

FWO DMP Template - Flemish Standard Data Management Plan

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	Josep Balasch – ORCID 0000-0002-6066-8710 (Project Supervisor)
Contributor name(s) (+ ORCID) & roles	Songqiao Cui – ORCID 0000-0001-9407-1050 (FWO PhD Fellow)
Project number ¹ & title	Architectural Defenses against Hardware Attacks on RISC-V cores
Funder(s) GrantID ²	1SF8223N
Affiliation(s)	<input checked="" type="checkbox"/> KU Leuven <input type="checkbox"/> Universiteit Antwerpen <input type="checkbox"/> Universiteit Gent <input type="checkbox"/> Universiteit Hasselt <input type="checkbox"/> Vrije Universiteit Brussel <input type="checkbox"/> Other: Provide ROR ³ identifier when possible:

¹ “Project number” refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. Applicants can only provide one project number.

² 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.

³ Research Organization Registry Community. <https://ror.org/>

Please provide a short project description	<p>This objective of this project is to design, implement and evaluate architectural hardware extensions with built-in resistance against hardware attacks. These attacks allow to extract secret data from electronic circuits (cryptographic keys, passwords, proprietary code or models), either by monitoring physically observable signals at run-time or by actively tampering with an electronic device.</p> <p>The research activities of the project are centered around RISC-V, a free and open instruction set architecture that can be extended and customized with application-specific modules. The project investigates how to prevent data leakages exploited in hardware attacks by using hardened modules, that is, by adding protection to critical hardware components such as execution units or instruction handlers. The performance and security guarantees of the resulting architectures are evaluated and demonstrated using FPGA platforms.</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 data⁴.

				ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR PHYSICAL DATA
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
Hardware Files	Source code files that describe digital hardware modules using a Hardware Description Language (HDL)	<input checked="" type="checkbox"/> Generate new data	<input checked="" type="checkbox"/> Digital	<input checked="" type="checkbox"/> Software	<input checked="" type="checkbox"/> other: .v (Verilog), .scala (Scala)	<input checked="" type="checkbox"/> < 100 MB	
Firmware Files	Source code files that implement software routines to test the RISC-V hardware extensions	<input checked="" type="checkbox"/> Generate new data	<input checked="" type="checkbox"/> Digital	<input checked="" type="checkbox"/> Software	<input checked="" type="checkbox"/> other: source code files (.c, .asm)	<input checked="" type="checkbox"/> < 100 MB	
Power Measurement Files	Sets of power consumption measurements obtained with an oscilloscope when prototyping our hardware and firmware designs on an FPGA platform	<input checked="" type="checkbox"/> Generate new data	<input checked="" type="checkbox"/> Digital	<input checked="" type="checkbox"/> Observational	<input checked="" type="checkbox"/> .csv	<input checked="" type="checkbox"/> < 100 GB	
Data Analysis Files	Source code files to process the measurement data	<input checked="" type="checkbox"/> Generate new data	<input checked="" type="checkbox"/> Digital	<input checked="" type="checkbox"/> Software	<input checked="" type="checkbox"/> other: python files (.py)	<input checked="" type="checkbox"/> < 100 MB	
RISC-V Design Files	Source codes files available in the open domain that describe RISC-V processor cores	<input checked="" type="checkbox"/> Reuse existing data	<input checked="" type="checkbox"/> Digital	<input checked="" type="checkbox"/> Software	<input checked="" type="checkbox"/> other: .v or .sv (Verilog), .vhdl (VHDL), .scala (Scala), .py (Python)	<input checked="" type="checkbox"/> < 100 MB	

