



# netidee

PROJEKTE

## OpenBioLink

Zwischenbericht | Call 15 | Projekt ID 5171

Lizenz: CC-BY

*(Nicht zutreffendes bitte löschen!)*

# Inhalt

|     |  |   |
|-----|--|---|
| 1   | Einleitung.....  | 3 |
| 2   | Status der Arbeitspakete.....  | 3 |
| 2.1 | Arbeitspaket 1 - <i>Detailplanung und Formales am Projektstart</i> .....                     | 3 |
| 2.2 | Arbeitspaket 2 - <i>Benchmark Erweiterung</i> .....  | 3 |
| 2.3 | Arbeitspaket 3 - <i>OpenBioLink Explorer tool</i> .....                                      | 3 |
| 2.4 | Arbeitspaket 4 - <i>Durchführung und Dokumentation der "OpenBioLink 2021 Challenge". ...</i> | 4 |
| 3   | Umsetzung Förderauflagen.....  | 4 |
| 4   | Zusammenfassung Planaktualisierung .....   | 4 |
| 5   | Öffentlichkeitsarbeit/ Vernetzung.....   | 4 |
| 6   | Eigene Projektwebsite.....   | 4 |

## 1 Einleitung

Unser Projekt hilft biomedizinischen ForscherInnen, mittels AI und Webtechnologie schneller Zusammenhänge im Netzwerk biomedizischer Forschungsergebnisse zu finden, und dadurch vielversprechende Hypothesen zu formulieren und zu testen (z.B. 'Medikament X hilft bei Krankheit Y'). Andererseits hilft unser Projekt AI EntwicklerInnen, ihre Modelle zu testen und zu verbessern.

Im Projekt entwickeln wir auch generell domänenübergreifend anwendbare Algorithmen und Tools für Explainable AI und Link Prediction in Wissensgraphen.

## 2 Status der Arbeitspakete

### 2.1 Arbeitspaket 1 - *Detailplanung und Formales am Projektstart*

Der Projektstart wurde erfolgreich abgewickelt (Vertragsprüfung, Unterzeichnung des Vertrages, Projektplanung basierend auf Excel-Vorlage, Start der Projektwebsite, Förderabruf für Förderrate 1).

### 2.2 Arbeitspaket 2 - *Benchmark Erweiterung*

Eine umfassende Analyse von weiteren Datenquellen wurde erstellt und Anknüpfungspunkte bzw. Erweiterungen des derzeitigen Wissensgraphen wurden identifiziert. Dieses Mapping hilft Anwendern, unseren Wissensgraphen mit anderen Wissensquellen zu verknüpfen und zu erweitern.

Der Code für die Durchführung des Benchmarks (generieren von Vorhersagen, generieren von Scores) wurde neu implementiert und erlaubt nun eine bessere Vergleichbarkeit verschiedener Methode. Dies ist von grundlegender Bedeutung auch für die faire Durchführung der „Challenge“ (Arbeitspaket 4).

Dokumentation wurde erweitert.

### 2.3 Arbeitspaket 3 - *OpenBioLink Explorer tool*

Das OpenBioLink Explorer Interface Softwarepaket und User Interface wurde entwickelt, ist bereits benutzbar und beinahe fertig gestellt. Dieses Interface erlaubt die Explorierung von Zusammenhängen in großen Wissensbasen, sowie eine Darstellung der Vorhersagen, die durch unseren Explainable Link Prediction Algorithmus generiert werden.

Der Explainable Link Prediction Algorithmus (SAFRAN) wurde im Projekt weiter entwickelt und liefert State-of-the-Art Ergebnisse. Eine wissenschaftliche Publikation zu dem Algorithmus wurde fertig gestellt und publiziert; eine weitere Publikation zum OpenBioLink Explorer Interface ist in Vorbereitung

#### 2.4 Arbeitspaket 4 - *Durchführung und Dokumentation der "OpenBioLink 2021 Challenge"*.

Das Regelwerk, Prozedere und Tools für die Durchführung der Challenge wurden fertig gestellt. Eine Website für die Challenge wurde erstellt (<https://openbiolink.github.io/>). Eine Bewerbung der Challenge in sozialen Netzwerken wurde kurz vor Komplettierung des Berichtes gestartet; weitere Bewerbung der Challenge wird folgen.

### 3 Umsetzung Förderauflagen

Aufgrund des Feedbacks in den Förderunterlagen wurde mehr Augenmerk auf die domänen-übergreifenden Entwicklungen im Projekt gelegt (d.h., auf jene Komponenten, die auch außerhalb der biomedizinischen Domäne relevant sind). Wir haben mit SAFRAN einen State-of-the-Art Algorithmus für Explainable Link Prediction geschaffen. Der ‚Explorer‘ wird nun ebenfalls für nicht-biomedizinische Use-Cases entwickelt und kann auch hier ein global wichtiges Tool darstellen.

### 4 Zusammenfassung Planaktualisierung

Aufgrund von COVID kam es zu einer Verzögerung der Stellenbesetzung, wesentliche Arbeitspakete wurden daher erst im Mai gestartet und verdichtet. Der spätere Start kann erfolgreich kompensiert werden (mehr Arbeitsstunden pro Monat, mehr Involvement von Senior Researchers/Developers).

### 5 Öffentlichkeitsarbeit/ Vernetzung

Eine Bewerbung der OpenBioLink Challenge wurde begonnen. Der im Projekt weiterentwickelte Algorithmus wird später in diesem Jahr bei der AKBC 2021 Konferenz präsentiert werden.

### 6 Eigene Projektwebsite

<https://github.com/OpenBioLink/OpenBioLink>  
<https://openbiolink.github.io/>