ライフサイエンス統合データベースの横断検索サービスの現状

○杉崎太一朗[1 3]、牧口大旭[3]、高橋順子[3]、平井信一[1 3]、今西麗子[3]、奥村利幸[3]、谷口仁[2]、畠中秀樹[1]、川本祥子[1]、高木利久[1]
([1]情報・システム研究機構ライフサイエンス統合データベースセンター、 [2]株式会社日立ソリューションズ、[3]三井情報株式会社総合研究所)

Database Center for Life Science (DBCLS), Japan
ライフサイエンス統合データベースセンター

生命情報データベースの横断検索サービスの提供目的

文部科学省「統合データベースプロジェクト」では、ライフサイエンス関連データベースの利用を促進するために多種多様なシステムを開発し、サービスとして外部に提供しており、その中のひとつである生命科学データベースの横断検索システムは、国内外274個のデータベースを横断的にテキスト検索できるシステムです。(http://lifesciencedb.jp/dbsearch/)

専門性の高いデータベースをシステムの検索対象とすることにより、利用者である研究者が知りたい情報を迅速に取得出来、ライフサイエンス研究の現場で活用されることを目的としています。

システム概要

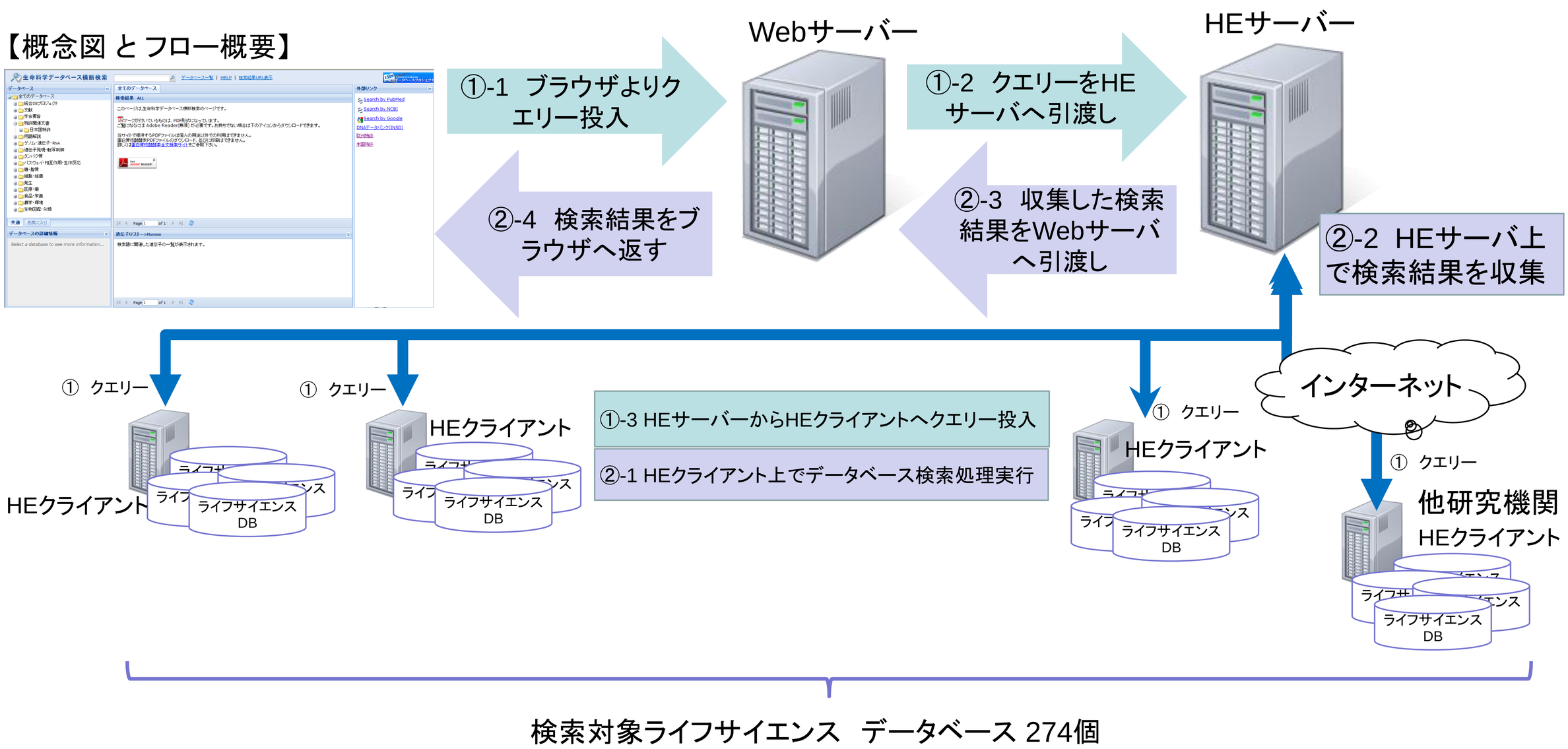
検索エンジンはオープンソースであるHyperEstraier(以下HE)を利用しています。

本システムは、プロトコルがHTTPプロトコルベースのクライアントサーバ方式で構成されており、HEクライアントの増設などスケーラビリティを向上させることが比較的容易です。

尚、通信路がHTTPプロトコルベースなので、インターネット経由で外部の研究機関のマシンと連携することが可能です。

また、各HEクライアントに組み込んだデータベースのメタ情報の管理は、当センターで開発した専用の管理ツールで行っています。

【概念図 とフロー概要】



横断検索サービスの最近の主なトピックスと今後について

1. 国内外のライフサイエンス系データベース274個を検索対象としています。
2. 2010年6月に、経済産業省の統合データベースプロジェクトと連携して相互でデータベース検索が可能となりました。
3. 2009年度に、システム内の検索辞書機能(日英翻訳、同義語)の整備を行ないました。
4. 今後は、更に検索対象データベース数の拡張を予定しています。
   また横断検索のユーザーインターフェースの機能改良・追加、英語対応等を予定しています。

Licensed under a Creative Commons表示2.1日本 license ©2010 ライフサイエンス統合データベースセンター