobility or other health restrictions/issues:

- move to the **first** elevator to the **left** of COST office entrance
- **wait** for fire marshals to pick you up



I. Welcome and Frame



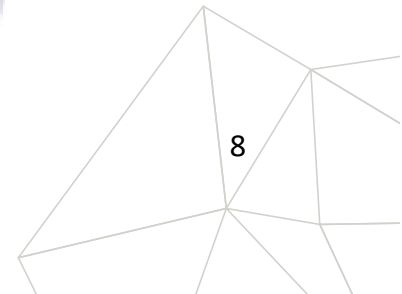
House rules

- No recording / pictures to be taken without consent
- Use this chance to engage and interact
- Keep your statements short and to the point



Who is attending?

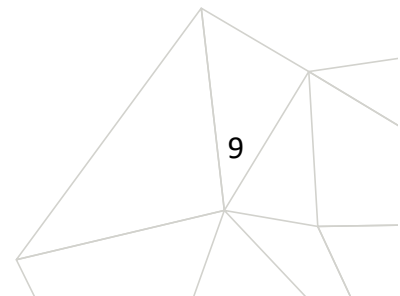
- Nominated MC members by their COST National Coordinator
- Beneficiaries to the Leadership Program
- Agreed WG applicants that are candidate for leadership positions



Goals vs Non-Goals



- ✓ Get familiar with COST framework and COST Action management
- ✓ Common understanding and taking ownership of the Action challenge and objectives
- ✓ Initial networking and community building
- ✓ MC Decision-making: election of mandatory leadership positions, proposal for Work and Budget Plan
- ✗ Detailed explanation of the Annotated Rules
- ✗ Deciding on internal processes



Agenda 9.00 – 17.30 (CET)



- I. Welcome and Frame
- II. Getting to know each other
- III. COST presentation + Q&A
- IV. Action Chair Candidate presentation + Q&A

Break

- V. COST Action Café (parallel discussion sessions)

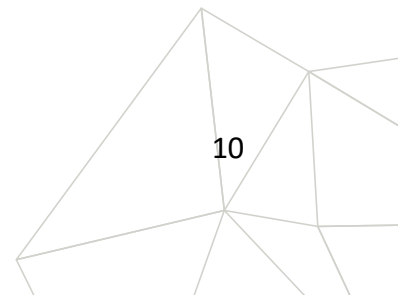
Lunch

Group Photo

- VI. Debriefing from the COST Action Café

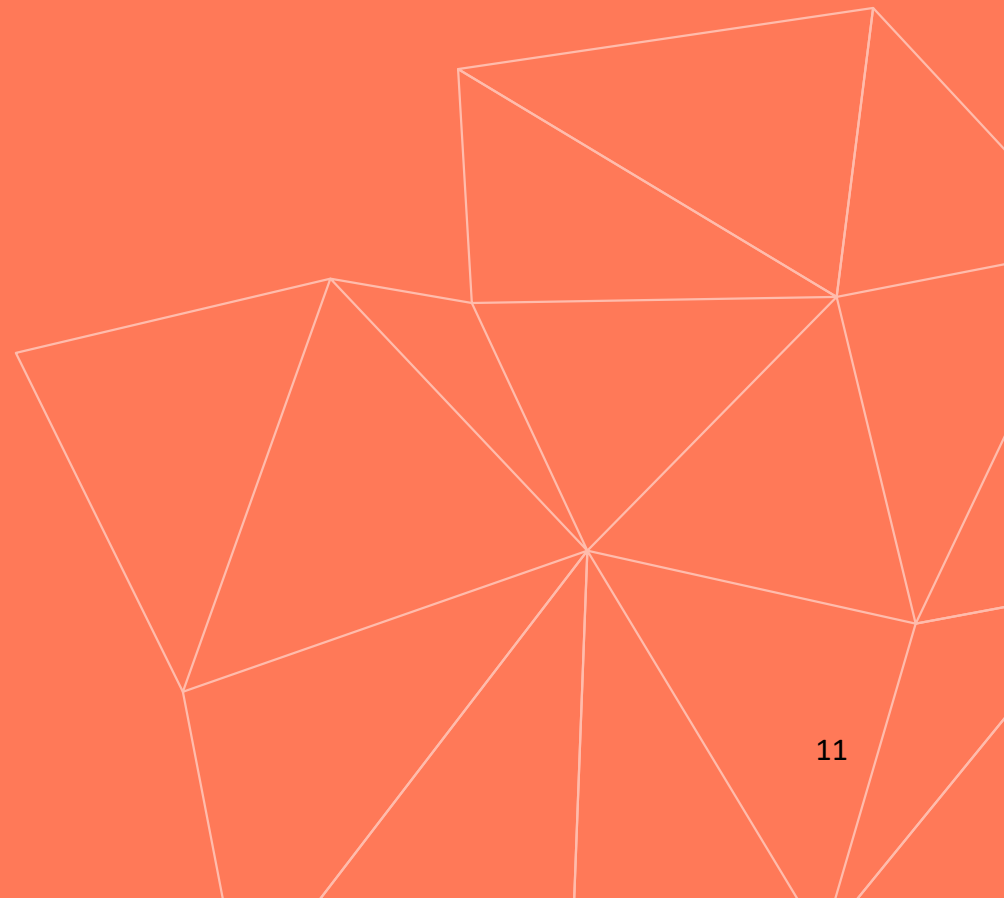
Break

- VII. Election of mandatory leadership positions (Chair, Vice-Chair, Grant Holder)
- VIII. MC Discussion and Decisions (moderated by elected Chair)



Check-in

II. Getting to Know each other

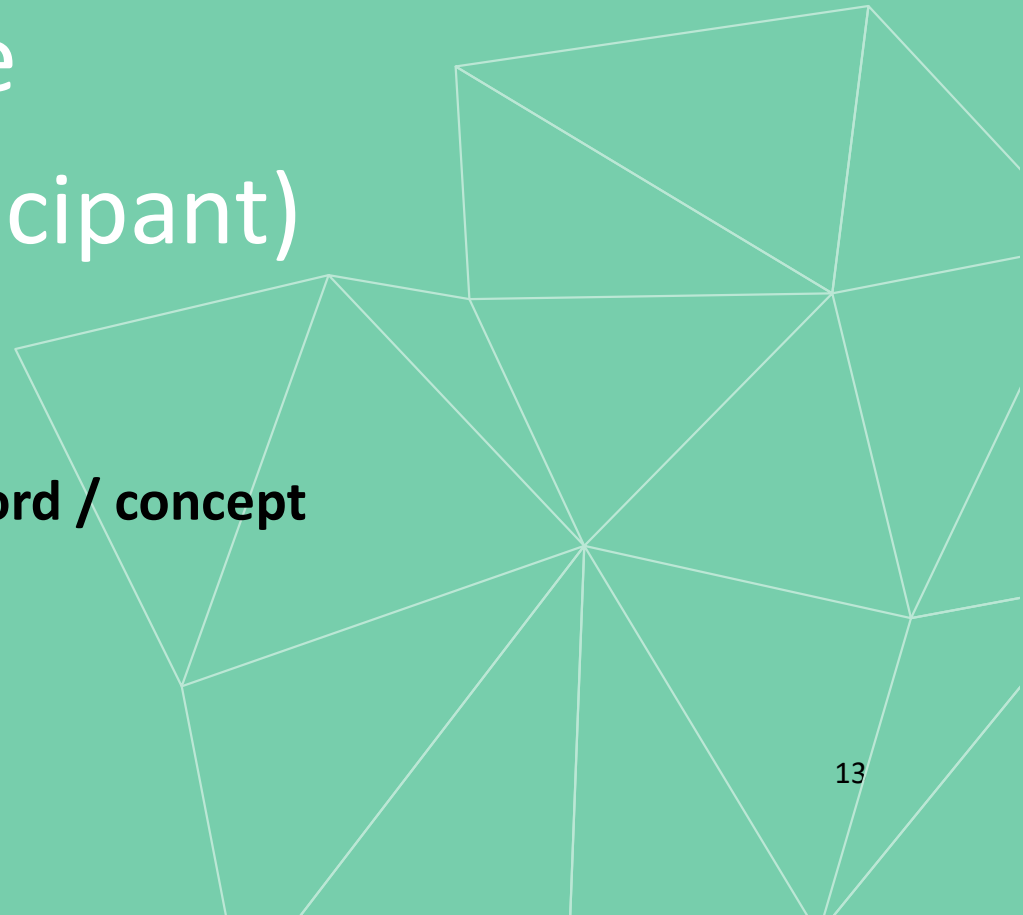


1. Have I already participated in a COST Action?



2. 'Speed' Tour de Table (max 30 seconds / participant)

- **Name**
- **Country**
- **Your expectation from the Action in 1 word / concept**



3. What expertise do I bring to the Action?

(5 min, talk to participants next to you)



III. COST Presentation



COST - European Cooperation in Science & Technology

MISSION

COST provides **networking opportunities** for researchers and innovators

to strengthen Europe's capacity to **address scientific, technological and societal challenges**

VISION

Europe's most **empowering** research programme

STRATEGIC PRIORITIES

1. Promoting and spreading **excellence**
2. Fostering **interdisciplinary** research for breakthrough science
3. Empowering and retaining **young researchers and innovators**

COST Excellence and Inclusiveness Policy



Special measures to Young Researchers and Innovators

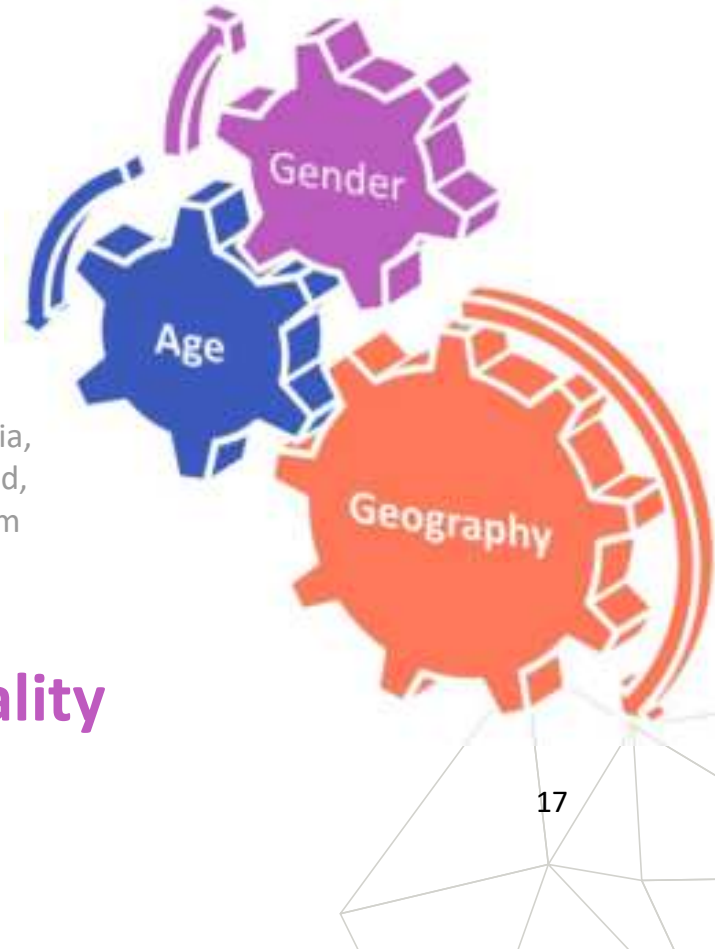
A researcher or innovator under the age of 40

Special measures to Inclusiveness Target countries in Horizon Europe

Albania, Bosnia and Herzegovina, Bulgaria, Cyprus, Croatia, Czech Republic, Estonia, Georgia, Greece, Hungary, Lithuania, Latvia, Malta, Moldova, Montenegro, North Macedonia, Poland, Portugal, Romania, Slovenia, Slovakia, Serbia, Turkey and Ukraine + Outermost regions from Portugal, Spain and France

