

EDWARD MOYSE

SURVEY OF DOCUMENTATION TOOLS AND PRACTICES IN OTHER EXPERIMENTS

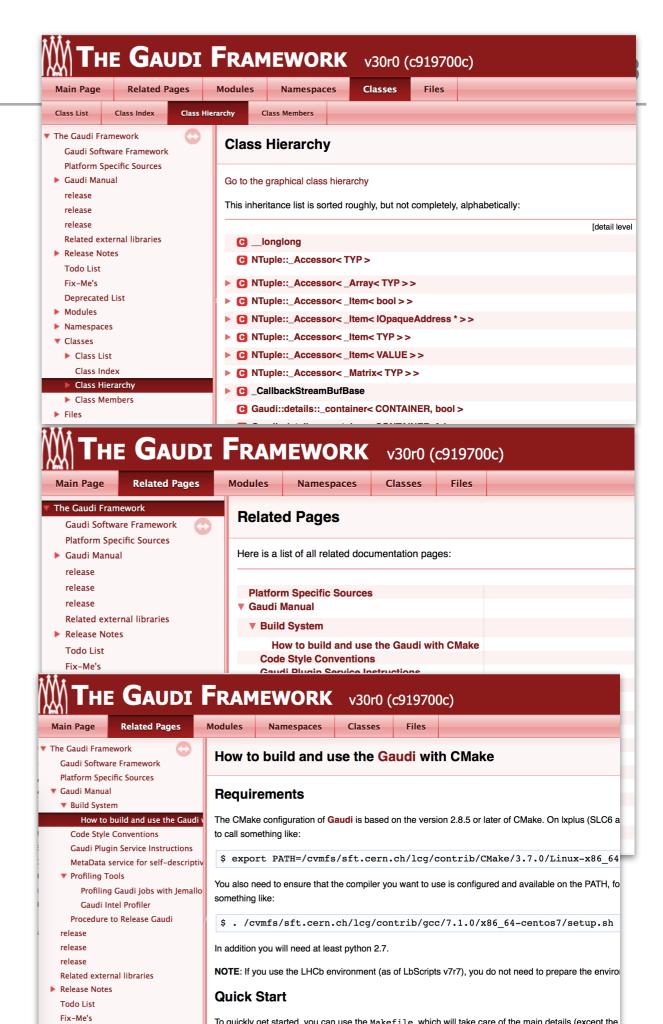
INTRODUCTION

- We thought it would be instructive to try to get a feeling for what other experiments and opensource projects use
 - Including code documentation, user guides (e.g. git tutorial, our workbooks) and wikis
- In the next few pages I will show a few examples
 - (And add some subjective opinions)
- Then will summarise a few of the discussions (thanks to Graeme for the introductions!) we had with experts from other communities, in particular:
 - ▶ Linear Collider Frank Gaede
 - SuperNemo & Geant4 Ben Morgan
 - ▶ Belle II Martin Ritter

DOCUMENTATION SURVEY

GAUDI

- Main webpage looks to be a simple static homebuilt one
- Extensively use doxygen to document code
 - however they make a lot more use of 'related pages' than ATLAS does
 - have instructions for how to profile, code style conventions, how to build
 - i.e. they use doxygen for user guides
 - Subjective bit I find them a bit ugly, but at least they're full of content and shipped 'for free' with the code
 - They build documentation for every release



LINEAR COLLIDER

- Moved from SVN to github just over a year ago
 - https://github.com/iLCSoft.
 - Moved/moving to exclusively use github for issues, feature requests and general discussion
 - "Advantage is everything is in one place"
 - Even documentation is in github (via README.md files), and plan is to stick to this for most smaller packages...
 - Still in progress main <u>webpage</u> is simple HTML, generated by <u>zms</u>(?)
 - Still have doxygen for <u>some</u> packages
- One exception is <u>DD4HEP</u>, which has a wider audience and so it was considered worthwhile to give it a dedicated (and fancier webpage)
 - Created using <u>hugo</u> (faster? alternative to jekyll). Uses markdown + extensions ('shortcodes') for content

■ README.md

Marlin (Modular Analysis and Reconstruction for the LINear Collider)

uild passing cov

The idea is that every computing task is implemented as a processor (module) that analyzes data in an LCEvent and creates additional output collections that are added to the event. The framework allows to define the processors (and their order) that are executed at runtime in a simple steering file. Via the steering file you can also define named parameters (string, float, int - single and arrays) for every processor as well as for the global scope. By using the framework users don't have to write any code that deals with the IO they simply write processors with defined callbacks, i.e. init(), processEvent(), end().

