|  |  |
| --- | --- |
| codatalogo-ungrouped.eps | **Task Group Proposal for Presentation to the 30th CODATA General Assembly**  **Denver, CO., USA, 11 September 2016** |

This document invites proposals for Task Groups to be approved by the 30th CODATA General Assembly which will precede International Data Week (comprising our conference SciDataCon 2016).

**The deadline for submissions of Task Group Proposals is 11 July. Proposals will be reviewed by the CODATA Executive Committee and by external experts. Recommendations will be presented to the CODATA General Assembly which will select the Task Groups.**

**Existing CODATA Task Groups may apply for renewal and are subject to the same criteria, as described below.**

### Task Groups and the CODATA Strategic Plan

Task Groups are an important means through which CODATA delivers on its mission and Strategic Plan.[[1]](#footnote-1) It is imperative that Task Groups should have clearly defined outputs that contribute to one or more of the strategic priorities that are laid out in that plan.

Those three priority areas are as follows:

1. **To promote Open Data[[2]](#footnote-2) principles, policies and practices**

Although science is an intrinsically international enterprise, and although international programmes that address global challenges are increasingly important, researchers work within national systems of organization, funding and priority setting. There is a need for: a framework of international agreements, practices or standards codified in the context of subject-specific conventions; national policies and practices for funding and incentivising research, and for physical and software infrastructure; and institutional policies and practices through which open data curation and support for researchers are managed.

1. **To advance the frontiers of data science[[3]](#footnote-3) and its adaptation to scientific research;**

Data science is an increasingly widely and, therefore, ambiguously used and misused term. The profound transformations in the conduct of research associated with the ‘Data Revolution’ mean that there is an enormous and urgent need for a systematic study of the role of data in research. The transformations raise issues that affect science policy, the conduct and methods of research and the data systems, standards and infrastructure that are integral to research. The evidence-based study of these things can correctly be called ‘data science’.

Two areas of activity are particularly important: 1) research into the basic science of data analysis, to ensure that the inferences drawn from data are statistically valid; and, 2) research into those processes in relation to data that enable the research lifecycle to work efficiently and that ensure that the burden of data management does not fall excessively on the shoulders of domain researchers.

1. **To build capacity for improving skills and the functioning of science systems (particularly in low and middle income countries - LMICs)**

Exploiting the data revolution to maximum national and international benefit requires substantial investment in human capital, both in the training of a new cohort of data professionals and in the education of domain researchers. This need is particularly important in low and middle income countries in order that the ‘data revolution’ does not result in even greater divides between these countries and those with strongly resourced research systems.

**Proposals should state clearly and specifically how the Task Group will contribute to the realisation of one or more of these objectives.**

## High Priority Topics

To ensure that Task Groups contribute as much as possible to the realisation of CODATA’s strategy, the following high priority topics are suggested as examples. Task Group proposals addressing these or similar issues are particularly encouraged. This does not mean that Task Group proposals addressing areas not described below will be considered out of scope: we invite and encourage innovative ideas from the CODATA community.

### Open data principles, policies and practices

**Further development of good practice and recommendations around data principles and policies.** Building on the principles laid out in the [Science International Accord on *Open Data in a Big Data World*](http://www.icsu.org/science-international/accord) and the [CODATA Guide to *Best Practice for Research Data Management Policies*](http://dx.doi.org/10.5281/zenodo.27872) there is important work to be done to compile, analyse and promote good practice in data policies for funders, research institutions, research disciplines, journal editorial policies etc. Such work relates to and could augment the activities of the CODATA Data Policy Committee.

Other important areas requiring work include:

1. examining the **limits of openness** (whether to do with commercial, personal or security concerns);
2. synthesizing and advancing work on the **benefits of data sharing** (for example the economic benefits, or examples of societal impact and benefits);
3. exploring the theme of **data diplomacy** to consider what international agreements can promote the more open availability of data, particularly in relation to research issues of major global importance or to data sharing in countries with less well resourced science systems.

