

# FWO DMP Template - Flemish Standard Data Management Plan

## Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following [link](#).

1. General Project Information	
Name Grant Holder & ORCID	Nicolas Scepi Orcid: 0000-0003-3909-2486
Contributor name(s) (+ ORCID) & roles	Fabio Bacchini Orcid: 0000-0002-7526-8154
Project number <sup>1</sup> & title	Project number: 12AVW24N Towards the most realistic radiative models of luminous black hole accretion
Funder(s) GrantID <sup>2</sup>	
Affiliation(s)	KU Leuven  ROR identifier KU Leuven: 05f950310

---

<sup>1</sup> "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

<sup>2</sup> Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

Please provide a short project description	<p>Black holes (BHs) are typically surrounded by plasma (a tenuous, hot gas of charged particles) that accretes onto them, i.e. falls inward. As accretion proceeds, plasmas are heated and emit light, which we observe on Earth. Most studies of BHs focus on a minority population of luminous BHs. Even for these luminous BHs, we have reliable interpretive tools only for the dense emitting gas. Interpreting emission from low-density, optically thin and/or collisionless regions requires modeling processes at both the macro-scale of the accretion flow and the micro-scale of the particles ultimately radiating. In this proposal, I explain how I will create the most realistic radiative models of luminous BHs to date, including the low-density regions that produce the highest energy emission, by developing a new version of the public code BHAC able to self-consistently treat the physics at the fluid scale and the particle scale using the best radiation treatment in the community. This work represents a huge step towards the resolution of outstanding problems such as BH spectral state changes, jet formation in luminous disks or the origin of quasi-periodic oscillations but will also revolutionize interpretative tools using fast photometric and spectral variability. These revolutionary methods will help interpreting existing and future observations from missions such the all-sky X-ray survey eROSITA, the gamma-ray detector CTA, the X-ray polarimeter IXPE or the Athena X-ray observatory.</p>
--	--

## 2. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data <sup>3</sup>.

Dataset Name	Description	New or Reused	Digital or Physical	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR PHYSICAL DATA
				Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
Simulation outputs	The research in this proposal will produce output files from large 3D simulations and restart files that allow to restart the simulation at a given point. The data from the simulation will be analyzed to produce time and space averaged data files. The more compact averaged data will be shared with the community on demand. The rest of the data will be kept for the entire duration of the project and for 5 years thereafter. The last restart files of each simulation will also be stored to be able to restart the simulations if needed.	<input type="checkbox"/> Generate new data <input type="checkbox"/> Reuse existing data	<input type="checkbox"/> Digital <input type="checkbox"/> Physical	<input type="checkbox"/> Audiovisual <input type="checkbox"/> Images <input type="checkbox"/> Sound <input type="checkbox"/> Numerical <input type="checkbox"/> Textual <input type="checkbox"/> Model <input type="checkbox"/> Software <input type="checkbox"/> Other:	.Vtk .dmp	<input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> > 5 TB <input type="checkbox"/> NA	

<sup>3</sup> Add rows for each dataset you want to describe.

<p><b>GUIDANCE:</b></p> <p><i>The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should be described under documentation/metadata.</i></p> <p><a href="#">RDM Guidance on data</a></p>	
<p>If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type.</p>	
<p>Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.</p>	<p><input type="checkbox"/> Yes, human subject data; provide SMEC or EC approval number:</p> <p><input type="checkbox"/> Yes, animal data; provide ECD reference number:</p> <p><input type="checkbox"/> Yes, dual use; provide approval number:</p> <p><input type="checkbox"/> No</p> <p>Additional information:</p>
<p>Will you process personal data<sup>4</sup>? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).</p>	<p><input type="checkbox"/> Yes (provide PRET G-number or EC S-number below)</p> <p><input type="checkbox"/> No</p> <p>Additional information:</p>
<p>Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)? If so, please comment per dataset or data type where appropriate.</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p>If yes, please comment:</p>

<sup>4</sup> See Glossary Flemish Standard Data Management Plan

<p>Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)?</p> <p>If so, please explain to what data they relate and what restrictions are in place.</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p>If yes, please explain:</p>
<p>Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use?</p> <p>If so, please explain to what data they relate and which restrictions will be asserted.</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p>If yes, please explain:</p>

