

自発的な学びを育む 連想的情報アクセス技術 — 検索から連想へ —

高野明彦 (NII 連想情報学研究開発センター)

1

自発的な学びを育む 連想的情報アクセス技術の研究

・「連想の情報学」を基礎づける連想計算機構

- 連想計算＝情報空間における関連情報の探索・分析・提示
- 検索から連想へ：文字列検索と根本的に異なる情報処理原理
- 「人の連想」と「電子情報空間における連想」の創造的相互作用
- スケーラブルな分散処理（文書1億件）、規模・性能とも世界一

・「連想する情報サービス」の構築方式

- コンテンツ（情報の蓄積）に相互運用性を付与
 - 連想検索、関連語抽出、関連性フィードバック
- 基本システム一式をオープンソース形態で公開予定

・「連想する場」としてのコンテンツ利用環境「想・IMAGINE」

- 多数の異なる情報源を動的に融合して、情報に文脈を与える技術
- ユーザの主体的な情報収集・発見を支援、世界的に類似技術なし

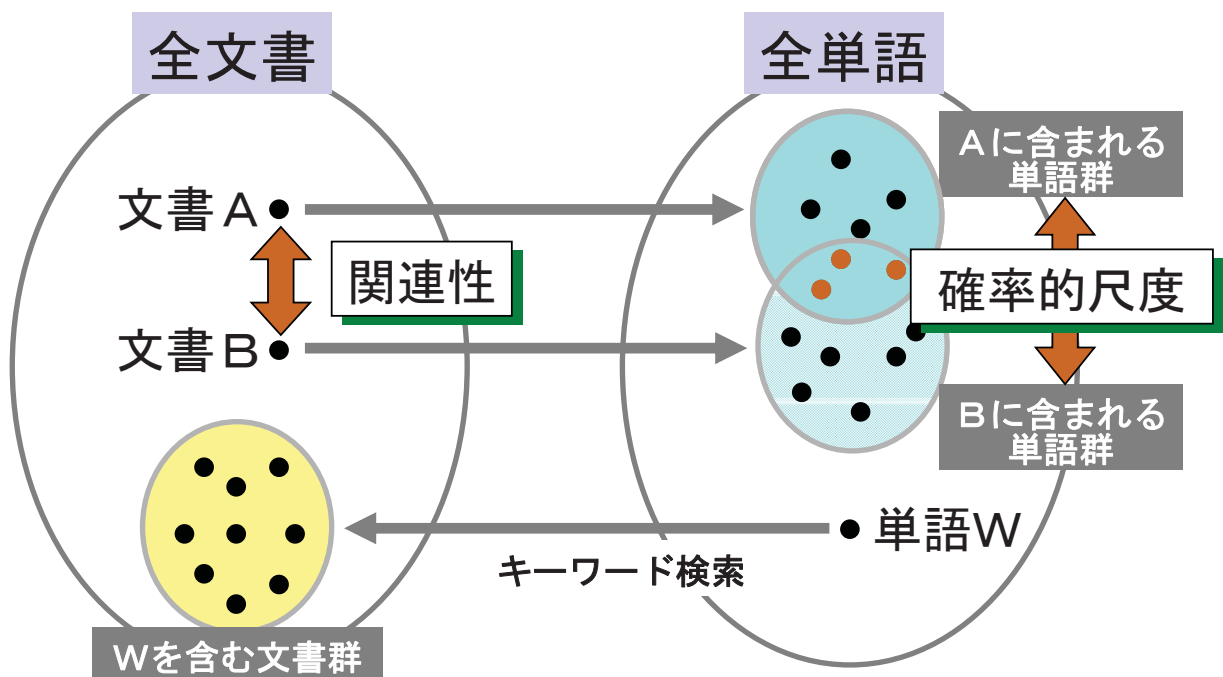
2

「連想の情報学」

- DualNAVI 1996.12～
 - 文書連想検索・特徴語グラフ・双対空間メタファー
- GETA(汎用連想計算エンジン) 2002.7～
 - 連想計算の高速実装・連想計量の動的変更
 - コンテンツの相互運用化コンパイラ
 - 連想計算Webサービスを発信
- 「想・IMAGINE」2006.6～
 - 分散管理されたコンテンツ群の統合利用環境
 - 分割コンパイルされたコンテンツの動的リンカー
 - 連想計算Webサービス同士の“相互作用”