<p><i>GUIDANCE:</i></p> <p><i>DATA CAN BE DIGITAL OR PHYSICAL (FOR EXAMPLE BIOBANK, BIOLOGICAL SAMPLES, ...). DATA TYPE: DATA ARE OFTEN GROUPED BY TYPE (OBSERVATIONAL, EXPERIMENTAL ETC.), FORMAT AND/OR COLLECTION/GENERATION METHOD.</i></p> <p><i>EXAMPLES OF DATA TYPES: OBSERVATIONAL (E.G. SURVEY RESULTS, SENSOR READINGS, SENSORY OBSERVATIONS); EXPERIMENTAL (E.G. MICROSCOPY, SPECTROSCOPY, CHROMATOGRAMS, GENE SEQUENCES); COMPILED/AGGREGATED DATA⁵ (E.G. TEXT & DATA MINING, DERIVED VARIABLES, 3D MODELLING); SIMULATION DATA (E.G. CLIMATE MODELS); SOFTWARE, ETC.</i></p> <p><i>EXAMPLES OF DATA FORMATS: TABULAR DATA (.POR, .SPSS, STRUCTURED TEXT OR MARK-UP FILE XML, .TAB, .CSV), TEXTUAL DATA (.RTF, .XML, .TXT), GEOSPATIAL DATA (.DWG, .GML, ..), IMAGE DATA, AUDIO DATA, VIDEO DATA, DOCUMENTATION & COMPUTATIONAL SCRIPT.</i></p> <p><i>DIGITAL DATA VOLUME: PLEASE ESTIMATE THE UPPER LIMIT OF THE VOLUME OF THE DATA PER DATASET OR DATA TYPE.</i></p> <p><i>PHYSICAL VOLUME: PLEASE ESTIMATE THE PHYSICAL VOLUME OF THE RESEARCH MATERIALS (FOR EXAMPLE THE NUMBER OF RELEVANT BIOLOGICAL SAMPLES THAT NEED TO BE STORED AND PRESERVED DURING THE PROJECT AND/OR AFTER).</i></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>The existing data we will reuse in the project consists only of software source code. These are hardware description files of RISC-V processors available in the open domain and written in standard Hardware Description Languages (HDL) formats. The most important designs we will reuse and/or extend during the project are VexRiscv (https://github.com/SpinalHDL/VexRiscv), CV32E40P (https://github.com/openhwgroup/cv32e40p), CVA6 (https://github.com/openhwgroup/cva6) and LiteX (https://github.com/enjoy-digital/litex). All of them have a permissive license (BSD-2, MIT, Apache) such that they can be used in our research.</p> <p>Besides this software reuse, we will not reuse any datasets within the project.</p>

⁴ Add rows for each dataset you want to describe.

⁵ These data are generated by combining multiple existing datasets.

<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, please describe these issues further and refer to specific datasets or data types when appropriate.</p>	<p> <input type="checkbox"/> Yes, human subject data <input type="checkbox"/> Yes, animal data <input checked="" type="checkbox"/> Yes, dual use <input type="checkbox"/> No If yes, please describe: An application has been submitted to the KU Leuven Ethics Committee on Dual Use, Military use & Misuse of Research (EC DMM). The EC DMM has evaluated our project proposal on January 10th 2023 and has given our project a positive advice (Ref. no.: D-20230110.f). </p>
<p>Will you process personal data⁶? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.</p>	<p> <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes: <ul style="list-style-type: none"> - Short description of the kind of personal data that will be used: - Privacy Registry Reference: </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 checked="" type="checkbox"/> Yes <input type="checkbox"/> No If yes, please comment: the design of hardened digital modules (hardware source files) could have potential for tech transfer or valorisation. This is however not the case for the remaining data (source files and measurements). </p>

⁶ 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 checked="" 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 checked="" 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 data understandable and usable, 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>All source code files generated in the project (hardware, firmware, data analysis) will include comments to aid readability and interpretation. They will be accompanied by README files that explain their structure and dependencies. For hardware and firmware files, the README will additionally include compilation instructions and/or scripts to facilitate its (re)use.</p> <p>The power measurements files will be stored as CSV files consisting of a series of sequential bytes, where each byte corresponds to a sample in the time series. They will be accompanied by a README file detailing the acquisition settings of the measurement (FPGA model, clock frequency of design, sampling frequency of oscilloscope) and Python scripts to load the data. If relevant, the input data associated to the measurement will be provided also as CSV file.</p>
---	---

<p>Will a metadata standard be used to make it easier to find and reuse the data?</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</p> <p><input checked="" 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>The measurement data will be stored in CSV format. We are not aware of any metadata standard for power measurements, so we will ensure the formatting of our data is described in detail.</p>
---	---

