



# 創薬・疾患研究のためのデータベース検索システム Sagace

長尾知生子<sup>1</sup>、〇五十嵐芳暢<sup>2</sup>、森田瑞樹<sup>1,3</sup>、陳怡安<sup>1</sup>、深川明子<sup>4</sup>、坂手龍一<sup>5</sup>、水口賢司<sup>1</sup>

1. 医薬基盤・健康・栄養研究所 バイオインフォマティクスプロジェクト 2. 同研究所 トキシコゲノミクス・インフォマティクスプロジェクト 3. 岡山大学大学院医歯薬学総合研究科 クリニカルバイオバンクネットワーク事業化研究講座 4. 医薬基盤・健康・栄養研究所 政策・倫理研究室 5. 同研究所 難病資源研究室

## Sagace <http://sagace.nibiohn.go.jp>

- Sagace(さがす)は、2012年からJSTバイオサイエンスデータベースセンターと連携して開発している、創薬・疾患に関する国内のデータベースを対象とした横断検索システム。
- より効率的に創薬・疾患研究を支援する検索システムの構築を目指している。

### Sagaceの3つの特徴

#### 1. 医学/生命科学に特化したデータベースの選定と分類

- NBDCと共同で作られたデータベース一覧より、創薬・疾患関係のデータベース、約180個を選定・分類。
- 検索結果の表示にアクセス解析の結果を活用。

#### 2. ファセットによる検索結果の絞り込み

医薬品など

- すべて
- 化合物 (5334)
- 漢方薬 (1098)
- 薬効・効能 (11196)
- 副作用・有害情報 (6127)
- 臨床試験・治験 (11480)
- その他 (698093)

データベースの種類

- すべて
- 生物資源 (3175)
- 特許 (158507)
- 実験プロトコル (76)
- 文献・用語集・辞書 (346552)
- レギュレーション・ガイドライン (13460)
- 学会・研究会 (163251)
- その他のデータベース (220729)

生物種

- すべて
- ヒト (370644)
- 動物 (ヒト以外) (116761)
- 植物 (25703)
- 微生物 (18483)
- 特定生物なし (354510)

- データベースを医薬品・データベースの種類・生物種の各カテゴリーに分類。
- 検索結果画面左側のラジオボタンによって、検索結果を絞り込み。

#### 3. メタデータを利用した検索結果の効果的な表示

エンタリーID

データベース名(説明がpop up)

生物種(画像がpop up)

- データベース名、エンタリーID、生物種、疾患名、最終更新日等の情報をmicrodata、RDFaLite、JSON-LD形式でマークアップすると、それらの情報が検索結果画面に反映される。
- ユーザが目的のデータを迅速に探すことができる。

Sagaceが対応しているボキャブラリやマークアップ方法は、以下のリンク先をご参照ください。

- マークアップ方法 <http://sagace.nibiohn.go.jp/press/metadata/markup/>
- マークアップスキーマ <http://sagace.nibiohn.go.jp/schema/schema.html>

### Sagace 2016年の改良点

#### 1. 治験情報に関する検索の強化

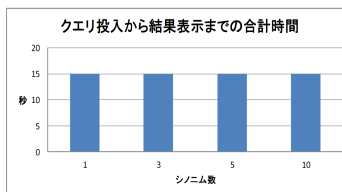
- 新規データベースの追加  
ClinicalTrials.govに加え、臨床研究情報ポータルサイト、New England Journal of Medicine日本語国内版、EU-CTRを新たに追加。
- 治験に関するメタデータの追加。  
(schema.orgよりタームを選択)

メタデータ		例
試験ID	@MedicalEntity_code	UMIN000019843
対称疾患	@MedicalStudy_studySubject	進行扁平上皮肺がん
最終情報更新日	@Date_dateModified	2015-10-22
試験開始日	@Date_startDate	2015-10-22
進捗状況	@MedicalStudy_status	限定募集中

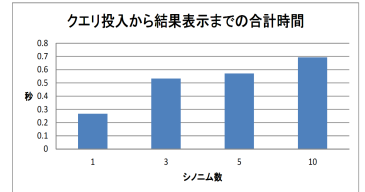
#### 2. 検索速度の向上

- 検索エンジンをHyper EstraierからElastic Searchへ変更。
- 検索速度が21倍向上。

Hyper Estraier



Elastic Search



SagaceのHE, ESサーバに対して、同義語を3、5、10個含む100種類のキーワードを用意し、2秒間隔で100回リクエストを送った。リクエストごとに毎回異なるキーワードを用いて検索を実行し、平均をとった。

#### References

- Sagace 創薬・疾患研究のためのデータベース横断検索サービス, 実験医学増刊, 2014 32, 3212.
- Sagace: a web-based search engine for biomedical databases in Japan., BMC Res Notes, 2012 Oct 31, 5:604.



Licensed under a Creative Commons 表示4.0国際 (c)2016 長尾知生子