Marlin is distributed under the GPLv3 License



Instructions for building Marlin with CMake:

. path_to_ilcsoft_installation/v01-XX/init_ilcsoft.sh

Return to website



DD4hep Manual

2 User Manual

This chapter describes how supply a physics application developed with all the information related to the detector which is necessary to process data from particle collisions and to qualify the detecting apparatus in order to interpret these event data.

The clients of the detector description are the algorithms residing in the event processing framework that need this information in order to perform their job (reconstruction, simulation, etc.). The detector description provided by DD4hep is a framework for developers to provide the specific detector information to software algorithms, which process data from particle collisions.

In the following sections an overview is given over the various independent elements of *DD4hep* followed by the discussion of an example which leads to the description of a detector when combining these elements. This includes a discussion of the features of the *DD4hep* detector description and of its structure.

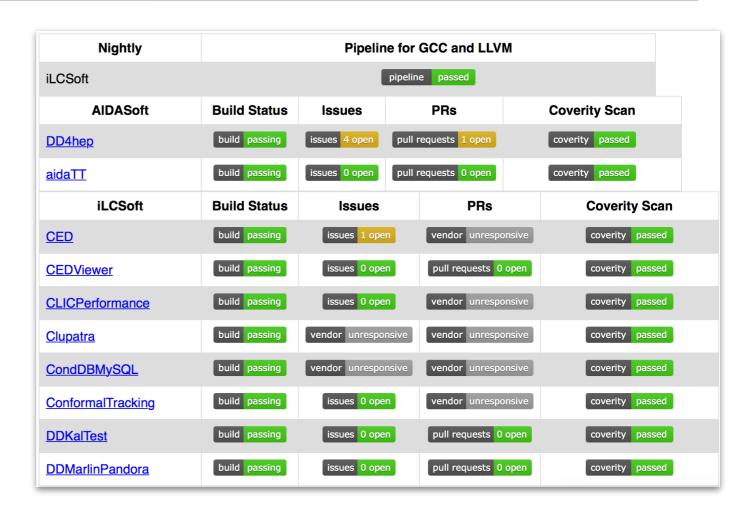
2.1 Building DD4hep

The *DD4hep* source code is freely available. See the licence conditions. Please read the Release Notes before downloading or using this release.

LINEAR COLLIDER (2)

- Use github pages for some status pages
 - http://ilcsoft.github.io

Main webpage - plain HTML





DOCUMENTATION SURVEY

SUPERNEMO

- Use jekyll for <u>main page</u>
- Main code documentation is in doxygen. Also use doxygen + markdown for the <u>user guides</u>
- And also have an internal wiki...
 - Apparently not used much yet



The Experime

The Collaborat

News

Publications and Media



How about a SuperNE presentation or poster

Request a spe



Searching for Neutrinoless Double Beta Decay

The NEMO (Neutrino Ettore Majorana Observatory) collaboration is an international physics effort including the experiment SuperNEMO and its predecessor, NEMO-3.

The SuperNEMO demonstrator module is currently being assembled at the LSM underground lab, located in the Fréjus tunnel near Modane, France. NEMO-3, also at the LSM, ran from 2003-11. Its rich repository of data is still being analysed today.

Both SuperNEMO and NEMO-3 are designed to study extremely rare double-beta decay processes, and in particular are looking for evidence of neutrinoless double beta decay. This is a rare type of radioactive decay which has been predicted, but has never been observed. If this process was seen, it would prove that neutrinos were their own antiparticles, which could be a clue to the matter-antimatter

Falaise 3.1.1

SuperNEMO Software Toolkit

Main Page Related Pages Namespaces ▼ Classes ▼ Files ▼

Welcome to the Falaise Documentation

Falaise is the software system for the SuperNEMO experiment. These pages document the usage of the core applications, tog functionality and the full C++ API of the Falaise extension system.

Getting Started

If you're reading this online and don't yet have an install of Falaise, a Quickstart Guide is available. Note that at present Falais and even here not all variants are quaranteed to work!