Promoting Gender Balance and Gender Equality

<https://www.cost.eu/about/strategy/gender-equality>



COST Membership

40 Members

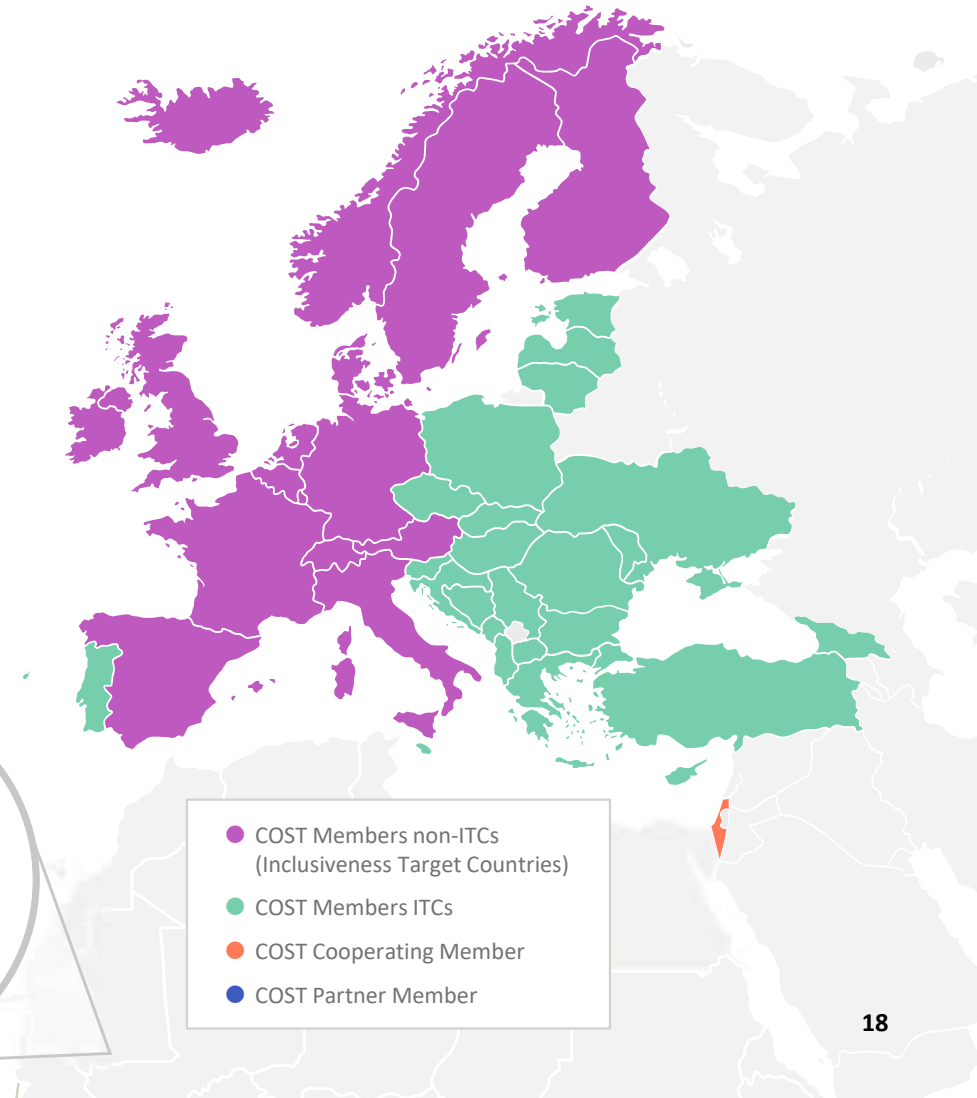
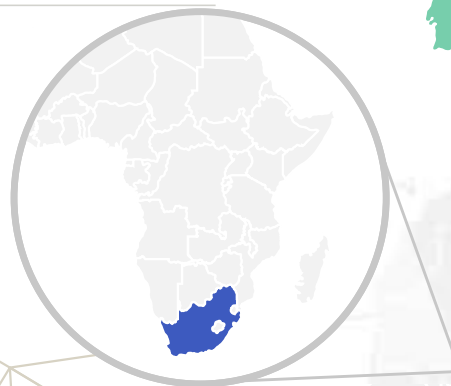
- | | | |
|--------------------------|-----------------------------------|------------------|
| ● Albania | ● Greece | ● Norway |
| ● Austria | ● Hungary | ● Poland |
| ● Belgium | ● Iceland | ● Portugal |
| ● Bosnia and Herzegovina | ● Ireland | ● Romania |
| ● Bulgaria | ● Italy | ● Serbia |
| ● Croatia | ● Latvia | ● Slovakia |
| ● Cyprus | ● Lithuania | ● Slovenia |
| ● Czech Republic | ● Luxembourg | ● Spain |
| ● Denmark | ● Malta | ● Sweden |
| ● Estonia | ● The Republic of Moldova | ● Switzerland |
| ● Finland | ● Montenegro | ● Turkey |
| ● France | ● The Netherlands | ● Ukraine |
| ● Georgia | ● The Republic of North Macedonia | ● United Kingdom |
| ● Germany | | |

1 Cooperating Member

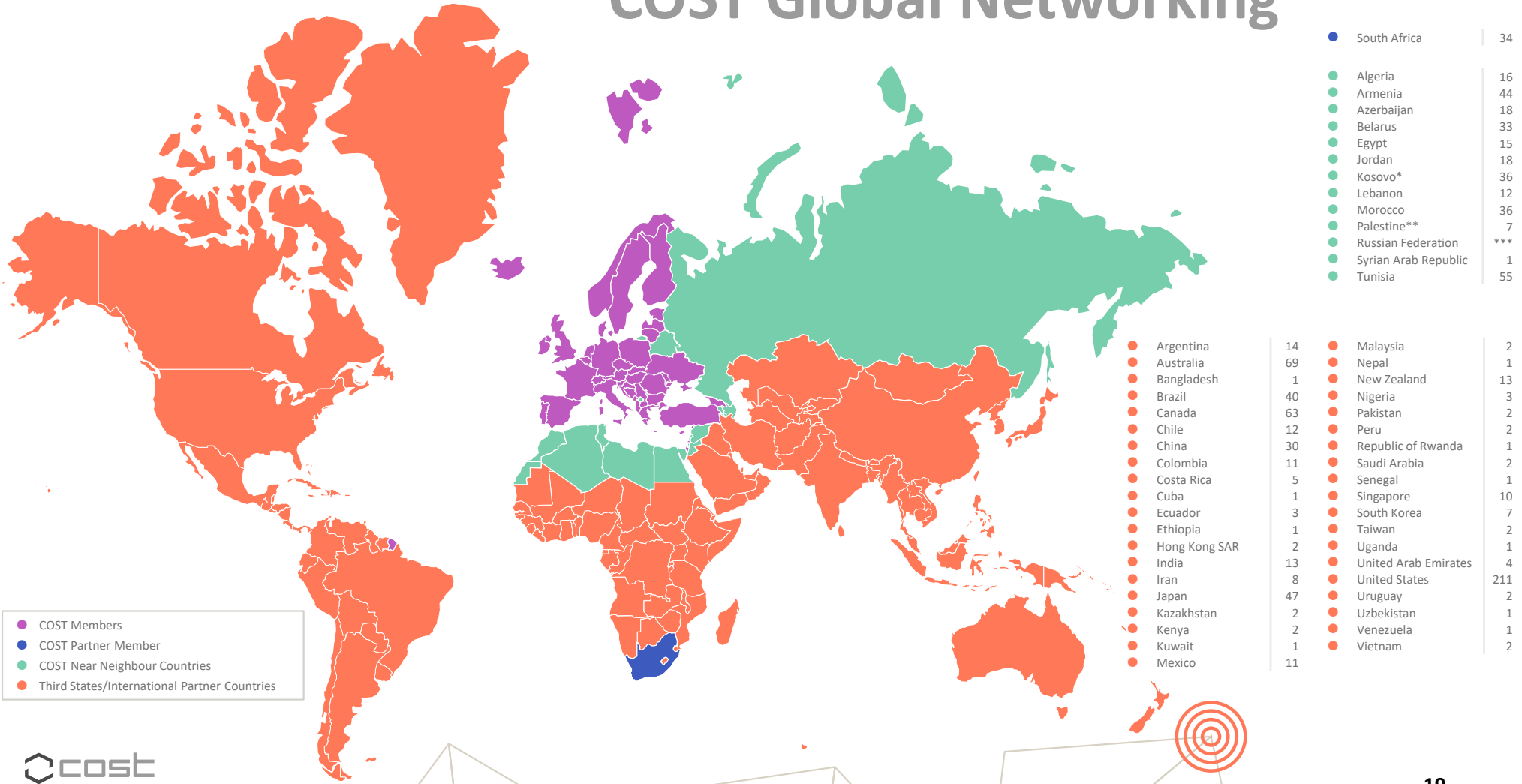
- Israel

1 Partner Member

- South Africa



COST Global Networking

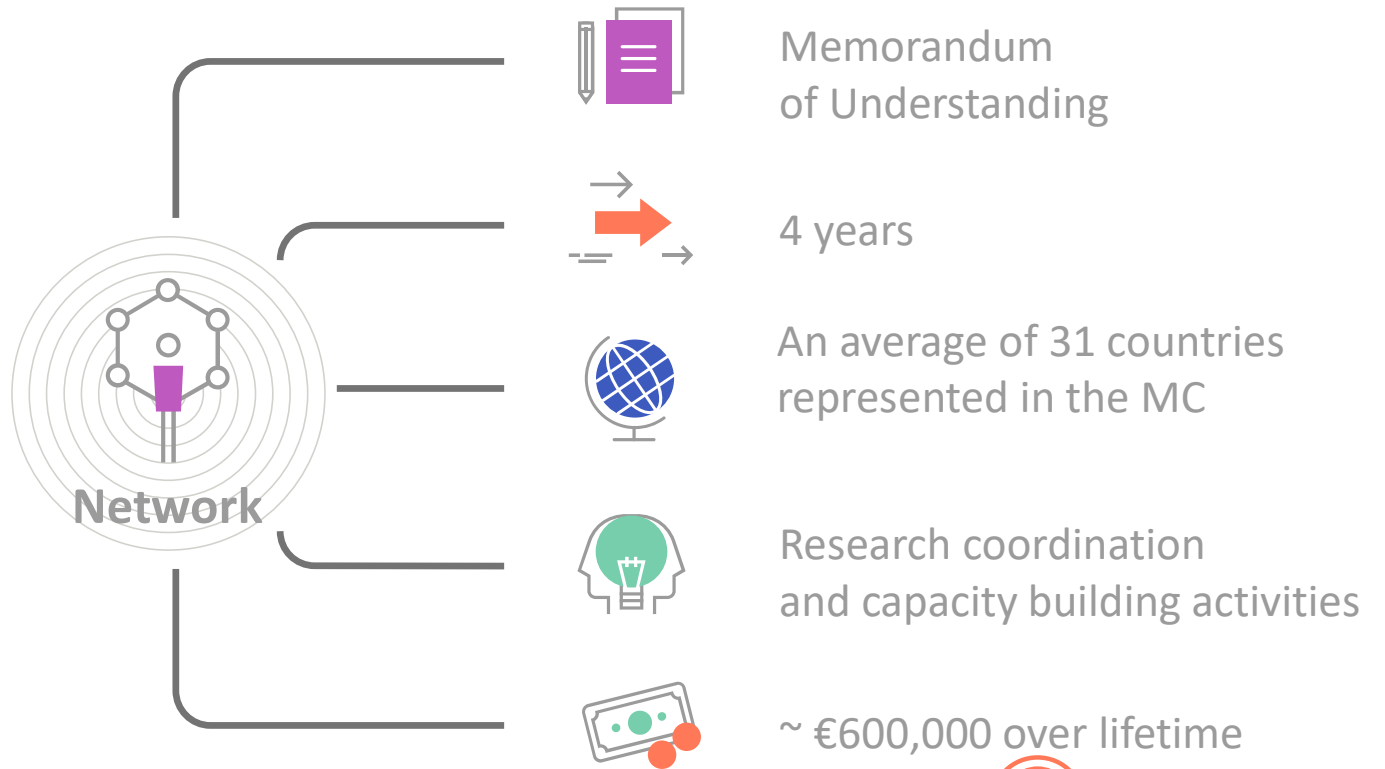


*** As of 6 May 2022 until further notice, measures are taken to suspend cooperation with Russia for the execution of the COST Actions and COST activities. Amongst these measures, participation and eligibility of COST Action participants affiliated to a legal entity established in Russia are suspended for all COST Actions and COST activities.