### 3. Documentation and Metadata

<p>Clearly describe what approach will be followed to capture the accompanying information necessary to keep <b>data understandable and usable</b>, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).</p> <p><a href="#">RDM guidance on documentation and metadata.</a></p>	<p>We expect to mostly analyse either 2D snapshots or time-averaged and azimuthally averaged versions of our numerical 3D simulations. These 2D cuts and those averaged versions files will be made available to the public on the journal website as 2D vtk files sizing ~1-2 MB for our highest resolution runs. We will also provide the Jupyter notebooks that were used to make the plots and analyse the data to ensure complete reproducibility of our results.</p>
--	--

<p>Will a metadata standard be used to make it easier to <b>find and reuse the data</b>?</p> <p>If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse.</p> <p><i>REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.</i></p>	<p><input type="checkbox"/> Yes  <input type="checkbox"/> No</p> <p>If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used:</p> <p>If no, please specify (where appropriate per dataset or data type) which metadata will be created:</p> <p>We do not expect to have such a large number of simulations so that we would need to use a metadata standard. We will simply store the data in the few (&lt;10) folders corresponding to our production runs.</p>
---	--

4. Data Storage & Back-up during the Research Project	
<p>Where will the data be stored?</p> <p><i>Consult the <a href="#">interactive KU Leuven storage guide</a> to find the most suitable storage solution for your data.</i></p>	<p><input type="checkbox"/> Shared network drive (J-drive)  <input type="checkbox"/> Personal network drive (I-drive)  <input type="checkbox"/> OneDrive (KU Leuven)  <input type="checkbox"/> Sharepoint online  <input type="checkbox"/> Sharepoint on-premis  <input type="checkbox"/> Large Volume Storage  <input type="checkbox"/> Digital Vault  <input type="checkbox"/> Other:</p>



<p>How will the data be backed up?</p> <p><i>WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?</i></p>	<p><input type="checkbox"/> Standard back-up provided by KU Leuven ICTS for my storage solution</p> <p><input type="checkbox"/> Personal back-ups I make (specify)</p> <p><input type="checkbox"/> Other (specify)</p>
<p>Is there currently sufficient storage &amp; backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p>We expect that our production runs + test runs will amount to a maximum of ~10 TB of data at the end of the FWO. There is enough space on the Large Volume Storage to store this for long-term storage. During the FWO, we will use the shared network drive for data we are working on and the Large Volume Storage for data that is not used actively. We will strive to have a maximum of 2TB of data on the shared network drive at any time and move the rest to the Large Volume Storage.</p> <p>If no, please specify:</p>
<p>How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?</p> <p><i>CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.</i></p> <p><a href="#">Guidance on security for research data</a></p>	<p>Only I and Fabio Bacchini will have access to the data on the Large Volume Storage and the shared network drive.</p>

<p>What are the expected costs for data storage and backup during the research project? How will these costs be covered?</p>	<p>We estimate to need only 500GB of memory on the shared network drive for the first year since it will be mainly dedicated to development and so not very data heavy. For the second and third year, we will purchase 2TB on the shared network drive. We will also increasingly need space on the Large Volume Storage as we accumulate data from simulations. We estimate 5TB for the second year and 10TB for the third year. Hence, we estimate the costs to be:</p> <p>1<sup>st</sup> year : ~250 euros  2<sup>nd</sup> year : ~1,500 euros  3<sup>rd</sup> year : ~2,000 euros</p> <p>These expenses will be covered by the bench fee of 4,000 euros yearly that has already been approved by the FWO. We might ask for an extra 2,000 euros (over three years) to the FWO to ensure that the bench fee is enough to cover for travel and data storage.</p>
--	---

## 5. Data Preservation after the end of the Research Project

<p>Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).</p> <p><a href="#">Guidance on data preservation</a></p>	<p><input type="checkbox"/> All data will be preserved for 10 years according to KU Leuven RDM policy</p> <p><input type="checkbox"/> All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans</p> <p><input type="checkbox"/> Certain data cannot be kept for 10 years (explain)</p> <p>We will keep the data to a maximum of 5 years for budget issue. Moreover, with the always increasing performance of numerical codes and computing hardware, it is more costly (money-wise and environmentally/energy-wise) to store the data than to redo a simulation with a code and hardware whose performances might have doubled or tripled over 5 years. Of course, we will keep the configuration files of each simulations to ensure that we are able to rerun the exact same simulation. Those files are of the order of the KB.</p>
<p>Where will these data be archived (stored and curated for the long-term)?</p> <p><a href="#">Dedicated data repositories</a> are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the <a href="#">interactive KU Leuven storage guide</a>.</p>	<p><input type="checkbox"/> KU Leuven RDR</p> <p><input type="checkbox"/> Large Volume Storage (longterm for large volumes)</p> <p><input type="checkbox"/> Shared network drive (J-drive)</p> <p><input type="checkbox"/> Other (specify):</p> <p>For storage during the five years after the FWO, we will keep only 1TB of data with only the last restart (.dmp format) files of our ~10 production runs and some 3D snapshots for each simulation (.vtk format). The restart files are the biggest files with an estimated size of ~20-30 GB per simulation for our highest resolutions while the 3D snapshots have an estimated size of ~1-2 GB. We will also keep 2D snapshots and time and azimuthally averaged snapshots of the data as well as the configuration files for each simulation. Those are respectively of the order of ~MB and ~KB and so are not an issue to store.</p>