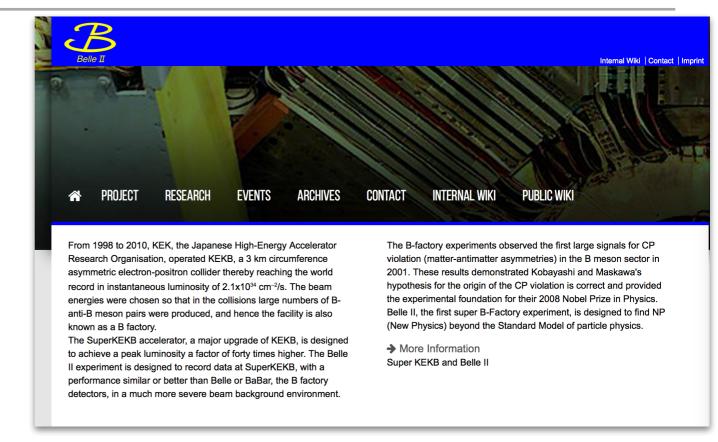
User Guides for Core Applications and Libraries

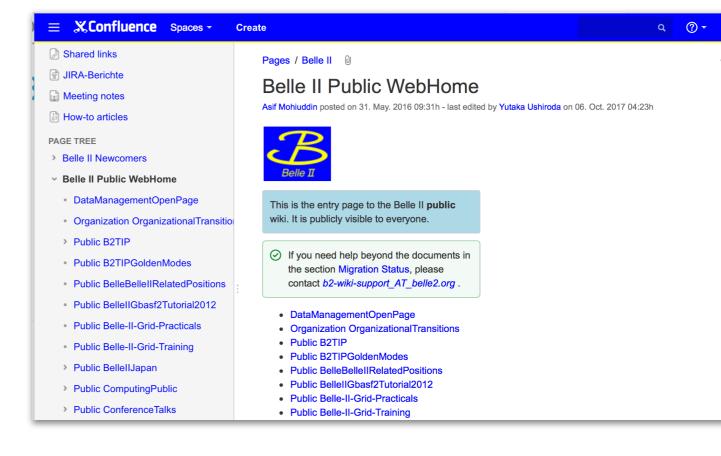
Once you have an install of Falaise, the following pages provide introductions to using and extending the core applications:

- FLSimulate: guide for using the Geant4 based simulations of the SuperNEMO Demonstrator and Commissioning detections.
- · FLReconstruct guide for using the pipeline application for reading, reconstructing and outputting data generated by the
 - Data Structures output by FLReconstruct
 - Converting FLReconstruct BRIO Format to ROOT
 - Customizing the pipeline for flreconstruct.

BELLE II

- C++ documentation is done in Doxygen, but they run a 'preexec' (INPUT_FILTER=) to insert some documentation to group it
 - Has consequence that line numbers don't entirely match source
 - They also parse 'algorithms' and add a diagram showing the inputs and outputs (would be trivial for us to do with Read/WriteHandles I guess!)
- In process of switching python documentation from doxygen to sphinx
 - Later will shift analysis code too
 - "Personally I think it's very nice how it allows to easily mix manual and automated documentation"
- Seems like they're using <u>Confluence wiki</u>, both for internal use, and <u>public</u>
 - Looks like quite desy has a contract with atlassian?
 - Has anyone checked





LHCB

- LHCb <u>starterkit</u> is very nice manual style documentation
- Written in <u>gitbook</u>
 - Markdown based
 - Is free (subject to some <u>limitations</u>) for opensource projects
- Also use <u>twiki</u> extensively
- Not sure if they use doxygen(!)
 - Couldn't really find any code documentation at all
 - starterkit does link to the gaudi documentation though

LHCb Starterkit



Starterkit is a group of physicists who want to improve the working lives of young researchers working on the LHCb experiment.

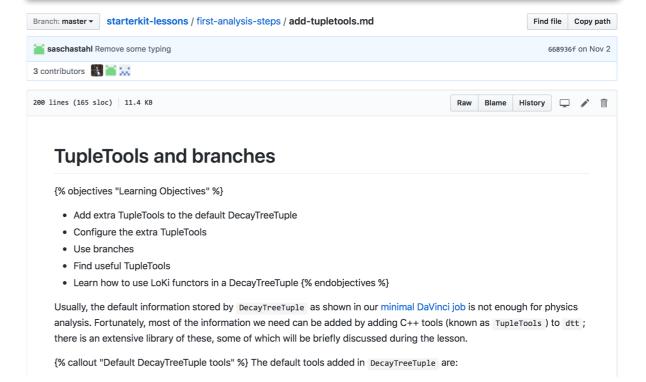
We create and organise workshops that cover topics ranging from basic programming skills, such as using Bash and Python, up to in-depth explorations of the LHCb software.





