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	**Bart Preneel https://orcid.org/0000-0003-2005-9651**
Contributor name(s) (+ ORCID) & roles	**Yu Long Chen https://orcid.org/0000-0002-0369-2423**
Project number [1] & title	**Design and Analysis of Block Cipher Modes of Operation (1264825N)**
Funder(s) GrantID [2]	FWO
Affiliation(s)	X KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310

[1] "Project number" refers to the institutional project number. This question is optional. 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.

Please provide a short project description	One of the most important primitives in cryptography is the block cipher. For a natural number n, block ciphers can be seen as permutations on the set of n-bit strings parameterized by a secret key. The publication the Advanced Encryption Standard (AES) block ciphers by the National Institute of Standards and Technology of America has generated much more interest in symmetric-key cryptography. The impact of the AES standard cannot be underestimated: for example, a conservative lower bound estimate of the AES block cipher standard for the US economy is $250 billion. However, block ciphers only process messages with a length equal to the block size n, so in order to support encryption of messages of size larger than n, such block ciphers are usually embedded in a mode of operation. The purpose of this proposal is to both develop new models and techniques for the generic security analysis of block cipher modes, as well as the analysis of existing standards and new submissions to NIST. The first goal is quite theoretical: the models should cover real-world applications, while the techniques will be used to identify weaknesses in existing modes of operation and to construct secure and efficient alternatives. Although the second goal is more practical, but it uses the results obtained from the first phase. If it is possible to break existing standards, the models and techniques will also be used to design new constructions that meet the requirements of these applications.

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 [3].

				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
		☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:		☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
	Scientific articles (open access)		x Digital	☒ Textual	Latex, MS Word format, PDF	5 GB	
	PhD thesis		x Digital	☒ Textual	Latex, MS Word format, PDF	5 GB	
	Code (will be made available as open source)		x Digital		in several programming languages (C, C++, magma, Matlab, Python, VHDL, Verilog)	10 GB	
	Project		x Digital	☒ Textual	Latex, MS Word	5 GB	

[3] Add rows for each dataset you want to describe.

	deliverables				format, Markdown, Plaintext files, ODT,		
	Presentations		x Digital	☒ Textual	Latex, PPT, ODP, PDF	5 GB	
	Figures, graphs, media		x Digital	☒ Images	TIF, DRAWIO, JPG, PNG, SVG, Inkscape/GIMP formats	10 GB	
	Scripts		x Digital		.m,.tcl	10 GB	

GUIDANCE:
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 described under documentation/metadata.
RDM Guidance on data

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.	N/A
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.	☐ Yes, human subject data; provide SMEC or EC approval number: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☒ No Additional information:

Will you process personal data[4]? 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).	☐ Yes (provide PRET G-number or EC S-number below) ☒ No Additional information:
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.	☒ Yes ☐ No If yes, please comment: It is not very likely, but it may be that a patent will be file to enable commercial valorization; this will bring a limited delay in publishing the research results (3-6 months)
Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:
Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. Documentation and Metadata

[4] See Glossary Flemish Standard Data Management Plan

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). *RDM guidance on documentation and metadata*.	When data are stored at the KU Leuven RDR platform to upload, describe, and share research, it includes all necessary documentation that will help others understand our data and make it fully reusable: information about instruments of data collection, codebooks, methods reports, field reports, protocols, interviewer guidelines, and so on. Documentation is available in standard word processing tools: MS Word, latex or PDF documents. The project data complies with community norms and is clearly licensed, so others know what kind of reuse is permitted. Licensing of the research results is organized through the tech transfer office of the KU Leuven (LRD). All the project peer-reviewed publications are made openly accessible through the OpenAIRE repository. In accordance with KU Leuven's policies, our publications are publicly available through Lirias, which is the official institutional repository of the university: https://research.kuleuven.be/en/lirias Our publications, presentations, artifacts are also available from the database of the research group: https://cosicdatabase.esat.kuleuven.be/
Will a metadata standard be used to make it easier to **find and reuse the data**? 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. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☒ Yes ☐ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: To make sure the provided data and metadata is commonly understood, my project uses recognized standards to allow potential users to combine or exchange them. Controlled vocabularies, keywords, thesauri or ontologies are used to enhance interoperability. All reports are in standard text processing tools (MS Word, Latex, PDF) documents. All code is in standard programming languages (Python, C, C++, Rust, Verilog, VHDL) and stored in standard repositories: gitlab or github. For example, project's reports are trivially interoperable, as they can be edited and reused with any compatible editor. Moreover, open-source repositories are also interoperable because anyone can download and inspect them. If no, please specify (where appropriate per dataset or data type) which metadata will be created:

