

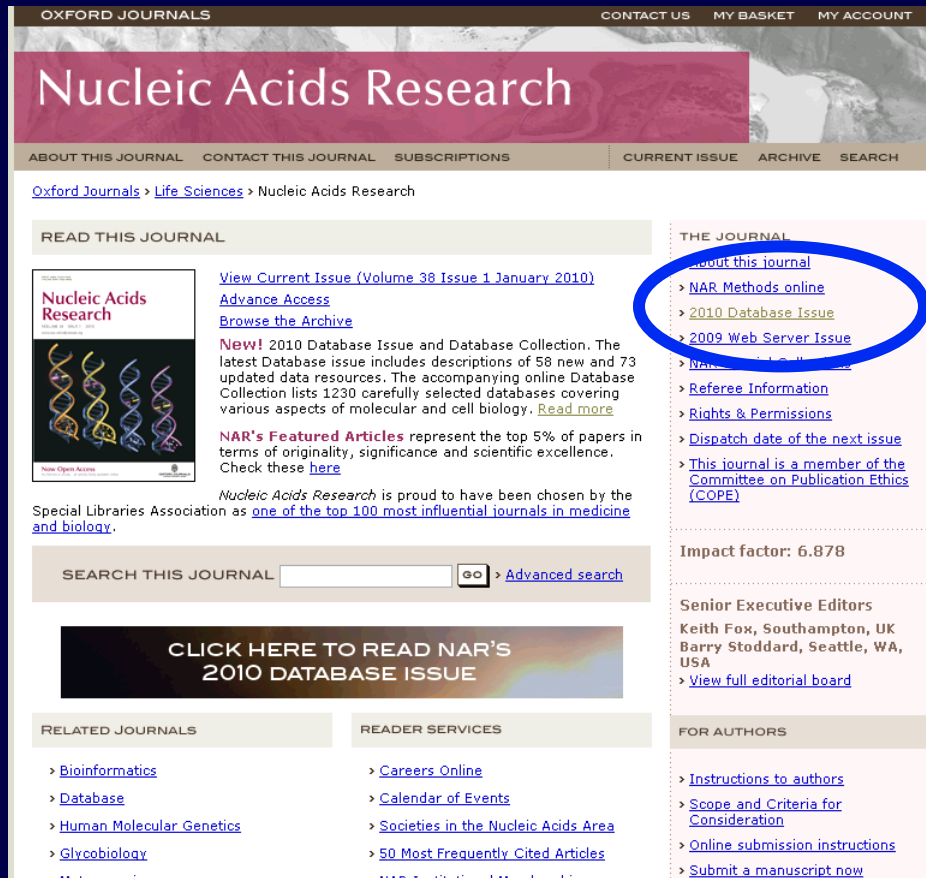
統合データベース講習会@筑後 生命科学情報の統合に向けて とくに情報検索

大学共同利用機関法人 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
特任准教授
川本 祥子

世界に生命科学関係のデータベースは どれくらいあるでしょう



Nucleic Acids Research Database issue



OXFORD JOURNALS CONTACT US MY BASKET MY ACCOUNT

Nucleic Acids Research

ABOUT THIS JOURNAL CONTACT THIS JOURNAL SUBSCRIPTIONS CURRENT ISSUE ARCHIVE SEARCH

[Oxford Journals](#) > [Life Sciences](#) > Nucleic Acids Research

READ THIS JOURNAL

Nucleic Acids Research
View Current Issue (Volume 38 Issue 1 January 2010)
[Advance Access](#)
[Browse the Archive](#)

New! 2010 Database Issue and Database Collection. The latest Database issue includes descriptions of 58 new and 73 updated data resources. The accompanying online Database Collection lists 1230 carefully selected databases covering various aspects of molecular and cell biology. [Read more](#)

NAR's Featured Articles represent the top 5% of papers in terms of originality, significance and scientific excellence. Check these [here](#)

Nucleic Acids Research is proud to have been chosen by the Special Libraries Association as [one of the top 100 most influential journals in medicine and biology](#).

SEARCH THIS JOURNAL > [Advanced search](#)

CLICK HERE TO READ NAR'S 2010 DATABASE ISSUE

THE JOURNAL

[About this journal](#)

- > [NAR Methods online](#)
- > [2010 Database Issue](#)
- > [2009 Web Server Issue](#)
- > [What's new in NAR](#)
- > [Referee Information](#)
- > [Rights & Permissions](#)
- > [Dispatch date of the next issue](#)
- > [This journal is a member of the Committee on Publication Ethics \(COPE\)](#)

Impact factor: 6.878

Senior Executive Editors
Keith Fox, Southampton, UK
Barry Stoddard, Seattle, WA, USA
> [View full editorial board](#)

RELATED JOURNALS

- > [Bioinformatics](#)
- > [Database](#)
- > [Human Molecular Genetics](#)
- > [Glycobiology](#)
- > [Mutagenesis](#)

READER SERVICES

- > [Careers Online](#)
- > [Calendar of Events](#)
- > [Societies in the Nucleic Acids Area](#)
- > [50 Most Frequently Cited Articles](#)
- > [NAR Institutional Membership](#)

FOR AUTHORS

- > [Instructions to authors](#)
- > [Scope and Criteria for Consideration](#)
- > [Online submission instructions](#)
- > [Submit a manuscript now](#)

