Case Statement – Neuroimaging Data WG
Dr. Etienne Roesch (Univ. Reading)

1. **WG Charter**: A concise articulation of what issues the WG will address within a 18 month time frame and what its "deliverables" or outcomes (including a Recommendation) will be.

The Neuroimaging Data WG fulfils the RDA's mission to build the social and technical bridges that enable open sharing and re-use of data in the domain of neuroimaging. The WG envisions a neuroimaging research landscape in which knowledge is generated in a reproducible fashion (in terms of data, analysis and computation) and coupled with the ability to reuse and extend these studies by others in the community. We aim to shift the way neuroimaging research is performed and reported, with the development and implementation of technology that supports reproducibility at the levels of data management, analysis and utilisation for both within- and between-lab opportunities, including the use of widely distributed and/or large populations to address basic and clinical research questions. In short, our vision is to help neuroimaging researchers: Find and Share data in a FAIR fashion (discover resources); Comprehensively describe their data and analysis workflows (describe research processes); and Manage their computational resource options (do analysis).

The stark realisation that scientific results do not always readily replicate has led some to investigate the root causes of the so-called "reproducibility crisis". Such self-critical appraisal has been so far more prevalent in Psychology and Neuroscience than in other disciplines, and typically highlight statistical issues, like inadequate statistical designs, as well as poor computational training; problems that are only likely to worsen as data grow larger, become more widely shared, and advanced techniques are imported from fields of engineering, like machine learning.

Specifically, neuroimaging data, in both clinical and fundamental research, have the particularity that they involve a large number of processing steps on a very heterogeneous set of equipment and infrastructures, from the moment they are gathered in proprietary devices (magnetic resonance imaging scanners, electro-encephalography systems, etc) through pre-processing, analysis to annotation, curation and finally deposited into open repositories for others to use in upstream research. A lot of this pipeline remains an error-prone, manual process that relies on the researcher's voluntary (and unpaid) efforts to acquire an understanding of the infrastructure available and their technical knowledge to use it, to ensure the traceability and provenance of the data, the reproducibility and replicability of the work, and the production of FAIR open datasets, and ultimately train others.

The successful integration of such data into routine neuroimaging practice thus requires neuroscientists to develop skills that fall outside of ordinary training curricula, which should also include data curation, data handling, high performance and on-demand computation (in the "cloud"), semantic web annotation, as well as statistics suitable for large scale inference. The researchers who have been the most receptive to exploring and developing such techniques are typically early career researchers, motivated by the desire to learn,

apply and share robust practices. The Neuroimaging Data WG seeks to alleviate some of the biggest challenges they face: They are not formally trained and teach themselves these new data practices following online resources, in isolation and on a voluntary basis. The WG thus fills this gap of support, by pooling interests, experiences and expertises into a platform available globally.

Objectives: A specific set of focus areas for discussion and action.

The Neuroimaging Data WG takes a train-the-trainer approach and aims to create a WG focusing on 1) creating, consolidating and providing formal training to at least fifteen aspiring trainers ("EOSC/RDA Neuroimaging Data WG Fellows"), and 2) engaging the community in an interdisciplinary examination of the gaps that remain, and identifying the resource and infrastructures available in Europe, including EOSC and JISC services. Our project is intended to be partnered with the Psychological Data IG, coordinated by Jim Grange, Andrew Stewart and Etienne Roesch, as well as national efforts, like national Reproducibility Networks (UKRN, FRRN, Swiss RN, etc) and international partners, like TESS/ELIXIS or the North American sister project ReproNim.

2. **Value Proposition**: A specific description of who will benefit from the adoption or implementation of the WG outcomes and what tangible impacts should result.

Within the year, at least fifteen EOSC/RDA Neuroimaging Data WG Fellowships will be awarded to early career researchers (ECRs) in Neuroscience engaged with open and transparent research activities at their home institutions within the EU and the UK. These fifteen Fellows complement additional fellows from the Psychological Data IG. The RDA fellowships will allow these fellows to 1) fund time for software and data management training for themselves, and 2) organise training on the principles of software and data management at their home institutions.

Fellows are selected according to their current role and experience of training students and researchers at their local institutions, and within the scope of their 1-year fellowship, we expect the Fellows to build from the support of the WG to develop a curriculum of training activities they will supervise and conduct themselves. Additionally, the Fellows will participate in the design and elaboration of online resources, to build the online presence of the WG.

Moving forward, we expect the WG to grow as the train-the-trainer approach leads to more researchers being able to consolidate their understanding of a robust data practice, and skills in teaching it, thereby growing the community within the WG.

3. **Engagement with existing work in the area**: A brief review of related work and plan for engagement with any other activities in the area.

This WG has been proposed by Dr Roesch (U. Reading, UK), who acts as Coordinator of the project. He is helped by Prof Andrew Stewart (U. Manchester, UK, Co-coordinator of the

Psychological Data IG, Dr Michael Dayan (U. Geneva, CH), Dr Karolina Finc (Nicolaus Copernicus U., PL), Dr Camille Maumet (Inria, FR), Dr Nicholas Hedger (U. Reading, UK) and Dr Romain Valabrègue (Paris Brain Institute, FR). An Advisory Board provides friendly advice, and includes Prof Carole Goble (U Manchester, UK), Prof David Kennedy (U. Mass, USA) and Prof Jean-Baptiste Poline (McGill, CA).

All partners on the WG are involved in both national and international efforts, complementary and in support of the WG. These include national Reproducibility Networks, national and international programmes.

