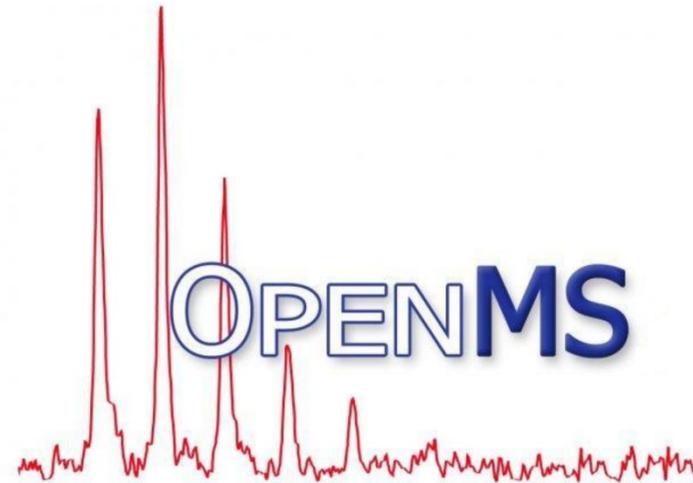


Projektmanagment im Softwarebereich



Faster Bioinformatics with C++

Chris Bielow

Proteomics

Genomics

Transcriptomics

“OMICS”

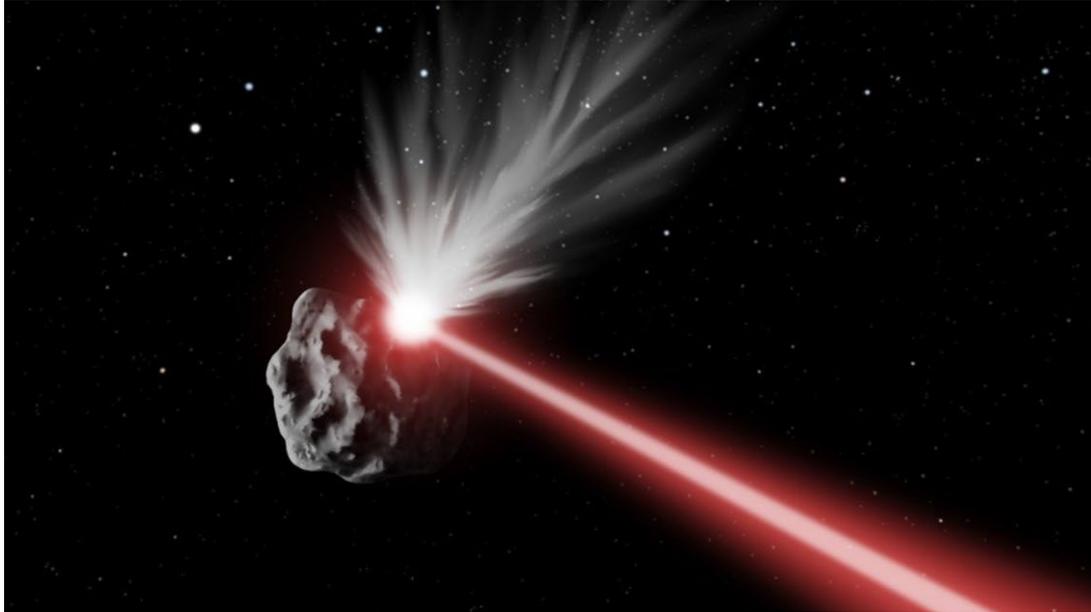
Connectomics

Foodomics

Metabolomics

OMICS – Ganzheitliche Charakterisierung und Quantifizierung der Moleküle einer Zelle/Probe/Organismus