•1993年、24のデータベースに関する論文から始まったデータベース特集号

•2010年度版で
合計1230データベースに

•オープンアクセスで全部読めます！

•<http://nar.oxfordjournals.org/>

生命科学のデータベース

- ◎ 毎年、膨大なデータと、数多くのDBやツールが公開されます。

→ 必要とするものがどこにあるかわからない、使い方もわからない

- ◎ 数多くのDBが作られているが、バラバラに構築されていて使い勝手が悪い。

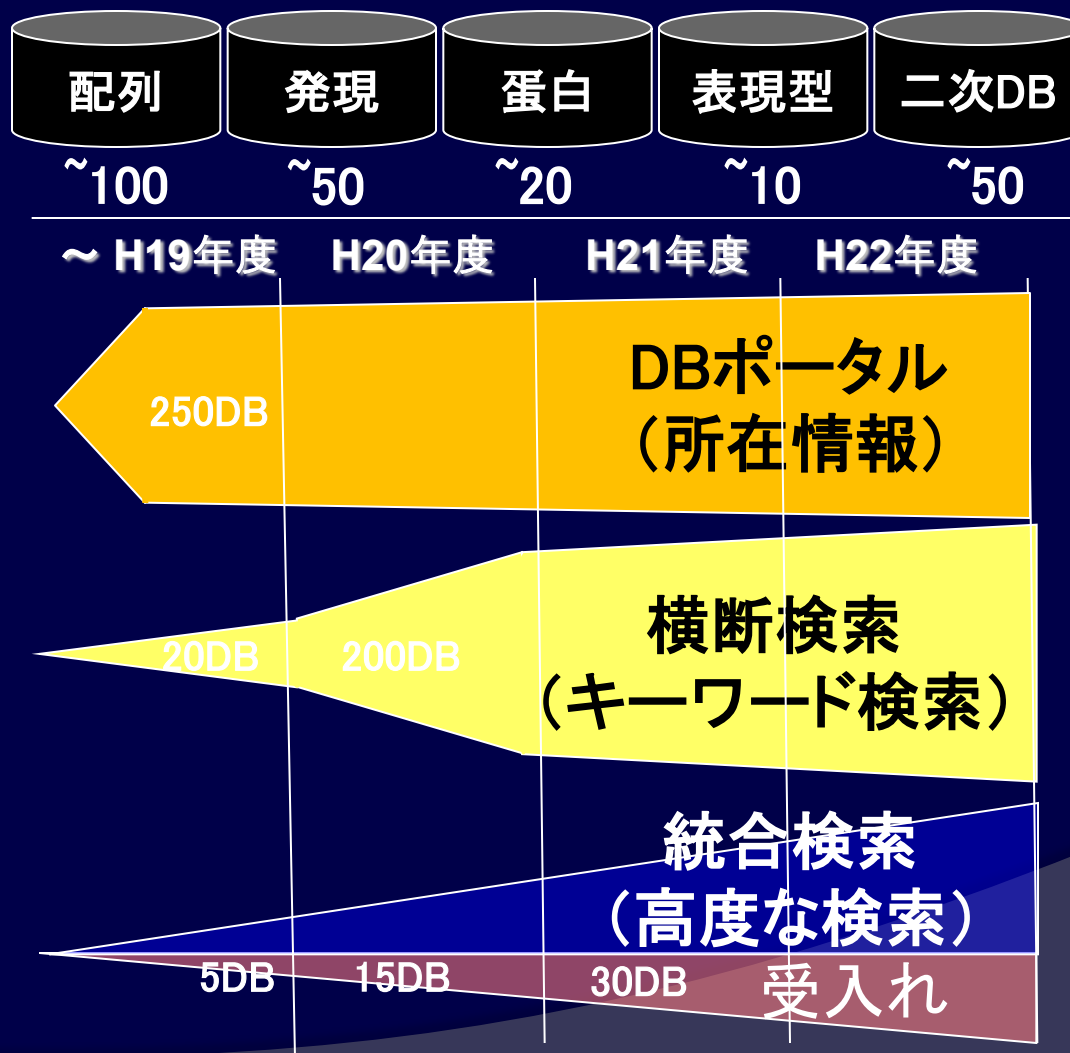
→ 一度に検索したり利用したり使い方を説明したりする窓口が必要

- ◎ 統合DBプロジェクトがサービスとして取り組んでいます

統合年次計画：縦割り日本のデータベースを統合

文科ライフ委DB戦略WS報告

H18年度
戦略立案
ポータル
統合技術



産業・臨床・育種
従事者の課題
解決への展開

4
省庁
統合
DB

環境変化と知識
データ増大への
絶え間ない対応

統合DBプロジェクトのサービス概要

- ◎ 分子系- DNA, RNA, 蛋白、糖鎖、化合物
- ◎ テキスト系（自然言語）
- ◎ ツール系、教材
- ◎ 全て一カ所から公開

この後、使い方を
実習します



統合DBプロジェクトのサービス以外の活動

- ◎ 生命科学情報の利用を促進するために、権利やルールを整備する

例1: 会社からDBを使おうとしたけど営利目的で使ってよいかわからない
共通の利用許諾をつける: DBアーカイブサービス
→ クリエイティブコモンズライセンス

例2: 病院で匿名でゲノムを解析してもよいかどうかインフォームドコンセントを尋ねられた

ヒトゲノム公開倫理指針の作成: ヒトゲノムバリエーションDB
→ 個人ゲノムの時代に対応、倫理と情報セキュリティ

データが多いのは生命科学に限ったことでは
ありませんね

情報爆発時代：

人類が有史以来作製した文書数の推移

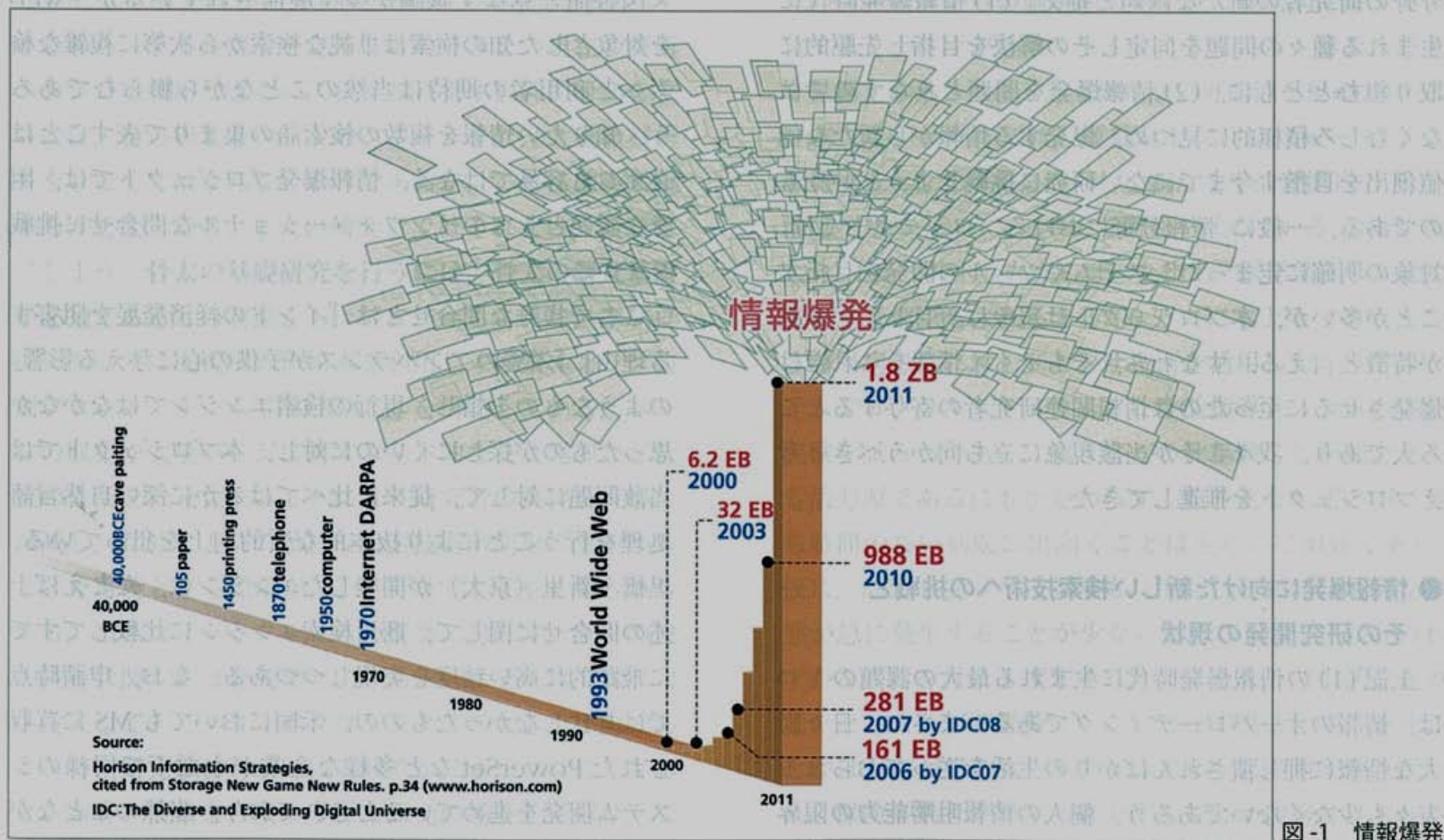


図-1 情報爆発

☆¹ IDC の 2007 年予測値は 2008 年に修正されたが、本図は、両者とさらに 2004 年特定申請当時 Horizon 社から公開された数字を合わせて描いた。

現代は＝研究は
情報検索の能力が要求される

世界中の情報を検索できる便利な時代