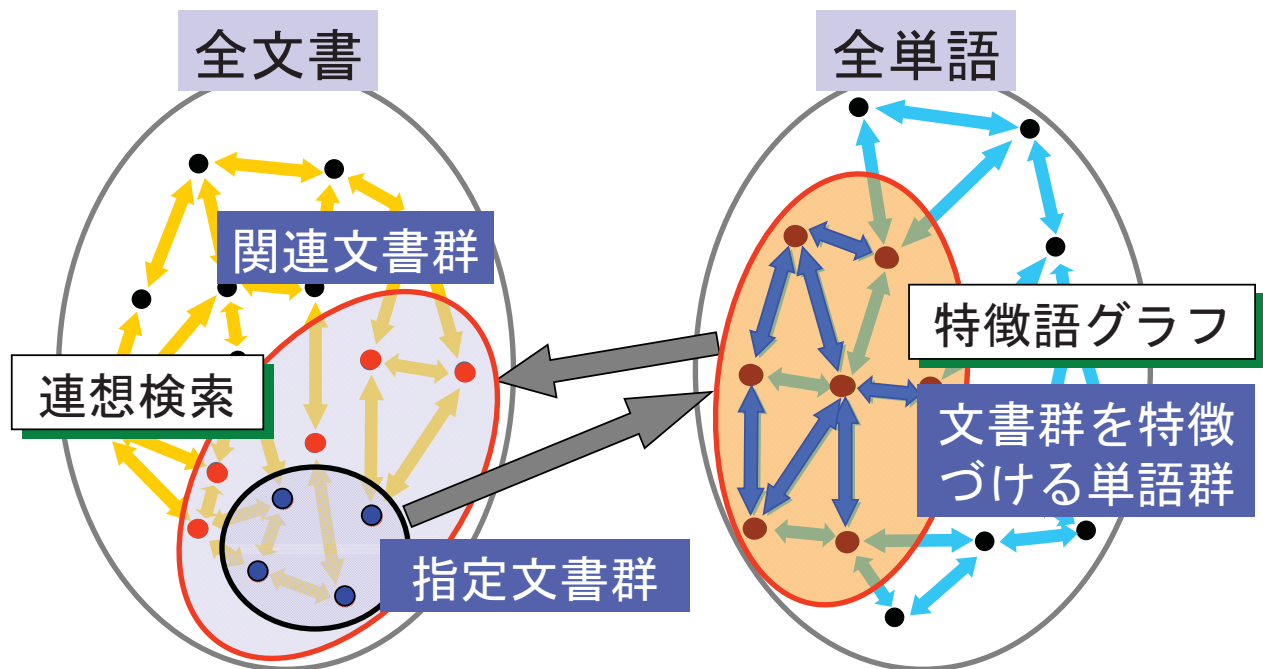
3

文書空間と単語空間の双対性



4

双対計量による連想計算の例



5

汎用連想計算エンジン GETA

(Generic Engine for Transposable Association)

- 1億件規模の文書DBの連想計算を実現
 - 各種PCクラスタ(1~64ノード)で動作可能
 - 連想検索、文書クラスタリング、重要語抽出などの基本機能を提供するライブラリ
- オープンソース形態で無償公開中
 - ホームページ: <http://geta.ex.nii.ac.jp>
 - オープンソースのOS上で動作可能
- ウェブ・サービス版 GETAssoc の公開準備中

