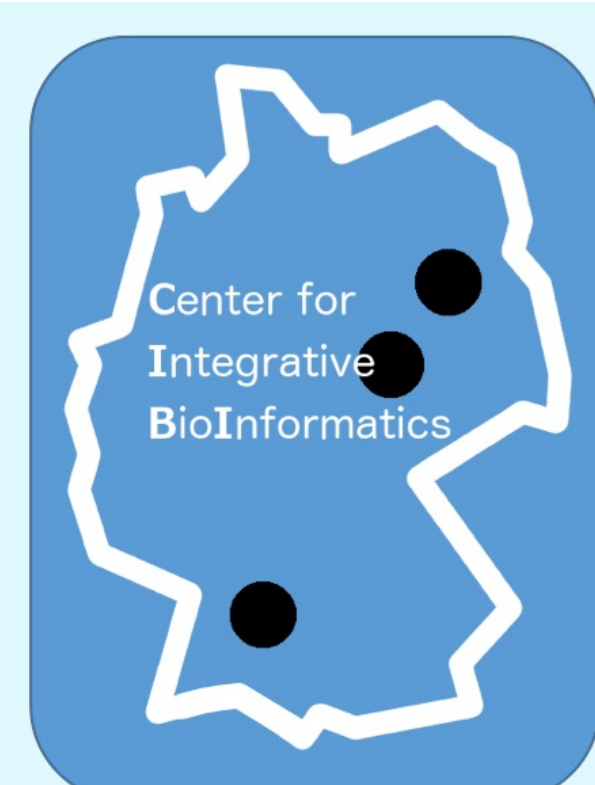




Center for Integrative Bioinformatics SeqAn

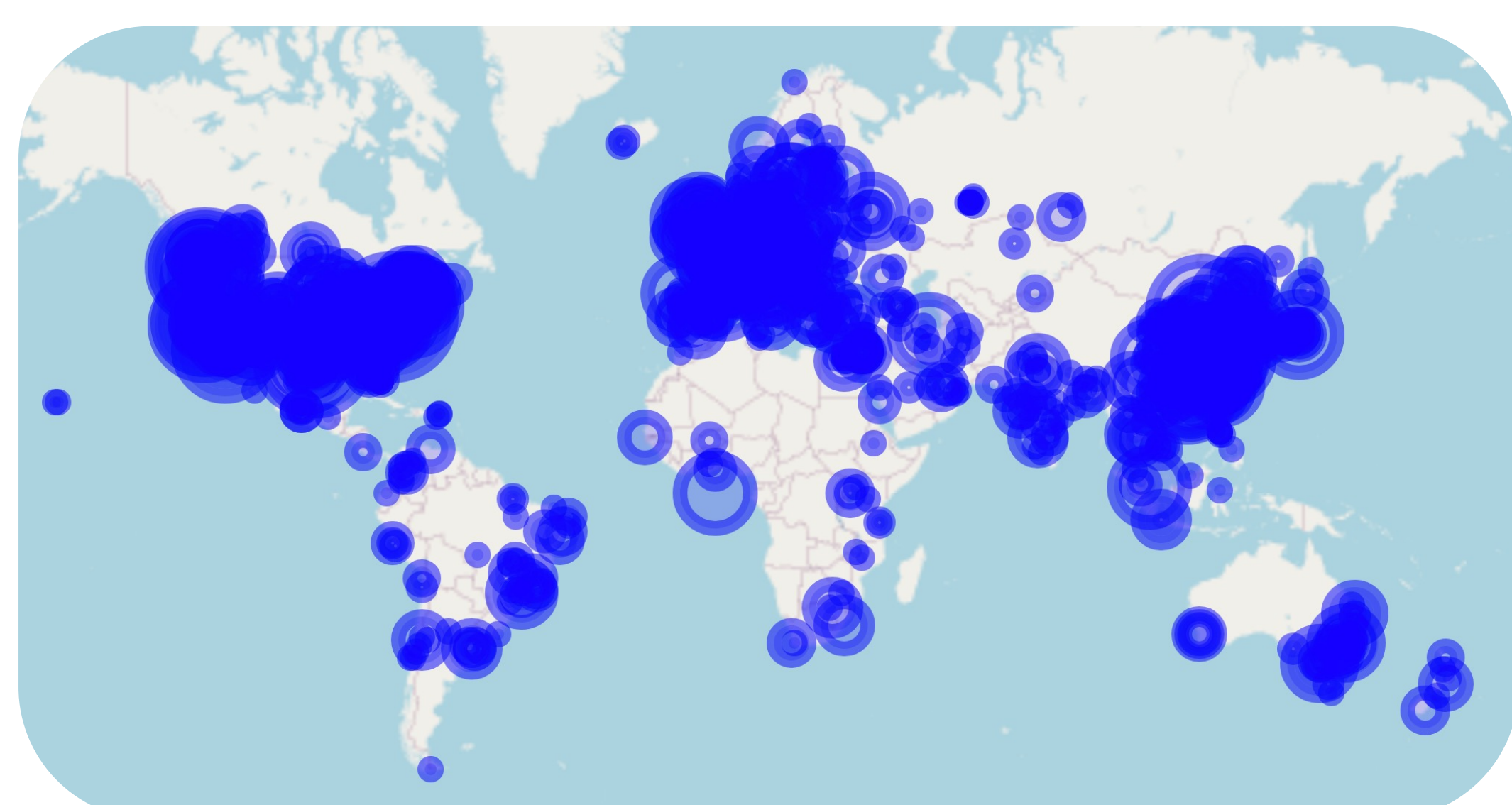