Specifically, Dr Roesch and Prof Stewart are the main partners involved. They are both co-investigators on the 5-year Research England Development (RED) Fund programme of research "Growing and Embedding Open Research in Institutional Practice and Culture". This £8.5m project aims to increase the skills base amongst UK researchers (across disciplines) to allow them to adopt transparent research practices in their own programmes of research. Andrew is also a Fellow of the Software Sustainability Institute (whose focus is on the recognition of the importance of research software, including the provision of best-practice software development skills, in research in the research environment). Dr Roesch is on the UK Reproducibility Network steering group. Both Dr Roesch and Prof Stewart are UKRN Institutional Leads representing their home institutions, and both are qualified Carpentry Instructors. Dr Stewart is also a Fellow of the Sustainable Software Institute. This WG will interact closely with the UKRN and other national and international programmes. It will also interact with partners such as ELIXIR (the co-lead for which sits on our advisory board). Initial activity of this WG will be funded via the RDA/EOSC Community of Practice award, and via interactions with other Psychologists within the UKRN.

Both the Neuroimaging Data WG and the Psychological Data IG will be hosted on a common platform, to consolidate material and engagement with the community in Europe and globally. Both projects will engage with their Fellows and respective communities according to their own schedule of work, composed of tailored training sessions and workshops, and at least once a year, will organise a General Assembly with members of projects, to assess activities and make plans for the future.

4. **Work Plan**: A specific and detailed description of how the WG will operate including:
   - The form and description of the final Recommendation of the WG,
   - The form and description of milestones and intermediate documents, code or other deliverables that will be developed during the course of the WG's work,
   - A description of the WG's mode and frequency of operation (e.g. on-line and/or on-site, how frequently will the group meet, etc.),
   - A description of how the WG plans to develop consensus, address conflicts, stay on track and within scope, and move forward during operation, and
   - A description of the WG's planned approach to broader community engagement and participation.

Fellows will meet monthly for sustained interaction and community building. Tailored activities, workshops and training will be organised as and when possible, and include

software carpentry workshops, interventions by partners and relevant IGs/WGs/CoPs and research projects.

<u>Timeline: Describe draft milestones and goals for the first 18 months, including the mandatory Public report to be submitted to Council for review every 18 months.</u>

Date	Milestone Description	Notes
Start date as agreed in contract (Month 0)	Start of the project	
M+1	Adverts circulated	Starts 8 weeks period for applicants to submit (including motivations, experience, scope of community, and letter of support by host institution)
M+4	Decision announced	Email invitations, institutions send invoices to Manchester and Reading to receive Fellowship funding
M+5	All fellowships paid	
M+5	Kick off event	
Every month	Meet up	Alternating between round table and expert-led discussion (together psych + neuro) using Gather Town. These experts will include representatives from RDA, EOSC, UKRN, SSI, and related organisations.
Ongoing	Keeping in touch via Gather Town	
Ongoing	Fellow organised (PI, Co-I facilitated) training sessions	
M +17	Reports from each of the funded fellows on training experience, evidence of engagement, feedback etc	
M + 18	Our report to RDA/EOSC	

5. **Adoption Plan**: A specific plan for adoption or implementation of the WG Recommendation and other outcomes within the organizations and institutions

represented by WG members, as well as plans for adoption more broadly within the community. Such adoption or implementation should start within the 18 month timeframe before the WG is complete.

The Neuroimaging Data WG Fellows will be selected on the basis of their willingness and capacity to engage and share the output of the WG, and the funding requested will enable the allocation of their time to do so, in the period of the project. The Train-the-Trainer scheme proposed will assure the sustainable growth of activity in the WG, and maximise exposure of the WG's deliverables. All deliverables will be made available on a dedicated website, reviewed and consolidated by the community at large. Further, activities throughout the year will be organised to establish explicit links with similar projects, by co-hosting seminars and training sessions.

The WG is supported by several senior academics, who have established links with similar projects nationally and internationally. The Neuroimaging Data WG will thus be joining a growing community of researchers committed to evaluation, design and propose better data practices in the field of neuroimaging. Particularly, the WG will be engaging with national reproducibility networks, currently established in Australia, Brazil, Finland, German, Italy, Norway, Portugal, Slovakia, Sweden, Switzerland and the United Kingdom. We will reach out to related research centres in countries not represented, such as the Cape Universities Brain Imaging Centre (Cape Town, South Africa) and Cairo University, as well as provided training to institutions that do not have imaging facilities but are likely to want to access and analyse secondary datasets.

Additionally, initial participants in the WG include individuals with projects in Canada, the United States of America, South Africa, as well as international organisations such as UNESCO. Reproducibility in psychology and neuroimaging being a global issue, we thus anticipate the WG to grow a global reach accordingly.

6. **Initial Membership**: A specific list of initial members of the WG and a description of initial leadership of the WG.

This WG has been proposed by Dr Roesch (U. Reading, UK), who acts as Coordinator of the project. He is helped by Prof Andrew Stewart (U. Manchester, UK, Coordinator of the WG "Building a Psychology RDA Community of Practice"), Dr Michael Dayan (U. Geneva, CH), Dr Karolina Finc (Nicolaus Copernicus U., PL), Dr Camille Maumet (Inria, FR), Dr Nicholas Hedger (U. Reading, UK) and Dr Romain Valabrègue (Paris Brain Institute, FR). An Advisory Board provides friendly advice, and includes Prof Carole Goble (U Manchester, UK), Prof David Kennedy (U. Mass, USA) and Prof Jean-Baptiste Poline (McGill, CA).
The University of Reading is the main supportive institution, and will manage finances. Leveraging the roles of the Coordinators in partnering institutions and projects, and that of the Advisory Board, we will ensure the WG is engage with the relevant communities and aware of developments in the field.