4. Data Storage & Back-up during the Research Project	
<p>Where will the data be stored?</p>	<p>Source code files will be stored in a Gitlab repository with access control hosted and managed by KU Leuven ICTS. In a later phase, after the publication of the research article concerned, the relevant source codes will be made available in a public Github repository.</p> <p>The measurement data will be stored in KU Leuven OneDrive for Business, as it is not suited to version-controlled systems such as Gitlab.</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? DESCRIBE THE LOCATIONS, STORAGE MEDIA AND PROCEDURES THAT WILL BE USED FOR STORING AND BACKING UP DIGITAL AND NON-DIGITAL DATA DURING RESEARCH.⁷</i></p> <p><i>REFER TO INSTITUTION-SPECIFIC POLICIES REGARDING BACKUP PROCEDURES WHEN APPROPRIATE.</i></p>	<p>Both storage options (Gitlab and OneDrive) as available through KU Leuven ICTS have adequate backup strategies in place.</p>
<p>Is there currently sufficient storage & 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 checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>If yes, please specify concisely: The default Gitlab repository storage capacity of 1 GB is sufficient to store and backup all source files generated by the project. For the measurement files, the default 2TB storage capacity of OneDrive for Business is sufficient as well.</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.⁷</i></p>	<p>The project will not use or store any personal data. Nevertheless, both the Gitlab repository and OneDrive for Business storage have access control mechanisms to ensure secure storage behind proper authentication. Only members of the research team will have access to the stored data.</p>

⁷ Source: Ghent University Generic DMP Evaluation Rubric: <https://osf.io/2z5g3/>

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	There are no costs expected for data storage and backup, as we do not see a need for additional storage beyond what is already covered by the service offered by KU Leuven ICTS.
---	--

5. Data Preservation after the end of the Research Project

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...).	All data will be retained for 10 years after the end of the project, in lines with the KU Leuven RDM policy.
Where will these data be archived (stored and curated for the long-term)?	Data will be archived using the Server File Storage hosted in the KU Leuven ICTS data centers.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The expected cost for the volume of data for 10 years is below 1000 EUR and will be covered by the project supervisor.

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><i>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 & RESTRICTED ACCESS. FOR MORE INFORMATION:</i> HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS</p>	<p><input checked="" type="checkbox"/> Yes, in an Open Access repository <input type="checkbox"/> Yes, in a restricted access repository (after approval, institutional access only, ...) <input type="checkbox"/> No (closed access) <input type="checkbox"/> Other, please specify:</p> <p>A selection of the source codes necessary to the reproduce project results will be made available via Github. Measurement data can be made available upon request if necessary.</p>
<p>If access is restricted, please specify who will be able to access the data and under what conditions.</p>	<p>N/A</p>
<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 checked="" 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>The data necessary to reproduce project results will be made available via Github. Additional measurement data can be provided upon request.</p>

<p>When will the data be made available?</p> <p><i>THIS COULD BE A SPECIFIC DATE (DD/MM/YYYY) OR AN INDICATION SUCH AS 'UPON PUBLICATION OF RESEARCH RESULTS'.</i></p>	<p>Data will be made available after publication of the research results</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 LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.</i></p> <p><i>EXAMPLE ANSWER: E.G. "DATA FROM THE PROJECT THAT CAN BE SHARED WILL BE MADE AVAILABLE UNDER A CREATIVE COMMONS ATTRIBUTION LICENSE (CC-BY 4.0), SO THAT USERS HAVE TO GIVE CREDIT TO THE ORIGINAL DATA CREATORS." ⁸</i></p>	<p>The source codes will be provided under a permissive license (MIT or Apache) allowing maximum reuse.</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 <input checked="" type="checkbox"/> No If yes:</p>
<p>What are the expected costs for data sharing? How will these costs be covered?</p>	<p>No costs are expected for sharing the data.</p>

⁸ Source: Ghent University Generic DMP Evaluation Rubric: <https://osf.io/2z5g3/>

7. Responsibilities

Who will manage data documentation and metadata during the research project?	The project supervisor Josep Balasch
Who will manage data storage and backup during the research project?	The project supervisor Josep Balasch
Who will manage data preservation and sharing?	The project supervisor Josep Balasch
Who will update and implement this DMP?	The project supervisor Josep Balasch