Publication of the Belle II Software

Fabrizio Bianchi, Giacomo de Pietro,
David Jaffe, Thomas Kuhr,
Frank Meier, Martin Sevior,
Philipp Urquijo (and many more)

Start in 2008

- Belle II Collaboration officially founded in 2008
- Coordinated software development by a few people (using svn) started about then
- Some early discussions about licensing and making the software open source, but kept internal to not be limited by external constraints

Belle II Licensing Task Force

- Active 2017/18, Members: Carl Rosenfeld (chair), Martin Ritter (externals), Martin Sevior, Michel Jouvin, Rok Pestotnik, Thomas Kuhr, Torben Ferber (generators), Thomas Hauth, Frank Meier, Marko Bracko
- Overview of licenses of used packages
  → many! (see https://github.com/belle2/externals#readme)
  → difficult to be compatible with all of them
- Discussions with LHC colleagues
  → host lab as software owner?
- KEK, DESY, CERN considered, but not viable

Software Package	Version	License
Artistic Style (astyle)	2.05.1	MIT
bazel	0.26.1	Apache 2.0
BELLE_FLC	?	-
belle_legacy	1.10	-
GNU Binutils	2.36.1	GPLv2
Boost C++ Libraries	1.72	Boost License
bzip2	1.08	BSD-style
clhep	2.4.1.3	LGPLv3
CMake	3.19.7	BSD-style
Cosmic-ray Shower Library (CRY)	1.7	-
cppzmq	4.7.1	MIT-style
curl	7.75.0	MIT-style (see also: License Mixing)
davix	0.7.6	LGPLv2
Eigen	3.3.9	MPLv2
EOS DOI 10.5281/zenodo.1304320	0.3.3	GPLv2
EPICS	7.0.5	EPICS License
EvtGen	R02-00-00	GPLv3
FANN	2.2.0+git8ae6bb0	LGPL
FastBDT	5.2	GPLv3
FastJet	3.3.4	GPLv2 (and others)
FFTW3	3.3.9	GPL
GCC, the GNU Compiler Collection	10.2	GPLv2/GPLv3 + Runtime library Exception
GDB: The GNU Project Debugger	10.1	GPLv2
Geant4	10.06.p03	Geant4 License
git	2.31.0	GPLv2
git-lfs	2.13.2	MIT
GSL - GNU Scientific Library	2.6	GPLv3
Google Test	1.10.0	BSD
Hammer	1.1.0	GPLv3
HDF5	1.12.0	BSD
HepMC	2.06.11	GPLv2
jemalloc memory allocator	5.2.1	BSD-style
libffi	3.3	BSD-style
Libxml2	2.9.10	MIT
Libxslt	1.1.34	MIT
LZ4	1.9.3	BSD + GPLv2

Belle II Open Source Task Force

Open Source of Belle II Software Task Force

10 May 2019

The Belle II Executive Board establishes the "Open Source of Belle II Software Task Force" and mandates it to draft an Open Source of Belle II Software Policy Document that:

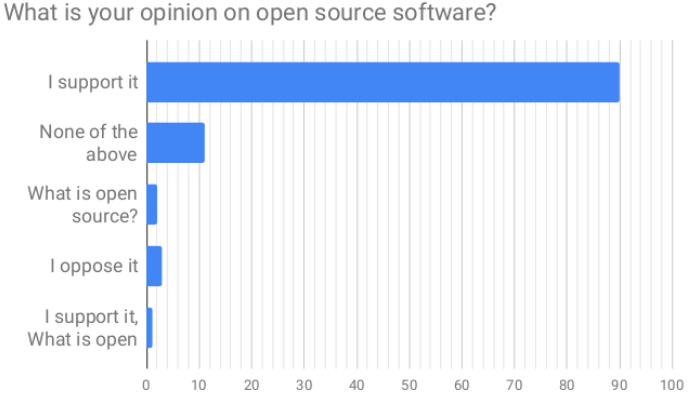
1) Defines categories of Belle II software that falls under the Open Source of Belle II Software for consideration of the policy.
2) Documents advantages and disadvantages of sharing Belle II software with entities outside Belle II as well as advantages and disadvantages of making Belle II software fully open publicly.
3) Includes procedures for of sharing Belle II software with entities outside Belle II and/or with making Belle II software fully open publicly.

- Members: Fabrizio Bianchi, David Jaffe, Thomas Kuhr, Martin Sevior, Phillip Urquijo

Advantages and Disadvantages

- Doing right by young people
- Proper reuse of code
- Collaboration on software
- Scientific value of software
- Software quality
- Access to open source services and support
- Access without authentication, searchability
- Prerequisite for open data
- Return for public funding

- Risk of exposing scientific intellectual property
- Security risk
- Licensing and copyright issues
- Risk of making inappropriate content public
- Possible reluctance to commit code



Approved Policy

- Ratified by Institutional Board in June 2020:
  - Belle II strongly profits from, recognizes the value of, and supports open source software. By committing code to a Belle II repository the right to be recognized as original author is granted and the author allows the members of the collaboration to use and modify it.