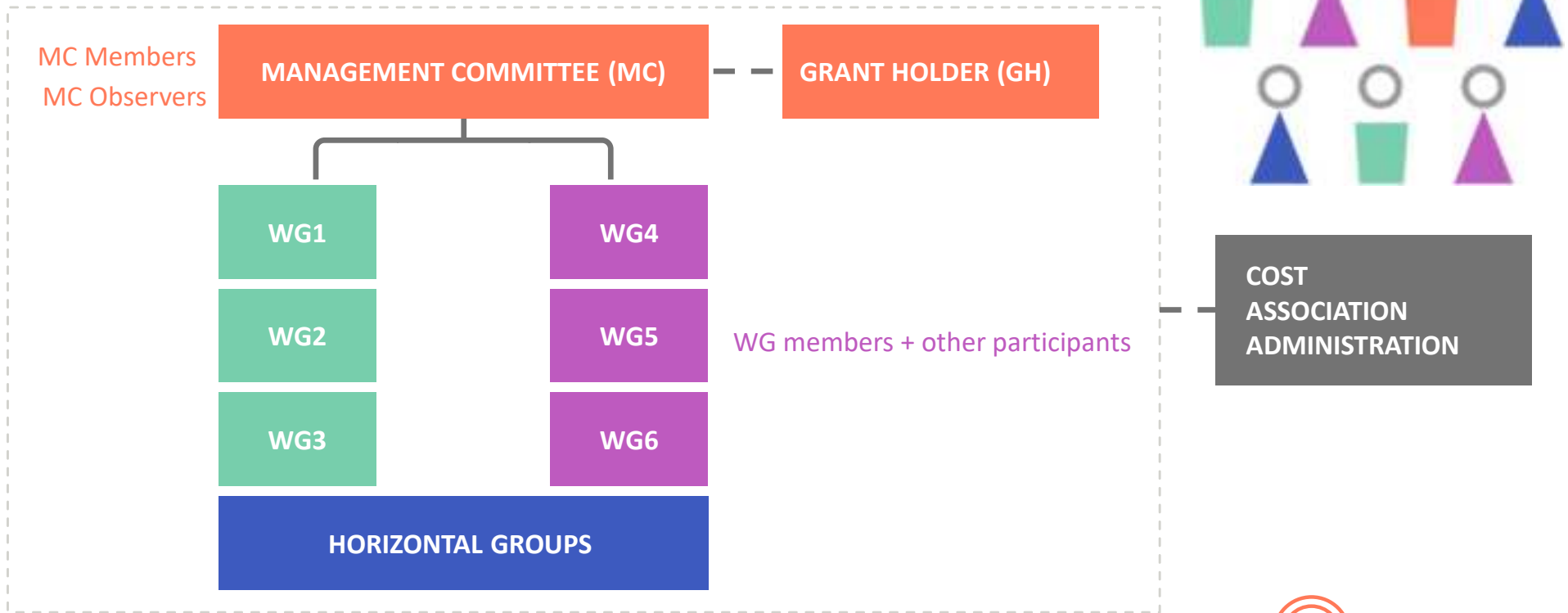
COST Actions

COST provides funding for research coordination and capacity building activities.

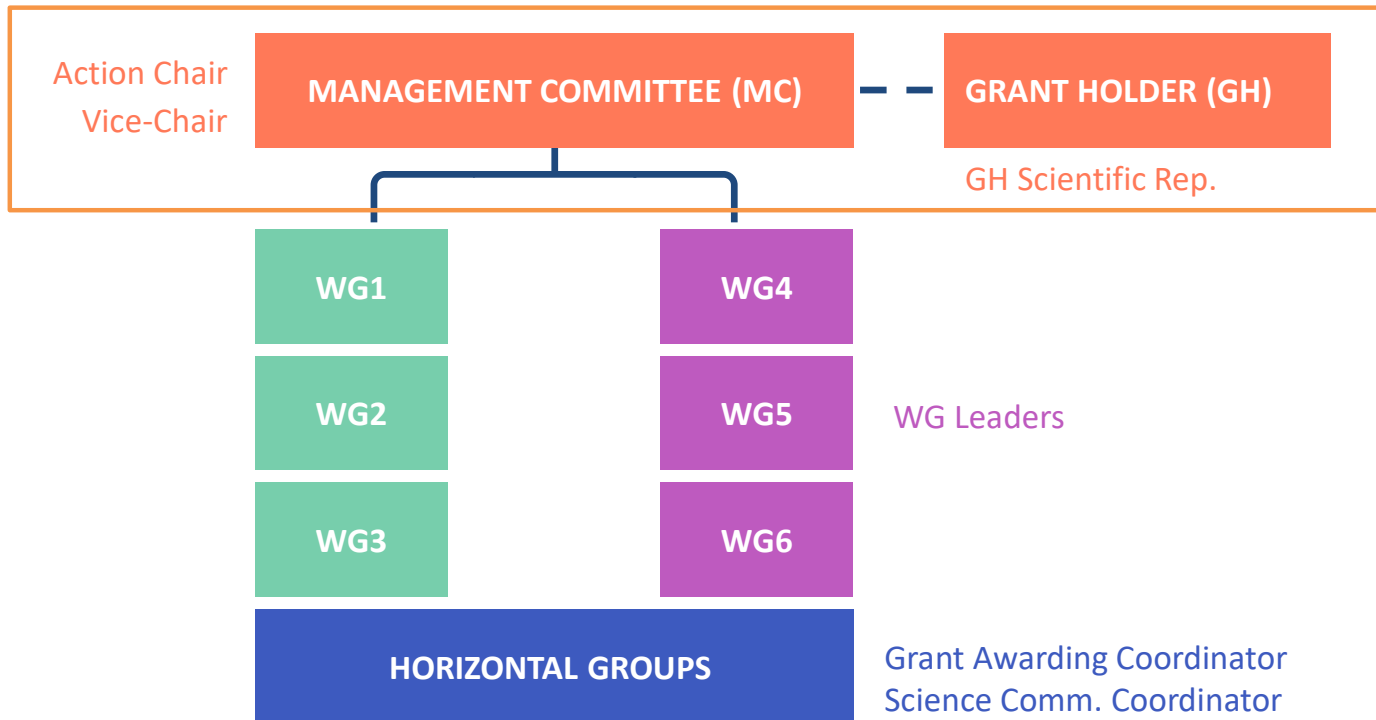
COST is not funding research itself, but pooling resources and research results by networking.



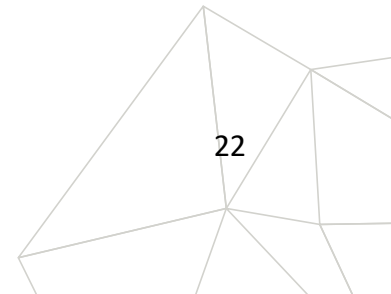
Action Structure and Participation



Leadership Positions and Core Group



Must be elected at MC1



MC Members



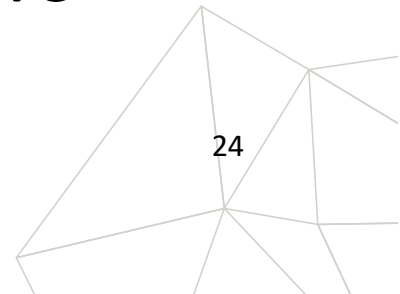
- The Action Management Committee (MC) is the decision-making body.
- Manage, vote and fulfil obligations regarding:
 - **Content:** monitoring achievement of MoU objectives/deliverables and implementation of activities specified in the Work and Budget Plan
 - **People:** appoint Action Leaders, approve new MC members and Observers (after MC1), as well as Working Group Members
 - **Processes:** e.g. criteria for WG membership, definition of grant awarding process, develop communication & Dissemination strategy
- Represent the COST Member State's national community of researchers and innovators on the Action's topic



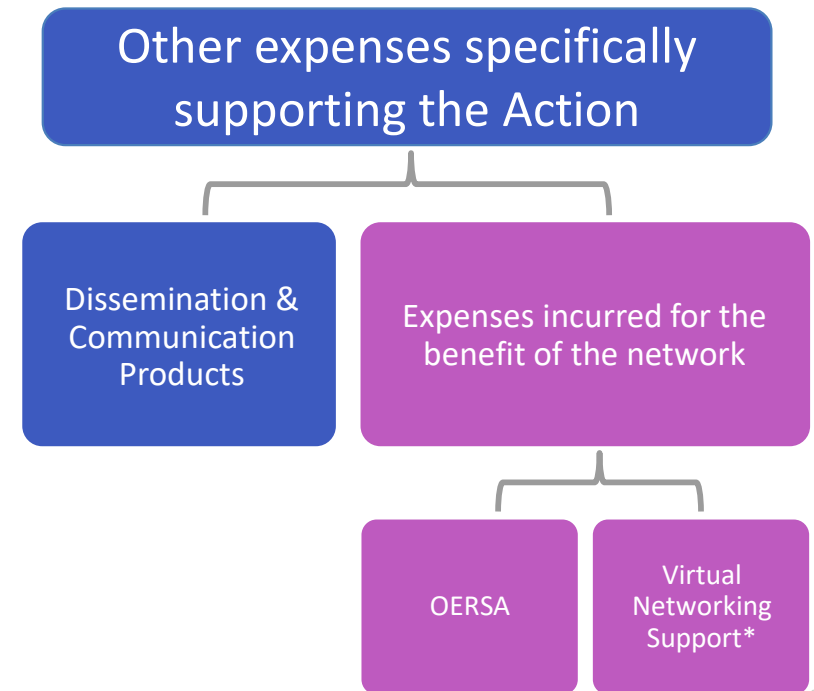
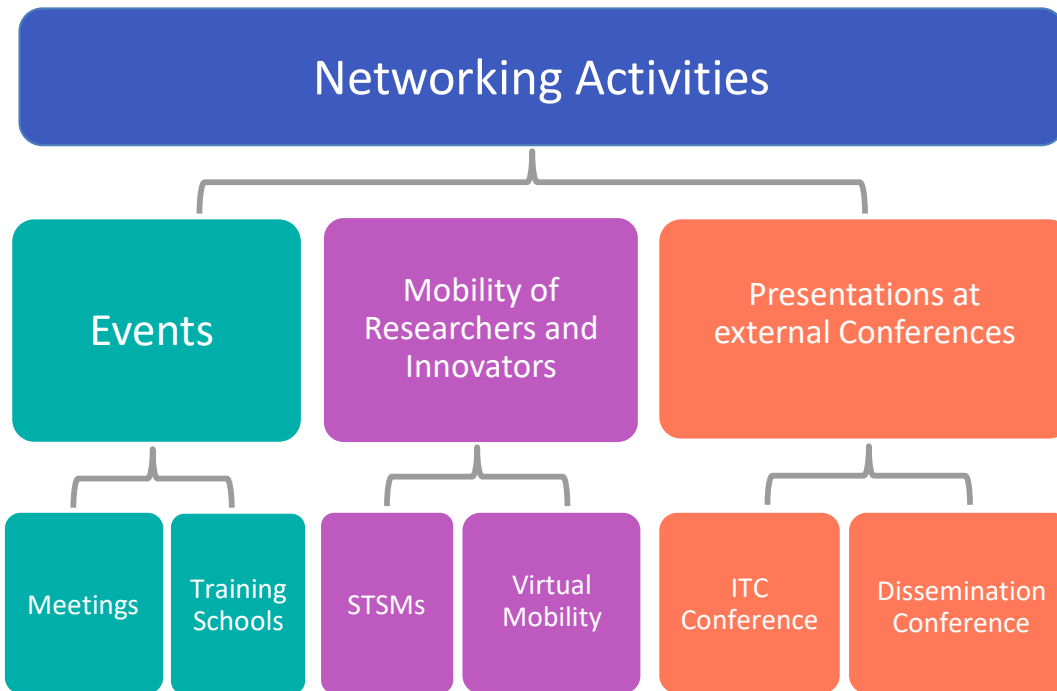
WG members