From the bottom up

The goal of Starterkit is to demystify the LHCb software. If you've ever felt like you're not sure why you're writing the code you're writing, or not sure what code you even need to write, we can help you. We start from the very basics, each lesson building on the last. Once you understand the details and how things fit together, you have the power to learn more on your own.



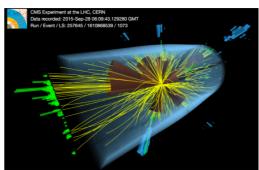
DOCUMENTATION SURVEY

CMS

- Code is on github
- Use github pages for user guides
- Doxygen to document C++
 - They have a hand-build 'splash page' for all releases
 - (implies they build doxygen for many releases)

CMSSW on GitHub Project page Feedback

Welcome to CMS and CMSSW



The LHC smashes groups of protons together at close to the speed of light: 40 million times per second and with seven times the energy of the most powerful accelerators built up to now. Many of these will just be glancing blows but some will be head on collisions and very energetic. When this happens some of the energy of the collision is turned into mass and previously unobserved, short-lived particles – which could give clues about how Nature behaves at a fundamental level - fly out and into the detector. Our work includes the experimental discovery of the Higgs boson, which lead to the award of a Nobel prize for the underlying theory that predicted the Higgs boson as an important piece of the standard model theory of particle physics.

CMS is a particle detector that is designed to see a wide range of particles and phenomena produced in high-energy collisions in the LHC. Like a cylindrical onion, different layers of detectors measure the different particles, and use this key data to build up a picture of events at the heart of the collision.

Many of CMS Software components (CMSSW) are hosted on Github. You can find here a

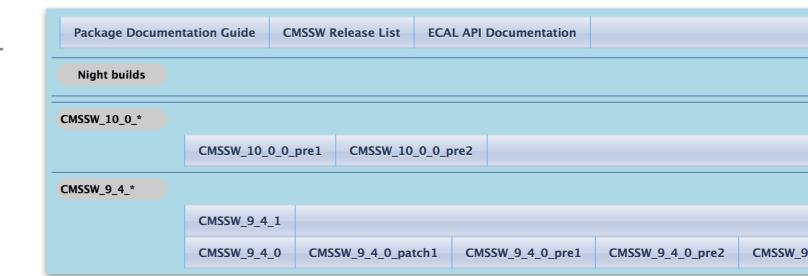
Quick start

- Main page
- FAQ: CMSSW on Github
- Proposing changes to CMSS
- CMSSW pull request approvements workflow
- Collaborating with peers
- Working with CMSSW and UserCode
- Resolving conflicts & porting features
- How to use git through a pro

Release

Management

- Building a CMSSW release
- Integration Builds results
- Historical plots
- Starting a new release cycle
- Forward ports
- Managing users and categories
- List categories and package
- Latest available IBs
- List of pending PRs
- CMS-bot documentation



ASIDE: THIS IS WHAT ATLAS HAS

ATHENA DOCUMENTATION

Release:

atlas_22.0.X-DOX

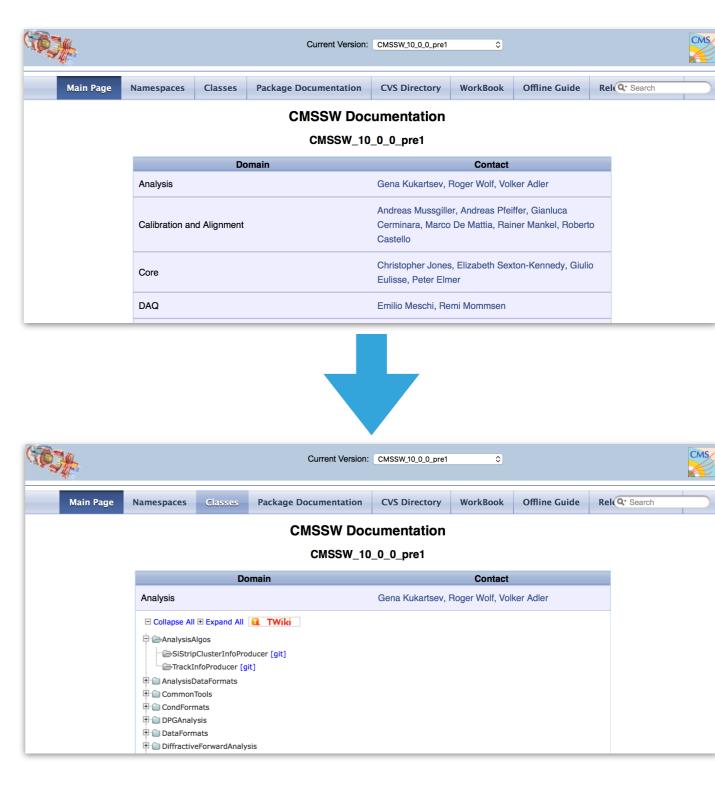
https://atlas-sw-doxygen.web.cern.ch/atlas-sw-doxygen/index.php