## Frontiers of Data Science

**Reproducibility in research.** The ability to reproduce results is a fundamental feature of the scientific method. It is a feature that depends on the availability of data for reanalysis. Work is needed at an international level to develop definitions, standards and criteria for reproducibility and to designate the practices around data that promote reproducibility. The specificity of particular disciplinary approaches may mean that this work should be done in partnership with particular International Scientific Unions.

**Statistical, epistemological and ethical challenges of Big Data:** Relatedly, there is a need for a clear sighted and international examination of the statistical, epistemological and ethical challenges which can be involved in Big Data. The focus might relate to the methodological issues of drawing valid inferences from Big Data or the ethical challenges posed by ubiquitous sensors and the Internet of Things.

**Criteria for selection and preservation of research data.** What data should be kept and why? One of the ongoing major challenges in research data is to develop applicable criteria for the selection of data to be retained, published and preserved. Individual data repositories have collection policies and some very generic guides exist, however, there is need for further research into the criteria for selection. Such work might start with particular disciplines and could involve International Scientific Unions. Outputs would include published guides to selection and retention criteria in particular research areas.

**Development of glossaries, data dictionaries or ontologies.** Through a Task Group, CODATA can assist the development and curation of glossaries, data dictionaries or ontologies. A Uniform Description System for Materials on the Nanoscale (v01 at <http://dx.doi.org/10.5281/zenodo.20688>) has recently been developed as part of a CODATA Working Group. With support from ICSU, the CODATA Working Group convened a series of meetings that brought together experts from different disciplines and standards bodies: a description of the process at <http://www.codata.org/nanomaterials> and <http://www.codata.org/nanomaterials/history-and-workshop>. Similarly, CODATA will be collaborating with CASRAI and Research Data Canada and other partners to advance and curate the Research Data Glossary <http://dictionary.casrai.org/Category:Research_Data_Domain>. Such activities should be core business for CODATA, particularly in partnership with International Scientific Unions, ICSU sponsored programmes or other international bodies like GEO. Proposals in this area for joint Task Groups with International Scientific Unions are particularly strongly encouraged.

**Data integration and data strategies for international, transdisciplinary research programmes:** Research programmes that tackle the most pressing environmental and societal issues facing humanity are necessarily transdisciplinary and involve gathering data from many diverse sources. This raises a number of important challenges relating to data. These include 1) the challenge of integrated analysis across diverse datasets, and 2) the ongoing question of the availability of appropriate data of sufficient quality to address the research questions. For the latter question, data audits and gap analysis around the research of particular international programmes or in relation to the scientific input to the Sustainable Development Goals or the Sendai Framework are important issues.

## Capacity building, data skills and education

CODATA has an ambitious programme of capacity building for data skills and education, including a data science capacity mobilizing initiative (to be announced) in Africa and the work of the CODATA-RDA Research Data Science Short Courses activity <http://www.codata.org/working-groups/research-data-science-summer-schools>. There are considerable opportunities for Task Groups to contribute to this strategic objective: by analyzing requirements for data education; through the development and refinement of the curriculum for ‘research data science’ and its adaptation to particular disciplinary or geographical needs; through the articulation of processes and standards for the creation of reusable online materials and the development of those materials.

**Data and Research Software.** A set of issues relating to capacity building (as well as reproducibility) are those around research software and data. There can be issues of transparency but there are also capacity building challenges relating to access to research software, the maintenance and development of Open Source research software. Such issues can be particularly acute in Lower and Middle Income Countries and are an important factor in research inequalities.

### Requirements of CODATA Task Groups

The following attributes are required of CODATA Task Groups and will be assessed in the selection process and in periodic evaluation.

**Tangible and achievable mission:** Task Groups must have clear objectives and a tangible and achievable mission. What is the issue that the TG will address? What are the outputs that the TG will produce? What will the outcomes of this activity be (i.e. how precisely will this advance the CODATA strategy)?