- Contribute to the realisation of the Action's objectives and deliverables
- An individual with an affiliation in a legal entity anywhere in the world approved by the Management Committee (via e-COST)
- MC to review and decide on WG applications within 3 months



COST Networking Activities



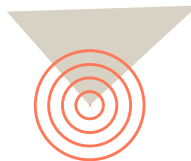
Eligibility to receive funding



Affiliation to		Meetings		Training Schools			Other networking activities	
		Invited individual	Local Organiser	Trainer	Trainee	Local Organiser	ITC Conf Grant*	All others
COST Members	Legal entity in COST F/C Member	✓	✓	✓	✓	✓	✓	✓
	Legal entity CPM	At COST Partner Member's costs (dedicated budget line)						



Affiliation to		Meetings		Training Schools			Other networking activities	
		Invited individual	Local Organiser	Trainer	Trainee	Local Organiser	ITC Conf Grant*	All others
Specific Organisations	EC/EU	At EC, EU body, office or agency cost(minimisation of cascade funding)						
	Eur. RTD Org.	✓	✓	✓	✓	✓	✗	✓
	IO	Once in a lifetime of the Action	✗	✓	✗	✗	✗	✗



Affiliation to		Meetings		Training Schools			Other networking activities	
		Invited individual	Local Organiser	Trainer	Trainee	Local Organiser	ITC Conf Grant*	All others
NNC	Legal entity in a NNC	✓	✓	✓	✓	✓	✓	✓
IPC/ Third states	Legal entity in a IPC	Once in a lifetime of the Action	✗	✓	✗	✗	✗	✗

*ITC Conference Grants are Limited to Young researchers and innovators from COST Inclusiveness Target Countries and NNC.

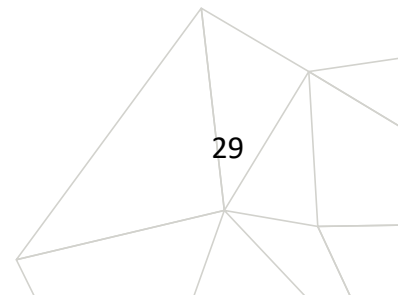
Clarification of affiliation



Any form of recognised relationship between the individual and a legal entity.

Examples:

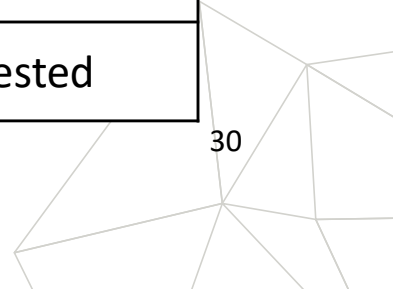
- A work contract;
- Enrolment in a research performing study programme (for example, PhD or Post-Doctoral programme);
- Voluntary service in a NGO;
- Emeritus professorship.



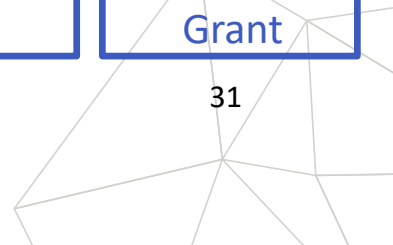
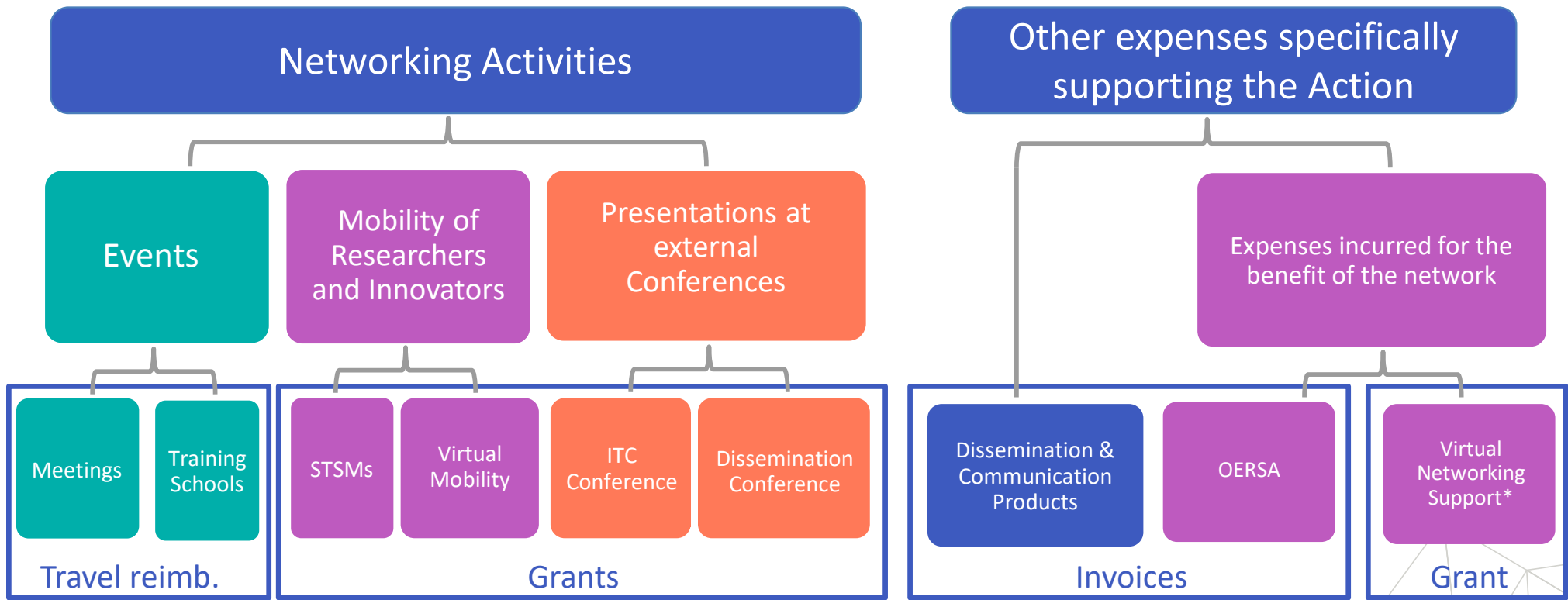
Location of Activities



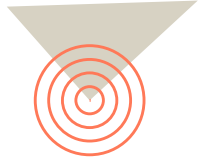
		Meetings	Training Schools and other network activities
Virtual Environment		✓	✓
COST Members	Full or Cooperating Members	✓	✓
	Partner Member	✗	✓
NNC		✓	✓
IPC/Third State		✗	✓
Specific Organisations		Rule applied to territory where it is vested	



Reimbursement modalities



Travel Reimbursement rules



Daily allowance (DA) covers:

- Accommodation
- Meals
- Short distance travel (<100 km one way)

Based on:

- Meeting country
- Travel dates and times



Long distance travel (≥ 101 km one way) covers total travel expenses up to 1500 EUR

- Train, bus
- Plane
- Ferry
- Car – limited to 2000 km, 0.35 €/km

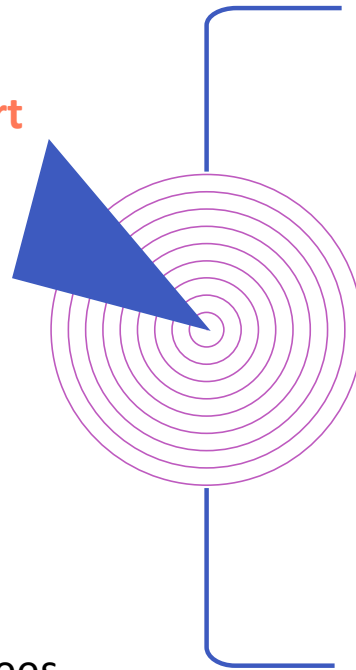
Others:

- Visa costs
- Travel cancellation insurance

Local Organiser Support (LOS) Grant



Local Organiser Support



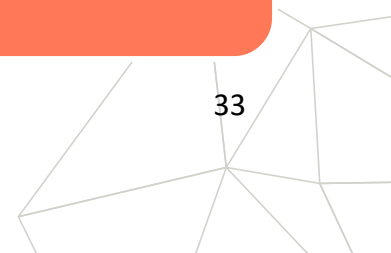
e-COST invitation to all attendees
(regardless of eligibility)

Daily attendance list signed daily

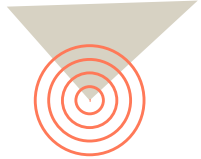
Virtual meeting log as generated
from virtual platform (e.g. Zoom,
Teams, etc)

Grant parameters:

- Format of meeting (F2F, Hybrid, or Virtual)
- Number of **unique** attendees
- Duration



Short-term scientific missions (STSM)



WHERE:

To a host organization located in a **different** country than the country of affiliation

Grant AMOUNT:

Up to 4000 EUR
Paid after report approval (after the STSM)

ITC or Dissemination Conference Grants

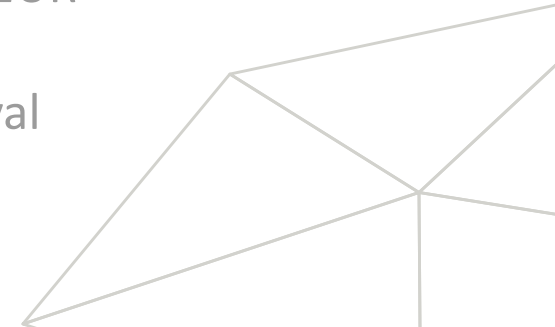


WHERE:

To high-level conference **not organised by the Action itself**

Grant AMOUNT:

Up to 2000 EUR or 500 EUR (virtual)
Paid after report approval



Virtual mobility grant

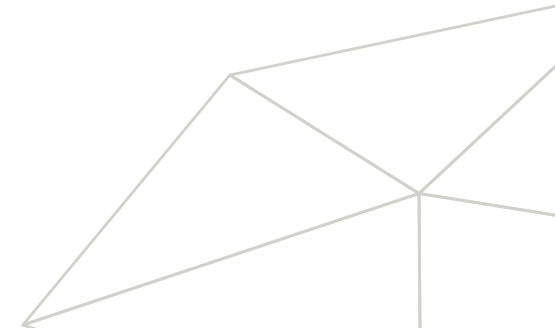
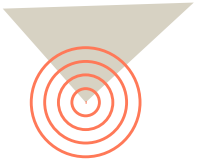
WHAT:

A collaboration in a virtual setting among Action researchers or innovators to exchange knowledge, learn new techniques, etc.

