

— Incorporating Natural Language Processing (NLP) Workflows into an open-source genomics cancer portal



Quest Lover / Teams welcome

Support clinicians in decision-making processes through easy access of mutation dependent targeted therapy options by using NLP.

Apply until **Nov 30, 2020** / Pitch on **Dec 8, 2020**

Question to be solved

Are you curious to deep-dive into one of the most commonly used genomics cancer portals in the scientific community and investigate how to incorporate self-created NLP workflows to get further evidence for molecular-guided therapy options?

General Background

The cBioPortal is one of the most commonly used cancer genomics platforms. Some hospitals adapt the platform and customize it to use it in their clinical decision making process. In this challenge we want to include results from natural language processing that are able to provide further evidence for molecular guided therapy options into the platform. Part of the task will be to understand the architectural setup in order to derive options on how to best integrate NLP results. The challenge will be conducted in a real clinical setting of a German hospital.

Data Types & Technologies

- Genomics and mutation data
- Platform and software architecture
- NLP

Supporting Material or Links

- Cerami et al., The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data, AACR 2012
- Hinderer et al., Implementing Pharmacogenomic Clinical Decision Support into German Hospitals, Stud Health Technol Inform. 2018
- Hoefflin et al., Personalized Clinical Decision Making Through Implementation of a Molecular Tumor Board: A German Single-Center Experience, JCO Precision Oncology 2018 :2, 1-16

Needed Skills

- Self-starter and organisational talent
 - Programming skills (e.g. R and/or Python)
 - Software/platform architecture know how is a plus
 - NLP technologies is a plus, not a must
 - Interest in biological and clinical data
-

Mentor

Dr. Lars Hummerich

Head of Oncology Innovations, Roche Diagnostics Deutschland GmbH

Dr. Markus Bundschus

Head RDG Data Office

Form of Cooperation

Preferred scale: 3-6 months full-time (flexible models are also possible), several persons could work simultaneously on the topic Possible formats: from working student to internship to bachelor or master thesis...and even bigger, if your idea requires a larger engagement.

How to present your Idea

- Show us how you would approach the problem in 3 to 5 slides. We do not expect a bullet-proof solution to the problem.
 - Be aware you only have 10 minutes to sell your idea.
 - Ideally, your presentation features different approaches how to include NLP workflows into the cBio portal.
 - Please be prepared that we will ask questions that will check your coding skills.
-