Google 久留米 ラーメン 検索 検索オプション

☒ ウェブ全体から検索 ☐ 日本語のページを検索

ウェブ 検索ツールを表示 久留米 ラーメン の

他のキーワード: 久留米ラーメン ランキング 久留米ラーメン 太砲 久留米ラーメン 通販 久留米ラーメン 清陽軒 久留米ラーメン

久留米駅(福岡) 付近のラーメンの検索結果



A map of the Kurume area in Fukuoka, Japan, showing several ramen shops marked with red pins. The map includes labels for various locations like Kurume Station, Kurume University, and several ramen shops. The pins are labeled with letters A through G, corresponding to the list of results on the right.

- A 大茶ラーメン
daieiramen.fc2web.com - 0942-39-3977 - レビュー (8)
- B 丸幸ラーメンセンター原古賀店
www.marukoura-men.com - 0942-40-3737 - レビュー (7)
- C 大龍ラーメン ベルモール東町店
www.dairyuramen.com - 0942-36-0446 - レビュー (4)
- D 来福軒
maps.google.co.jp - 0942-33-7957 - レビュー (2)
- E ラーめん亭
maps.google.co.jp - 0942-35-7472 - レビュー (6)
- F 太砲ラーメン 長門石店
www.taiho.net - 0942-39-7424 - レビュー (3)
- G 清陽軒本店
maps.google.co.jp - 0942-35-2237 - レビュー (5)

久留米駅(福岡) 付近のその他の検索結果 >

久留米ラーメン - Wikipedia

久留米ラーメン(くるめラーメン)は、福岡県久留米市を中心に作られる、豚骨スープとストレートの細麺をベースにしたラーメンである。豚骨ラーメンの元祖ともいわれているラーメンで、多くの九州ラーメンだけでなく、様々な地域のラーメンにも影響を ...