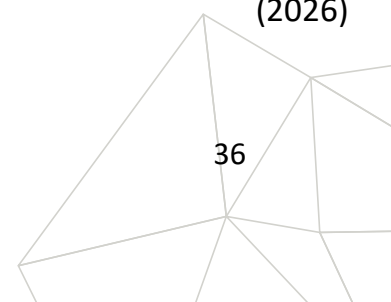
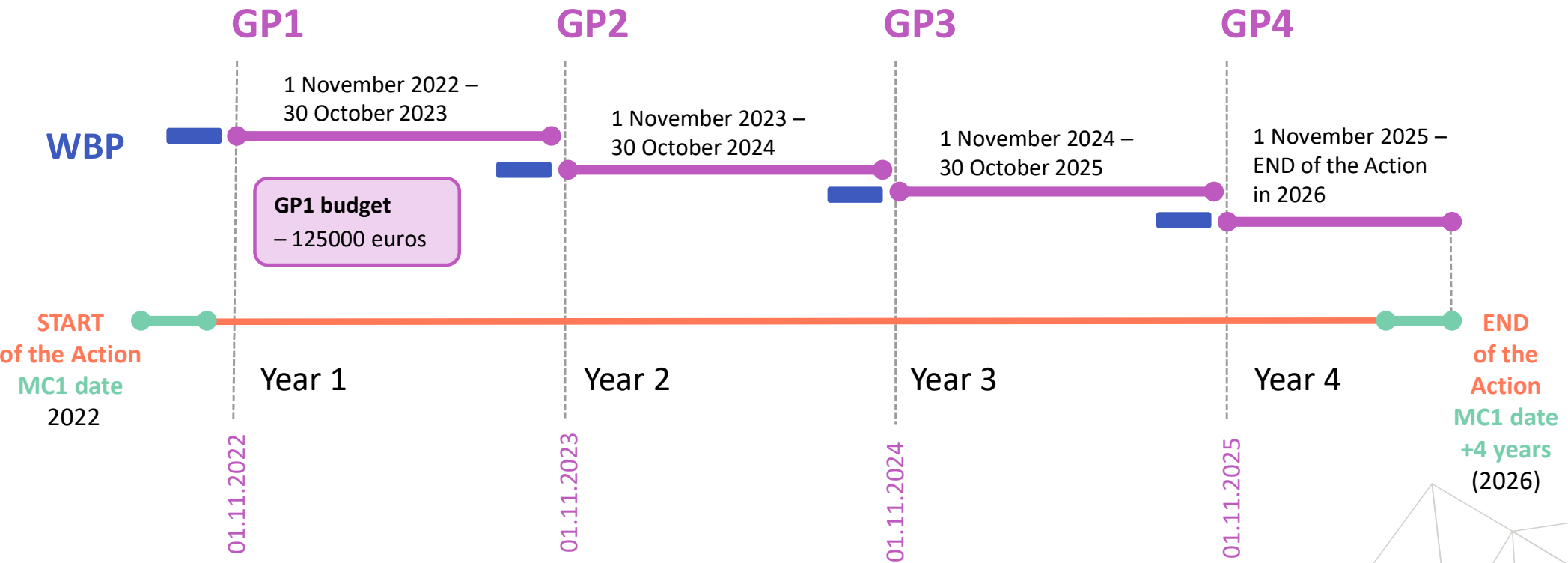
Grant AMOUNT:

Up to 1500 EUR
Paid after report approval
(after the STSM)

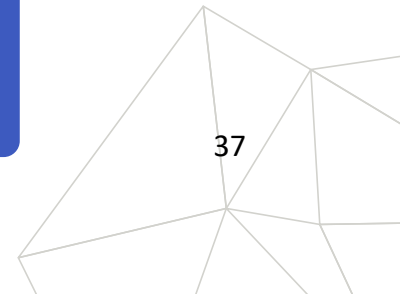
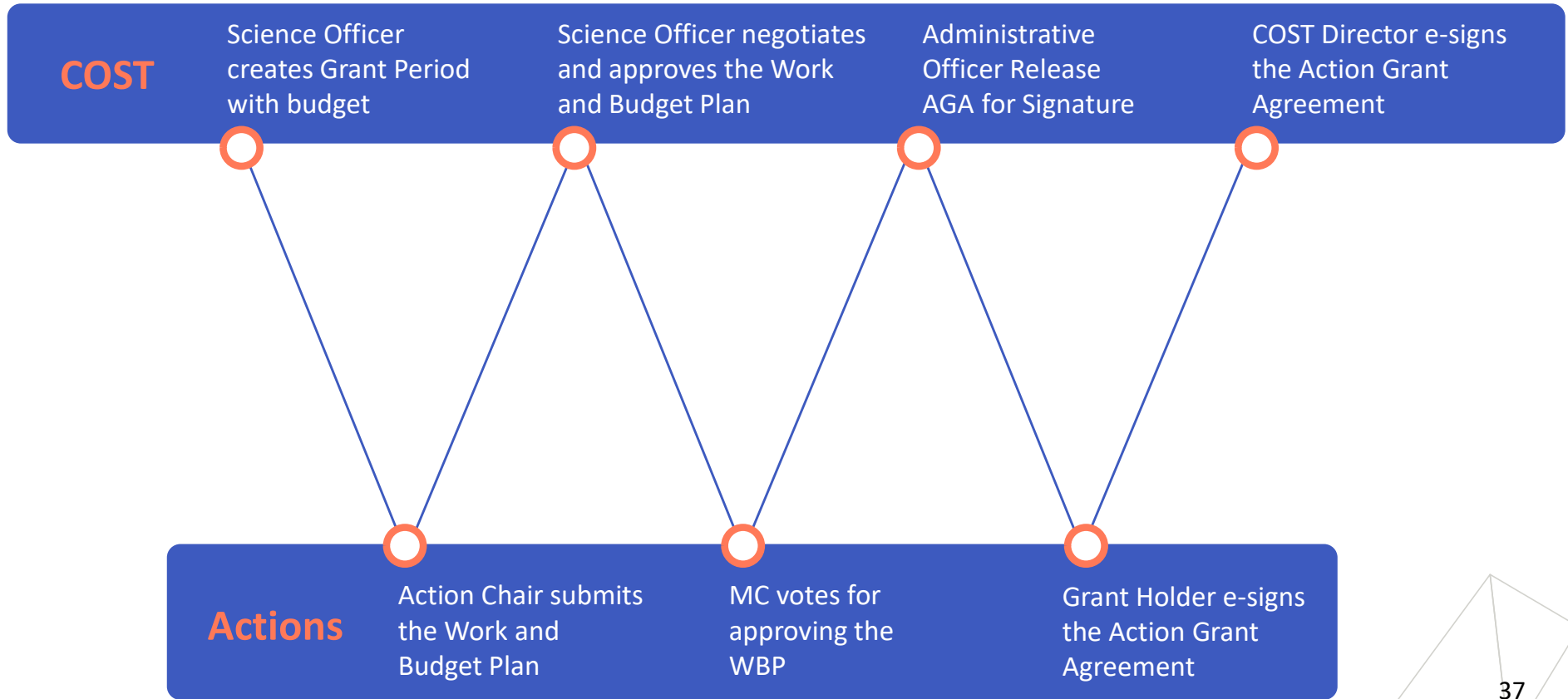
***All grants must take place within a single Grant Period.**



Grant Periods and budgets

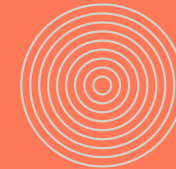


Action Grant Agreement and the Work and Budget Plan



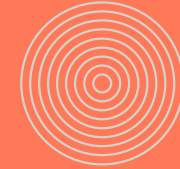
III. COST Presentation

Q&A session



Action Chair candidate

IV. Presentation of the Action



UniDive

Universality, diversity and idiosyncrasy in language technology

COST Action CA 21167

Agata Savary and colleagues

23 Sept 2022

Landscape

Language and technology

- Language diversity – vital heritage to be preserved
- Language technology (NLP) – flourishing data science branch

Inter-language diversity in NLP

- Overwhelming dominance of **English** (Bender 2011)
- Booming **multilingual** NLP
- Data **scarceness** (low-resourced and endangered languages)

Intra-language diversity in NLP

- Most linguistic phenomena have a **Zipfian** distribution
- Data **sparseness** - most phenomena are individually infrequent
- NLP favors **few frequent** phenomena and under-performs in the **many rare** ones
- **Rare** phenomena are **interesting** (unbounded dependencies, idiosyncrasies in multiword expressions)

Landscape

Fragmentation

- Linguistic phenomena appear "on the surface" in **language-specific** ways
- An expert can only reason in terms of **a few languages** she is familiar with
- This leads to **divergent** theories, terminologies and methods

Universality

- Major linguistic debate:
 - Long-standing tradition of **language universals** (Greenberg, 1966; Chomsky, 1976; Tallerman, 2009)
 - vs. no property is universal, **diversity** is the very nature of language (Evans & Levinson, 2009)
- **Universality** in NLP
 - **Cross-linguistically** consistent and **applicable** language descriptions
 - Focus on "**statistical universals**" while leaving room for true **peculiarities**
 - Universal Dependencies (UD, Nivre et al. 2020), PARSEME (Savary et al. 2018, Ramisch et al., 2020) and UniMorph (Kirov et al. 2018)

Universality vs. diversity

Universality (convergence of different points of view on similar phenomena) might be perceived as opposed to diversity but ...

Contributions of universality-driven initiatives to diversity

- Unification makes true **peculiarities more visible**
- Promoting **inclusiveness**:
 - centralized infrastructures
 - easy integration of new experts, grassroots contribution
 - open licenses
 - UD: 35 **endangered/vulnerable** languages, 12 **extinct**, 20 **underrepresented** in NLP although demographically strong
- **multilingual** tools - universal algorithms, language-specific input data
- **cross-lingual** tools - tools for a language benefit from resources in other languages
 - annotation transfer (Yarowsky et al., 2001; Hwa et al., 2005)
 - model transfer (Zeman and Resnik, 2008; McDonald et al., 2011; Agić et al., 2014; Ponti et al., 2018)
 - corpus sampling for variety (de Lhoneux, 2017)
- **multilingual** language **models** (Devlin et al. 2019), expected to encode statistical universals, **fine-tuned** for languages with few or no data

Objectives

General aim

To reconcile language diversity with rapid progress in language technology.

Approach

- both **inter-** and **intra-linguistic diversity** addressed
- NLP-applicable **universality** of terminologies and methodologies
- **quantifying** inter/intra-linguistic **diversity**
- universality- and diversity-driven development of **language resources and tools** for both low- and well-resourced languages

Research coordination objectives

- **quantifying** inter- and intra-linguistic diversity
- **common understanding of language universals** (across 70 languages)
- **coordinating** unified language **resources** (i.a. in UD and PARSEME)
- better coverage of inter-/intra-linguistic **diversity** in NLP **tools**
- **raise awareness** about **diversity preservation** in NLP

Capacity-building objectives

- **network** of experts (morphology, syntax, semantics) in many languages
- promotion of **young** researchers and **inclusiveness** countries
- coordinating **universality**-driven initiatives (also outside Europe)
- **roadmap** for the joint efforts of the universality-driven NLP community

WG1: Corpus annotation

Work program

- Studies and community discussions on morphology, syntax and semantics, with special attention to **idiosyncrasy**
- Unification and enhancement of cross-lingual **annotation guidelines** for morpho-syntax and MWEs
- Paving the way towards annotating **constructions**
- Coordinating the development and maintenance of centralized **software** for universality-based **corpus construction**
- Defining **file formats** for corpora annotated according to the unified guidelines
- Construction of **annotated corpora**

WG2: Lexicon-corpus interface

Electronic lexica are complementary to corpora because they aim at **holistic** language modelling, describing possibly many linguistic objects, whereas in corpora many phenomena occur **rarely or never**.

Work program

- Cross-language **unification of lexical properties**:
 - definition of a "syntactic word"
 - lemmatization rules (for words and MWEs)
 - set of lexical features
 - lists of **closed categories** (auxiliaries, pronouns and determiners)
- **Interlinking** MWE lexicon entries with their occurrences in corpora
- Standards for **lexical encoding of MWEs**

WG3: Multilingual and cross-lingual language technology

Work program

- **Syntactic parsers** with better coverage of rare phenomena
- Prototypes of **semantic parsers** which resolve idiosyncrasies
- **MWE discovery** tools to compensate for MWE sparseness in annotated corpora
- **MWE identifiers** which coupled with MWE discovery and lexica to better handle rare/unseen data
- Prototypes of **identifiers** for **idiosyncratic constructions**

WG4: Quantifying and promoting diversity

Work program

- **Networking** for diversity:
 - bringing together pre-existing groups dedicated to NLP-applicable universality
 - integrating and training experts of **new languages** and in **linguistic typology**
- **Quantifying** diversity:
 - designing **measures of inter- and intra-linguistic diversity**
 - quantify diversity in UD and PARSEME **corpora**
- **Promoting** diversity:
 - diversity-driven **corpus** split, sampling and augmentation,
 - diversity-driven selection of **new data** to be annotated
 - diversity-driven **evaluation** scenarios (favour tools performing well on rare and diverse phenomena and low-resourced languages)
 - coordinating corpora, lexica and NLP tools for **low-resourced and endangered languages**
 - ...

