
Plan Overview

A Data Management Plan created using DMPonline.be

Title: NETWORKS AS EMERGENT WIRELESS SYSTEMS: A PROPOSAL FOR ELEMENTARY AUTOMATA TO ADDRESS CONSTRAINTS (NEWSPEAC)

Creator: Jonathan Oostvogels

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Project abstract:

Wireless Cyber-Physical Systems (CPS) typically communicate using network stacks comprising reliable, layered services. Such reliability-driven design, however, is proving increasingly hard to justify: recent “cross-layer” or “semantic” network designs show that abandoning said reliability focus enables several novel application classes that formerly faced prohibitive resource constraints. This project hypothesises that the research efforts surrounding cross-layer, semantic design can be accelerated, automated, and systematised by re-thinking networks as interconnections of “simple” automata: recent cross-layer optimisations appear to converge on a set of necessary characteristics expressible within this formalism. Motivated by the commercialisation of hardware enablers (i.e. low-power, radar-like interfaces at >3 GHz frequencies), this project hence aims to (1) build a low-power software-defined networking platform to facilitate automata-based networking, (2) use that platform for the automated search and validation of automata that, when interconnected, exhibit emergent properties useful for communication, (3) establish a software package that schedules the right automata based on application semantics, and (4) validates these contributions in consultation with application-oriented research focusing on distributed audio analysis or robot swarm control. The project targets 2-4x improvements to the latency, power, throughput, density, and accuracy trade-offs in today's CPSs.

ID: 211900

Start date: 01-10-2024

End date: 30-09-2027

Last modified: 02-01-2025

NETWORKS AS EMERGENT WIRELESS SYSTEMS: A PROPOSAL FOR ELEMENTARY AUTOMATA TO ADDRESS CONSTRAINTS (NEWSPEAC)

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.

| Dataset Name | Description | New or reused | Digital or Physical | Digital Data Type | Digital Data format | Digital data volume |
|-------------------------------|--|---------------|---------------------|-------------------|-------------------------|---------------------|
| Source code | Software implementing systems under design | New | Digital | Software | Code | < 1 GB |
| Simulation data | Time-series output of simulation data | New | Digital | Experimental | CSV | < 100 GB |
| Time series evaluation data | Second-by-second traces of relevant events in experimental set-ups | New | Digital | Experimental | CSV | < 100 GB |
| Processed evaluation data | Aggregated analyses, higher-level metrics describing time series | New | Digital | Aggregate | Relational (PostgreSQL) | < 1 GB |
| Third-party computer programs | Helper software | Reuse | Digital | Software | Machine code | < 100 GB |
| Documentation | Supporting documents | New | Digital | Other | PDF or plaintext | < 1 GB |
| Design documents | Diagrams, pseudocode detailing design decisions | New | Digital | Other | PDF | < 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:

Third-party computer programs; details to be decide throughout the project.

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

Newly developed software may be published on public code repositories under licensing constraints or directly licensed to industry. Alternatively, methods implemented in these software artefacts may be patented and subsequently licensed to industrial and/or academic partners. Appropriate licensing terms will be discussed with KU Leuven's Tech Transfer Office on a case-by-case basis.

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

The use of third-party software modules in newly developed software will be subject to licensing restrictions, and their impact on possible valorisation will be evaluated on a case-by-case basis (e.g. when choosing between MIT or GPL-licensed software libraries). No restrictions on the dissemination of evaluation data are to be expected.

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).

The source code of the computer programs will be annotated with documentation, as is standard practice. Where necessary for the publication of the scientific papers or publication of the newly created software itself, more generally accessible documentation (i.e. text documents, manuals) will be created and provided alongside the actual data.

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

The data types used in this project are either self-descriptive (the use of time series and relational databases implies the existence of metadata, i.e. database schemas), heavily standardised (explicit definitions of the semantics of programming languages exist), or both.

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

Where will the data be stored?

- Source code and computer programs: KU Leuven's GitLab instance
- Simulation and evaluation data: The research group's AWS cloud computing platform equipped with PostgreSQL and InfluxDB instances
- Documentation, design documents: KU Leuven OneDrive or equivalent

How will the data be backed up?

All of the above storage solutions come with automated back-up tools.

**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

Cloud storage can be expanded if necessary, but initial size estimates indicate this should not be necessary throughout the project.

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

All of the above storage solutions come with existing access control mechanisms. Local copies on e.g. the researcher's personal devices are protected by a password.

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

All data is expected to fit within existing storage infrastructure; no explicit data storage cost is expected to be attributed to 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 data will be preserved for 10 years according to KU Leuven RDM policy.

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

KU Leuven Large Volume Storage.

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

Costs related to long-term archival will be covered by the PI's reserve funds and at the current rate should not exceed 95 EUR/year (<https://icts.kuleuven.be/storagewijzer/en>).

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 an Open Access repository
- Yes, in a restricted access repository (after approval, institutional access only, ...)
- Documentation (i.e. graphs, papers etc) will be published by the publisher and through the university's repositories.
- Source code for computer programs will in general be published through publicly accessible code repositories (e.g. GitHub, Bitbucket), though further restrictions may be imposed to protect valorisation interests (e.g. through availability on request only, or through licensing terms to be discussed with the Tech Transfer Office).
- Raw evaluation data can be made available upon request.

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

See above.

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, Intellectual Property Rights

See above.

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

See above.

When will the data be made available?

Upon publication of accompanying paper(s).

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

MIT license unless that cannot be reconciled with valorisation interests or third-party license agreements. In that case, the details will be worked out on a case-by-case basis.

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

Publications may receive a DOI, depending on the publisher.

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

No cost is expected.

6. Responsibilities

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

Jonathan Oostvogels (fellow), together with Danny Hughes (PI).

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

Jonathan Oostvogels (fellow), together with Danny Hughes (PI).

Who will manage data preservation and sharing?

Jonathan Oostvogels (fellow), together with Danny Hughes (PI).

Who will update and implement this DMP?

Jonathan Oostvogels (fellow), together with Danny Hughes (PI).