

---

## Plan Overview

*A Data Management Plan created using DMPOnline.be*

**Title:** IGNITE: Injecting Global iNformation In local opTimization mEthods

**Creator:** Dominik Bongartz

**Affiliation:** KU Leuven (KUL)

**Template:** KU Leuven BOF-IOF

### Project abstract:

Nonconvex optimization problems and in particular nonconvex nonlinear programs have countless applications in science and engineering, but remain challenging to solve due to suboptimal local optima. Common derivative-based (e.g., Newton-type) methods rely only on local information and have no way to recognize whether a local optimum is global. Deterministic global optimization methods do utilize rigorous global information in the form of convex relaxations. However, the branch-and-bound algorithms these are used in often have prohibitively long runtime. In this project, we aim to integrate the rigorous global information of convex relaxations into the strategies that underlie derivative-based methods for unconstrained optimization. The goal is to construct methods that (a) still converge to local optima, while (b) having a higher chance to locate good – rather than just any – local optima, and (c) avoiding the long runtime of branch-and-bound, but rather have runtimes comparable to existing derivative-based solvers.

**ID:** 213265

**Start date:** 01-10-2024

**End date:** 30-09-2028

**Last modified:** 11-03-2025

## IGNITE: Injecting Global iNformation In local opTimization mEthods

### 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 / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		<i>Indicate: N(ew data) or E(xisting data)</i>	<i>Indicate: D(igital) or P(hysical)</i>	<i>Indicate: Audiovisual Images Sound Numerical Textual Model Software Other (specify)</i>		<i>Indicate: &lt;1GB &lt;100GB &lt;1TB &lt;5TB &gt;5TB NA</i>	
1. Theoretical Results	Description of developed optimization algorithms, convergence proofs	N	D	T	.tex	<1 GB	
2. Code	Implementations of developed algorithms	N	D	SO	.cpp, .h, .py	<1 GB	
3. Benchmark Problems	Optimization problems on which the performance of the developed algorithms is tested	N, E	D	M	.txt, .py, .gms	<1 GB	
4. Log Files	Solver log files documenting the performance of the implemented algorithms (ID 2) on the benchmark problems (ID 3)	N	D	N, T	.log, .txt, .csv	<100 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:

3. Benchmark Problems: In addition to newly created problems and problems curated from the engineering literature during the project, data from the following benchmark libraries will be used:

- MINLPLib: <https://www.minlplib.org/download.html>
- COCONUT library: <https://arnold-neumaier.at/glopt/coconut/Benchmark/Benchmark.html>
- CUTEst: <https://github.com/ralna/CUTEst>
- Floudas, C.A., Pardalos, P.M., Adjiman, C., Esposito, W.R., Gümus, Z.H., Harding, S.T., Klepeis, J.L., Meyer, C.A. and Schweiger, C.A., 2013. Handbook of Test Problems in Local and Global Optimization (Vol. 33). Springer Science & Business Media.

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.

- No

Will you process personal data? 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).

- 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

2. Code: If the developed algorithms demonstrate good performance, there may be potential to develop them further into a commercial software product, and/or provide consulting based on an open-source code.

**Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or 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.**

- No

**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

#### **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).**

Theoretical Results (Dataset 1) will be documented in the form of reports, thus ensuring understandability. These serve as basis for structuring the remaining datasets. Algorithm variants will be given unique names / identifiers that will be used to label corresponding implementations and results in the other datasets.

Code (Dataset 2) will be developed within the Gitlab instance of KU Leuven. Documentation will be based on best practices for (C++) programming, including descriptive names for all objects, short comments in the code where necessary, and automatically generated html or pdf documentation using, e.g., Doxygen. Additional information about the algorithms and how to use them will be included in a Readme.md as well as manual pages generated using Doxygen. These can be hosted via Gitlab pages to enable online access. The documentation will make reference to the algorithm identifiers from Dataset 1.

Benchmark Problems (Dataset 3) will be given unique identifiers that will be used to label corresponding results.

Log Files (Dataset 4) will make reference to the unique algorithm identifiers from Dataset 1 and the unique problem identifiers from Dataset 3, to clarify which algorithm was used and which problem was solved, respectively.

**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.**

- No

Benchmark Problems (Dataset 3) will be complemented with meta-data in Readme files detailing the data source, and problem statistics (number of variables etc.).

Log Files (Dataset 4) will be complemented with meta-data in Readme files detailing the algorithm and implementation version used, problem, and computer hardware used.

## Data Storage & Back-up during the Research Project

### Where will the data be stored?

- Other (specify below)
- OneDrive (KU Leuven)
- Code: Gitlab instance of KU Leuven

### How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

### Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

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

The OneDrive for Business environment offered by KU Leuven is access-controlled via multifactor authentication with the KU Leuven Authenticator app.

The same applies for the GitLab instance of KU Leuven, as long as the project is set as private.

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

OneDrive for Business and Gitlab are provided free of charge by KU Leuven.

## Data Preservation after the end of the Research Project

### Which data will be retained for 10 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)?

- Large Volume Storage (longterm for large volumes)

- KU Leuven RDR
- Other (specify below)

All research data will be archived on the large volume storage.

Benchmark Problems may additionally be published via RDR to make them readily accessible to other developers of optimization algorithms.

Code will also remain available on the Gitlab instance of KU Leuven.

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

For this project, required storage capacity of <1 TB is required, resulting in a cost of <100€ / year, which can be covered by the grant and beyond.

#### 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.**

- Yes, as open data

Code (Dataset 2) will be available via the Gitlab instance of KU Leuven.

New Benchmark Problems (Dataset 3) may be made available via RDR.

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

Not applicable

**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.**

- No

**Where will the data be made available?**

**If already known, please provide a repository per dataset or data type.**

- KU Leuven RDR (Research Data Repository)
- Other (specify below)

KU Leuven Gitlab instance, see above

**When will the data be made available?**

- Upon publication of research results

**Which data usage licenses are you going to provide?**

**If none, please explain why.**

- CC-BY 4.0 (data)
- MIT licence (code)

**Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.**

- No

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

No costs are expected

## **Responsibilities**

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

The PhD student working on the project will be responsible for data documentation and metadata during the project under supervision of the promoter (Dominik Bongartz).

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

The PhD student working on the project will manage data storage and backup during the project under supervision of the promoter (Dominik Bongartz).

**Who will manage data preservation and sharing?**

The promoter (Dominik Bongartz) will manage data preservation and sharing, with support from the PhD student.

**Who will update and implement this DMP?**

The promoter (Dominik Bongartz) will be responsible for updating and implementing this DMP, with support from the PhD student.