Deliverables

Coordination

- annotation guidelines
- centralized documentation of software infrastructures, file formats, annotated corpora, lexicon prototypes, NLP tools
- diversity benchmarks for NLP

Dissemination

- action's website
- reports, event material, proceedings
- joint publications
- dissemination material for large public

Recommendations from the COST Scientific Committee

From the UniDive proposal evaluation report

To comply with the COST Excellence and Inclusiveness Policy, in the implementation of the Action:

- the level of involvement of **Inclusiveness Target Countries** (ITCs) should be **maintained** and the **plans** described in the proposal for ensuring ITC involvement should be **implemented**;
- the level of involvement of **Young Researchers and Investigators** (YRIs) should be **increased** and the **plans** described in the proposal for ensuring YRI involvement should be **implemented**;
- the **gender** balance should be **improved** and the **plans** described in the proposal for ensuring gender balance should be **implemented**

Leadership positions and candidates

Leadership positions standard for COST:

- Chair, Vice-Chair, GH Scientific Representative, WG leaders, Grant Awarding Coordinator, Science Communication Coordinator

Leadership positions specific to UniDive:

- 1–2 WG vice-leaders
- Task leaders/coordinators
- Representative of Young Researchers and Investigators

Action Chair candidate

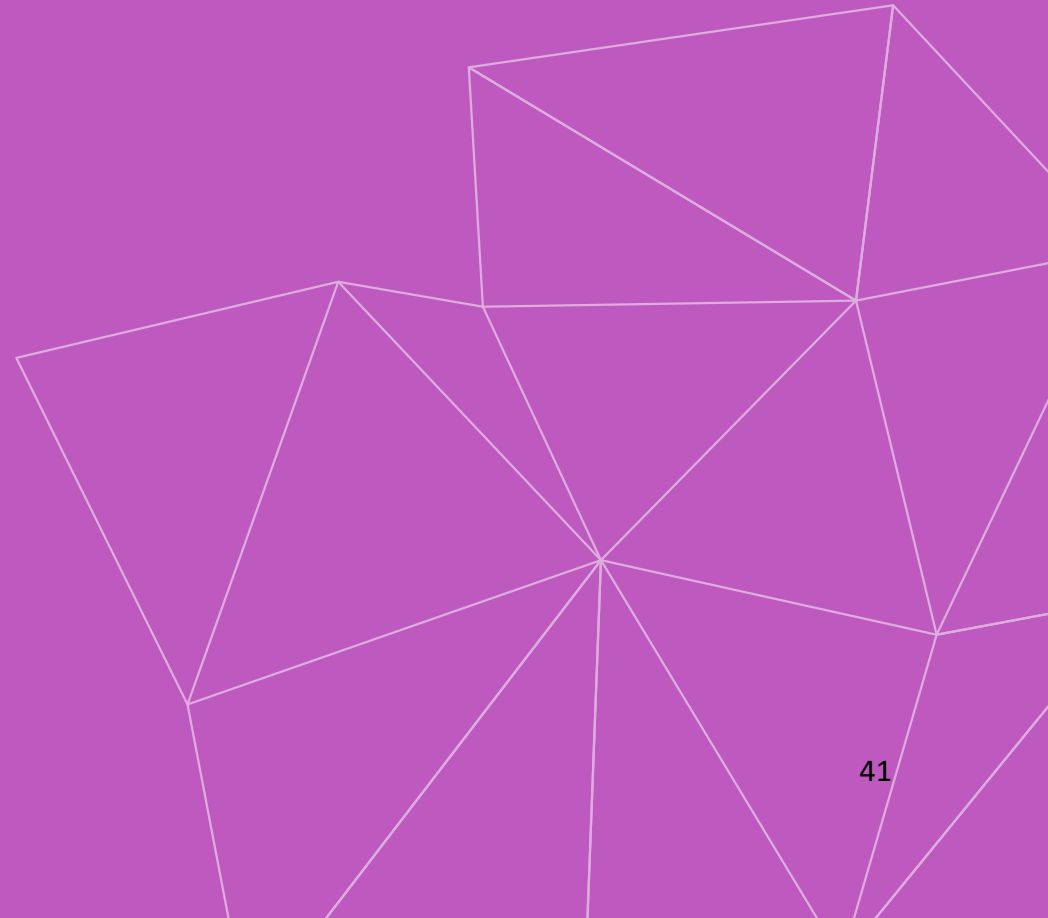
IV. Presentation of the Action

Q&A session



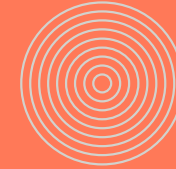
Break

We resume at 11:10

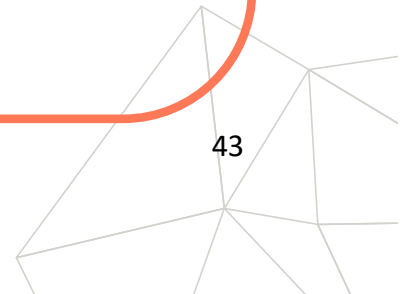
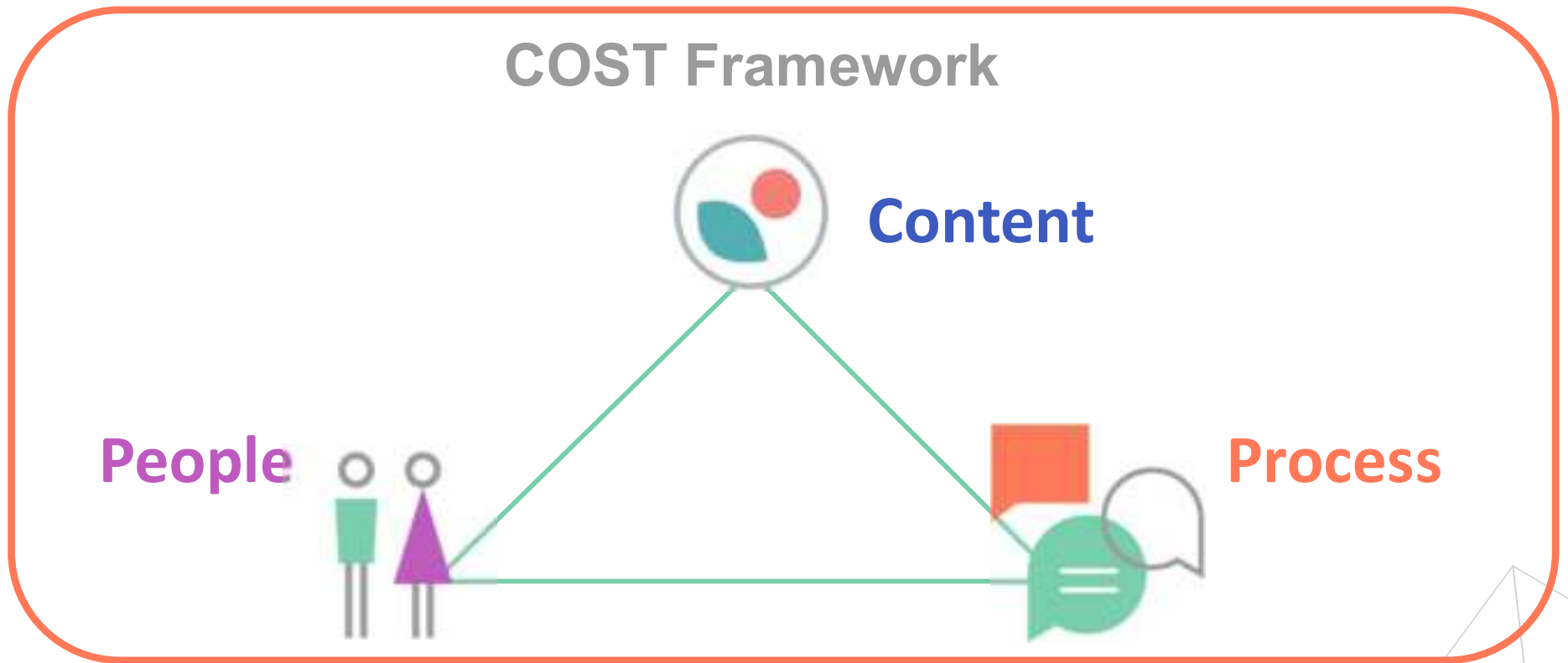


V. COST Action Café

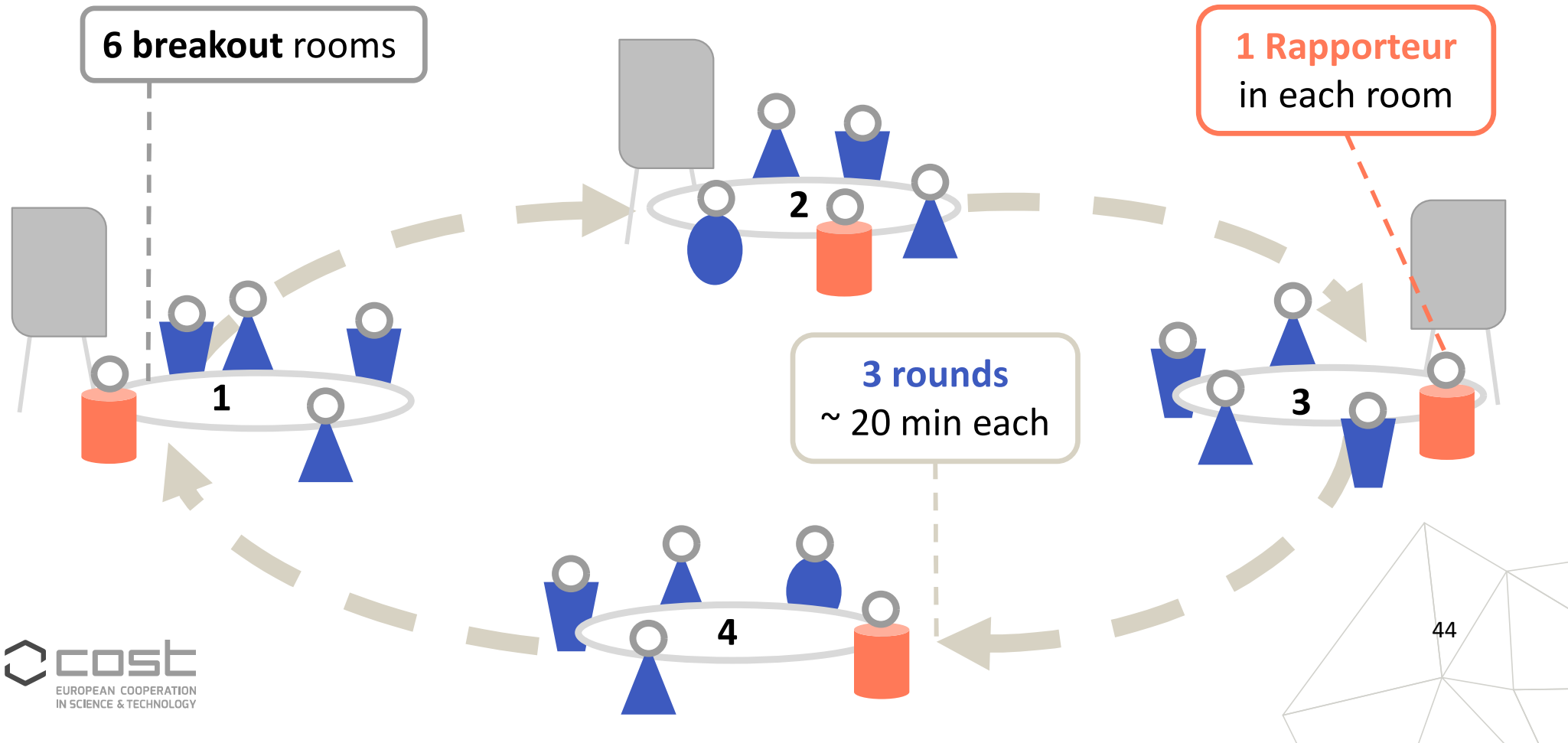
Planning for success



Key ingredients for a successful COST Action



Structured Discussion: COST Action Café

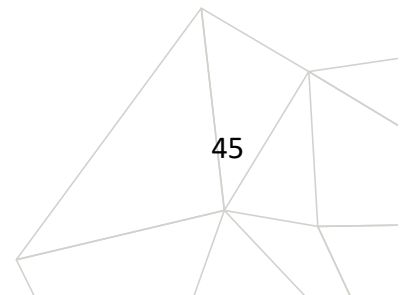


Structured Discussion: COST Action Café



Rapporteurs

- Remain in the same room for the three rounds
- Moderate the discussions (they can be assisted!)
- Keep the time & invite participants to move to other rooms after each round
- Use the standard template for guiding the discussions
- Report back in plenary after lunch break (5 minutes) on:
 - What are our next steps?
 - What do we need from the Action?

