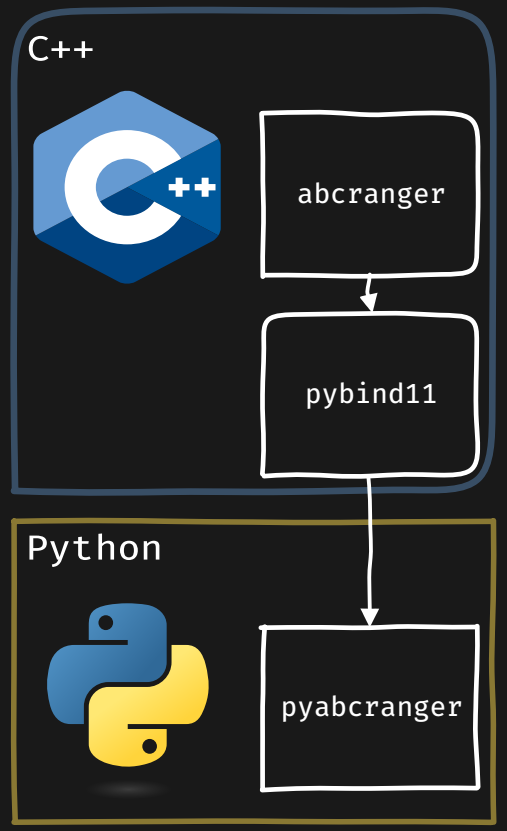
Pybind11/reticulate as an alternative to Rcpp