6

「連想する情報サービス」の構築

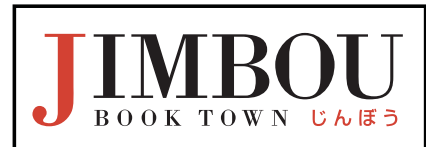
・専門知識を要求しない情報サービス

- 連想検索: 自然文・例示による検索
- 関連語抽出: 収集された情報の概観・自動要約
- 関連性フィードバック: 正解例を選んで再検索

⇒情報空間に奥行きと安心感を与える対話技法

・連想による情報サービス間の動的連携機能

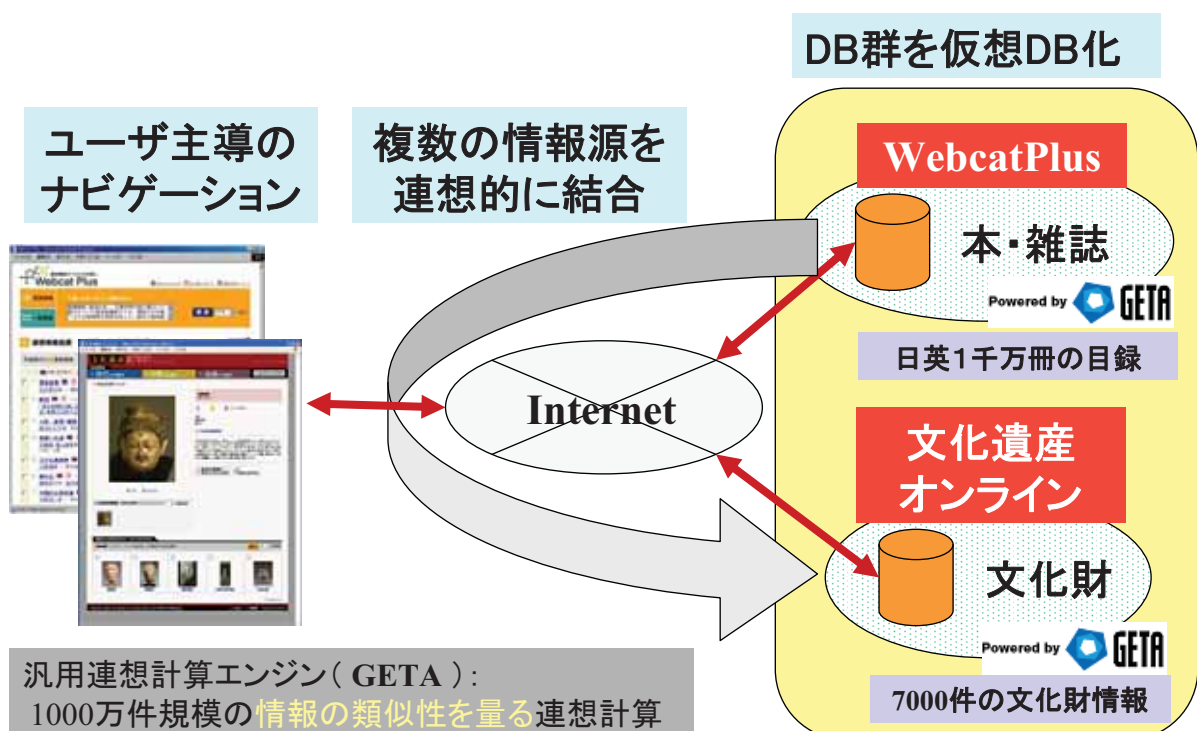
- 自然文・要約語群を仲介とする相互連想実現



新千代田図書館

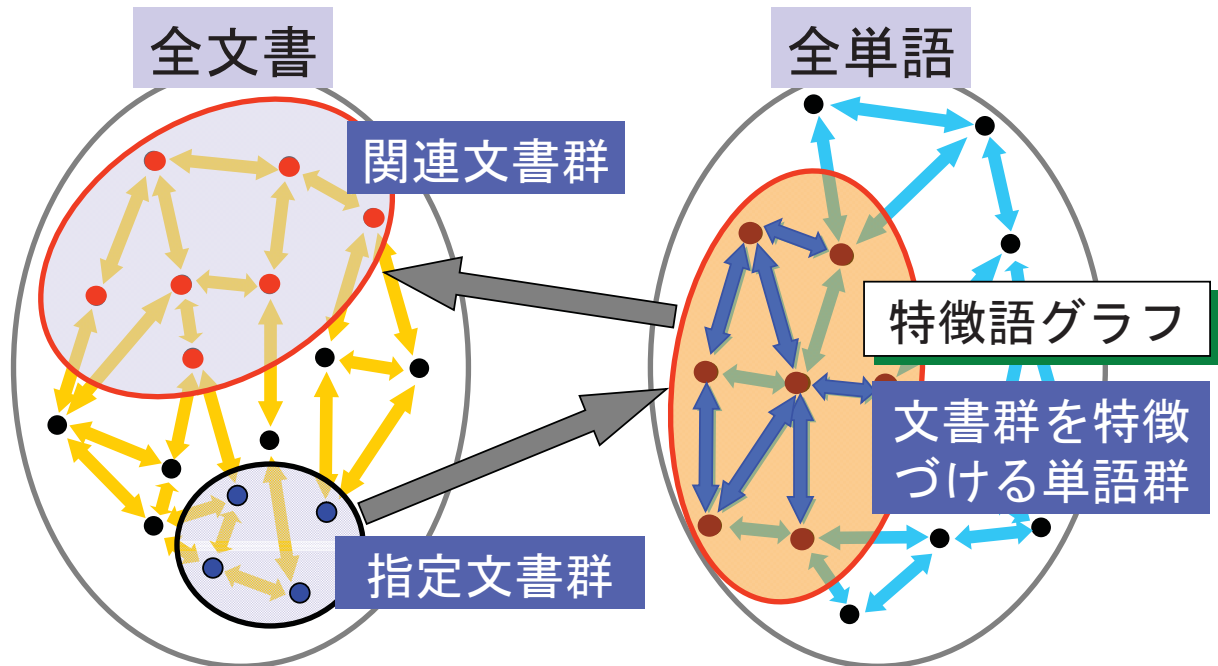
7