経緯 - 特徴 - 影響 - 関連項目

ja.wikipedia.org/wiki/久留米ラーメン - キャッシュ - 類似ページ - 印刷

久留米 太砲ラーメン - [イントロをスキップ]

呼び戻しスープ発祥の老舗。とんこつラーメンの神髄。

www.taiho.net/ - キャッシュ - 類似ページ - 印刷

ちょっと待った

Googleの処理対象URL 1兆突破

日本のブログ数 1690万

記事総数13億

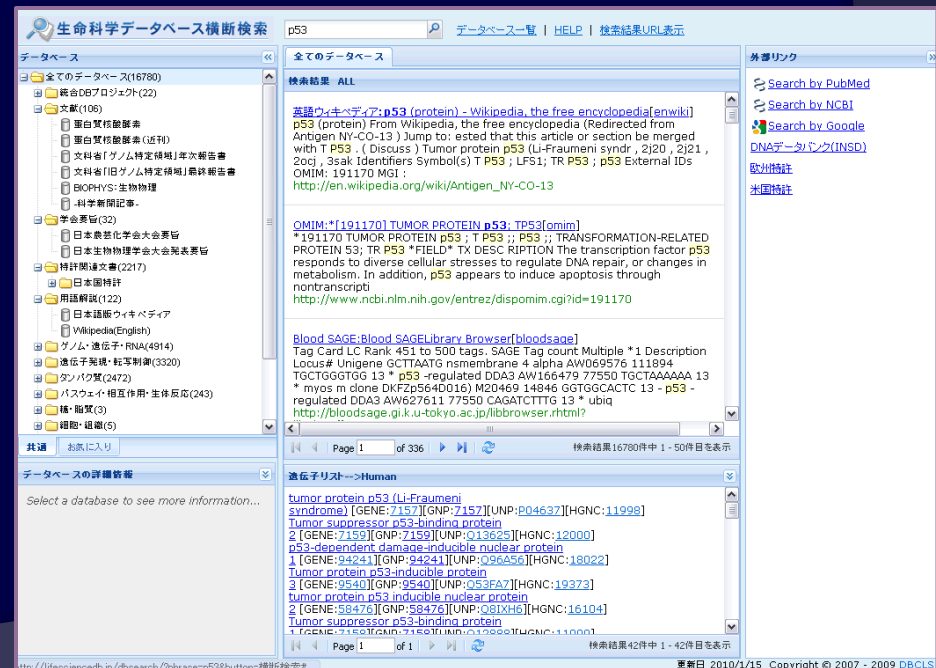
Pubmed 1859万

Genbank 核酸 7500万

(2008年データ)

