**Application form for Fostering Open Science Practice Fund**

**Closing date: 20 May 2022**

The Open Science Fund is an opportunity for **Utrecht University** and **University Medical Centre Utrecht** employees to access small grants with which they can apply Open Science principles into their research. This funding amounts to € 10.000 (minimum) - € 15.000 (maximum) per application.

**Contact and information**

If you are considering an application and you would like to discuss this with a member of the Open Science Programme team, please send a mail to [openscience@uu.nl](mailto:openscience@uu.nl) or contact [Judith de Haan](https://www.uu.nl/medewerkers/JJdeHaan), programme director.

More information, such as selection criteria, who can apply and the selection process, can be found on the [fund website](https://www.uu.nl/onderzoek/open-science/fostering-open-science-fund).

|  |  |  |  |
| --- | --- | --- | --- |
| **Names** | Gerko Vink and Hanne Oberman | | |
| **Position/role** | Associate Professor / Junior Researcher | | |
| **Department** | Methodology and Statistics | | |
| **Faculty** | Social and Behavioural Sciences | | |
| **Email address** | [g.vink@uu.nl](mailto:g.vink@uu.nl) | | |
| **Telephone number** | 0624111054 | | |
| **Title of proposed project** | automatic report generation for incomplete data analysis | | |
| **Project start date** | 01/09/2022 | **Project end date** | 30/08/2023 |
| **WBS number** | KEVIN VRAGEN | | |

**Please provide a summary of your project (max. 100 words):**

*(to describe the project on our website)*

|  |
| --- |
| Imputation has become the backbone of contemporary incomplete data analysis. In imputation, algorithms are used to effectively replace missing values with data that could have been. The proper evaluation of the used imputation procedure is always omitted from scientific manuscripts, as it would take up too much space. This leads to a sub-optimal science and poses challenges for authors, reviewers, publishers and readers. We implement a method that standardizes the necessary evaluations by automatically generating a report for a given analysis model. The resulting report can be added as a supplement to any manuscript wherein imputation has been used. |

**Please outline the proposed project, including the *purpose* of Open Science Practice, the specific** [***topic***](https://www.uu.nl/en/research/open-science/topics) ***it addresses*, the *approach* being taken and the *links* to research’ (max. 500 words):**

|  |
| --- |
| We develop an automated report generator for studies wherein imputation techniques are used to solve for missing values. With imputation, the process of evaluating the validity and plausibility of the resulting imputations and algorithms is vital for the interpretation of submitted manuscripts and scientific results. This poses challenges for authors, reviewers, publishers and readers, as a proper evaluation process is often lengthy and would take up valuable space in manuscripts. We implement a method that standardizes such evaluations by automatically generating a report for a given analysis model. The resulting report can be added as a supplement to any manuscript wherein imputation has been used. With our method, anyone can interpret the utilized imputation procedure and evaluate its validity, even when access to the original data or the computer code is restricted.  Purpose of open science practice   * The report protects privacy of respondents * The report promotes the core values of open science and dissemination * The resulting methodology will be open, *community-driven and implemented in the open-source software package mice*, which is the de facto standard for the analysis of incomplete data sets. Voortschrijdend inzicht kan de reports beter maken. * Nog iets meer over purpose met \*dit moet je echt vertellen bla-bla praatje\*   Approach being taken   * The resulting methodology will be open, community-driven and implemented in the open-source software package mice, which is the de facto standard for the analysis of incomplete data sets. * Deliverables formuleren: convergentie/densityplots/marginalen en conditionals, correlatietabbellen/nieuwe plaatje/analysemodelplot/pooled en individual results * Welke software ondersteunen we vanaf begin: mice::mids object en .txt in long format en iteration\_history. Dan hebben we alle software zo’n beetje te pakken * In R; evt met shiny facility for ease of use. * TOPIC is FAIR data and software |

**How will you evaluate the progress, outcomes and impact of your project? How will these results be shared? (max. 300 words)**

|  |
| --- |
| Progress en outcomes   * Deliverables formuleren: convergentie/densityplots/marginalen en conditionals, correlatietabbellen/nieuwe plaatje/analysemodelplot/pooled en individual results   Impact:   * Downloads on CRAN en GitHub   Dissemination   * In R – perhaps not in mice because of impact monitoring. Perhaps separate package to aid visibility with function in mice that runs the relevant package code: REPORTER package with REPORTER::imputation() as function. REPORTER allows for more standardized implementations than imputation alone. |
|  |

**Please describe the potential for learning and/or development for researchers (max. 150 words):**

|  |
| --- |
| Nog doen |

**Please detail the amount of funding applied for and justify the costs requested:**

|  |
| --- |
| Urenraming maken |

**Please send the completed application form to** [**openscience@uu.nl**](mailto:openscience@uu.nl) **by 20 May 2022.**