A humble non-fluent R user's perspective

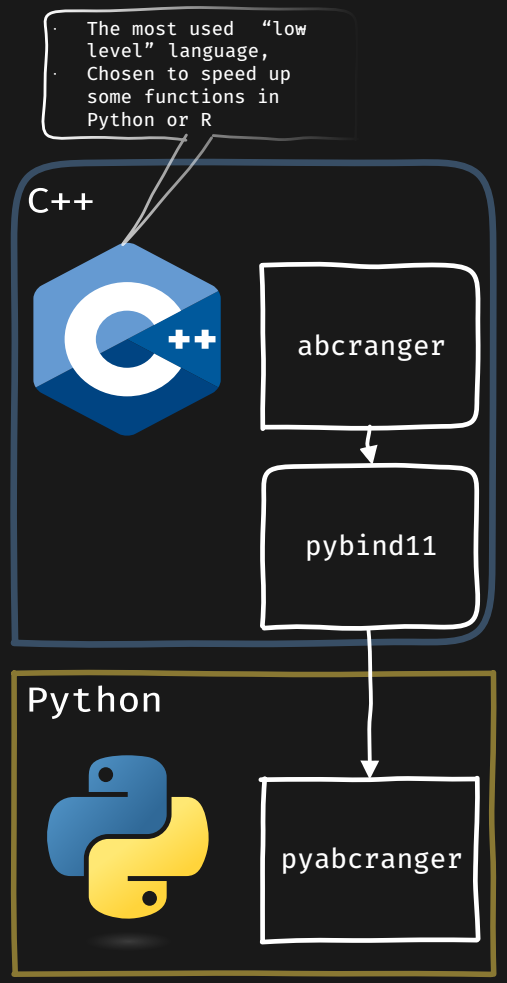
François-David Collin
IMAG Montpellier

2023-06-18

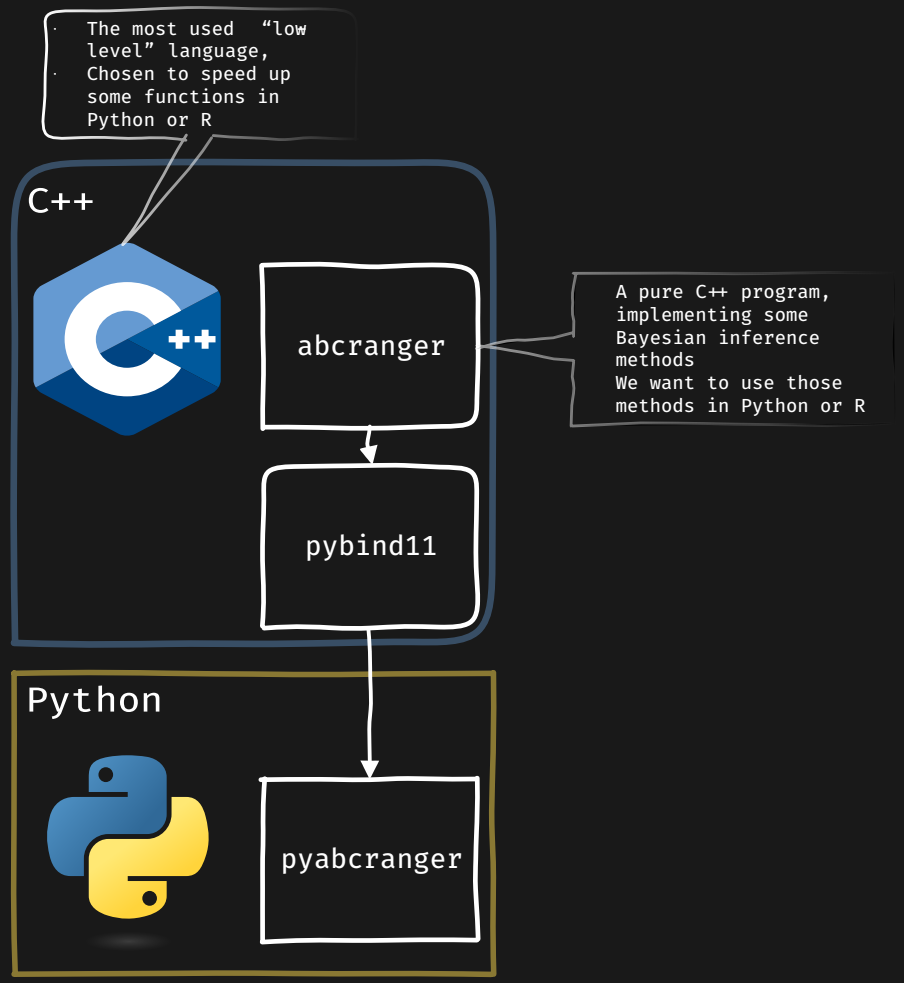
Python/C++ Development Workflow



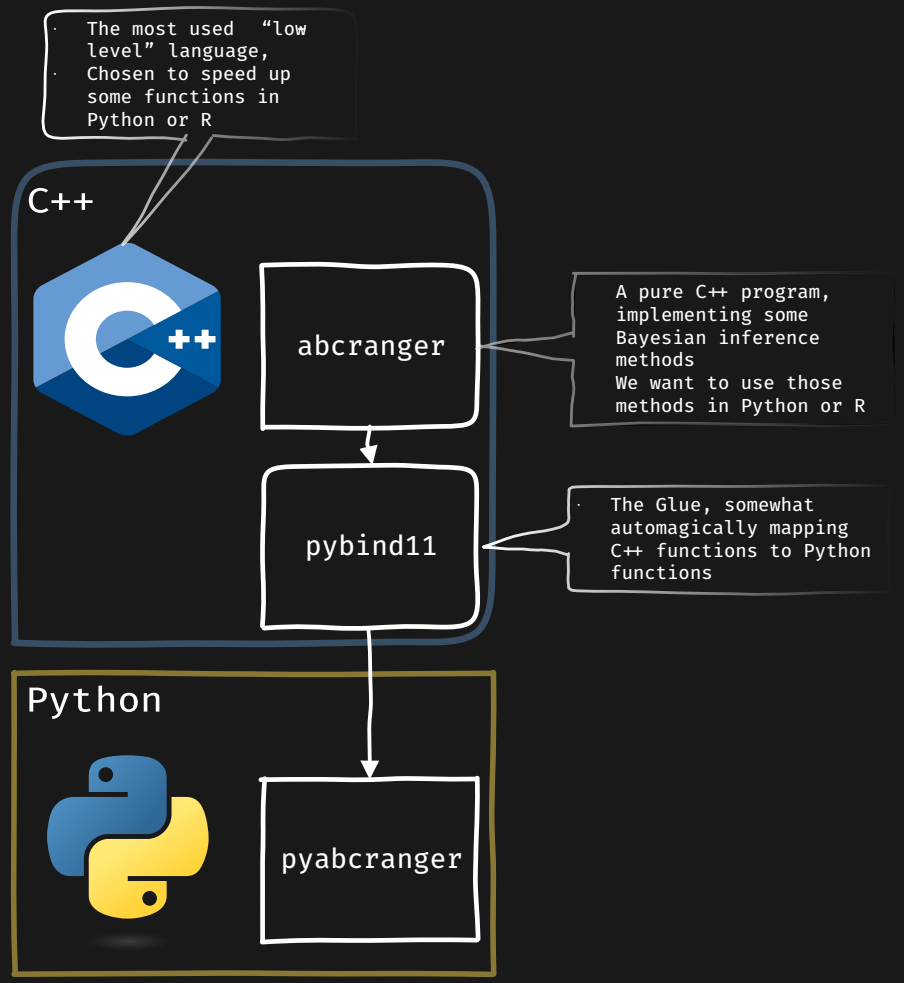
Python/C++ Development Workflow



Python/C++ Development Workflow



Python/C++ Development Workflow



Python/C++ Development Workflow

- The most used "low level" language,
- Chosen to speed up some functions in Python or R

C++

abcranger

- A pure C++ program, implementing some Bayesian inference methods
- We want to use those methods in Python or R