Proteomics



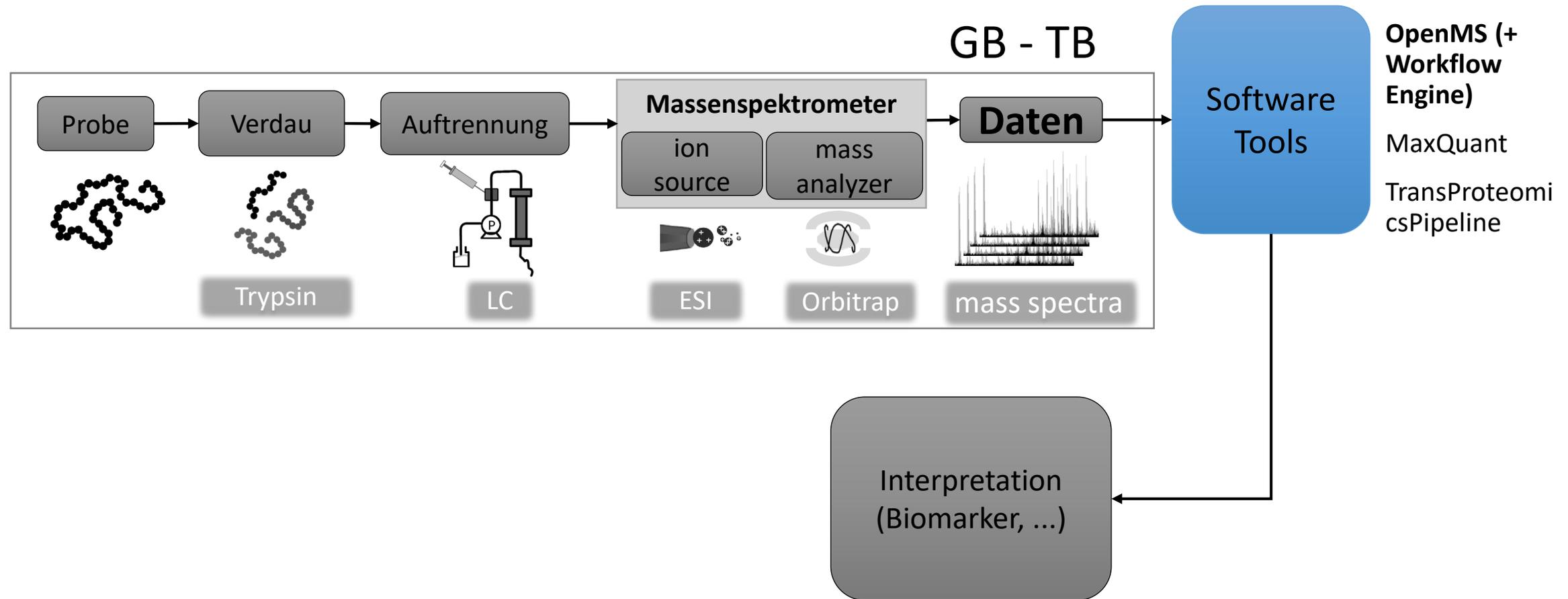
“Shooting at stuff
with lasers in space”!

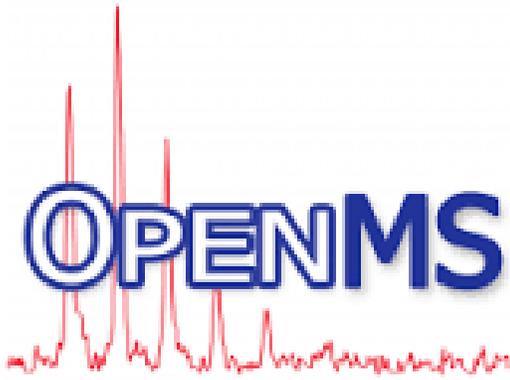
Method of choice:

Liquid Chromatography – Mass Spectrometry (LC-MS)

Proteomics-Pipeline

Liquid Chromatography (LC) + Massenspektrometrie (MS)