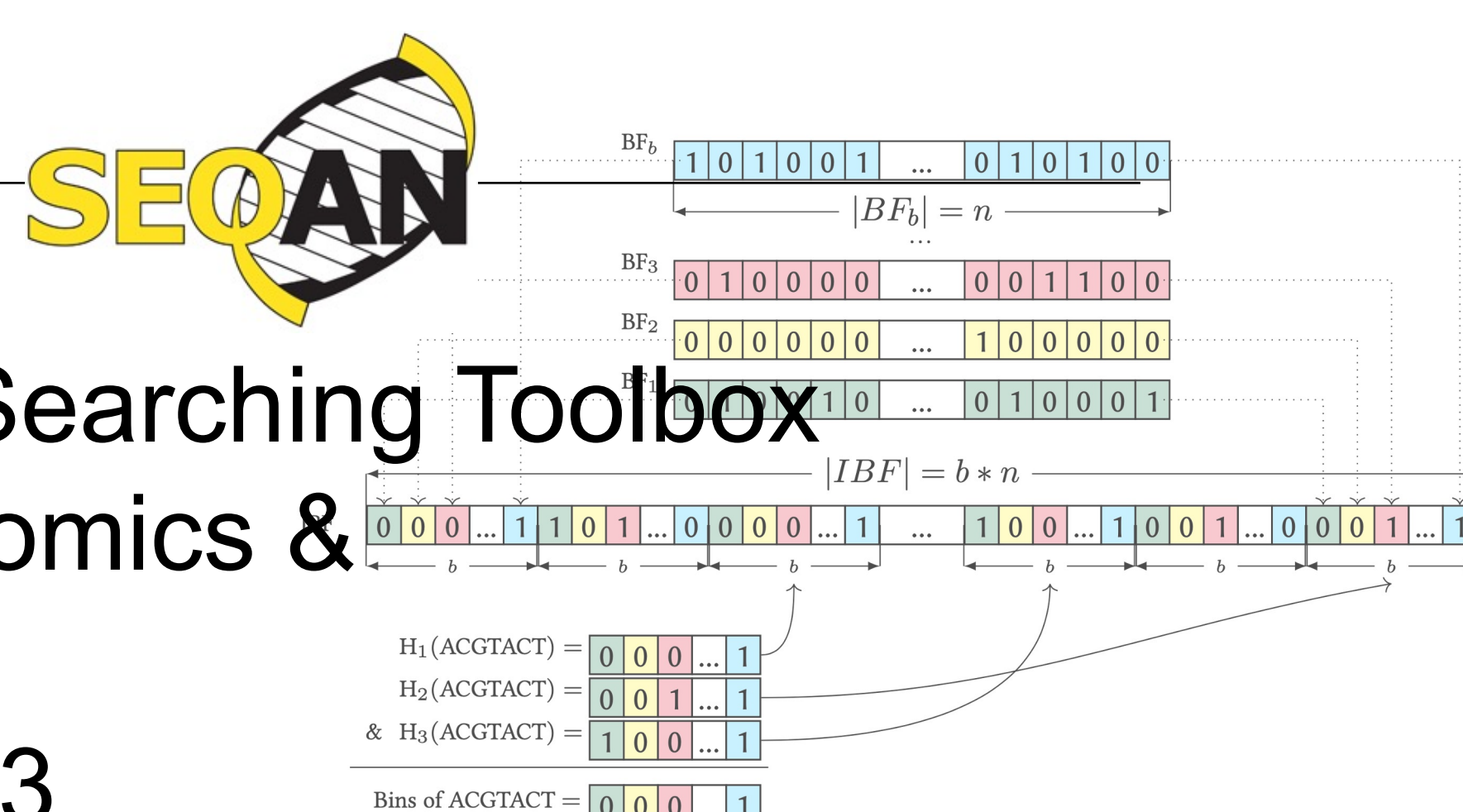
René Rahn, Knut Reinert
Algorithmic Bioinformatics, Free University Berlin