- I think we should make this look a bit nicer.
 - Adam did a great job with the functionality! But can we add some CSS?
- I think we should be building Doxygen for multiple releases.

CMS

- Doxygen cont'd
 - They divide up into domains with contacts
 - Can expand and shows content from twiki(?) and links to git
 - Code to generate this <u>here</u>
 - Not perfect some links are broken, but I do think this linking between git and doxygen is very nice, and I also think that dividing the software up into Sections, makes it more comprehensible.
- Also:
 - ▶ LXR to search code
 - Make extensive use of Twiki, much as we do

http://cmsdoxygen.web.cern.ch/cmsdoxygen/CMSSW_10_0_0_pre1/doc/html/index.html



Q Search

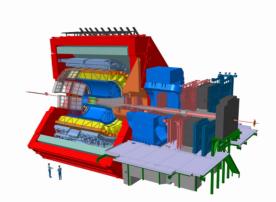
ALICE

- User guides
 - Use github pages for top level software webpages & user guides
 - They also use doxygen for this (see right)
- And of course, use doxygen for documentation of code

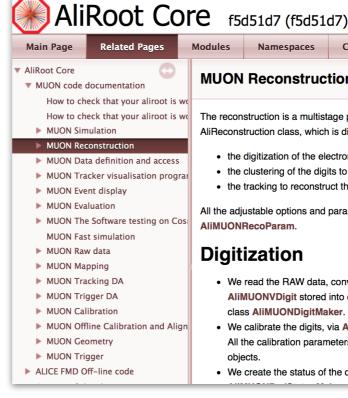


ALICE Software on Github

The ALICE Collaboration has built a dedicated detector to exploit the unique physics potential of nucleus-nucleus collisions at LHC energies. Our aim is to study the physics of strongly interacting matter at the highest energy densities reached so far in the laboratory. In such condition, an extreme phase of matter called the quark-gluon plasma - is formed Our universe is thought to have been in such a primordial state for the first few millionths of a second after the Big Bang. The properties of such a phase are key issues for Quantum



Chromo Dynamics, the understanding of confinement-deconfinement and chiral phase transitions. For this purpose we are carrying out a comprehensive study of the hadrons, electrons, muons and



MUON Reconstruction

The reconstruction is a multistage process, driven by the AliMUONTracker and AliMUONReconstructor classes via the AliReconstruction class, which is divided into three parts:

· the digitization of the electronic response,

Namespaces

. the clustering of the digits to locate the crossing point of the muon with the chamber,

Files

the tracking to reconstruct the trajectory of the muon in the spectrometer from which we can extract the kinematics

