Bioc Technical Advisory Board Minutes

3 October 2024

Present: Vince Carey, Charlotte Soneson, Michael Lawrence, Wolfgang Huber, Laurent Gatto, Helena Crowell, Erdal Cosgun, Jacques Serizay, Hervé Pagès, Levi Waldron, Ludwig Geistlinger, Henrik Bengtsson, Marcel Ramos, Kasper Hansen, Alexandru Mahmoud, Davide Risso, Lori Kern, Jen Wokaty, Rafael Irizarry
Apologies: Stephanie Hicks

:03 - :04 Previous minutes approved.

:04 - :07 Brief discussion on Quarto vs R Markdown.
- Some advantages of Quarto:
    - The book concept - integrated search.
    - Collapsible elements (not only code).

:07 - :18 Welcome to new TAB members
- Michael Lawrence
    - Returning TAB member, brings perspective from R Core, R Foundation, Industry.
- Jacques Serizay
    - Interests include chromatin landscape, chromatin folding, background on molecular/cellular biology.
    - Aims: Improve cross-package documentation (BiocBook), develop/homogenize local/cloud-based data containers for epigenomics, teaching.
- Big thanks to outgoing members Michael Love and Sean Davis.

:18 - :38 Comments from existing members on goals/perspectives for coming year
- Some topics of interest:
    - Spatial omics data
    - Single-cell data
    - Data repositories / ExperimentHub
    - Build system move to GitHub Actions
    - (Generative) AI
    - Performance improvement of packages
    - Interactions with other communities (CRAN, scverse)
    - Multi-modal data
    - Infrastructure building
    - Making sure that Bioconductor stays competitive for users
    - Cloud infrastructure/on-disk representations of mass spec data
    - Perturb-seq
    - Long read genomics
    - Long term survival of the project

    - How to justify to developers submitting to Bioconductor
    - Teaching, outreach, conferences
    - Workflows spanning multiple packages - solving practical problems
    - GPU computing
- Possible working groups
    - Outreach
    - Cloud-based data representation (within cloud working group)

:38 - :60 Transition of build system to GitHub

- 1) hosting: moving git.bioconductor.org to a GitHub-based platform. Looking for package developer volunteers for (passive) testing: please apply via https://forms.gle/bU8jWCee9PkfY4AfA
- 2) building: with a focus on GitHub actions to deal with checking – improved DX, actions in collaborative effort with r-universe, https://bioc.r-universe.dev
- 3) distributing: “master.bioconductor.org” ec2 instance to GitHub releases (via nginx reroute for endpoint)
- Comments/questions
    - https://bioc.r-universe.dev/builds - clone of devel branch by r-universe.
    - GitHub Enterprise account (with few paid seats, affordable) - 500 concurrent GitHub Actions runners (compared to 20 for the free plan), to use for building artifacts. Could be complemented by local runners.
    - Package developers added as external contributors, with write access to their repositories.
    - Both release and devel branches to be built.
    - Packages must pass check to be released (changes to be made at r-universe to allow this).
    - Fork r-universe to be more independent (can still pull updates).
    - Idea: upon change to a package, the package is rebuilt, but not immediately propagated. When built/checked cleanly, propagate to Bioconductor, trigger rebuild of reverse dependencies. Full rebuilds of entire ecosystem can still be done e.g. once per week.
    - Makes Bioc dependent on GitHub - contingency plans. Don't commit to GitHub specifically, code can be moved to self-hosted runners. GitLab uses the same workflow syntax.
    - Separate the building from the hosting. For the hosting, to start with will probably push to the current place. In the long run, artifacts may be hosted directly as GitHub releases. Possibly with multiple mirrors.
    - R repositories working group in R Consortium: lots of groups need to provide cohorts of packages (e.g. pharma companies providing a stable set of packages for FDA submissions). R multiverse. Would be good to have a Bioconductor representative in the working group (Alex is planning to join the upcoming meeting).