

# PhD proposal

## Open Science for the scalability of a new generation search technology

PhD advisor : Lucas Nussbaum, Madynes team (Lucas.Nussbaum@loria.fr)

**Keywords :** CIFRE PhD, Big Data, infrastructure, experimentation, large scale

### Context

The era of Big Data profoundly changes our way to think computer infrastructures, both on the hardware and software levels. The management of data becomes decentralized, or even dematerialized, and enables one to re-think algorithms of data processing and analysis to address increasing volumes.

Xilopix develops a new generation search technology, with a fully visual and tactile ergonomics, to avoid both the requirement of using a keyboard, and the language-related barriers, with an iterative refining of search results.

### Goals

This PhD project aims at pushing forward the scalability of Xilopix's search technology. The PhD student will develop a deep insight into available solutions (including Big Data ones) : strengths, weaknesses, impact and implied constraints on the using solution ; this will be achieved by performing experiments on those solutions. The PhD student will then analyze the specific needs of Xilopix's search technology, and possible evolutions on all sides.

In addition to contributions expected on the *infrastructure* aspects, this work will also serve as a basis for working on scientific and experimental methodology, in the context of *Open Science* and *reproducible research*. Specifically, methods and tools facilitating experimentation (orchestration of experiments, load injection, etc.) in the Big Data context, on a testbed such as Grid'5000, will be designed.

### Required skills

- Interest for research, specifically in interaction with industry
- Knowledge of systems and networking concepts
- Performance evaluation of systems and networks
- Mastering of the Linux environment (usage, basic administration)
- Software engineering (the Ruby language will be used, but its mastering is not a prerequisite)
- Good mastering of technical english

### Links and references

- Xilopix. <http://pro.xilopix.com/>
- Lucas Nussbaum's homepage. <http://www.loria.fr/~lnussbau/>
- The Grid'5000 testbed, which will be used for experiments. <https://www.grid5000.fr/>
- Some example tools for experimentation :
  - Distem (emulation). <http://distem.gforge.inria.fr>
  - XPFlow. <http://xpflow.gforge.inria.fr>

**Apply by email to [lucas.nussbaum@loria.fr](mailto:lucas.nussbaum@loria.fr)**