**TG Outputs:** Outputs should be tangible and public and should have as their intended outcome to advance a specific part of the CODATA strategy and to achieve a beneficial change in policy or practice, in the science of data or in matters relating to capacity, education and skills. Example outputs include: research papers, reports, white papers, policy documents, guides, manuals, teaching materials etc. Workshops and other activities for stakeholder engagement and dissemination are important but should have a lasting record, a public output and demonstrably lead to an outcome of the type described above.

**Partnerships and international collaboration:** CODATA TGs should concretely help improve collaboration between organizations with aligned objectives. TGs involving or in collaboration with one or more International Scientific Union, with ICSU-Sponsored Programmes and with other aligned international organizations (for example—but not limited to—the Global Young Academy, the Research Data Alliance, FORCE11, ICSTI, the World Data System, DataCite, etc, etc.) are strongly encouraged. Where this is the case, proposals should demonstrate that partner organizations have been contacted and that there is formal or in principle willingness to collaborate.

**Co-Chairs:** Task Groups must have 2-3 named co-chairs, with regional and gender diversity. Co-chairs provide direction for the TG and are the key point of contact with CODATA.

**Secretary:** In addition, to the co-chairs, it is strongly encouraged for the TG to have a named ‘secretary’ who will plan meetings, respond to requests for information, maintain the TG wiki, blog or other tool etc. This role could be taken by a co-chair, an early career member or a secondee / intern. The important issue is that there is a named person who will take primary responsibility for key aspects of TG coordination.

**Early Career Members:** It is encouraged for Task Groups to have at least two, named early career members who will be strongly involved in substantive aspects of the TGs activity. These members will be incorporated into the CODATA early career network and receive benefits from that engagement (including where available, training opportunities, bursaries, network benefits etc).

**TG Membership:** One of the benefits of CODATA TGs is to bring international perspectives and cooperation to bear on an issue. In addition to the criteria described above, therefore, TG membership should be geographically diverse. Gender diversity is also important and will be taken into account in the evaluation. TG co-chairs are expected to communicate frequently with and engage all named members.

**CODATA Funding:** CODATA designates part of its budget for Task Groups and at the discretion of the Executive Committee provides a small amount of seed funding: historically this has ranged between 1K and 10K euros per year. In recent years, allocations of 3-5K euros per year have been most common. Such funding must be used to achieve tangible TG outputs. For example, directly supporting the drafting, publication or testing of TG outputs, convening a workshop or similar is within scope; speculative attendance at a conference is not.

**External Funding:** TGs are strongly encouraged to look for funding from external sources and historically the status of being a CODATA TG has assisted with this. Proposals should indicate what steps will be taken to pursue external funding.

**Internships / Secondments:** TGs are strongly encouraged to increase capacity through internships / secondments. Effort might be lent to the TG by CODATA National Members or by partner organizations, for example. Similarly, the activity of the Task Group might be one that incorporates a topic for post-graduate or post-doctoral research. Proposers and co-chairs are strongly encouraged to explore such opportunities.

**Communication with CODATA Secretariat and TG Liaison:** CODATA provides practical support for TGs from the secretariat and in the form of a listserv, web-page, wiki, online note-taking tool, web-conferencing tools etc. TGs should take full advantage of this support to assist their activities. TGs should have regular communication and meetings to ensure progress in the delivery of outputs and mission. CODATA assigns to each TG, an Executive Committee liaison who will give advice, assist the secretariat and constructively monitor TG activities and progress. TGs must maintain good communication with the ExComm liaison, ensure that they are informed of activities and included in correspondence with the CODATA Secretariat.