    A Belle II repository can be declared public if all contributors agree that it can be distributed under an open source license. By committing code to a public Belle II repository contributors agree to the publication of their code and affirm compliance with the code of conduct. The copyright remains with the contributors.
  - https://confluence.desy.de/display/BI/Belle+II+Open+Source+Software+Policy

Contributors Agreement

- Asked contributors for approval to publish the Belle II Analysis Software Framework (basf2) via pull request and email
- ➔ 258 approval, 0 disapprovals, no reply from 94 (<10% of code)
  → average probability of disapproval < 0.35% at 90% CL
- IB vote: 72 in favor, 3 opposed, 9 abstained

CONTRIBUTING.md:

By contributing to this repository you

- affirm that your contribution to the Belle II software complies with the Belle II code of conduct (see https://confluence.desy.de/download/attachments/34997090/Belle%20II%20bylaws.pdf),
- grant the members of the Belle II collaboration the right to use and modify your contribution,
- grant the Belle II collaboration the right to distribute your contribution under an open source license of its choice,
- provide the contribution "as is" without warranty.

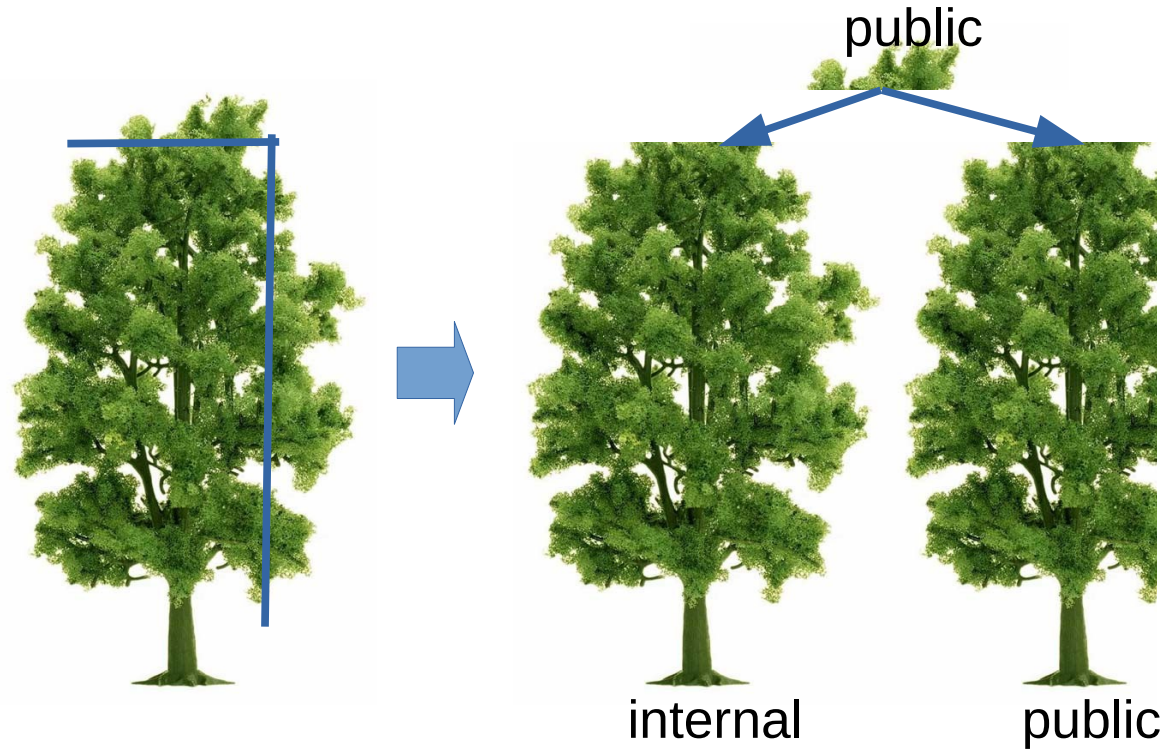
Repository Preparation

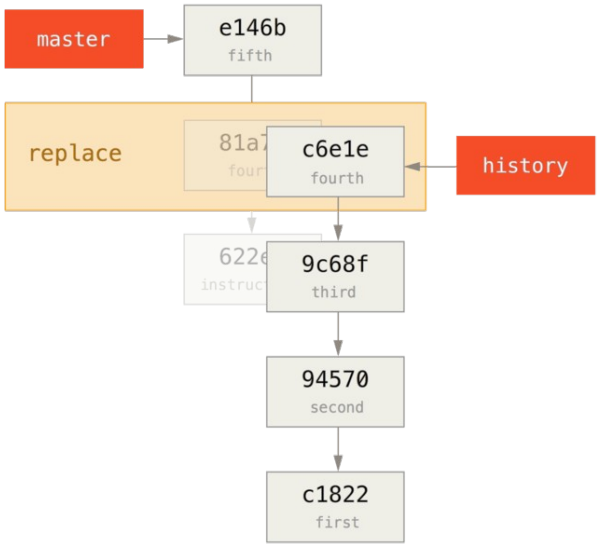
- License file (LGPLv3) and statement
- License issues
  - Generator authors contacted
    → CREDITS.md
  - Some (now unused) Xilinx code found
  - Sometimes licenses mentioned in contributed code
- Inappropriate content (in source code or commit messages)
  - Tools: profanity_check, better_profanity
    → many false positives
  - Searched for name of collaborators
  - Assessed by Diversity Officers

```
/**************************************************************************
 * basf2 (Belle II Analysis Software Framework)                           *
 * Author: The Belle II Collaboration                                     *
 *                                                                        *
 * See git log for contributors and copyright holders.                    *