4. Data Storage & Back-up during the Research Project	
Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☐ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☐ Teams ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ ManGO ☐ Digital vault ☒ Other: The publicly available software code and scripts will be primarily stored on the COSIC GitHub repository (https://github.com/KULeuven-COSIC). Code that is not publicly available will be stored on the COSIC research group's GitLab infrastructure, which is private and also maintained by the ESAT department of the KU Leuven.
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☒ Other (specify) The publicly accessible COSIC GitHub repository (https://github.com/KULeuven-COSIC) is backed up automatically by GitHub. The COSIC research group's GitLab infrastructure, is private and also maintained by the ICTS services of the ESAT department of the KU Leuven.
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.	☒ Yes ☐ No If no, please specify:

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *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.* *Guidance on security for research data*	Research results which remain closed will stay within the KU Leuven, with KU Leuven facilitating secure storage and exchange through the RDR for active and sensitive data. As explained above RDR is a Dataverse.org based KU Leuven platform to upload, describe, and share research data. RDR offers flexible access levels: open, restricted, or closed. RDR provides metadata fields for researchers to describe a research dataset and make it findable. Version control ensures tracking of changes and preservation of previous data versions, while regular backups data integrity. RDR is focused on the long-term preservation and sharing of research data. Accessibility is guaranteed beyond the project's lifetime, with project data stored for 10 years post-project.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Given the small size of the data (at most a few GBytes), the storage cost will either be free of charge, or covered by COSIC's or KU Leuven's IT costs. The cost for open access publications is being covered by project funding, in this case FWO funding.

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...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ 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 ☐ Certain data cannot be kept for 10 years (explain)

Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories 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 interactive KU Leuven storage guide.*	☒ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☐ Other (specifiy):
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Given the small size of the data (at most a few GBytes), the storage cost will either be free of charge, or covered by COSIC's or KU Leuven's IT costs. The cost for open access publications is being covered by project funding, in this case FWO funding.

6. Data Sharing and Reuse	
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. *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: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☒ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify: All the project peer-reviewed publications are made openly accessible through the OpenAIRE repository. In accordance with KU Leuven's policies, our publications are publicly available through Lirias, which is the official institutional repository of the university: https://research.kuleuven.be/en/lirias Our publications, presentations, artifacts are also available from the database of the research group: https://cosicdatabase.esat.kuleuven.be/
If access is restricted, please specify who will be able to access the data and under what conditions.	Research results which remain closed will stay within the KU Leuven, with KU Leuven facilitating secure storage and exchange through the RDR for active and sensitive data.

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.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☒ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☐ No If yes, please specify: Some of the software that will be developed may be used to generate collisions for widely used hash functions, which could be a concern for many security applications. Software will only be made available if it does not affect the security of widely used systems. If any vulnerabilities are found, a process of responsible disclosure will be followed.
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☒ KU Leuven RDR ☐ Other data repository (specify) ☐ Other (specify)
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)

Which data usage licenses are you going to provide? If none, please explain why. *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.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☒ CC-BY 4.0 (data) ☐ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify)
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No
What are the expected costs for data sharing? How will these costs be covered?	Given the small size of the data (at most a few GBytes), the sharing cost will either be free of charge, or covered by COSIC's or KU Leuven's IT costs.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	FWO mandate holder Yu Long Chen
Who will manage data storage and backup during the research project?	DMP officer at Cosic and general IT Service at ESAT and KU Leuven
Who will manage data preservation and sharing?	DMP officer at Cosic and general IT Service at ESAT and KU Leuven
Who will update and implement this DMP?	FWO mandate holder Yu Long Chen together with DMP officer at Cosic