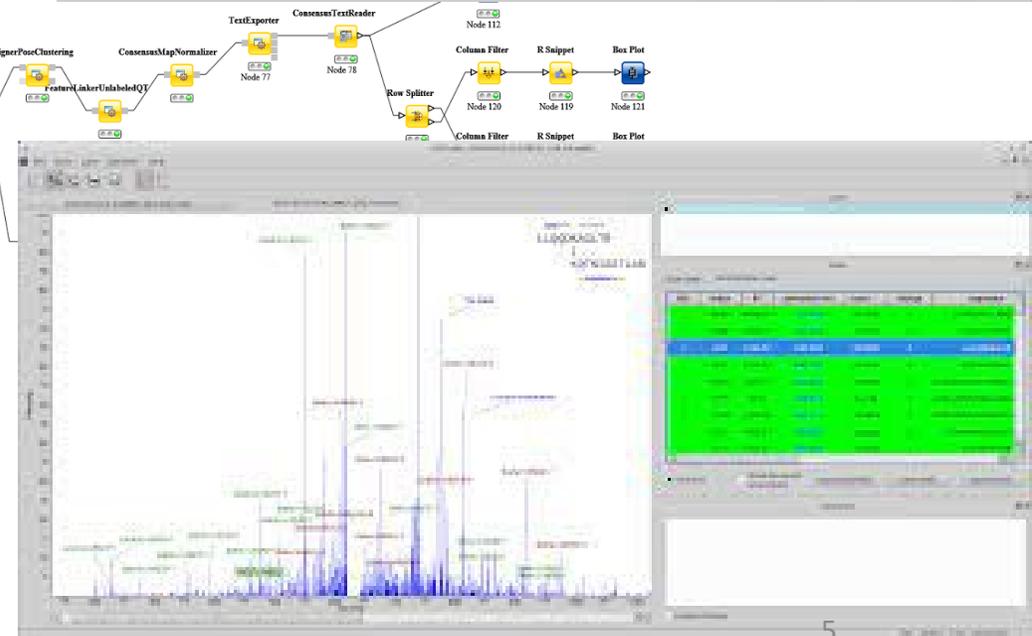
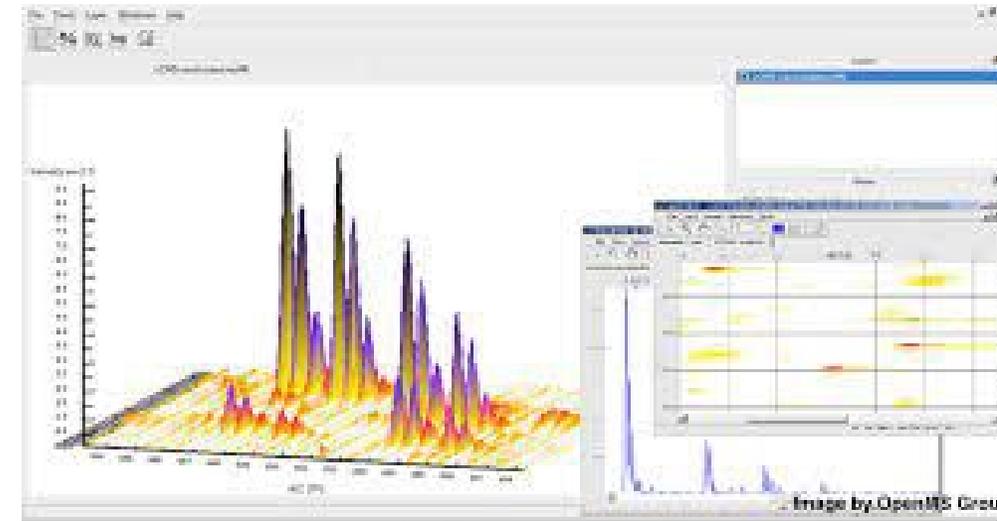
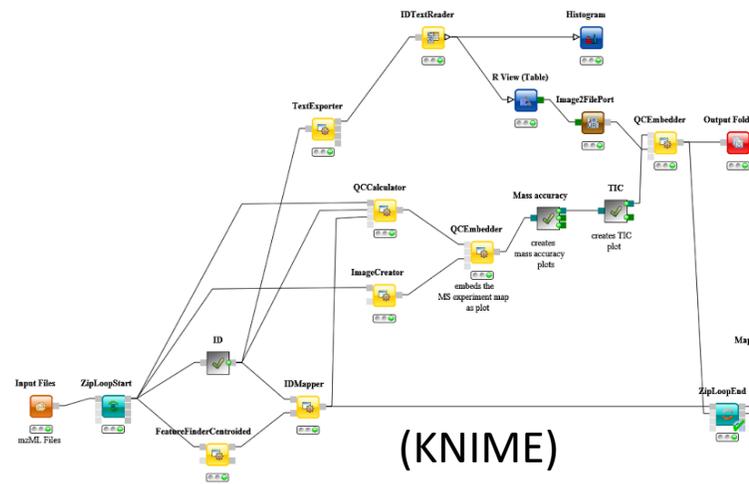




