



高月照江<sup>1</sup>、柳田達矢<sup>2</sup>、土肥栄祐<sup>3</sup>、山本泰智<sup>1</sup>、柵屋啓志<sup>2</sup>、藤原豊史<sup>1</sup>

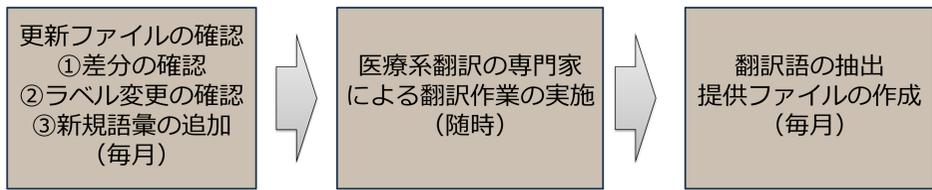
1：ライフサイエンス統合データベースセンター 2：理化学研究所バイオリソース研究センター 3：国立精神神経医療研究センター

## ● Abstract

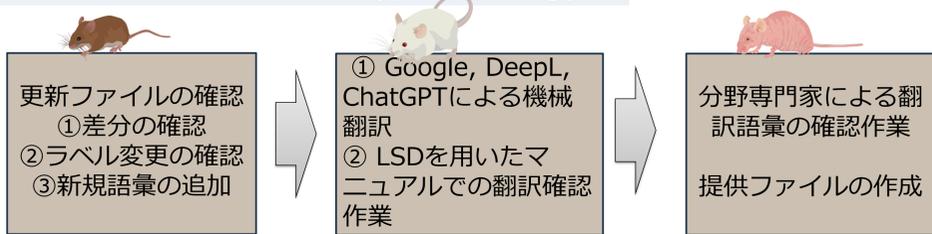
オントロジーは、特定の領域における概念や関係を明確に定義し、データ間の一貫性を保つための枠組みを提供するため、医療情報のデータ統合・共有のプロセスにおいて重要な要素となっている。疾患名オントロジーMondo (Mondo Disease Ontology) と哺乳類の表現型オントロジーMP (Mammalian Phenotype Ontology) は、基礎研究だけではなく臨床応用にも有用であり、世界的にその活用が進んでいる。しかし、これらのオントロジーは英語で作成されているため、英語が母国語でない国での利用拡大が課題とされている。そこでライフサイエンス統合データベースセンターでは、MondoおよびMPの日本語化と国際共有に取り組んでいる。Mondoに収録される約28,000件 (2024年7月時点) の英語ラベルと、MPに収録される約15,000件 (2024年7月時点) の英語ラベルに対して、ライフサイエンス辞書を用いて自動で日本語訳を割り当て、翻訳家および医療関係者により訳語の追加および修正を行った。その結果、Mondoの約23,000件の英語ラベルに、MPの全ての英語ラベル (OBSOLETEクラスを除く) に日本語訳を割り当てた。また、Babelon形式で共有された日本語訳はそれぞれのオントロジーに自動で取り込まれるため、今後の日本国内における利用拡大が期待される。

## ● 日本語訳作成の過程

### MONDO (Mondo Disease Ontology)



### MP (Mammalian Phenotype Ontology)



## ● 翻訳ファイルの提供

### MONDO (Mondo Disease Ontology)

<https://github.com/dbcls/mondo-japanese>



### MP (Mammalian Phenotype Ontology)

[https://github.com/dbcls/MP\\_Japanese](https://github.com/dbcls/MP_Japanese)



それぞれの翻訳語ファイルはDBCLSのGithubのページでDL可能  
適宜更新中

## ● 翻訳語彙数 (2024年9月時点)

オントロジー名	総クラス数	OBSOLETE数	翻訳数	翻訳数割合 (OBSOLETEを除く)
MONDO	28,017	3848	21487	88%
MP	15222	1136	14086	100%

MPについては日々更新が行われているため、追加語彙についての時差は発生するが、毎月差分の確認を行い、新規追加語彙の翻訳を定期的に行っている。

## ● 国際連携 (日本語訳の提供)

MONDO及びMPの日本語訳ファイルは、MONDOを作成しているMonarch Initiative, MPを作成しているMGI (Mouse Genome Informatics)にそれぞれ提供している。Monarc及びMGIはオリジナルのファイルに日本語訳を取り込み、多言語ラベルを含むオントロジーのファイルを作成し提供している。この取り組みはBabelonというオントロジーの多言語化共有フォーマットを用いて実施しており、Babelon形式のファイルを提供することで、自動的にデータを取り込むパイプラインが構築されている。

**Ontology Lookup Service** <https://www.ebi.ac.uk/ols4>



言語切り替えで日本語を選べると、提供しているMPの日本語が表示される



## ● 翻訳語利用例の紹介

### MONDO (Mondo Disease Ontology)



希少・遺伝性疾患の診療・研究に役立つWebサービスを提供する「PubCaseFinder」(<https://pubcasefinder.dbcls.jp/>)では、疾患名および症状名に関連するデータベースIDを検索する機能を追加し、CaseSharing"mini"として公開。CaseSharing"mini"ではMONDOの日本語訳をサービスに取り込み、検索の利便性が向上している。

### MP (Mammalian Phenotype Ontology)



理研バイオリソースセンターでは、既に公開されているMPの日本語訳をリソースの検索システムに取り込むことで、マウスリソースに対してアノテーションされているMPに対して、日本語で検索を行い、アノテーションしたマウスリソースが表示される仕組みを構築している。日本語で検索を行えることで、国内のマウス利用者にとってより検索のしやすい仕組みを提供している。

## ● Future plan

- \* LLMにLife Science Dictionaryを用いた、機械翻訳の精度の向上
- \* 翻訳業務の継続
- \* 利用者拡大へ向けた活動
- \* 国際連携の継続
- \* MPについては日本先天異常学会・用語委員会とのコラボ

**ライフサイエンス用語に関わる翻訳業務の協力者を募集しています！  
興味のある方は是非お声がけ下さい。**