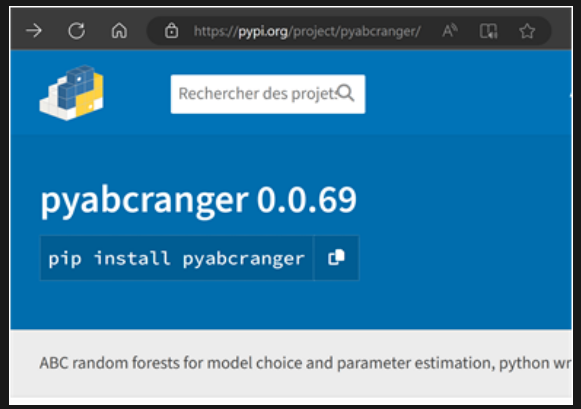
pybind11

- The Glue, somewhat automagically mapping C++ functions to Python functions

Python

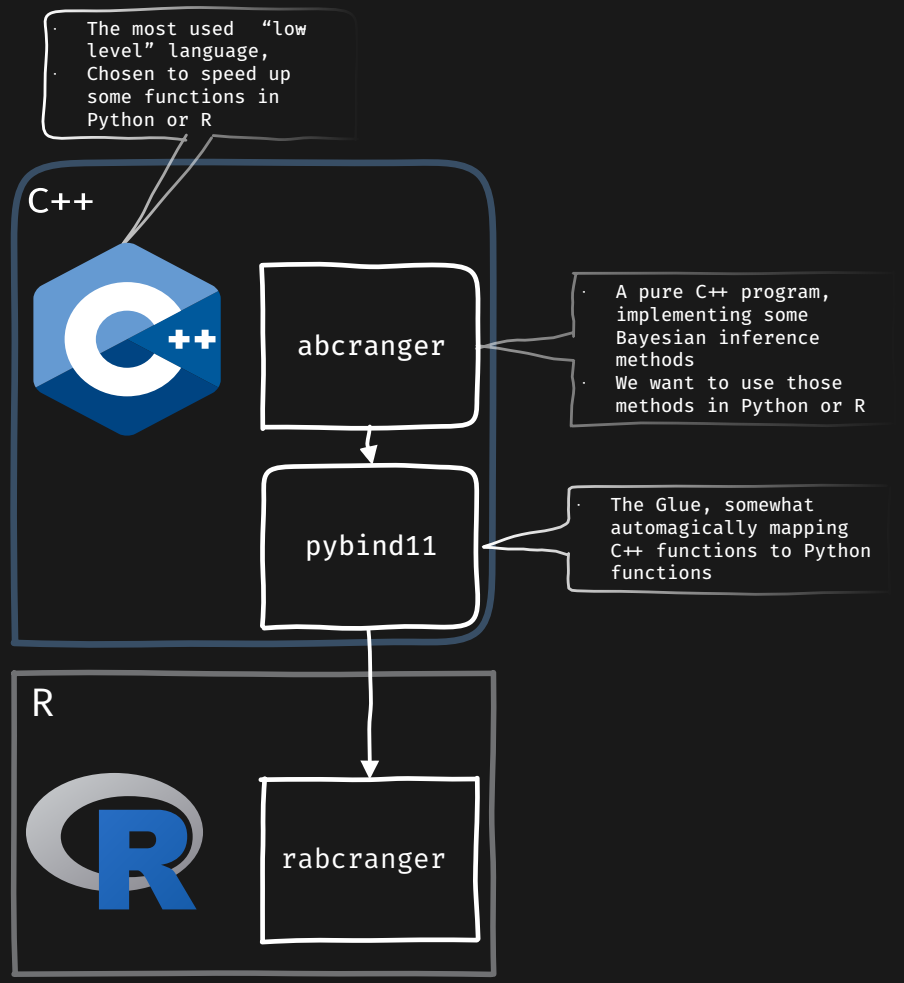
pyabcranger

- The actual Python package, available on PyPi

