Scalable Integrated Circuits for High-Throughput, Third-Generation DNA Sequencing using Solid-State Nanopores

Application DMP

Questionnaire

The questions in this section should only be answered if you are currently applying for FWO funding.
Are you preparing an application for funding?

- No

Scalable Integrated Circuits for High-Throughput, Third-Generation DNA Sequencing using Solid-State Nanopores

DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

Scalable Integrated Circuits for High-Throughput, Third-Generation DNA Sequencing using Solid-State Nanopores

GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Scalable Integrated Circuits for High-Throughput, Third-Generation DNA Sequencing using Solid-State Nanopores
FWO DMP (Flemish Standard DMP)

1. 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.

				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
Circuit designs		New	Digital	Experimental	*Virtuoso database*	*< 5 GB*	
Simulation data		New	Digital	Simulation data	SPF	< 1 TB	
Measurment results		New	Digital	Exerimental	Graphs	< 10 TB	
Papers		New	Digital	Other	Latex	< 1 GB	

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)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

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

The integrated circuit designs have large value and can be transferd to industry. The measurement results are used to prove the performance of the integrated circuit designs.

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 in the comment section to what data they relate and what

restrictions are in place.

- Yes

Agreements with TSMC limit the open distribution of the integrated circuit designs and simulations.

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 in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

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).

Integrated circuits are designed using the Cadence Virtuoso software which keeps the designs degistable as they get larger and classifies simulation results close to the simulated circuit. Measurement data will be labeled based on the type of measurement performed and under which conditions.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

3. Data storage & back-up during the research project

Where will the data be stored?

On the internal IMEC servers.

How will the data be backed up?

IMEC has internal processes for backing up data automatically.

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

IMEC has a large amount of storage available on its servers and this amount is expanded on a regular basis.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

IMEC servers are managed by IT employees who make sure that no unauthorised persons can infiltrate.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

All costs of data storage are covered by IMEC as overhead costs of the project.

4. 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 simulation and measurement results are saved on IMEC servers and backed up.

Where will these data be archived (stored and curated for the long-term)?

Depending on the relevancy of the measurement data, this data can be deleted after 5 years to free up storage space. Integrated circuit designs and simulations will be preserved.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

IMEC has large storage servers which get expanded on a regular basis to cover the growing data needs at no cost to the researcher.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, …)

If access is restricted, please specify who will be able to access the data and under what conditions.

IMEC researchers

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 in the comment section per dataset or data type where appropriate.

- Yes, Other

NDA agreements limit the ability to share data.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

On the internal filesystem of IMEC.

When will the data be made available?

All data is accessible by IMEC researcher as they are generated.

Which data usage licenses are you going to provide? If none, please explain why.

When data is published, accompanying data will be made availble using the MIT license in accordance with the MICAS KULeuven policy.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

What are the expected costs for data sharing? How will these costs be covered?

IMEC and KULeuven provide the necesarry file hosting options at no cost.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

The PhD Researcher, Sander Crols

Who will manage data storage and backup during the research project?

The PhD Researcher, Sander Crols and IMEC

Who will manage data preservation and sharing?

IMEC

Who will update and implement this DMP?

The PhD Researcher, Sander Crols