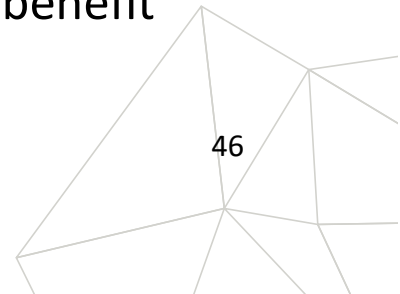


Structured Discussion: COST Action Café



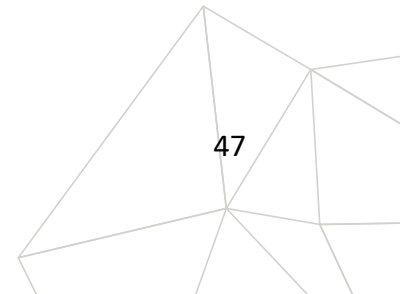
All other participants:

- Visit 1 different room per round.
- Provide active input to the overall success of the COST Action (do not go to a room only to listen).
- Maintain the number of participants in each room sufficiently big to allow different views but sufficiently small to allow everyone to speak (recommended 7-8 persons).
- Remember this is day 1 of the Action. You will be able to contribute and benefit from all topics even if you are not able to join today.



Action Number: CA21167 Universality, diversity and Idiosyncrasy in language technology (UniDive) Action MoU: link
Topic Title:
Rapporteur(s):
Meeting Participants interested in contributing to the topic: <i>(Only for the groups discussing management topics e.g. Grant Awarding, Science Communication plan or Stakeholder Engagement – WGs information on the e-COST platform.)</i>
What is our main objective and key deliverables (long-term, for the 4 years)? And our immediate goals (next 6-12 months) and desired outputs?
How can we achieve it?
Which resources do we need? <i>(e.g. budget, use of networking tools, different expertise etc.)</i>
What are the next steps?

Template for the break-out rooms

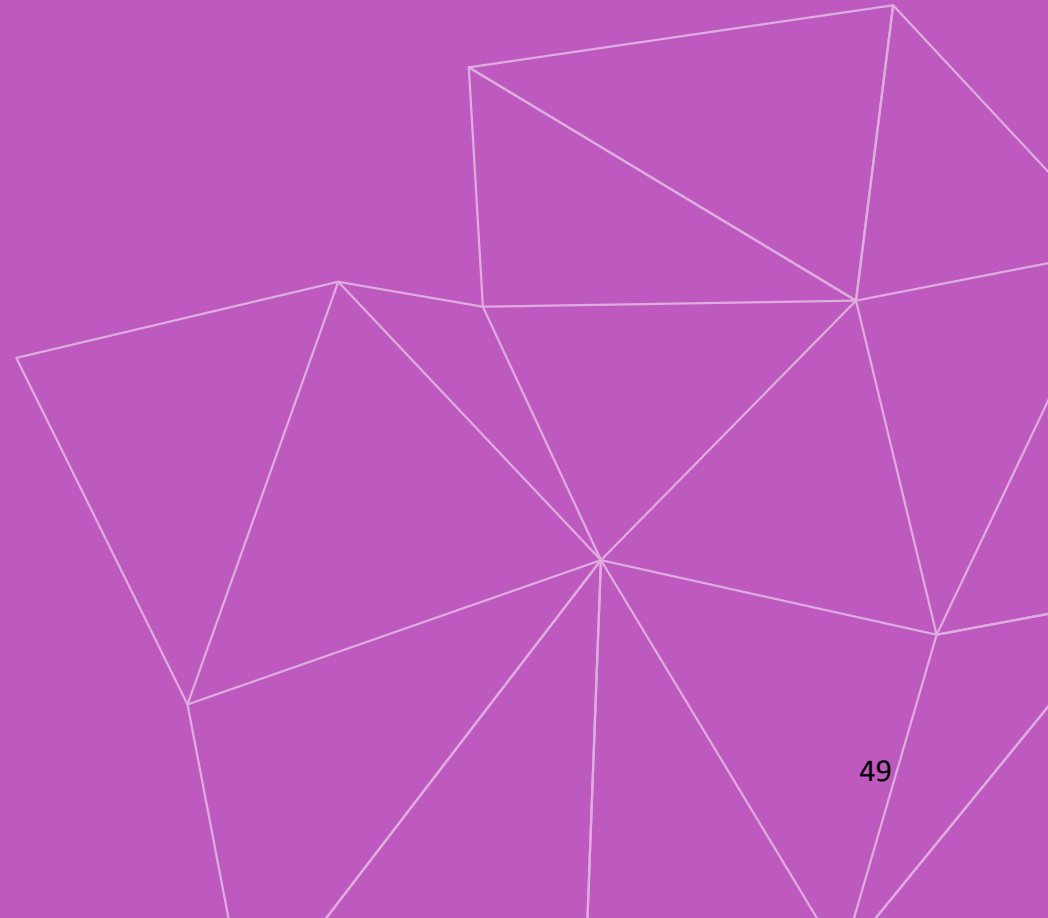


6 Breakout room topics

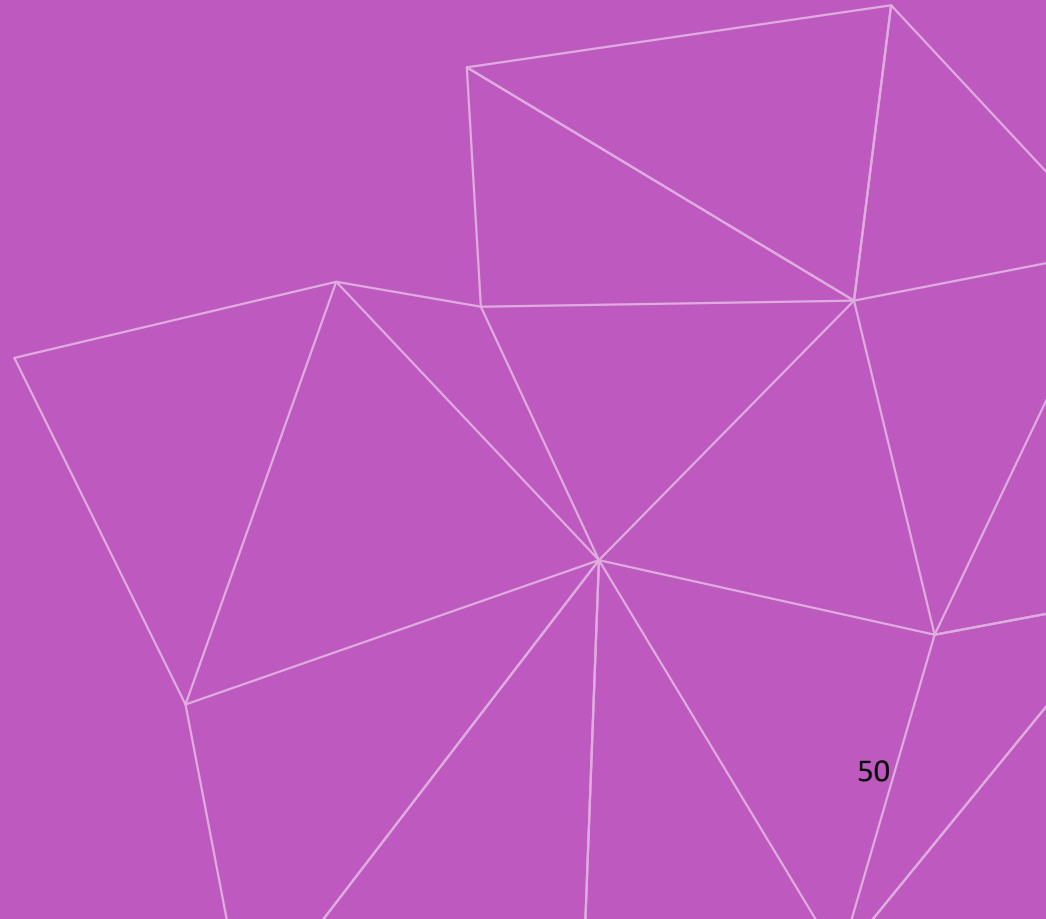


No	Topic and Room	Rapporteur Candidate
1	WG1: CORPUS ANNOTATION [<u>Marie Curie room</u>]	Carlos Ramisch, FR
2	WG2: LEXICON-CORPUS INTERFACE [<u>Marie Curie room</u>]	Verginica Barbu Mititelu, RO
3	WG3: MULTILINGUAL AND CROSS-LINGUAL LANGUAGE TECHNOLOGY [<u>Tesla room</u>]	Joakim Nivre, SE
4	WG4: QUANTIFYING AND PROMOTING DIVERSITY [<u>Tesla room</u>]	Marie-Catherine de Marneffe, BE
5	Defining a Communication Strategy [<u>Tesla room</u>]	Daniel Zeman, CZ
6	Grant Awarding - Implementing STSMs and ITC Conference Grants [<u>Jung room</u>]	Agata Savary, FR