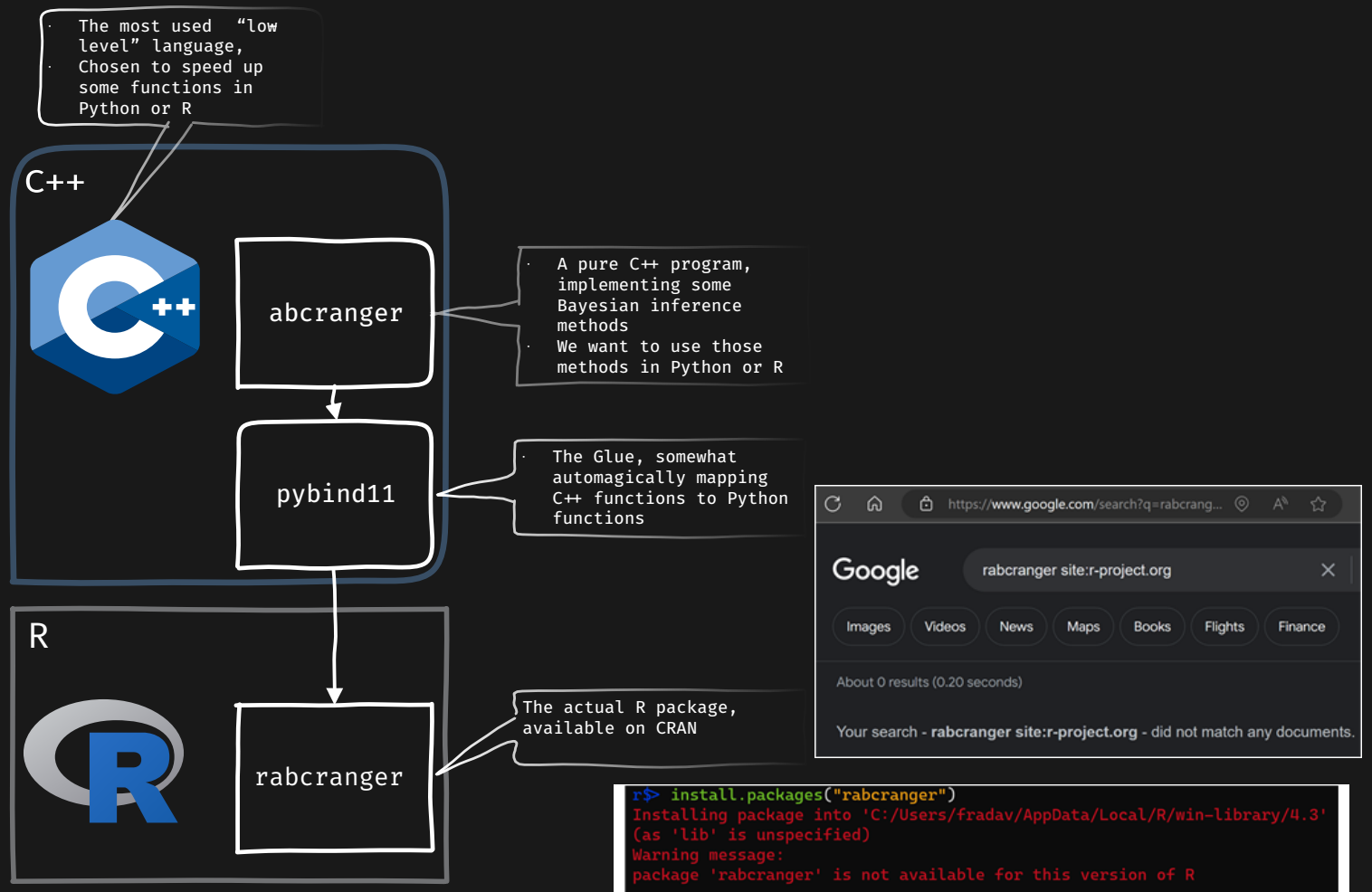



```
fradav ~ 2.538s pip install pyabcranger
Collecting pyabcranger
  Using cached pyabcranger-0.0.69-cp39-cp39-win_amd64.whl (628 kB)
Installing collected packages: pyabcranger
Successfully installed pyabcranger-0.0.69
fradav ~ 982ms
```