統合DBプロジェクトでは 生命科学に特化した検索サービスを提供

- ◎ 生命科学データベース横断検索 } 200のDB
 - ◎ OReFiL
 - ◎ Allie
 - ◎ inMeXes
- } Pubmed



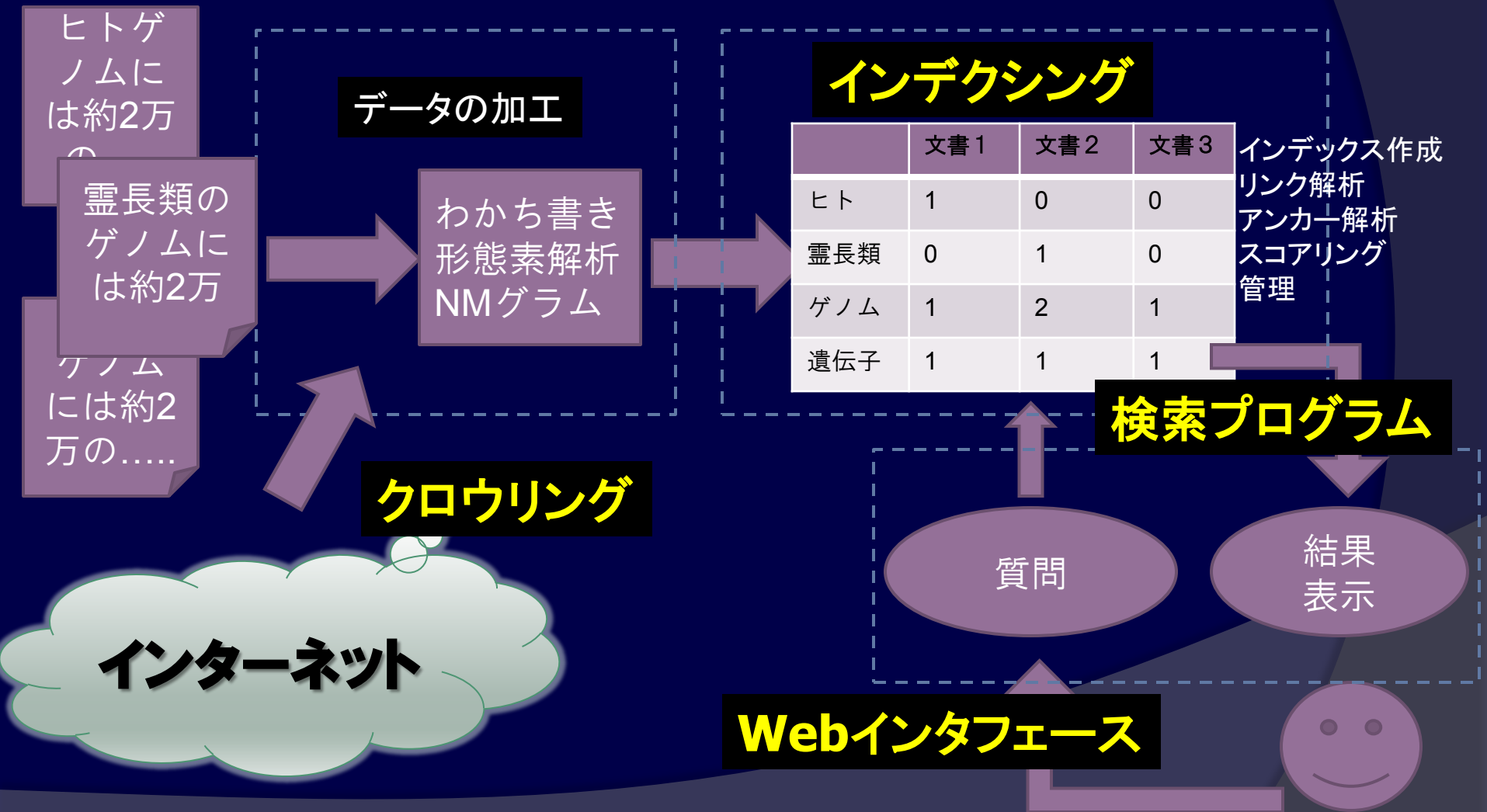
検索エンジンとは

- ◎ 情報の収集(クロウリング)、インデクシング、検索という3要素で成り立つシステム
- ◎ 大きく分けて二つのタイプ
 - ディレクトリ型---昔のYahoo, 統合DBのDBカタログ
 - ロボット型--- Google, 統合DBの横断検索

文章の中からキーワードを検索

◎ 1, 特定領域研究「生命システム情報」及び特定領域研究「比較ゲノム」と連携し、ゲノム研究関連特定研究領域で解析の対象となる真核生物であるメダカ（近縁種2種）、ホヤ（各種）、ナメクジウオ、ギボシムシ、**霊長類**（各種）、カイコ、コムギ、クラミドモナス等を中心に、オリゴキャップ法による完全長cDNAライブラリー、5' 端濃縮ライブラリー、5' 端SAGE(5' SAGE)ライブラリーを作製し、それらのcDNAライブラリーからのクローンの5' 端EST 配列を決定することで、遺伝子の転写開始点の詳細とそのコアプロモータ領域を同定する。

検索の基本的な仕組み

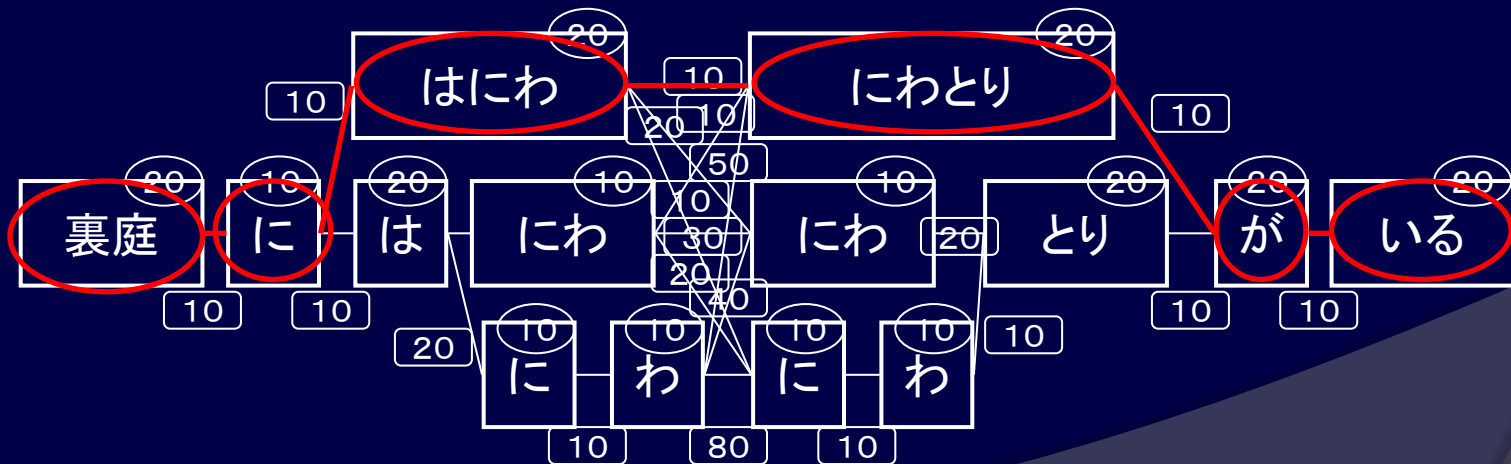


日本語の処理 形態素解析

(辞書引きアルゴリズム)

- 入力文に出現する全形態素の候補（辞書中にあるもの）を枚挙

裏庭にはにわにわとりがいる



配列検索

```
AGCTGTCTGTAGCTAGCTAGCAT  
CCTAAACCCTGTGTGGGGCATG  
AGCTGTCTGTAGCTAGCTAGCAT  
CCCTGTGTGGAAATAAGGCAT  
CTTTTGTATGTGTGTGTGTCCCC  
CTTTTGTATGTTAGTCCCC
```