### Responsibilities of CODATA Task Groups

Task Groups are expected to:

1. Contribute to the realisation of the CODATA Strategic Plan and CODATA’s overall mission;
2. Produce tangible public outputs (e.g. reports, white papers, publications, data collections, technical proposals etc) that contribute to the realisation of these objectives;
3. Help keep information about the TG and its activities on the CODATA website up to date;
4. Work with the Executive Director and Secretariat to communicate outputs and activities, including contributing material for blog posts and newsletters;
5. Take advantage of the Executive Committee Liaison for expert advice and links to the wider CODATA community;
6. Take advantage of the support offered by CODATA for Task Group management (online collaborative environments, web-based meetings etc) and provide public and transparent accounts of their activities;
7. Provide two brief progress reports (after 12 and 20 months).

### How CODATA Assesses Task Group Proposals

Task Group proposals are peer reviewed and are also assessed by the Executive Committee.

The criteria used include the following:

* Alignment of the proposal with the **Requirements of CODATA Task Groups** described above;
* The significance and merit of the proposed work and its contribution to CODATA’s mission and objectives as laid out in the Strategic Plan;
* The clarity and merit of the objectives, outputs and outcomes described;
* The feasibility of the workplan;
* The existence of appropriate collaboration with other organizations of groups, ensuring the proposed work does not duplicate other activities;
* The expertise and appropriateness of the Task Group membership, including gender balance, the participation of early career scientists, and appropriate representation from developing countries;
* The existence of other likely sources of funding and support.

On the basis of the peer review and its own assessment, the Executive Committee will make recommendations to the General Assembly.

Proposers will be required to give either a presentation or poster at the General Assembly. The presentation or poster may be given by a CODATA Member Delegate if proposers are unable to travel. The General Assembly takes this presentation and the recommendations of the Executive Committee into account when deciding which proposals to approve.

### How CODATA Supports Task Groups

Proposals for CODATA Task Groups are peer reviewed by the Executive Committee and by other experts in the community. The Executive Committee considers comments from peer reviewers and makes recommendations to the General Assembly, which may approve or reject the Task Group proposal. The decision of the General Assembly carries with it the endorsement and authority of the international CODATA community.

At the discretion of the Executive Committee, CODATA Task Groups are eligible for modest funding support (see above). It is intended that such seed funding and the endorsement of the CODATA community will help the Task Group find other support for its activities.

Task Groups will be provided with a space on the new CODATA website and tools to help collaboration. Task Groups are expected to liaise regularly with the Secretariat and to use the CODATA website to communicate their work and achievements.

CODATA Task Groups are provided with a Liaison from the Executive Committee. This Executive Committee member is considered an *ex officio* member of the Task Group and, as such, receives copies of all correspondence and outputs. The Liaison is in a position to provide expert advice and a link with the wider CODATA community in order to benefit the activities of the Task Group.

### Submission Procedure and Deadline

Applications should be made by completing the form below as completely as possible, without exceeding the page limits for each section. This form is the principal document used for evaluating proposals.

**For existing CODATA Task Groups that are applying for renewal, there is an additional section required, in which the Task Group should report on its activities, outputs and outcomes.**

Proposers are encouraged to liaise with Simon Hodson, CODATA Executive Director [simon@codata.org](mailto:simon@codata.org) as early as possible in the process for guidance, advice and to ensure criteria are met as far as possible.

### Deadline for Submissions, 11 July 2016

**The deadline for Task Group proposals is 11 July 2016. Proposals should be submitted on the form below to Simon Hodson, CODATA Executive Director at** [**simon@codata.org**](mailto:simon@codata.org)**.**

**Task Group Proposal for Presentation to the**

**30th CODATA General Assembly**

**Denver, CO., USA, 11 September 2016**

### Name of the Proposed Task Group

### Short summary of objective(s) of the Proposed Task Group

### Name and Contact Details of the Principal Proposer(s)

### What is the scientific merit of the proposed work and how does it contribute to CODATA’s mission and objectives as laid out in the Strategic Plan;

***The text should address the following questions:***

***What is the issue that the TG will address? Why is this significant and important? What are the outputs that the TG will produce (tangible things the TG has done or published)? What will the outcomes of this activity be (i.e. what is the impact and effect of these activities and how precisely will this advance the CODATA strategy, what will change as a result of the TGs work)?***

