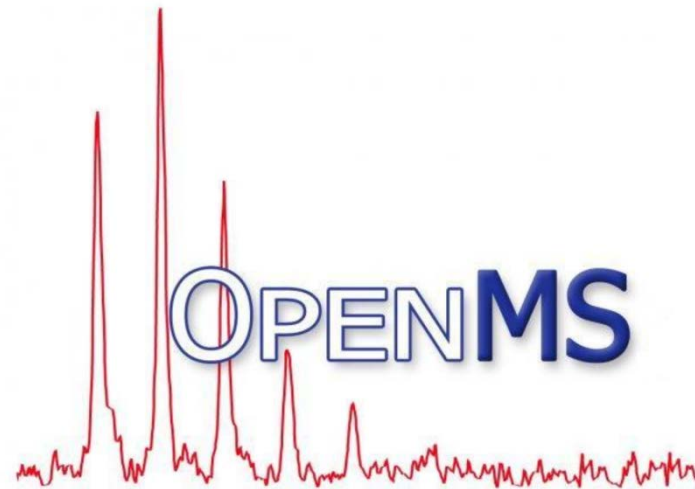


# Projektmanagment im Softwarebereich



Algorithms and Datastructures  
in  
High-Throughput Proteomics

**Chris Bielow**

Proteomics

Genomics

Transcriptomics

“OMICS”

Connectomics

Foodomics

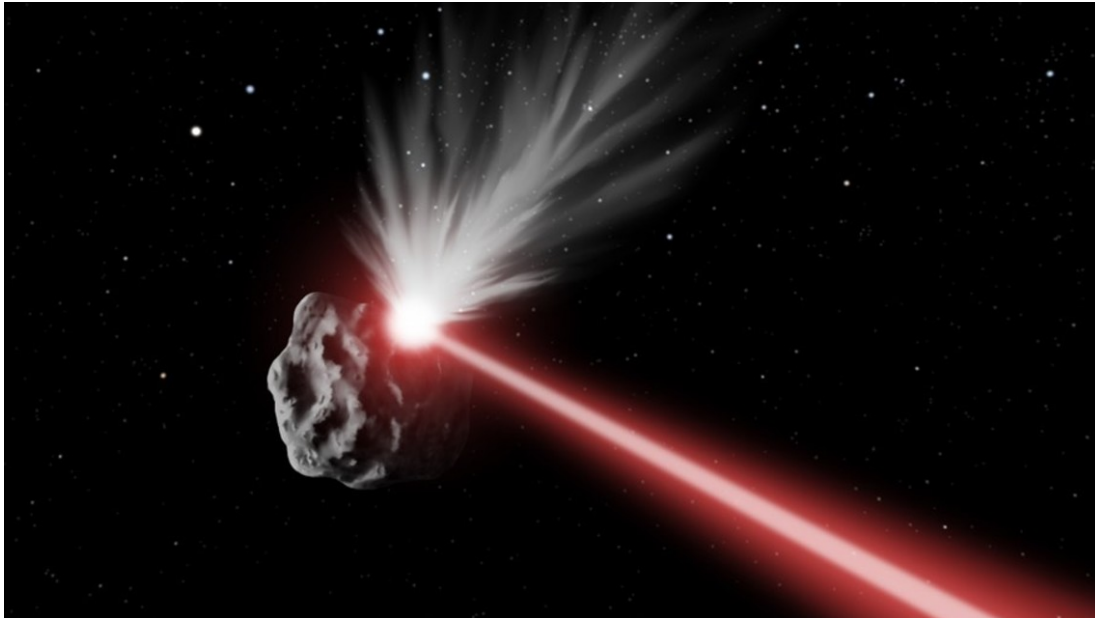
Metabolomics

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

# Proteomics

Method of choice:

Liquid Chromatography – Mass Spectrometry (LC-MS)

