

Bioc Technical Advisory Board Minutes

5 May 2022

Attending: Vince Carey, Lori Shepherd, Marcel Ramos, Laurent Gatto, Nitesh Turaga, Rafael Irizarry, Charlotte Sonesson, Michael Love, Stephanie Hicks, Jennifer Wokaty, Levi Waldron, Hervé Pagès, Alexandru Mahmoud, Aedin Culhane

Regrets: Wolfgang Huber, Robert Gentleman, Hector Corrada Bravo, Aaron Lun, Shila Ghazanfar, Kasper Hansen

:03 - :07) Previous [minutes](#) approved. For information, [CAB Minutes 04-14-2022](#).

:07 - :20) TAB membership

- Nomination process
- Step-down events, acknowledgments.
 - Review phrase in form on expectation of participation of TAB members.
 - Term expiration: Charlotte, Laurent, Stephanie, Aaron.
 - Discussion in TAB of new nominees.
 - Expectations for continuity and renewal - hopes that technical steps of the project are informed by TAB expertise.
- [Nomination form](#)
- Suggested timeline: announce this week, application deadline Tue May 31 (to have time to summarize before the next TAB meeting), discussion at June 2 TAB meeting, voting the week after that.
- How can we increase the interest/number of applicants, identify suitable candidates? Representatives of active working groups/specialists in particular areas (e.g., developer mentorship group, infrastructure/cloud, variants, spatial, single-cell). Input from a broader group representing Bioconductor. Identify larger areas and make sure there is a representative from each on the TAB. <http://workinggroups.bioconductor.org/>.

:20 - :30) Governance issues: suggestion that requirements for joining NumFocus be an impetus for changes - need a clear decision structure for use of funds and other assets.

- 16 May governance meeting (consider a working meeting with La Piana - might have to reschedule for time zone reasons).

:30 - :45) Short invited talks:

- Levi ([slides](#))
 - Q: Does the update happen even if there are no changes? Yes - there is at least the change in timestamps.
 - Q: Do you have to subscribe to MediaWiki? No, open system.
- Laurent ([slides](#))

:45 - :50) Discussion:

- 3.15 news (major thanks due to the core team), RJournal report needed.

- Overall pretty smooth.
- Landing pages, propagation of packages, etc should smooth out over the next week when additional builds (experiment data, workflows, etc) build on auto schedule.
- Post release tasks will commence (dockers, newly identified deprecate, switch SPB, additional hub resources, etc). Docker images and binaries available.
- BioC2022
 - If you plan to attend in person as part of CAB/TAB we have reserved in person spots. However due to limited in-person capacity (70 total) we will only hold the spot until May 15. Please confirm to Erica and Lori if you will need an in person registration spot before May 15!
 - Award deadline for nominations is May 20th. Please nominate. <http://bioconductor.org/about/awards/>
 - TAB.
 - Do we want our own dedicated slot or are we okay with sharing with CAB or Core? -> Keep separate
 - How long of a time slot would we prefer? -> 30 min TAB, 30 min core (back-to-back)
 - Do we plan on having a presentation or only open question/discussion?

:50 - :60) Goals for 3.16

- Discussion of older Bioconductor release availability and storage.
 - CloudFront vs S3 Bucket vs OSN vs other.
 - Current determinant of what is on CloudFront vs S3 bucket is the storage space on the instance we host on CloudFront. CloudFront however is more expensive. Would like to move all releases <= Bioc 3.9 to S3 bucket to reduce cost but this may mean slower download if not hosted through CloudFront. Eventually we think we would like to host old releases on OSN for free egress but need more storage.
 - Are there any concerns with making users have slightly slower download speeds to access older releases from not hosting on CloudFront to cut costs and resources?
 - <https://www.softwareheritage.org/>
- Distribution of assets: current role of AWS CloudFront, future roles of AWS Open Data, NSF/XSEDE Open Storage Network, mirrors, hybridization of diverse storage/distribution via Alex M geolocation service.
- Strong enhancement of community review corps, reduce engagement of core developers in package review - new batch of reviewers coming.
- Major introspection activities.
 - Identify core packages maintained by core developers.
 - Examples: Rhtslib, Rsamtools, VariantAnnotation, BSgenome, Biostrings, GenomicRanges, GenomeInfoDb [UCSC dependent, new genomes], AnnotationHub, BiocManager, DelayedArray, BiocFileCache, GenomicFiles, BiocParallel, SummarizedExperiment, S4Vectors, XVector,

ExperimentHub, AnnotationHubData, GenomicFeatures, Uniprot.ws, BiocCheck, BiocKubeInstall/binaries, GenomicAlignments, BiocStyle, biomaRt, rhdf5 (incl. rhdf5filters, Rhdf5lib).

- Identify APIs that govern user-facing resources and activities
 - Hub APIs need to be Swagger/OpenAPI ...
 - Cloud database resource (could use jetstream) (don't drop sqlite on user).
- Build system road map – Jen has established almost complete independence.
 - Containerization, configuration for the platforms.
 - Cadence of building.
 - SPB for new Windows toolchain ... simplifications.
- Role of zenodo communities as a simpler Hub contribution protocol (contribution, data unit test, adding new metadata) – durability seems good.
 - Question on slack: Is Figshare upload acceptable. Immediate reaction is yes but has anyone had any experience with Figshare reliability?
- Develop road maps for enhancement.
- Master Branch renaming - suggested to pick up the discussion again. Technically challenging, but may be easier if decoupling from other updates.