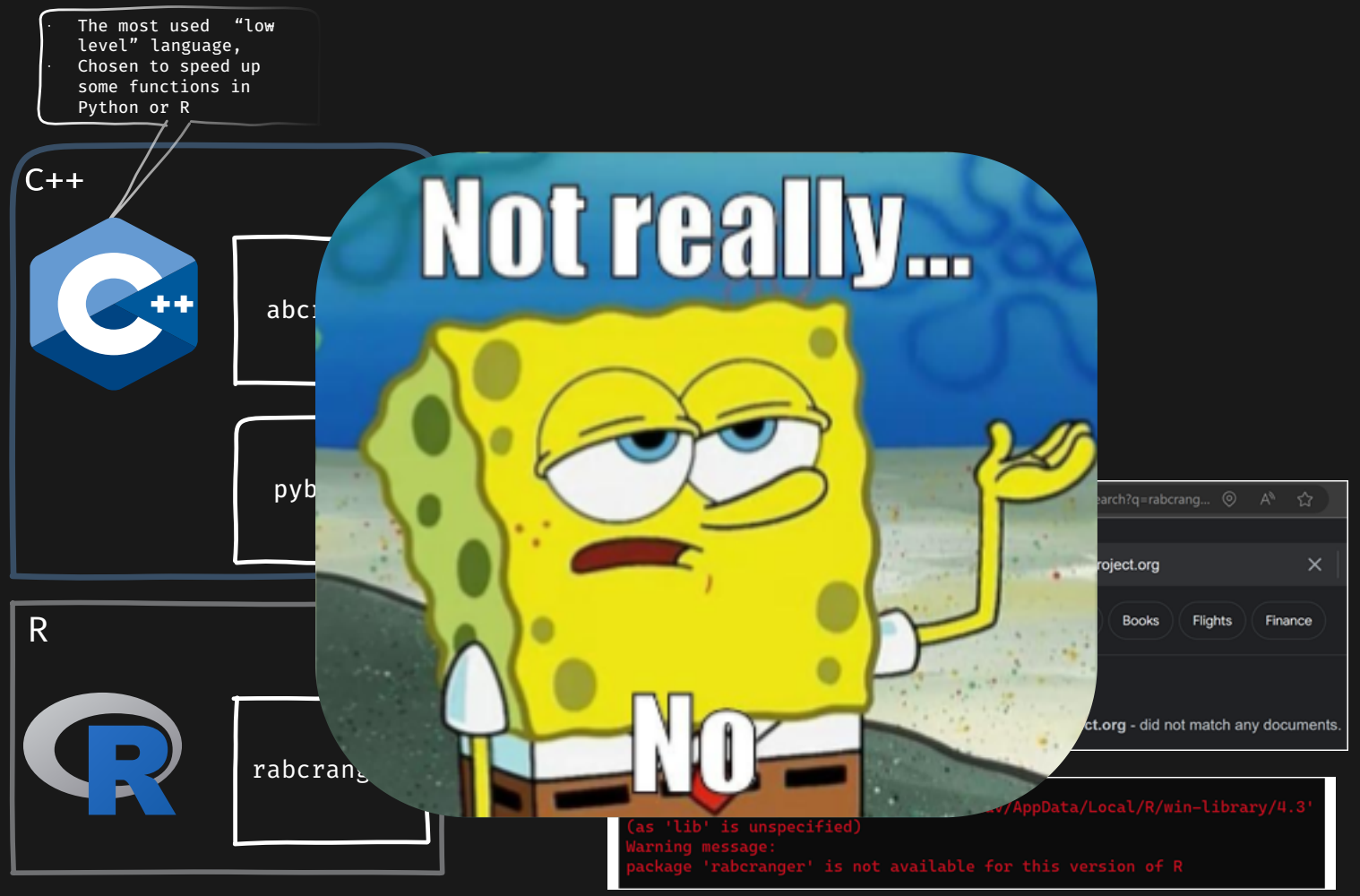
R/C++ Development Workflow



R/C++ Development Workflow



R/C++ Development Workflow



R package cancelled, why?

R package cancelled, why?

Heterogeneous *build* system between Rcpp and pybind11

- Rcpp based on makefile
- pybind11 on cmake

R package cancelled, why?

Heterogeneous *build* system between `Rcpp` and `pybind11`

- `Rcpp` based on `makefile`
- `pybind11` on `cmake`

Obsolete version of `Eigen` from `RcppEigen`

C++ part extensively using `Eigen`, conflicting C++ dependency management.

R package cancelled, why?

Heterogeneous *build* system between Rcpp and pybind11

- Rcpp based on makefile
- pybind11 on cmake

Obsolete version of Eigen from RcppEigen

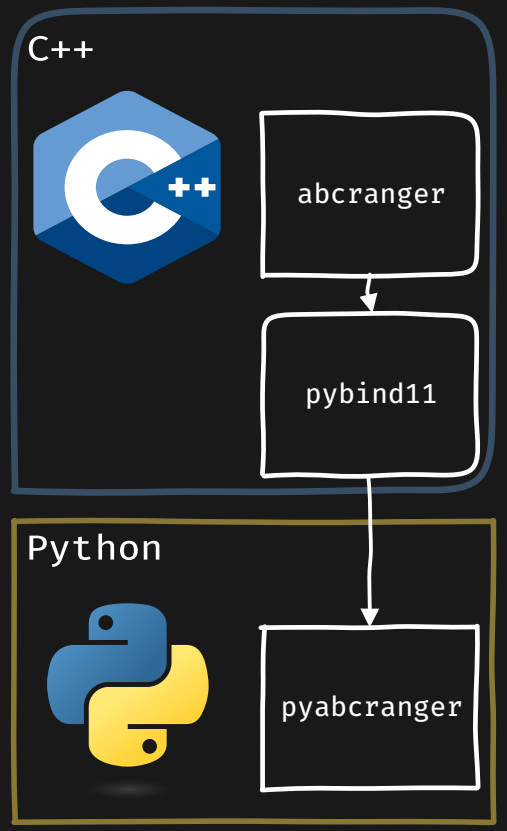
C++ part extensively using Eigen, conflicting C++ dependency management.

