# Short description of the project

**C++ library for biological sequence analysis**
- Open Source "BSD 3-Clause License"
- Support for Linux, Windows and OS X
- Support for GCC, CLANG, Visual Studio, ICPC (Intel)

**Key library functionalities**
- Full-text indexing
- Reference based alignment
- Formatted Input / Output
- Parallelized data structures and algorithms

**Tools for high-throughput sequencing**
- Flexbar – Read quality control
- Mason – Read simulation
- Yara – Read mapping
- Lambda – Protein BLAST
- And many more...

## de.NBI services

- Maintenance of build and test infrastructure for developers
- Integration of external code, data repositories or software contributions into SeqAn and SeqAn3 (325 issues solved, 676 pull requests integrated in 2019)
- Tool and algorithm development (358 citations of SeqAn publications in 2019)
- Consulting, dev. training and software support on daily basis (e.g. Bioinformatics Solution Center)
- Improvements to software documentation and teaching material
- Deployment of our tools (8,000+ unique users worldwide)
  - Source and binary releases (3663 downloads on github)
  - KNIME plugins to construct workflows for sequence analysis data processing and downstream processing
- Available on Ubuntu, Debian, Brew, Bioconda (11.4k downloads) and Biocontainer
- Ready for cloud computing (deNBI Cloud)

## General information on the project

- 1.5 full position (FTE)
- 1.0 student assistant (FTE)
- All positions are currently filled

# Progress report

## Infrastructure development

- Released SeqAn 3.0.0 and 3.0.1
- Started Github Actions for CI and CD
- More Architecture Support (e.g. Arm64, PPC, x64)
- Setup an Application Infrastructure
- Packaged on Major Distributions (Debian, BioConda, Brew, etc.)

## External Contribution

- Ranges Library: Bug Reports + Upstream Fixes
- SDSL Library: Became a Maintainer
- Reported Compiler Defects (e.g. Internal Compiler Errors, clang, gcc)
- Participated in C++ Standard Committee Meeting + Paper
- Collaboration with HD-HUB for Python Bindings

## Training and education

- Added new training materials
- Developer, user meetings and conferences
- Online lectures and tutorials
- Participated in Train a Trainer Workshop

## Publications

- DREAM-Yara: An exact read mapper for very large databases with short update time T. H. Dadé, E. Siragusa, V. Piro, A. Andrusch, E. Seller, B. Renard, K.K. Rahn - Bioinformatics
- Generic accelerated sequence alignment in SeqAn using vectorization and multi-threading R. Rahn, S. Budach, P. Costanza, M. Ehrhardt, J. Hanous, and K. Reiners – Bioinformatics

---

# de.NBI Training and education

**Training events at user/developer meetings and Conferences**

<table>
<thead>
<tr>
<th>Date</th>
<th>Event Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 Mar 2019 (Berlin):</td>
<td>SeqAn &amp; OpenMS (CIBI) Integration Workshop at the KNIME Spring Summit</td>
</tr>
<tr>
<td>22 Mar 2019 (Escola):</td>
<td>SeqAn3 Developer Retreat 2019</td>
</tr>
<tr>
<td>24 Feb 2020 (Malta):</td>
<td>Biodesign (Accepted)</td>
</tr>
<tr>
<td>25 Mar 2020 (Italy):</td>
<td>SeqAn3 Developer Meeting</td>
</tr>
<tr>
<td>26 Jul 2020 (Montreal):</td>
<td>ISMB (Accepted)</td>
</tr>
</tbody>
</table>

**Pathway:**

- A biotechnology walk through de.NBI tools...
- Algorithms and development of SeqAn and SeqAn3...
- The architecture and development of cloud computing platforms...

---

**Source and binary releases (3663 downloads on github)**

- KNIME plugins to construct workflows for sequence analysis data processing and downstream processing
- Available on Ubuntu, Debian, Brew, Bioconda (11.4k downloads) and Biocontainer
- Ready for cloud computing (deNBI Cloud)

**Deployment of our tools (8,000+ unique users worldwide)**

- Source and binary releases (3663 downloads on github)
- KNIME plugins to construct workflows for sequence analysis data processing and downstream processing
- Available on Ubuntu, Debian, Brew, Bioconda (11.4k downloads) and Biocontainer
- Ready for cloud computing (deNBI Cloud)

**Ready for cloud computing (deNBI Cloud)**

- Source and binary releases (3663 downloads on github)
- KNIME plugins to construct workflows for sequence analysis data processing and downstream processing
- Available on Ubuntu, Debian, Brew, Bioconda (11.4k downloads) and Biocontainer
- Ready for cloud computing (deNBI Cloud)