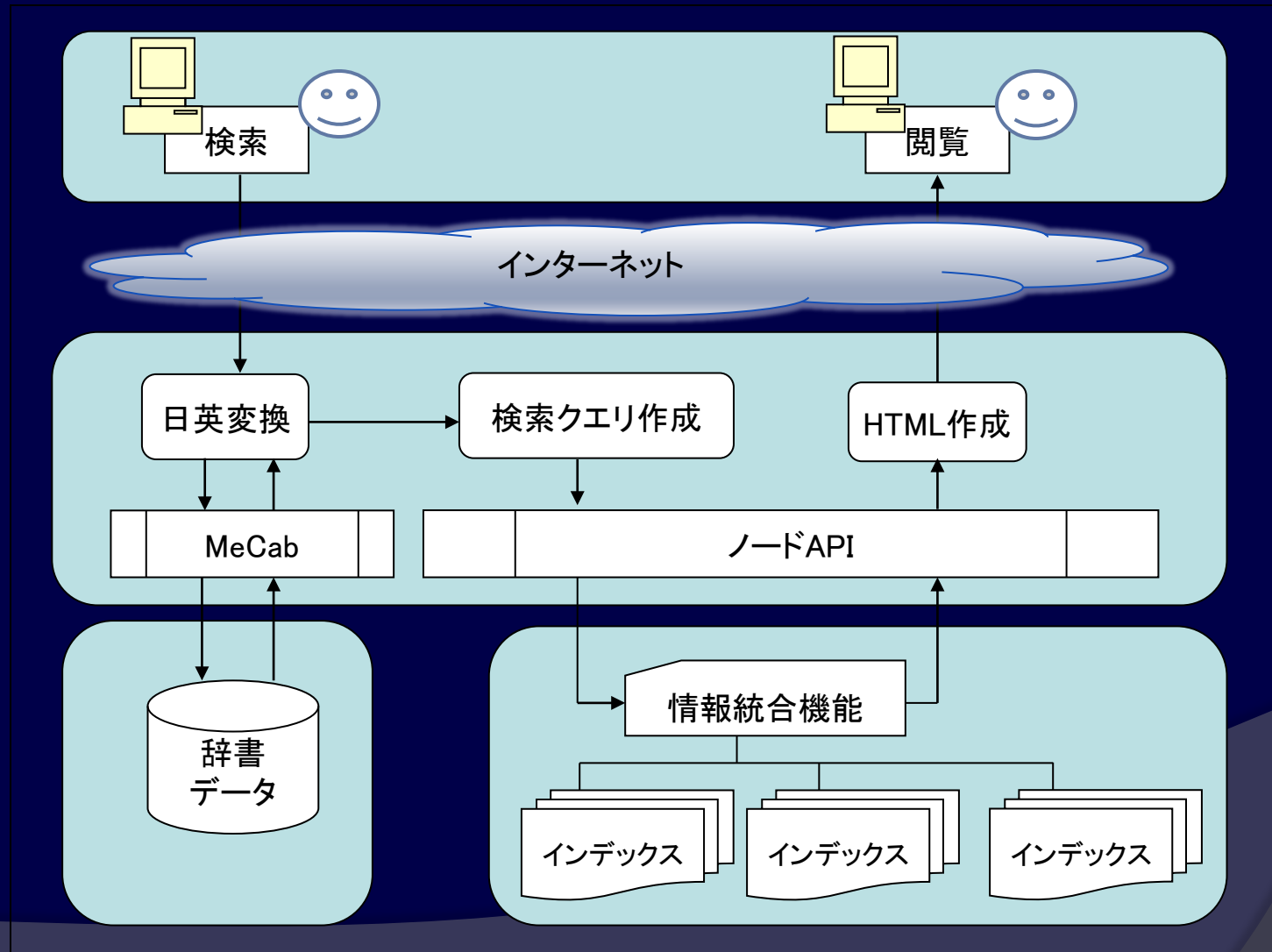
様々なアルゴリズムによるプログラムが開発された

生命科学データベース横断検索サービスの特徴

- ◎ 国内中心に集めた200以上の分子DBや文献をキーワードで検索できるGoogleライクな検索です
- ◎ 分子だけでなくPNEや特許などの文献も同時に検索できます
- ◎ 検索エンジンとしてhyperestraierを利用しています
- ◎ 日本語文書のインデックスはnグラム法で作製しています
- ◎ 日英翻訳が可能でライフサイエンス辞書(LSD)を使っています
- ◎ DBをカテゴリーごとに分けているので興味のある分野の絞り込みが簡単です

1月から著作権の一部が改正され検索エンジンが合法になりました

生命科学データベース横断検索エンジンの概略



様々な検索: J-GLOBAL

JSTが提供、日本の**科学情報**と**研究者情報**を統合検索

The screenshot displays the J-GLOBAL website, a platform for integrated searching of scientific information and researcher data in Japan. The header features the J-GLOBAL logo and the text "科学技術情報リソースセンター". Below the header, there is a search bar with a "検索" (Search) button and a "ホーム" (Home) link. The main content area is divided into several sections:

- お知らせ (Notice):** A section with a list of recent updates or announcements.
- 最新の検索履歴 (Latest Search History):** A section showing the user's recent search history.
- 統計 (Statistics):** A section displaying a line graph showing trends over time.
- 最新の検索履歴 (Latest Search History):** A section showing the user's recent search history.
- 最新の検索履歴 (Latest Search History):** A section showing the user's recent search history.