Memory copy

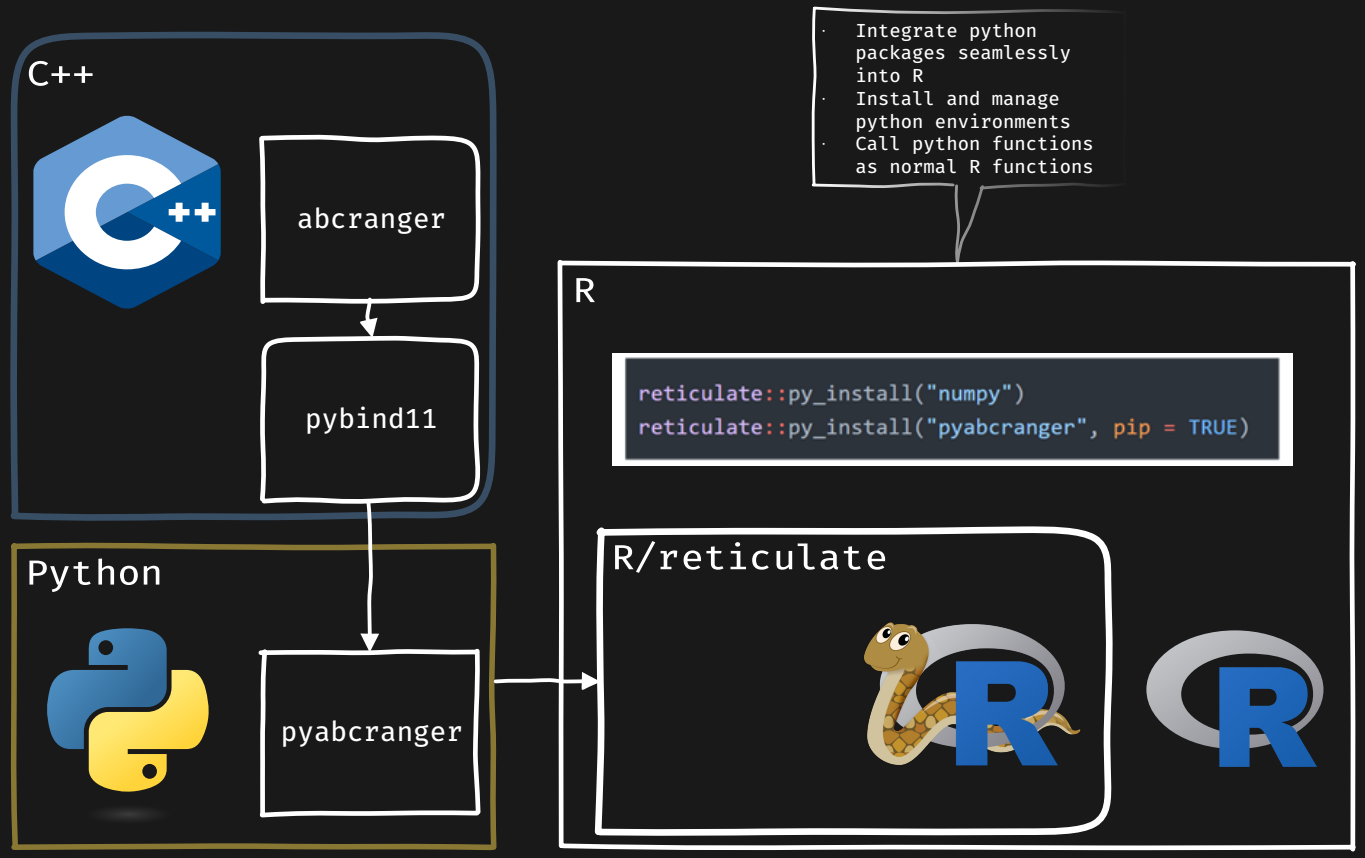
Copy of data between R and C++ (e.g. NumericVector to Eigen::VectorXd

$\Rightarrow$ Whereas pybind11 allows avoiding it, by using numpy arrays directly

Reticulate to the rescue



Reticulate to the rescue



Minimal example, C++ part

```
- CMakeLists.txt  ①
- main.cpp        ②
- main.h          ③
```