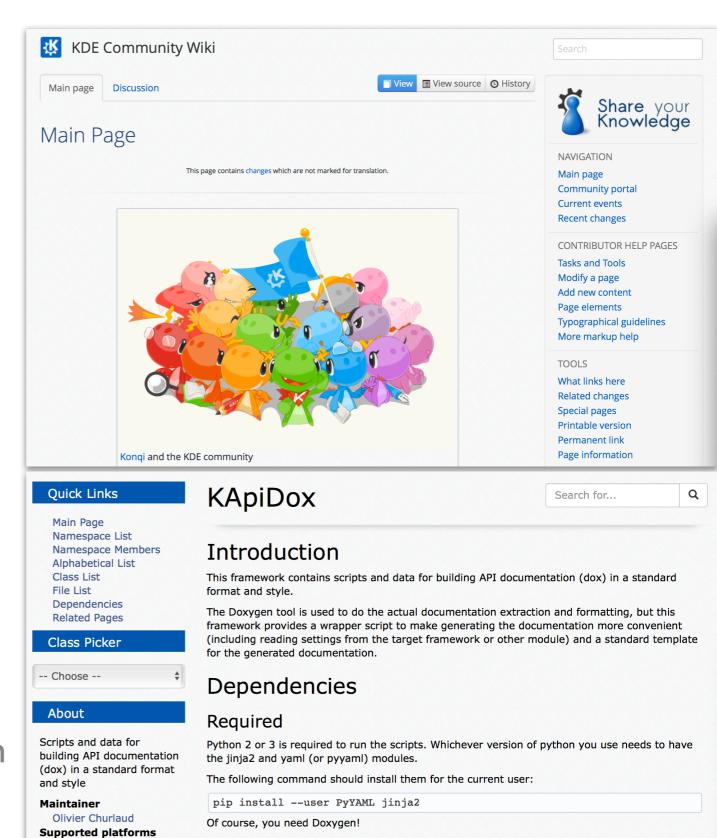
All the adjustable options and parameters used to tune the different part of the reconstruction are handled by the class AliMUONRecoParam.

Digitization

- . We read the RAW data, convert them (convert them back for simulated data) to digit (object inheriting from AliMUONVDigit stored into containers inheriting from AliMUONVDigitStore). This conversion is performed by the class AliMUONDigitMaker.
- . We calibrate the digits, via AliMUONDigitCalibrator, by subtracting pedestals and multiplying by a constant gain All the calibration parameters (pedestals, HV) are read from the OCDB and stored into AliMUONCalibrationData
- We create the status of the digit (e.g. pedestal higher than maximum or HV switched off), using

KDE

- Thought it might be interesting to look at a big opensource project
- They seem to use a lot of homegrown tools for their documentation
- For example, I think their wiki is custom
- They do, however, use <u>doxygen</u> (and lxr) for code documentation



Doxygml and doxypypy might be needed to let doxygen document respectevely qml and python

FreeBSD, Linux,

Community

MacOSX, Windows

IRC: #kde-devel on

Optional

sources.

SUMMARY OF TOOL USE

This is a bit of a guess in many cases...

Treat as illustrative rather than authoritative!

Project	KDE	Gaudi	Linear collider	Supern emo	Belle II	LHCb	CMS	Alice
Webpage s	html	html	github pages, html	Jekyll		twiki, html	twiki, html	github pages
Code	Doxygen + LXR	Doxygen	Doxygen	Doxygen	(filtered) Doxygen Sphinx	?	Doxygen	Doxygen
User guides	Userbas e <u>wiki</u>	Doxygen	github, hugo	Doxygen	confluen ce wiki	gitbook	github pages, twiki	github pages, Doxygen

HIGHLIGHTS FROM DISCUSSIONS

- No one had an alternative to using doxygen for C++ code, but no one I spoke to was completely happy with it
- ▶ We could add Breathe to generate Sphinx from doxygen xml, but this would slow down our documentation generation, and it isn't obvious its output is better
- One generate complaint of doxygen is that it doesn't make it easy to structure the code it basically dumps everything it is given into one massive list, which doesn't really work for Athena
 - We get around this to some extent by using packagedocs (our replacement for mainpages). Not ideal, because there is no automated link from the class to the @page (packagedoc)
- > Sphinx requirements more configuration but "that seems to result in documentation which is actually readable (as opposed to an alphabetic index of classes)"
- "The drawback is that users need to learn yet another markup language (reStructuredText)."

SUMMARY

- We seem to be using similar tools to everyone else, (and aren't doing too badly with our documentation - relatively speaking - from what I can see)
- Could try Hugo instead of Jekyll for static websites, (but Jekyll doesn't really seem slow to me)
- Almost everyone uses Doxygen for code documentation, and yet no one I spoke to is completely happy with it
 - Main complaints relate to problems navigating/understanding large codebase
 - Occasional Doxygen bugs (maybe Doxygen-clang will help?)
 - I think we could learn a bit from CMS about integrating it with git, and making it less monolithic
- Gitbook looks nice (and LHCb startkit is extremely good IMO), but I don't think nice enough for us to tie ourselves into a commercial product
- Wasn't able to look too much at other wikis (protected behind firewalls)