Lunch break
We resume at 14:00

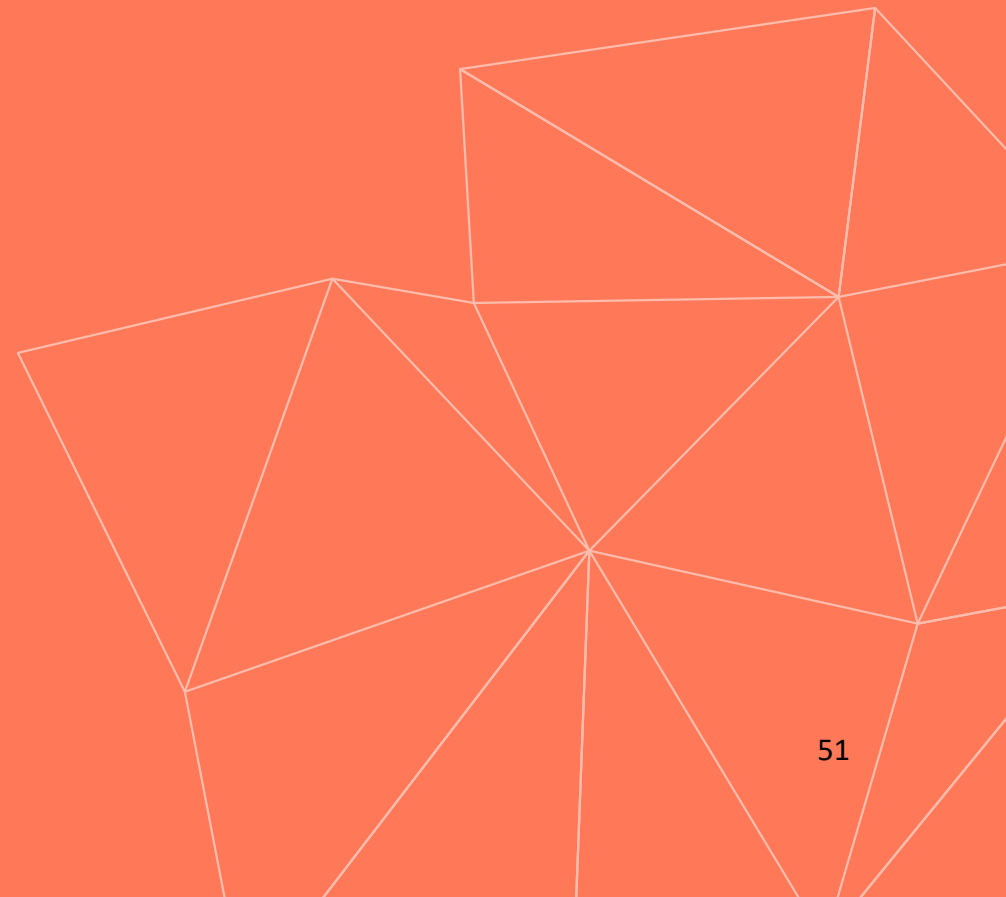


Group Photo



VI. Debrief by Rapporteurs

**What are your next steps?
What do you need from the Action?**



Debriefing on the 6 Breakout room topics



No	Topic and Room	Rapporteur
1	WG1: CORPUS ANNOTATION	Carlos Ramisch, FR
2	WG2: LEXICON-CORPUS INTERFACE	Verginica Barbu Mititelu, RO
3	WG3: MULTILINGUAL AND CROSS-LINGUAL LANGUAGE TECHNOLOGY	Joakim Nivre, SE
4	WG4: QUANTIFYING AND PROMOTING DIVERSITY	Marie-Catherine de Marneffe, BE
5	Defining a Communication Strategy	Daniel Zeman, CZ
6	Grant Awarding - Implementing STSMs and ITC Conference Grants	Agata Savary, FR



Break

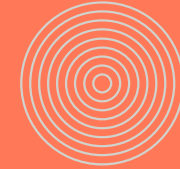
We resume at 15:30

***Please sit next to your country colleague when
back in the room to facilitate MC voting***

VII. MC Decision Making

Election of:

- Action Chair
- Action Vice Chair
- Action Grant Holder and its Scientific Representative in MC



Quorum

One vote per country
Must be done at every MC1 meeting!

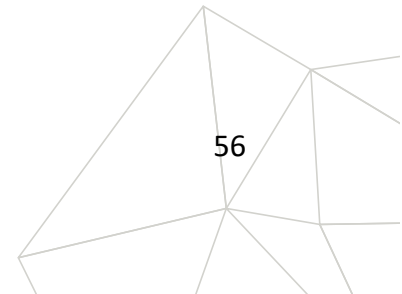


COST Action MC Decision Making



MC MEETINGS

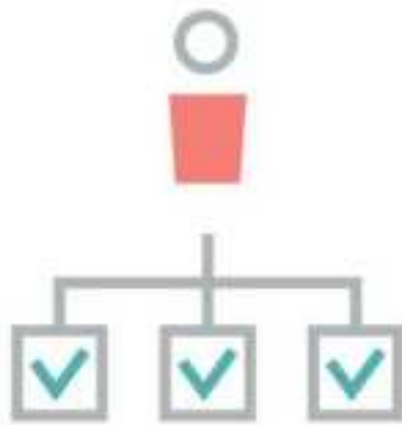
- Minimum once a year.
- Typical duration ½ day.
- Decisions only valid if at least 2/3 of the COST Countries represented in the MC are present.
- Simple majority vote of MC Members with one vote per COST Country represented at the MC.
- MC decision must be minuted and sent to COST Association.



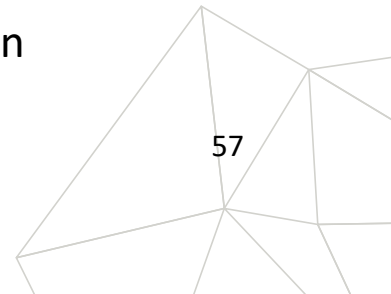
COST Action MC Decision Making



E-VOTE (in e-COST)



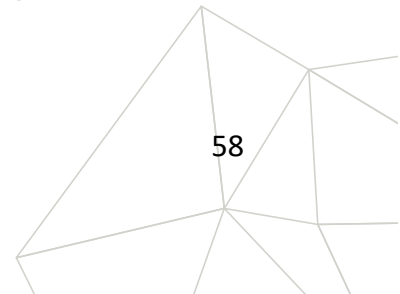
- Initiated and managed by the MC Chair (or Vice Chair if applicable)
- All MC members are automatically included.
- Vote open (7 days).
- Simple majority vote of MC Members with one vote per COST Country represented at the MC
- Abstention means tacit approval.
- MC decision automatically recorded in e-COST.



What are you voting for

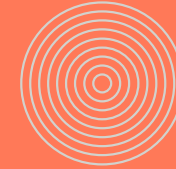


- **Action Chair:** responsible for the coordination and implementation of the Action.
- **Vice Chair:** assists in these activities where requested to do so by the Action Chair and can substitute for the Action Chair when required or mandated to do so.
- **Grant Holder:** the legal entity responsible for the administrative and financial implementation of the COST Action.



VIII. MC discussion and decisions

Elected Action Chair
Appointing Minute-taker



PART II

Plan

Minuting: Alina Wróblewska

Agenda (about 15 min. per item)

- Decision on other leadership positions
- Core Group and delegations of power
- Overview of WG applications
- Proposal for Work and Budget Plan
- Future tasks and events
- Any other business

Leadership positions

Criteria

- capacity of leaders/co-leaders to effectively work together
- balance:
 - variety of countries and affiliations,
 - Inclusiveness Target Countries (ITC) vs. other countries
 - Young Researchers and Investigators vs. senior researchers
 - female vs. male researchers

Selection of candidates

- Proposal core group (PCG) ⇒ candidates for mandatory positions
- PCG: WG applications ⇒ short list of candidates for WG leaders
- WG leader candidates: WG applications ⇒ short lists of candidates for WG co-leaders

Call for candidates

All positions are open. Spontaneous candidates are encouraged to express interest (see next slides).

Candidates

Already elected

Chair (France), Vice-Chair (Czech Republic) and GH Scientific Repr. (Poland)

Open positions (mandatory positions underlined, YRIs in **green**)

- WG1 (Corpus annotation):
 - Leader: **Carlos Ramisch** (France)
 - Co-leader: **Kaja Dobrovoljc** (Slovenia)
- WG2 (Lexicon-corpus interface):
 - Leader: **Verginica Barbu Mititelu** (Romania)
 - Co-leader: **Voula Giouli** (Greece)
- WG3 (Multilingual and cross-lingual language technology):
 - Leader: **Joakim Nivre** (Sweden)
 - Co-leader: **Gülşen Eryiğit** (Turkey)
- WG4 (Quantifying and promoting diversity):
 - Leader: **Marie-Catherine de Marneffe** (Belgium)
 - Co-leader: **Abigail Walsh** (Ireland)
- Grant Awarding Coordinator: **open position**
- Science Communication Coordinator: **open position**

Positions to remain open after MC1

- Second WG co-leaders (especially for large WGs)
- Representative of Young Researchers and Investigators
- Task leaders (whenever tasks are defined in or across WGs)

Candidates should contact the core group. Elections will be done by the MC.

Decision making

Other candidates for any position?

Discussion

Vote!

Establishment of the Core Group

Core Group

- Mandatory roles: Chair, Vice-Chair, GH Scientific Repr., WG leaders, Grant Awarding Coordinator, Science Communication Coordinator (9)
- + YRI representative? (10)
- + WG co-leaders? (18)

Core Group meetings

4 times a year, mostly online/hybrid

In favor of establishing the Core Group?

Vote!

Delegation of powers

Some MC rights and powers

- *Elect/replace Core Group members or the GH*
 - *Approval of the Work and Budget Plan*
 - *Exclude members due to a breach of conduct*
 - *Define criteria for WG membership*
 - *Tacitly validate or explicitly refuse (within 4 weeks) new MC Members after the kick-off*
 - Approve/refuse WG applications ⇒ **delegate to the Core Group?**
 - Selecting event participants who will be reimbursed ⇒ **delegate to the Core Group?**
 - Adjusting the location and dates of activities ⇒ **delegate to the Core Group (in close consultation with the Local Organizer)?**
- (*) Those in *italics* cannot be delegated.

Discussion ... Vote!

WG applications - overview

As of Sept 20, 2022

- 140 applications (huge potential!)
- 46 MC members, 94 others
- 133 from COST members countries, 7 from other countries
- 74 from ITCs, 1 from NNC
- 88 female, 51 male
- **47 YRIs (34%)**
- 96 WG1, 69 WG2, 80 WG3, 54 WG4

Work and budget plan for year 1

General aim and objectives (MoU)

Aims/primary Objective

reconcile language diversity with rapid progress in language technology

Secondary objectives

1. To develop methods for quantifying inter- and intra-linguistic diversity.
2. To develop a common understanding of language universals across 70 languages represented in the Action.
3. To coordinate the diversity-driven creation, merging and enhancement of language resources unified across over 100 languages from the UD and PARSEME collections.
4. To coordinate efforts towards a better coverage of inter-/intra-linguistic diversity in NLP tools.
5. To raise awareness of the international community about the importance of diversity preservation in language technology.
6. To disseminate the Action outcomes to stakeholders.
7. To create a network of experts in a large number of languages working on modeling and processing of morphological, syntactic and semantic phenomena within a common framework.
8. To foster the capacities of Young Researchers and Innovators (YRIs), with special focus on COST ITC participants.
9. To coordinate and boost universality-driven initiatives worldwide.
10. To set up a long-term roadmap for the joint efforts of the universality-driven NLP community.

Work and budget plan for year 1

Year 1 goals

- Establishing working relationships and structuring the community around the Working Groups
- Planning the activity of each Working Group
- Establishing internal and external communication means
- Extending the Action to new potential countries, languages and dialects
- Understanding the state of the art for the work program of each WG
- Facilitating and coordinating the development of language resources and tools for new under-resourced languages

Work and budget plan for year 1

Year 1 budget

A. COST Networking Tools	EUR
(1) Meetings	82,000.00
(2) Training Schools	0.00
(3) Mobility of Researchers and Innovators	18,200.00
(4) Presentation at Conferences organised by Third Parties	1,282.00
(5) Dissemination and Communication Products	0.00.00
(6) Other Expenses Related to Scientific Activities (ODRFA)	1,000.00
B. Total Science Expenditure (sum of (1) to (6))	102,482.00
C. Financial and Scientific Administration and Coordination (FSAC) (max. of 15% of B)	0.00
Total Grant (B+C)	102,482.00

Future events

Proposal for the on-site event of year 1

- On-site meeting, within the 1st general meeting: **13–14 February 2023**, University of Paris-Saclay, France
- Alternative dates: 20–21 Feb (Mon-Tue), 9–10 March (Thu-Fri), 13–14 March (Mon-Tue), 16–17 March (Thu-Fri)

Other events

- Dagstuhl Seminar: 8–12 May
- *Core Group meetings: November, February (on-site), May, August*
- WG meeting: August-September (hybrid)
- *webinar (online)*
- STSMs, conferences

In italic if no extra budget needed

First tasks

- Defining the **WG membership** rules
- Instruments to trigger contributions and manage expectations (e.g. a survey)
- Engaging members (webinars, ...)
- Calls for grant applications: STSMs, conference grants (late 2022)
- A meta paper about the Action

Any other business?

Calls for grant applications

- Informal dinner tonight
-
-

Check-out

Closing