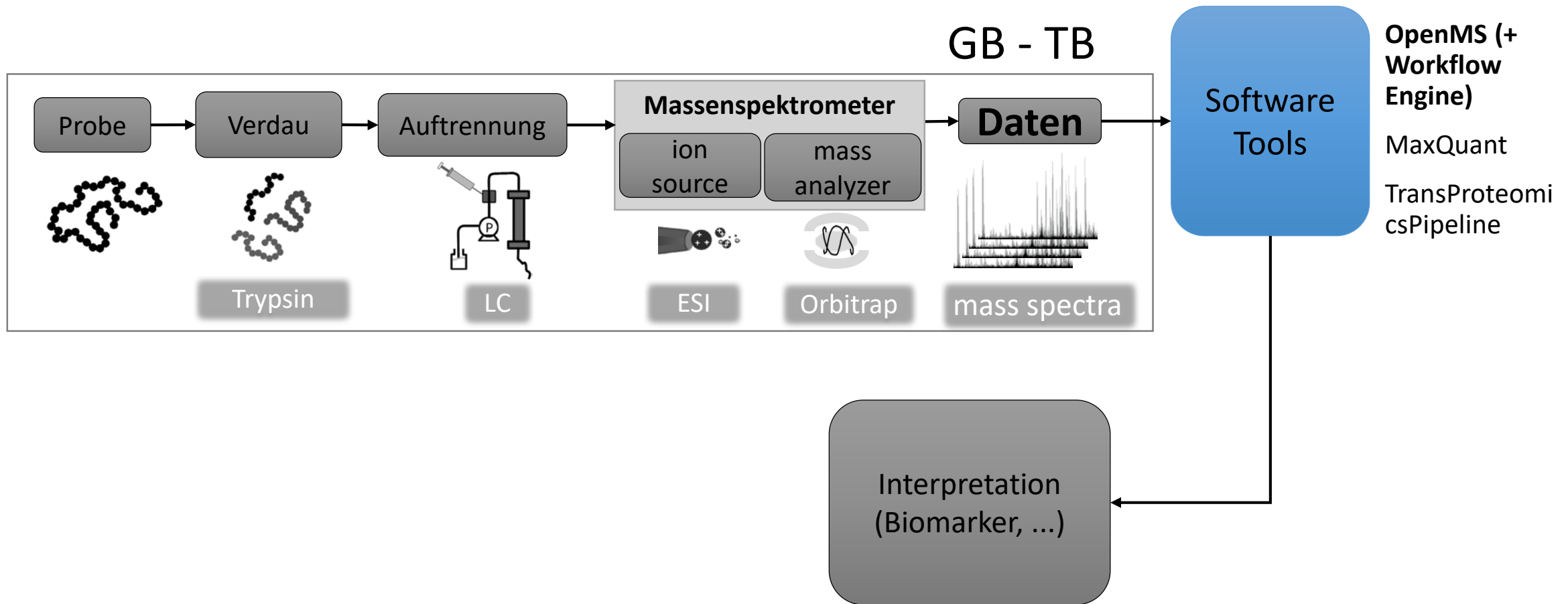


“Shooting at stuff  
with lasers in space”!

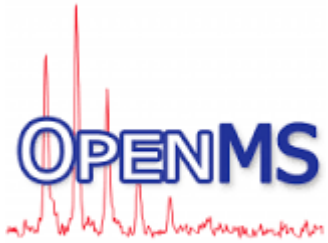
[https://www.syfy.com/sites/syfy/files/styles/1140x640/public/2017/06/170428-debris-laser-mn-1605\\_c382498d2eef0ff260fc6c8f70ca4806.focal-920x600.jpg](https://www.syfy.com/sites/syfy/files/styles/1140x640/public/2017/06/170428-debris-laser-mn-1605_c382498d2eef0ff260fc6c8f70ca4806.focal-920x600.jpg)

# Proteomics-Pipeline

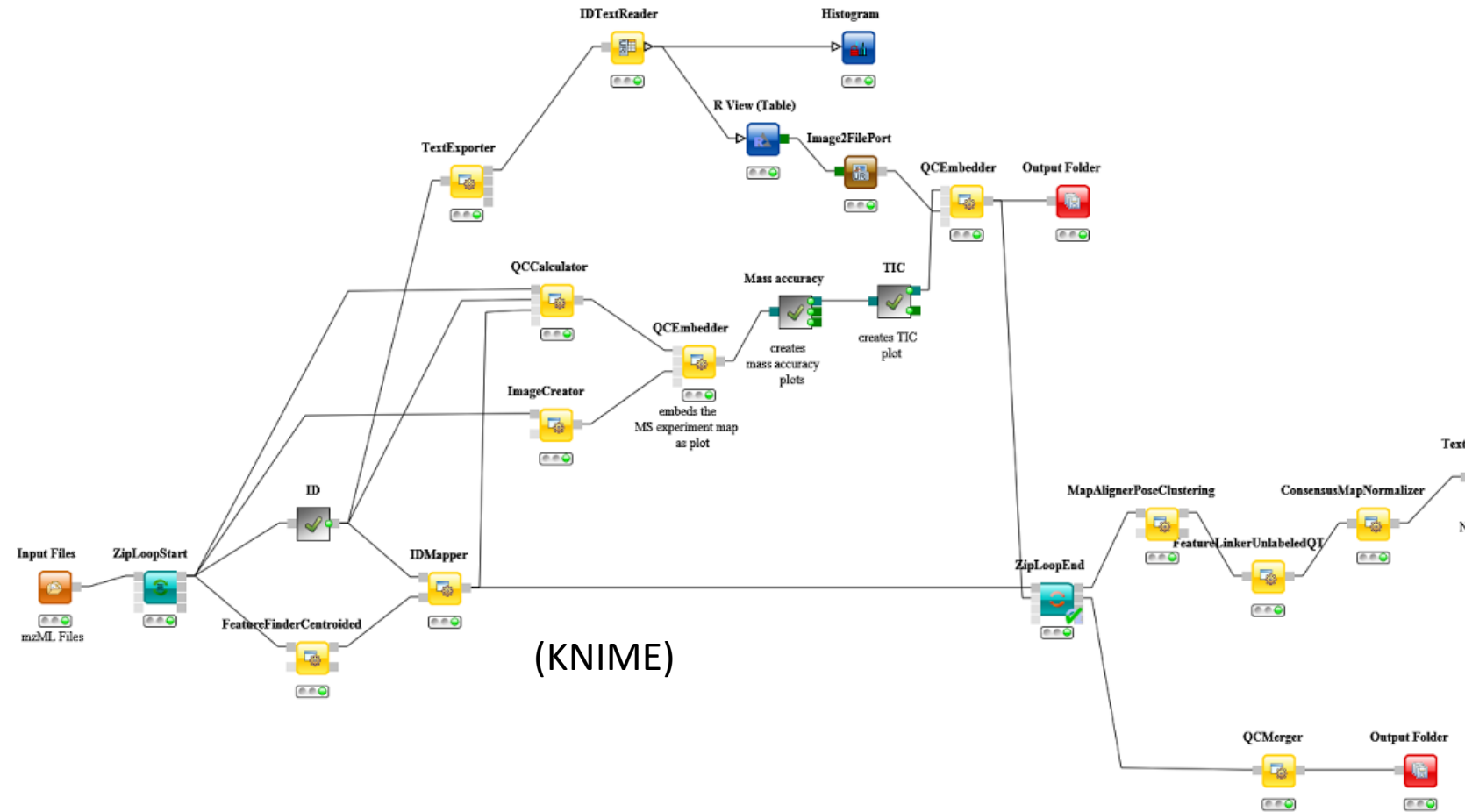
Liquid Chromatography (LC) + Massenspektrometrie (MS)