The footer contains copyright information: "Copyright © 2007 Japan Science and Technology Agency. All rights Reserved."

様々な検索:あの人検索 スパイシー

SPYSEE β version | ひと 669,677 | つながり 55,507,620

[スパイシー英語版](#) | [チアスパ! であのひとを応援しよう!](#) | [ユーザ登録](#) | [ログイン](#)

あの人検索 スパイシー



坊農秀雅

検索

アドバンス検索

今、話題のひと [石川知裕](#) [奥村公延](#) [小沢一郎](#) [もっとみる](#)

ユーザー登録

SPYSEE
テレビ

チアスパ!
日本代表チーム
応援キャンペーン



坊農秀雅



ウェブ上からみた人物像 [プロフィールをもっと見る・変更する](#)

坊農秀雅簡単な説明統合DBプロジェクトなかのひと&ライフサイエンス統合データベースセンター所属
市区町村名 (Town) 文京区弥生拡張プロフィールをクリックすると拡張プロフィールを表示します。公的なウェブ
サイトアドレス <http://lifesciencedb.jp/> 個人的なウェブサイトアドレス <http://DBCLS.rois.ac.jp/%7Ebano/>
会社/団体大学共同利用機関法人 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
<http://navi.kazusa.or.jp/bano/> より抜粋

[JMP Genomicsでゲノム解析](#)

データから意味のあるパターン発見 SASとJMP
の新しいソリューション

[DNA塩基配列解析は](#)

最安・最速・高品質 700~850円 DNA/次世代
塩基配列解析はこちら!

Ads by Google



坊農秀雅さんの相関図 [大きく表示する](#)



www.kanazawa-it.ac.jp

? 画像について x この画像を表示しない



選ぶ

[J-Bio21のSNP解析](#)

受託サービス・解析用試薬販売。高精度な解析を低価格で
提供!

www.j-bio21.co.jp/consal/snp.htm

[SQL Server 2008](#)

DBのコンプライアンス強化は企業にとって不可欠要素! 詳しく
はこちら

www.microsoft.com

[マスタデータ統合の真実](#)

MDMのノウハウを凝縮 無料マスタデータ(MDM)統合レポート
mdm-seminar.info/

[タンパク質複合体の解析](#)

複合体中のタンパク質をショットガン 解析で一斉同定!
www.cerij.or.jp

< >

Ads by Google



タグ

様々な検索：〒MAP ポストをひたすらマッピング

〒MAP

ポストをひたすらマッピング

現在 136,381 ポスト (全国192,213ポストの70.953 %)

マップ*

リスト

追加*

写真

効*

データ

コメント

ユーザー

POWERED BY Google

新着ポスト(上の地図は、新着10を表示)。一覧、検索は、[マップから](#)

お知らせ (一覧)

▼CSVダウンロード[エキスパート向け] コメント(24) 2010/1/13

▼タグBBSなどのタグ関係いくつか コメント(10) 2009/12/4

▼11/18(水) 9:00~18:00 メンテナンスで止まります。 コメント(3) 2009/11/11

ブログ (一覧)

▼第一回ポストマップ京都オフ会のご案内 コメント(5) hiro129 2010/1/17

▼再開の可能性が高い郵便局 コメント(0) 郵便局好き中学生 2010/1/14

▼厚木北支店再調査中 コメント(10) Hiroki sakurada 2010/1/14

▼〈徳島県〉阿南支店実地集配センター コンプリート！ コメント(3) kmana 2010/1/13

▼今日はとても寒いです!! コメント(4) なんと4 2010/1/13

アンケート (一覧): このタグどう?

検索

サイト内検索

ユーザー募集 | ログイン

地図

集合知

携帯版QRコード

iPhone版

I♥FX

夜の時間を有効活用!
今こそはじめるFX!

Point 1
1万円から
始められる!

Point 2
円高でも
円安でも
収益のチャンス!

Point 3
24時間
取引ができる!

▶今なら合計10,000円キャッシュバック

外為どっとコム

www.gaitame.com

Ads by Google

ポストマップについて

ポストは探すとなかなか見つからないもので、引越後などにポスト探しで消耗されている方をあちらこちらで見かけます。

というわけで、ポストマップは、みんなで郵便ポストの場所をマッピングしていくというプロジェクトです。

ユーザー登録することで、マッピングが可能になります。

▼マッピングには中毒性がありますので、やり過ぎにはご注意ください。視力低下、けんしゅう炎を招く恐れがあります。

▼ポストの写真を撮っていると不審がられることもありますが、自信を持って撮影して下さい。

▼このプロジェクトは、郵政民営化などのムーブメントとは一切関係がありません。

既にアカウントをお持ちの方は、ログインへ。
お持ちでない方は、ユーザー登録へ。

検索結果に付加情報をつける： レコメンデーション 推薦システム



改訂第2版 バイオデータベースとウェブツールの手とり足とり活用法
遺伝子の配列・機能解析、タンパク質解析、プロテオミクス、文献検索、検索エンジン...etc.真に役立つサイトを使い倒す！(大型本)
中村 保一(編集), 石川 洋(編集), 鎌合 敦(編集), 平川 美夏(編集), 坊間 秀雅(編集)
またカスタマーレビューはありません。最初のレビューを書く

価格: ¥ 4,410 国内配送料無料(一部例外あり) [詳細](#)
ポイント: 44pt (1%) [詳細はこちら](#)
在庫あり。 [在庫状況](#)について
この商品は、Amazon.co.jp が販売、発送します。ギフト包装を利用できます。