Contract Nr. W-de.NBI-002

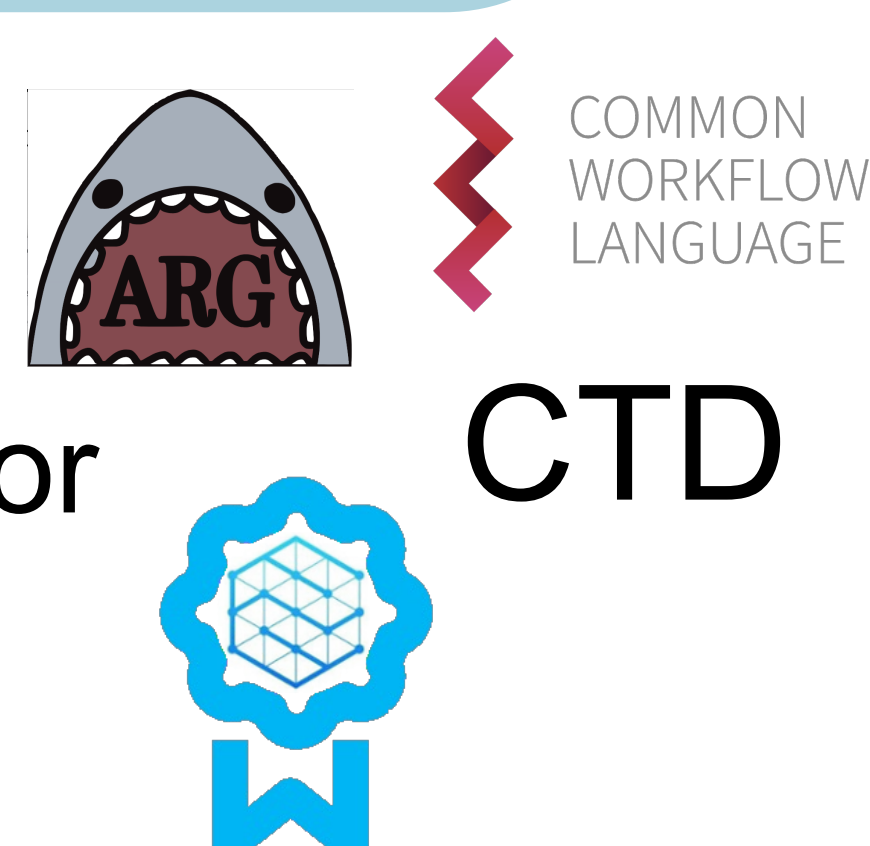
CIBI-SeqAn provides and maintains a C++ library for efficient biological sequence analysis (SeqAn). SeqAn is a globally used C++ library that enables the development of fast and targeted analysis applications for genomic and proteomic data. In addition, CIBI-SeqAn maintains and develops a range of ready-to-use analysis tools using SeqAn at their cores and made available to the bioinformatics community in Germany and Europe (Elixir).