- C++ Bibliothek für Analyse von LC-MS Daten in der Proteomik

- 804.000 Zeilen Code

- Tools/Executables
 - 150 Tools
 - „Lego“



Inhalt des Praktikums

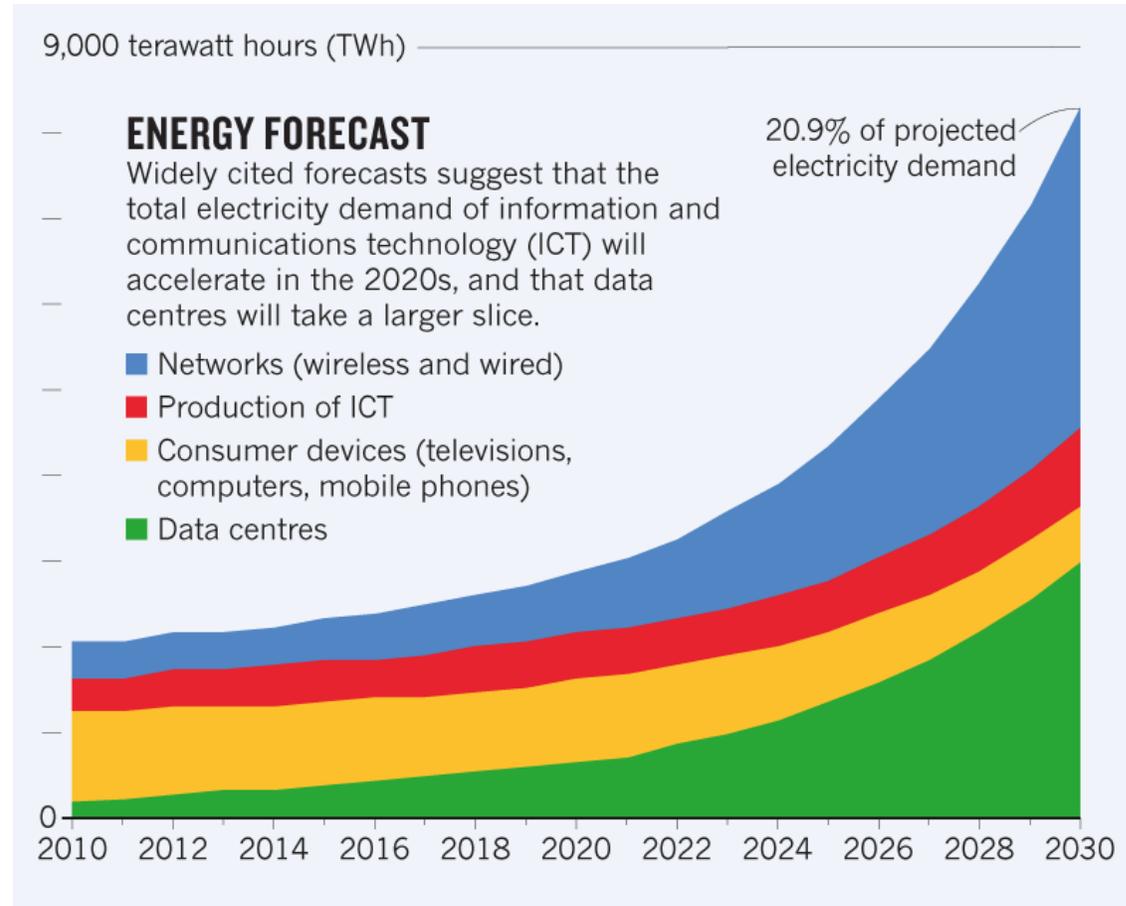
- Arbeit in bestehender open-source Bibliothek (BSD-Lizenz)
 - Erstellung von Pull-Requests ( **GitHub**) und Code Reviews
 - Aufnahme der neuen Funktionalität in die OpenMS-Entwicklerversion
- “Echte Daten, echte Probleme”
- Algorithmen in der Proteomik
 - Zum Beispiel:
 - Techniken zur Verbesserung der Laufzeit
 - Do/Don't Idioms in C++ (Wahl von Algorithmen und Datentypen, Datenstruktur, Allokatoren, Dispatching)
 - Optimierung der Programmlaufzeit
 - C++ Code Profiling: Intel V-Tune, perf, ...
 - CPU architecture: cache und SIMD/AVX
 - Benchmark-Umgebungen (Google Benchmark)
 - Tools zur Qualitätskontrolle von MS Daten (Schnittstelle mit R)
 - Visualisierung von MS Daten (Qt5) oder QC Daten (ggplot)
 - Anwendung auf Tools in der Proteomik-Pipeline von OpenMS

Warum Performance?

- Gut für den Akku
 - Laptops, Smartphones, ...

Warum Performance?

- Gut für den Akku
 - Laptops, Smartphones, ...
- Gut für unser Klima
 - ...



How to stop data centres from gobbling up the world's electricity, Nature **561**, 163-166 (2018)
doi: 10.1038/d41586-018-06610-y

Warum Performance?

- Gut für den Akku
 - Laptops, Smartphones, ...
- Gut für unser Klima
 - ...
- Gut für's Image
 - Quantum Supremacy
- Gut für den Benutzer
 - High-Performance-Computing

nature > articles > article

nature

Google

IBM

IBM Research Blog

*We argue that an ideal simulation of the same task can be performed on a classical system in **2.5 days** and with far greater fidelity.*

cy using a
perconducting

artinis 

article

660k Accesses | 26 Citations | 6028 Altmetric | Metrics

“a state-of-the-art supercomputer would require approximately **10,000 years** to perform the equivalent task.”