# Software Pipeline



- C++ Bibliothek für Analyse von LC-MS Daten in der Proteomik
- 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 instructions
      - Benchmark-Umgebungen (Google Benchmark)
    - Tools zur Qualitätskontrolle von MS Daten (Schnittstelle mit R)
  - Anwendung auf Tools in der Proteomik-Pipeline von OpenMS (im Team)

# Zeitplan - Seminar

## Ende Februar ...

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

## Mo, 22.03. - Fr, 26.03. – Projekteinführungswoche

- Projektmanagement, Zeit & Ressourcen-Planung
- Einführung in Flüssigkeitschromatographie & Massenspektrometrie (LC-MS)
- Refresher: C++ && Arbeit in einer großen Softwarebibliothek
- Tutorial OpenMS (C++ library und Tools)
- Workflow engines: TOPPAS, KNIME, ...
- Git (version control software)
- Einführung in Themen: z.B. 'perf' & Co; flame graphs, ...

# Zeitplan – praktischer Teil

Fr, 22.03.

Vorstellung der Pläne

Mo, 06.04. - Fr, 14.05.

## Woche 1(+2):

- OpenMS-develop installieren
- Projekt auf <https://git.imp.fu-berlin.de> clonen
- Einarbeitung in OpenMS
- Relevante Literatur lesen
- (Externe Programme testen)
- Gemeinsam Problemstellungen formulieren, Zeitplan erstellen, Vortrag

## Woche 3+4

- Implementierung (C++), Tuning, Testing
- Continuous Integration...

## Woche 5

- Code finalisieren (C++); evtl. Bugs fixen; finaler PR an OpenMS
- Abschlussbericht schreiben (Erreichte Verbesserungen, Plots)

Fr, 21.05. Abschlussbericht und Vorstellung der Ergebnisse

## wöchentliche Besprechung

Freitags: 10-12 Uhr,

Weiteres Vorgehen, Anpassung des Zeitplans, Problemlösung

Weitere Treffen nach Bedarf



# Zusammenfassung

**Vorraussetzung:** Erfahrung in objektorientierter Programmierung (Java, C++, ...)

- C++ Kenntnisse (wenig Templates)
- z.B. AIDaBi Praktikum oder C++-Kurs (Nachmeldung mgl., via Prüfungsbüro Formular)

## Empfehlenswert

- Eigener Laptop
  - Zum Entwickeln (50 GB frei, 4C)
  - Als thin-client
    - Zugang zu Compute-Servern 28C wird gewährt)

## Praktische Erfahrungen zum Mitnehmen

- Kompetenz in Proteomik/Massenspektrometrie
- Softwareentwicklung in großem Projekt
  - Continuous Integration
  - Pull Requests mit Git
  - Erweiterung der C++ Kenntnisse
- 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](mailto:chris.bielow@fu-berlin.de)

Ausbau zu B.Sc. Arbeit  
prinzipiell möglich

## Mehr Informationen

- OpenMS
  - [www.OpenMS.de](http://www.OpenMS.de)
- PMSB OpenMS
  - [https://www.bsc.fu-berlin.de/TeachingAndWorkshops/SoSe21/Software Praktikum OpenMS/index.html](https://www.bsc.fu-berlin.de/TeachingAndWorkshops/SoSe21/Software_Praktikum_OpenMS/index.html)

BTW:

Google Summer  
of Code