SeqAn services

- SeqAn C++ library
- Archive Indexing & Searching Toolbox for Genomics, Proteomics & Metagenomics
- 40K users since 2023



- Workflow Integration via Sharg Parser
- Reproducibility badging for data analysis workflows



General information on the project

- de.NBI Staff = 1 FTE

Former de.NBI staff

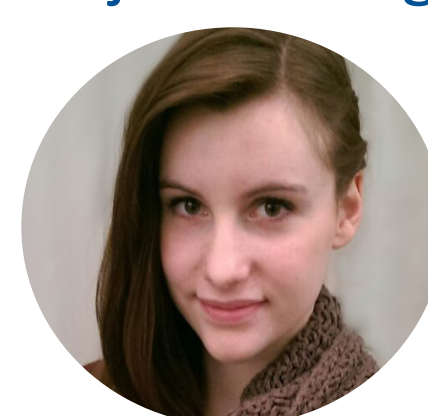
René Rahn

Knut Reinert

Marcel Ehrhardt

Simon Gottlieb

Svenja Mehringer



www.denbi.de

Progress report

SeqAn2

- Release v2.5.0
- Prioritizing compatibility with SeqAn3 by adapting namespaces and interfaces.
- Ongoing support via GitHub forums

SeqAn3

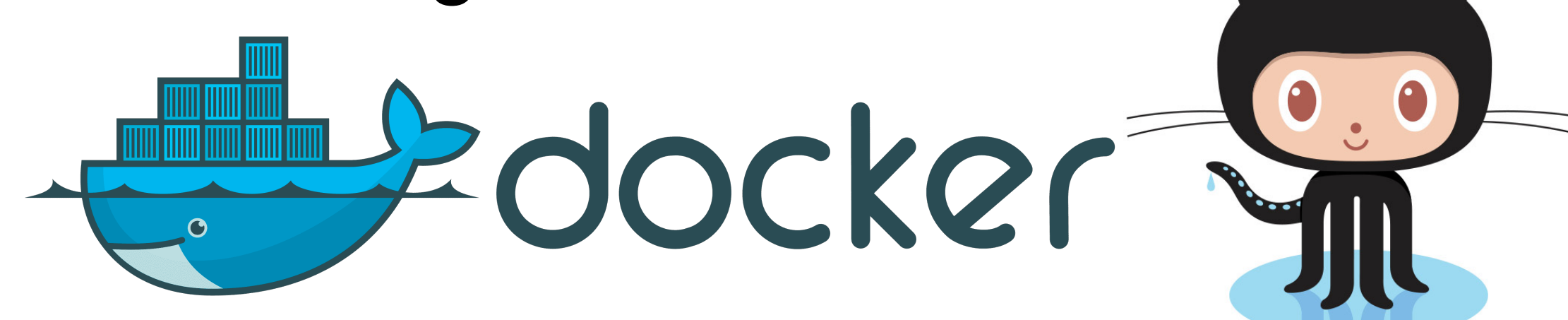
- Significant updates including release of v3.4.0

Supported platforms



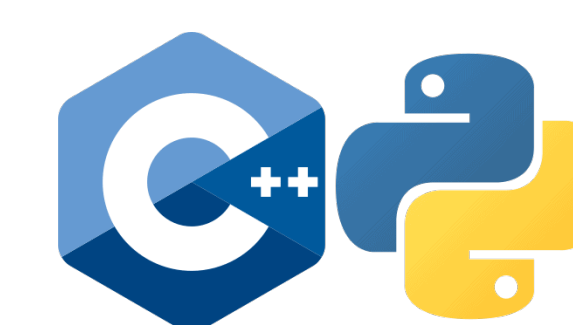
CI Infrastructure

- Automated linting and licensing tests
- Internal generation of doc previews
- Static code analysis
- Include what you use (IWYU) tests
- Docker images for CI builds



Language bindings

- Exploring possible strategies to add python bindings



Elixir officers

- contributing to the scientific activities of ELIXIR's Platforms, Communities and Focus Groups

de.NBI Training and education

- Expanded & improved online tutorials

Publications

- [Lambda3: homology search for protein, nucleotide, and bisulfite-converted sequences](#)
H Hauswedell, S Hetzel, SG Gottlieb, H Kretzmer, A Meissner, K Reinert; Bioinformatics 40 (3), btae097
- [Hierarchical Interleaved Bloom Filter: enabling ultrafast, approximate sequence queries](#)
S Mehringer, E Seiler, F Droop, M Darvish, R Rahn, M Vingron, K Reinert; Genome Biology 24 (1), 131
- [Linear: a framework to enable existing software to resolve structural variants in long reads with flexible and efficient alignment-free statistical models](#)
C Pan, R Rahn, D Heller, K Reinert; Briefings in Bioinformatics 24 (2), bbad071
- [ganon2: up-to-date and scalable metagenomics analysis](#)
VC Piro, K Reinert; bioRxiv, 2023.12. 07.570547