

Bioc Technical Advisory Board Minutes

6 February 2025

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

Apologies: Rafael Irizarry

:03 - :04 Previous meeting [minutes](#) approved

:05 - :07 Robustifying Bioc assets to challenges arising from federal government interventions/disruptions – risk of github.com disruptions as well

- Possibility to replicate assets outside US (EMBL/EBI could be explored).

:08 - :17 Review U24 renewal highlights

- Motivating support for the project
 - Usage metrics (downloads, publications, module reuse, documentation reuse, QA benefits).
 - Depth of community benefit (online content, slack, support.bioconductor.org).
 - Maintenance need (OSCA, BBS, ...).
 - Updated CITATION files in packages would be helpful to track publications, and information could be easily scraped.
 - Keyword search on PubMed, sentiment analysis?
 - Could be a community event - encourage maintainers to update their CITATION files, could highlight package on social media channels.
 - Explore formal corporate sponsorship?

:18 - :23 CAB/teaching update

- Joint CAB/TAB meeting should be done soon (March/April?). March meeting 10PM GMT, April one 2PM GMT (CAB hosted), second Thursday - vote in slack channel. TAB would host fall meeting.
- Training group preprint: <https://arxiv.org/abs/2410.01351>

:23 - :46 Working group updates

- spatial
 - events: informal hackathon (Ascona 2024-09), scverse hackathon (Basel 2024-11), OSTA hackathon (Zurich 2025-01)
 - resources - there's a lot in Bioc, but may be hard to find.
 - harmonization - lots of existing infrastructure, but already redundant, diverging, hard to consolidate.
 - scalability & flexibility - increasing data size and complexity.

- interoperability with python, cloud-based/interactive visualization.
- hackathons are great, but most people drop out shortly after (missing incentives?), feeding off scverse events (and funding) is not sustainable (can we get designated funding for Bioc events?), learn from scverse's devel strategy (design docs, task boards etc)?
- cloud
 - revisit the strategic plan draft (security, sustainability, usability of cloud resources in the Bioc community).
 - Bioc core-maintained cloud resources (containerized Bioc Docker images, Kubernetes deployment via Helm Chart).
 - Preconfigured Amazon Machine Image (deprecated), Bioc for AnVIL, Bioc in Galaxy, alternative data hubs on Terra or SevenBridges instances.
 - Developments in the past year: Storage Defender for Bioconductor Hubs, GHA workflow and Docker images for Bioc developers, explore alternative cloud-based platforms for building, checking and distributing packages.
 - Still to be done: needs outreach and workshops to address low usage and knowledge (Infosec is difficult), outreach for available cloud resources (survey options), survey number of Bioc packages built and deployed with GHA or alternative cloud infrastructure, costs for hosting infrastructure in the cloud are rising.
- Hervé array [benchmarking work](#)
- OOP classes
 - Michael Lawrence presentation of S7 in BiocClasses working group: https://www.youtube.com/watch?v=CxiBwiga_1A&ab_channel=Bioconductor

:46 - :59 Language-agnostic/cross-language methods

- Could alabaster::saveObject replace, e.g., saveHDF5SummarizedExperiment?
- alabaster.sfe
 - issue of serializing package_version instances
 - see [discussion](#)
 - [catalogue](#) of schemas; [use](#); example [schema](#)
- Application of alabaster to, e.g., EPIC array?
- alabaster development was motivated by data sharing, ups the game with respect to metadata/FAIRification.
- Used also for mass spec data (<https://github.com/rformassspectrometry/MsIO>)
- Schemas are manually generated (have to be maintained, several of them have multiple versions)
- Nomenclature (saveObject/readObject) too generic? Could add more semantically rich function names on top of these more generic ones.

:59 - :60 Other topics

- GBCC2025 (<https://gbcc2025.bioconductor.org/>)