***Up to three pages.***

### Summary of outputs to be completed by the 2016 General Assembly

***The headings below are intended as a guide and may be altered. One page.***

**Articles in the CODATA Data Science Journal** [**http://www.codata.org/dsj**](http://www.codata.org/dsj) **or other refereed Journals.**

**Reports, white papers, other publications, technical proposals etc.**

**Data products, web resources, software tools etc.**

**Conferences, workshops, meetings and associated reports.**

**Other products or accomplishments**

### Organization and schedule of proposed Task Group activities

***This should cover key milestones leading to deliverables, meetings/workshops, teleconferences, the organization of any subgroups etc. Be sure to specify who is responsible for particular activities/outputs. Use of the suggested table is optional. One page.***

**Overall Schedule of Activities and Milestones**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Dates** | **Activities, milestones and deliverables** | **Owner/Organiser** | **Participants** | **Comments** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

## Membership of the Proposed Task Group

***Note that participation by scientists from around the world as members of the Task Group and in Task Group activities is strongly encouraged. Particular attention should be given to gender balance, and representation from developing countries.***

***Please give institution, area of expertise, telephone, and e-mail of each member; indicate if the individual has been invited to participate and agreed to serve, or if the individual has not yet been contacted. Please add rows to the table as needed.***

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Name** | **Institution** | **E-Mail** | **Telephone** | **Invited? Agreed? Not yet contacted?** | **Contribution to and role in Task Group (inc. co-chair, secretary, etc).** | **Area of expertise** |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

### Please provide details of early career data professionals that will be involved in this Task Group.

## Budget Request from CODATA International

***CODATA designates part of its budget for Task Groups and at the discretion of the Executive Committee provides a small amount of seed funding: historically this has ranged between 1K and 10K euros per year. In recent years, allocations of 3-5K euros per year have been most common. Such funding must be used to achieve tangible TG outputs. For example, directly supporting the drafting, publication or testing of TG outputs, convening a workshop or similar is within scope; speculative attendance at a conference is not.***

**Financial request from CODATA International and justification, costings**

## Financial support from sources other than CODATA International (indicate if committed or expected)

***The headings below are intended as a guide and may be altered.***

**Grants, contracts, travel expense support, etc.**

**In kind support (equivalent salaried effort, infrastructure)**

### Other organizations or programs sponsoring or participating in the Task Group and their expected role(s)

### Summary of outputs and outcomes since the 2014 General Assembly

***ONLY FOR EXISTING TASK GROUPS THAT ARE SEEKING RENEWAL***

***Highlight outputs (e.g. publications, particularly those in the CODATA Data Science Journal, reports, white papers, workshops etc and tangible activities such as meetings, workshops etc). In this report, Task Groups should also describe the broader impact and outcomes achieved and how the TG has contributed to the CODATA Strategy (i.e. what has been the effect of what you have done, what has changed?). Up to two pages.***

### Other information or comments

***Optional. No more than 1 page***

## Suggested Referees (three).

***Do not include any current CODATA Executive Committee members*** [***http://www.codata.org/about-codata/executive-committee***](http://www.codata.org/about-codata/executive-committee)***. Please include the institution, relevant area of expertise, phone number and e-mail information of each referee.***

1. Mobilising the Data Revolution: the CODATA Strategy

   <http://dx.doi.org/10.5281/zenodo.50343>; Message from CODATA President, Geoffrey Boulton <http://www.codata.org/message-from-president-geoffrey-boulton> [↑](#footnote-ref-1)
2. CODATA uses the definition of intelligently open data contained in the 2012 Royal Society Report, Science as an Open Enterprise’ <https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/> [↑](#footnote-ref-2)
3. CODATA uses the broad and inclusive definition of data science used by the Data Science Journal <http://datascience.codata.org/about/> [↑](#footnote-ref-3)