Zeitplan - Seminar

Ende Februar

- Vorbesprechung; genaue Terminvereinbarung per E-Mail zur Vermeidung von Terminkollisionen

(C++ Blockkurs 27. Februar – 10. März)

Projekteinführungswoche

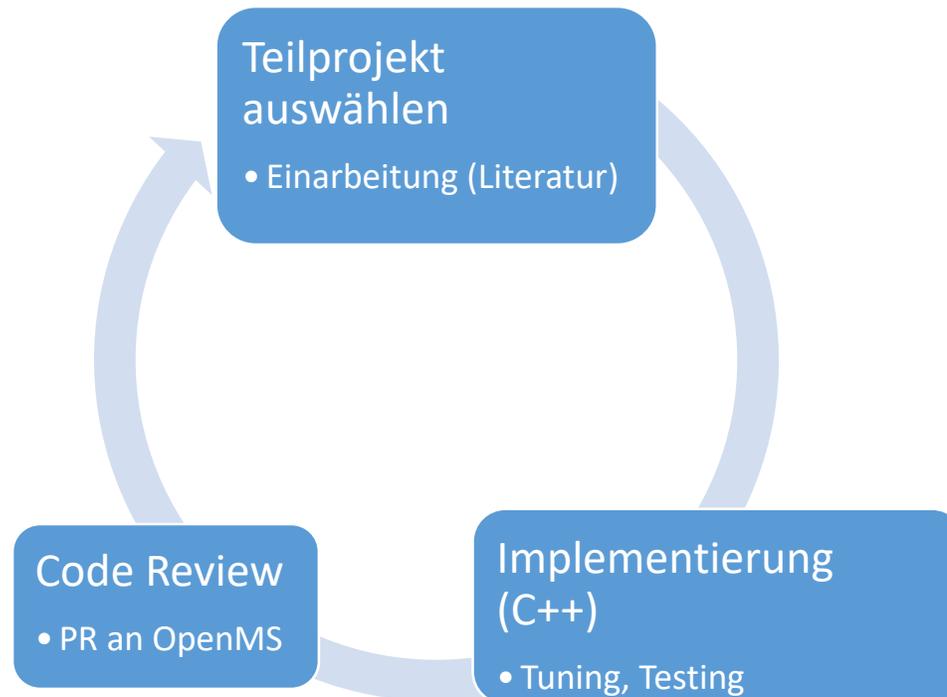
Mo, 20.03. - Fr, 24.03.

- Vormittags: 10-12 “Vorlesung”
- Nachmittags: kleine Übungen

- OpenMS installation
- Einführung in Flüssigkeitschromatographie & Massenspektrometrie (LC-MS)
- Refresher: C++ && Build-Systeme (Cmake, make, ninja)
- Tutorial OpenMS (C++ library und Tools)
- Workflow engines: TOPPAS, KNIME, ...
- Git für Fortgeschrittene
- Einführung in Themen: z.B. ‘perf’ & Co; flame graphs, ...

Zeitplan – praktischer Teil

Mo, 27.03. - Fr, 05.05.



Besprechung

Freitags: 10-12 Uhr,
Weiteres Vorgehen, Anpassung
des Zeitplans, Problemlösung

und immer wenn Bedarf
besteht

Fr, 12.05. Vorstellung der Ergebnisse (Zusammenfassung, Plots, Tabellen, ...)

Zusammenfassung

Vorraussetzung: Interesse an objektorientierter Programmierung (Java, C++, ...)

- Spass an C++
- ALDaBi Praktikum oder C++-Kurs (Nachmeldung mgl., via Prüfungsbüro Formular)

Praktische Erfahrungen zum Mitnehmen

- C++ Skills
- Kompetenz in Proteomik/Massenspektrometrie
- Softwareentwicklung in großem Projekt
 - Continuous Integration
 - Pull Requests mit Git
- Arbeit im Team

Schwierigkeitsgrad (Acht Sterne verteilt auf drei Bereiche):

A Programmieren: *****

B Biologie/Chemie: *

C Projektmanagement: **

Fragen?

- mehr Fragen
 - chris.bielow@fu-berlin.de

Ausbau zu B.Sc. Arbeit
prinzipiell möglich

Mehr Informationen

- OpenMS
 - www.OpenMS.de
- PMSB OpenMS
 - https://www.bsc.fu-berlin.de/TeachingAndWorkshops/SoSe23/Software_Praktikum_OpenMS/index.html

BTW:

Google Summer
of Code