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The purchase of 1TB of data on the Large Volume Storage for five years will cost 500 euros, which we will pay in advance with our bench fee provided by the FWO.
--	--

6. Data Sharing and Reuse	
<p>Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available.</p> <p><small>NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN &amp; RESTRICTED ACCESS. FOR MORE INFORMATION: <a href="https://wiki.surfnet.nl/display/STANDARDS/INFO-EU-REPO/#INFOEU-REPO-ACCESSRIGHTS">https://wiki.surfnet.nl/display/STANDARDS/INFO-EU-REPO/#INFOEU-REPO-ACCESSRIGHTS</a></small></p>	<p> <input type="checkbox"/> Yes, as open data  <input type="checkbox"/> Yes, as embargoed data (temporary restriction)  <input type="checkbox"/> Yes, as restricted data (upon approval, or institutional access only)  <input type="checkbox"/> No (closed access)  <input type="checkbox"/> Other, please specify:         </p> <p>We will share 2D snapshots and time-averaged and azimuthally averaged versions of our numerical 3D simulations. These 2D snapshots and averaged data will be made available to the public on the journal website as 2D vtk files sizing ~1-2 MB for our highest resolution runs. We will also provide the Jupyter notebooks that were used to make the plots and analyse the data to ensure complete reproducibility of our results.</p>
If access is restricted, please specify who will be able to access the data and under what conditions.	

<p>Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain per dataset or data type where appropriate.</p>	<p> <input type="checkbox"/> Yes, privacy aspects  <input type="checkbox"/> Yes, intellectual property rights  <input type="checkbox"/> Yes, ethical aspects  <input type="checkbox"/> Yes, aspects of dual use  <input type="checkbox"/> Yes, other  <input type="checkbox"/> No         </p> <p>If yes, please specify:</p>
<p>Where will the data be made available? If already known, please provide a repository per dataset or data type.</p>	<p> <input type="checkbox"/> KU Leuven RDR  <input type="checkbox"/> Other data repository (specify)  <input type="checkbox"/> Other (specify)         </p> <p>We will share the data on the publishing journal website.</p>
<p>When will the data be made available?</p>	<p> <input type="checkbox"/> Upon publication of research results  <input type="checkbox"/> Specific date (specify)  <input type="checkbox"/> Other (specify)         </p>

<p>Which data usage licenses are you going to provide? If none, please explain why.</p> <p><i>A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENSE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENSE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENSE THAT MIGHT PROHIBIT THAT.</i></p> <p>Check the <a href="#">RDR guidance on licences</a> for data and software sources code or consult the <a href="#">License selector tool</a> to help you choose.</p>	<p><input type="checkbox"/> CC-BY 4.0 (data)</p> <p><input type="checkbox"/> Data Transfer Agreement (restricted data)</p> <p><input type="checkbox"/> MIT licence (code)</p> <p><input type="checkbox"/> GNU GPL-3.0 (code)</p> <p><input type="checkbox"/> Other (specify)</p>
<p>Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here.</p> <p><i>INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.</i></p>	<p><input type="checkbox"/> Yes, a PID will be added upon deposit in a data repository</p> <p><input type="checkbox"/> My dataset already has a PID</p> <p><input type="checkbox"/> No</p> <p>The data will have the DOI of the journal publication.</p>
<p>What are the expected costs for data sharing? How will these costs be covered?</p>	<p>None submission to the Astrophysics &amp; Astronomy Journal is free for European researchers.</p>

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Nicolas Scepi
Who will manage data storage and backup during the research project?	Nicolas Scepi
Who will manage data preservation and sharing?	Nicolas Scepi

Who will update and implement this DMP?	Nicolas Scepi
---	---------------