

**Bsc project****DomainWorld - Algorithm development***Developing algorithms to analyze the evolution of domains*

Background: Protein domains, conserved sequence segments, allow to reuse existing functionality in different contexts. Over time the domain composition of a protein can change. We are developing algorithms and programs in our ongoing DomainWorld project (<http://domainworld.uni-muenster.de/>). These algorithms have the goal to support researchers in their efforts to analyze domains, domain arrangements and the evolution of them.

We usually have some projects in the context of DomainWorld available. These are usually with focus on the development and implementation of algorithms, for example developing and analyzing the clustering of domain arrangements. For currently available projects please contact Carsten Kemena.

Objectives: The goal of this project is to develop and benchmark new algorithms to analyze the evolution of domains.

- Implement an algorithm to help in the analysis of domains
- Benchmark the algorithm

Requirements:

- Interest in the general topic of domains and programming
- Knowledge of at least one programming language (Python or C++)

Methods:

- develop and implement new algorithms
- Compare the new algorithm to existing ones

Supervision: Carsten Kemena, Hüfferstr. 1, Raum 17, c.kemena@uni-muenster.de, Molecular Evolution and Bioinformatics Group (<http://bornberglab.org/>).

Selected Literature:

The Pfam protein families database: towards a more sustainable future: R.D. Finn, P. Coggill, R.Y. Eberhardt, S.R. Eddy, J. Mistry, A.L. Mitchell, S.C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G.A. Salazar, J. Tate, A. Bateman *Nucleic Acids Research* (2016) Database Issue 44:D279-D285