 * This file is licensed under LGPL-3.0, see LICENSE.md.                  *
 **************************************************************************/
```

```
0.999843742020064 [ecl/calibration/examples/muonPairAlignment/e1ProfilePlot
s.C:51]:     const int nCrysPerID[nThetaID] ={48, 48, 64, 64, 64, 96, 96, 9
6, 96, 96, 96, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144,
144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144,
144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144, 144,
144, 144, 144, 144, 144, 144, 144, 144, 96, 96, 96, 96, 96, 64, 64, 64};
0.9905618606318605 [generators/babayaganlo/Asu3.f:3099]:          den=mom(
1)*mom(1)-mom(2)*mom(2)-mom(3)*mom(3)-mom(4)*mom(4)
***: "zero suppression, hot strip masking not implemented"
*** [trg/cdc/examples/TrgCdcUnpacker.py:14]: argc = len(argvs)  # of arg
```

History Change

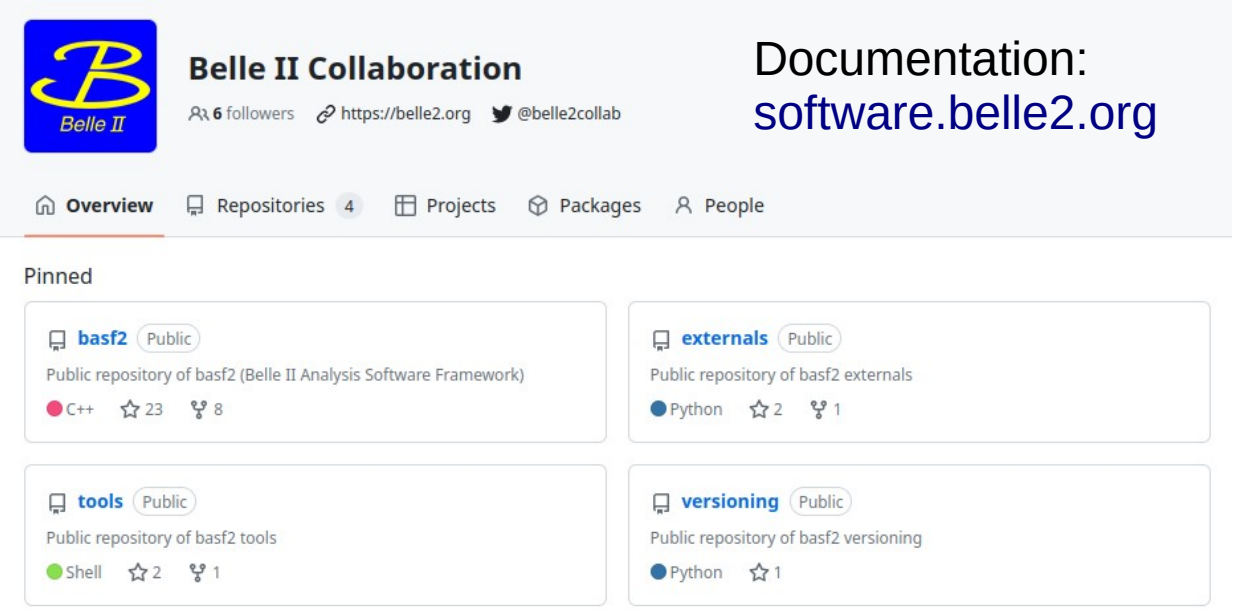




`b2code-history`
command to connect new main to history

Open Source (Finally)

- Released on github.com/belle2 in July 2021



- **Development still in internal repository**
- Manual update of main branch and releases from time to time
- ➔ To be automated (or move to public development)

Impact?

- ✗ No negative consequences observed so far
- ✔ Software citations

DOI 10.5281/zenodo.6949513

- ✔ Google Summer of Code
- ✔ Job application support?

main ▾ **basf2 / CITATION.cff**

tkuhr CITATION.cff added

1 contributor

9 lines (9 sloc) | 395 Bytes

```
1  cff-version: 1.2.0
2  message: "If you use this software, please cite it using the metadata from this file and
3  title: "Belle II Analysis Software Framework (basf2)"
4  authors:
5    - name: "The Belle II Collaboration"
6  license: LGPL-3.0-or-later
7  doi: 10.5281/zenodo.5574115
8  repository-code: "https://github.com/belle2/basf2"
9  url: "https://software.belle2.org"
```

Conclusions

- The longer one waits the harder it gets to publish software
- ➔ But it can be done
- I think it was worth the effort
- ➢ Many thanks to those who contributed!
- What's next?
  - Automated updates/releases
  - Public development?
  - Public data?