2009/1/24 土曜日にお届けします！ 今から20時間と25分以内にレジに進み、「お急ぎ便」オプション(有料)を選択して注文を確定されたご注文が対象です。詳しくは[こちら](#)
一部サイズが大きい商品の場合、上記の日付が適用されない場合があります。配送オプションを選択する画面、もしくは注文確定するとき必ず配達予定日をご確認ください。

新品1点 ¥ 4,410より

[Click here to see in English.](#)

数量: 1 [▼](#)
[ショッピングカートに入れる](#)
または
1-Clickで注文する場合は、[サインイン](#)をしてください。

こちらからも買えますよ
この商品を安く買いたい！
[中古商品を予約する](#)
この商品をお持ちですか？
[マーケットプレイスに出品する](#)

[ほしい物リストに追加する](#)
[お友だちに知らせる](#)

[自分のイメージを掲載する](#)
[出版社、著者の方へ「なか見」検索」で書籍を紹介させ](#)

これを買ったひとはこっちも買ってます

をあわせて買う



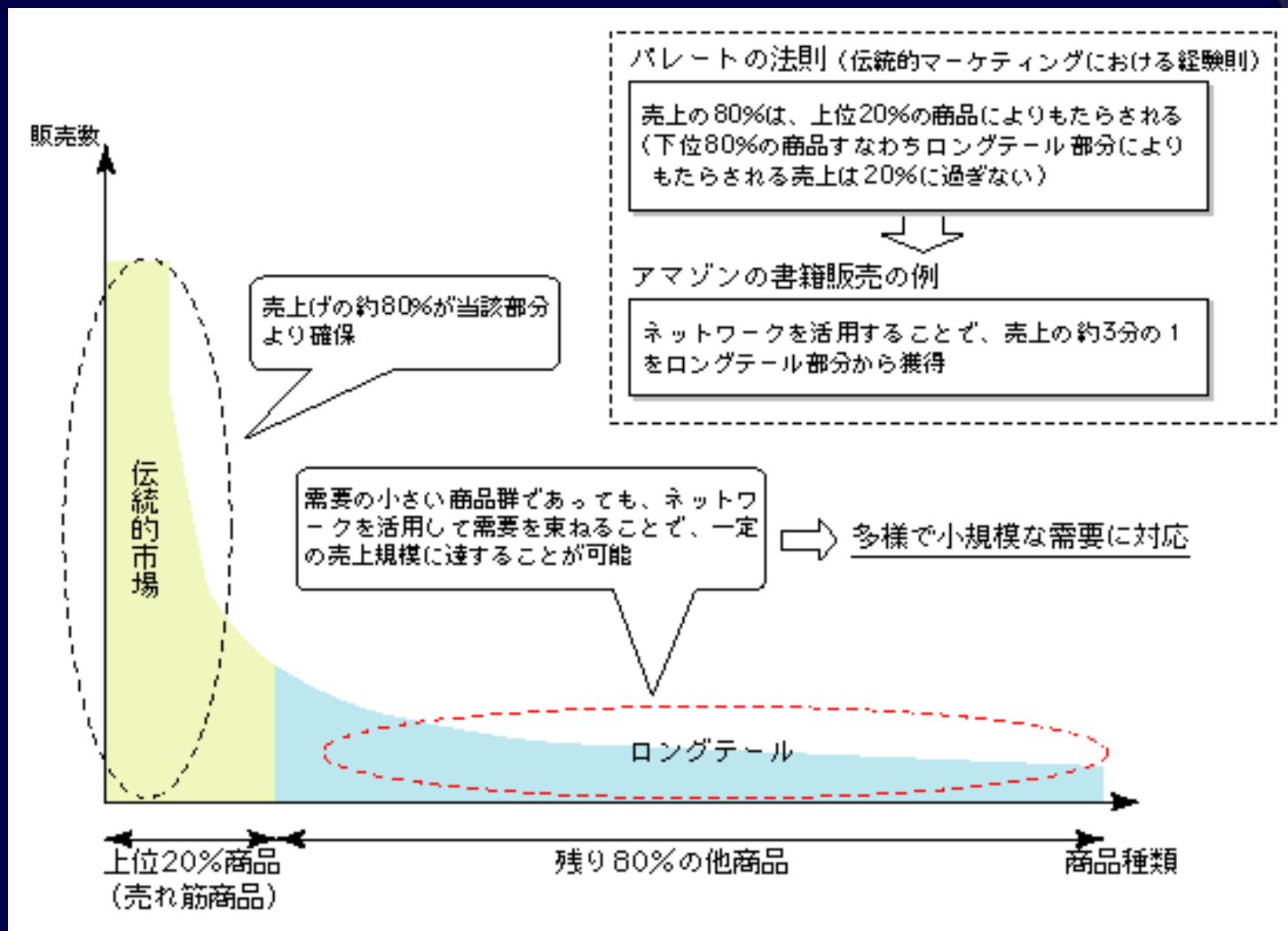
合計価格: ¥ 8,820
ポイント合計: 88pt (1%)
[両方ともカートに入れる](#)

この商品を買った人はこんな商品も買っています



ページ: 1 / 20

ロングテール現象



ウェブ技術を組み合わせた情報検索の例： マッシュアップ



◎ 感染症の動向を知らせるbiocaster

文書に意味をもたせる

検索者が欲しい文書をより良く探せるように

- ◎ Xmlなどのファイル形式
- ◎ RDFなどのメタデータ
- ◎ オントロジー・推論

セマンティック ウェブ

WWWの生みの親が提唱
Tim Berners Lee