① CMake file, to build the C++ library

② C++ file, containing the function to be called from Python/R

③ C++ header, containing the function declaration for pybind11

Minimal example, Python and R parts

`setup.py`

Python setup file, to call CMake and build the Python module

`script.R`

R script, to install/build/call the Python module

Minimal example, Python and R parts

`setup.py`

Python setup file, to call CMake and build the Python module

`script.R`

R script, to install/build/call the Python module

```r
library(reticulate)
config <- reticulate::py_config()
setupargs <- c("-m", "pip", "install", "--quiet", shQuote("."))
system2(config$python, setupargs)                                    ①
my_module <- import("my_module")                                     ②
my_module$my_function(1:10)                                          ③
```

① Compile/install the local Python module

② Load the Python module

③ Call the Python function

Conclusion

- `pybind11` is a modern C++ library: creating Python modules from C++ code
- `reticulate` is an R package: calling Python code from R
- `pybind11` and `reticulate` used together: using existing C++ code from R, without the need to rewrite it in Rcpp

Conclusion, continued

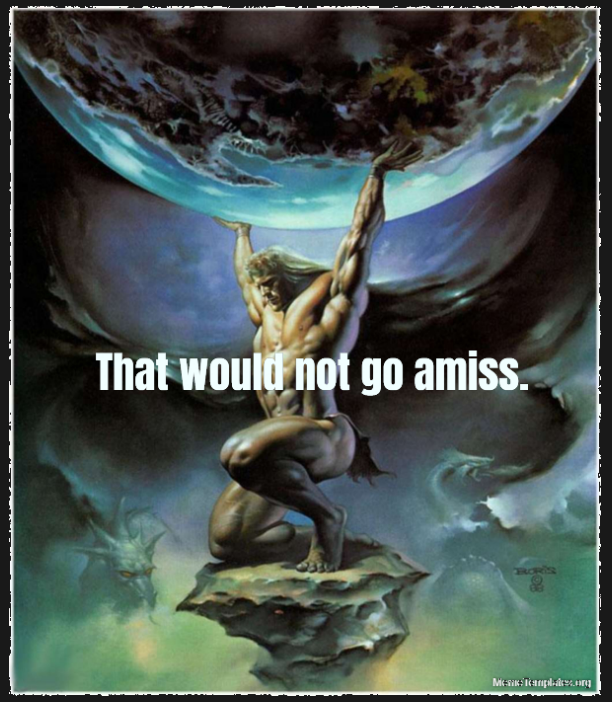
Finally

⇒ Only one interface and one package to maintain, the Python one

Conclusion, continued

Finally

⇒ Only one interface and one package to maintain, the Python one



References

[1] W. Jakob, J. Rhinelander, and D. Moldovan, “pybind11–seamless operability between c++ 11 and python,” *URL: https://github. com/pybind/pybind11*, 2017.

[2] K. Ushey, J. Allaire, and Y. Tang, *Reticulate: Interface to ’python’*. 2023.

[3] W. Tang *et al.*, “Towards understanding third-party library dependency in c/c++ ecosystem,” in *37th IEEE/ACM international conference on automated software engineering*, 2022, pp. 1–12.

[4] “GitHub Language Stats.” https://madnight.github.io/githut/ (accessed Mar. 23, 2023).

[5] “Stack Overflow Developer Survey 2022,” *Stack Overflow*. https://survey.stackoverflow.co/2022/?utm_source=social-share&utm_medium=social&utm_campaign=dev-survey-2022 (accessed Mar. 23, 2023).

[6] D. Eddelbuettel and R. François, “Rcpp: Seamless R and C++ integration,” *Journal of Statistical Software*, vol. 40, no. 8, pp. 1–18, 2011, doi: 10.18637/jss.v040.i08.

[7] F.-D. Collin, A. Estoup, J.-M. Marin, and L. Raynal, “Bringing ABC inference to the machine learning realm: AbcRanger, an optimized random forests library for ABC,” in *JOBIM 2020*, 2020, vol. 2020, p. 66.

[8] G. Guennebaud, B. Jacob, *et al.*, “Eigen v3.” http://eigen.tuxfamily.org, 2010.

[9] C. R. Harris *et al.*, “Array programming with NumPy,” *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020, doi: 10.1038/s41586-020-2649-2.