異種情報源への柔軟な情報アクセス



8

異なる情報源をつなぐ連想検索



9

意思決定で陥りやすい罠

- ・アンカリング: 最初に見つけた情報から過度に影響を受ける
- ・確証: 無意識に自分の既成概念を支持するデータを探し、それを覆す証拠は避ける
- ・記銘性: 直近の出来事や劇的な事件に過度に影響を受ける、複数の情報源から繰り返し同じ情報を受け取ると信用してしまう
- ・現状: 現状維持に役立つことを受入れ易い
- ・埋没費用: 過去の過ちをなかなか認めずに、これまでの選択を正当化する方向で意思決定を行う

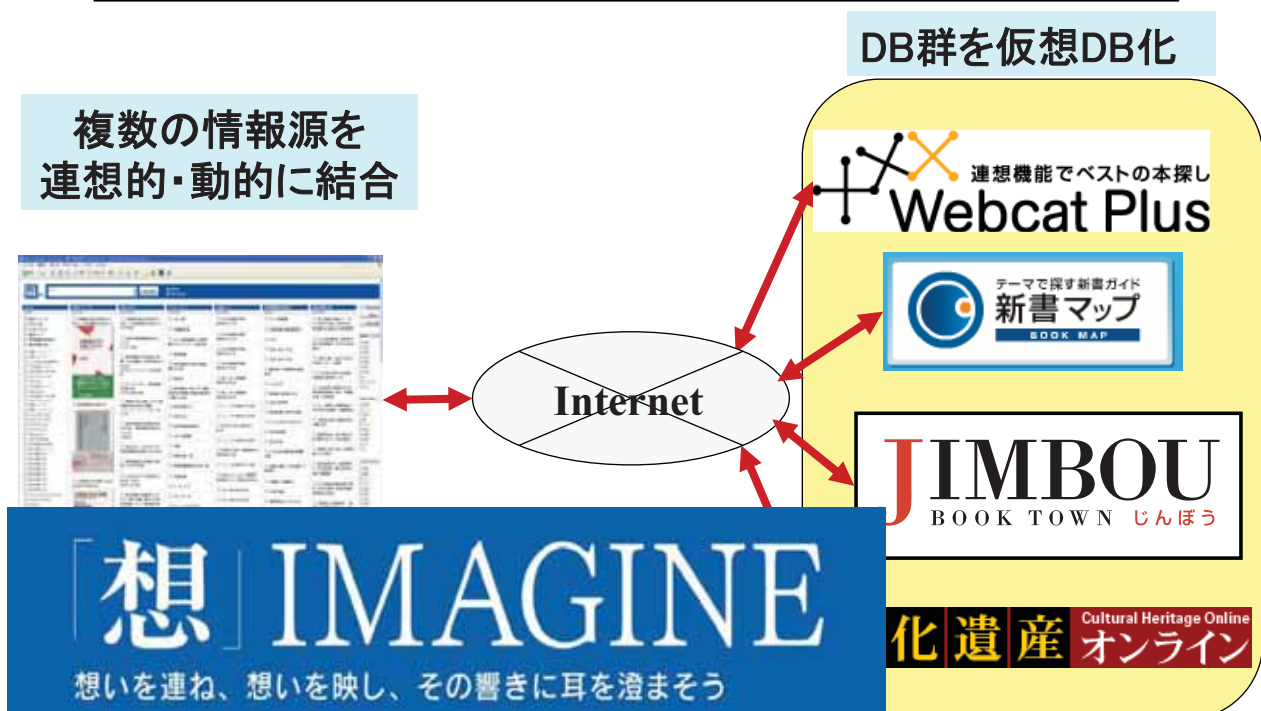
10

「連想の場」を提供する「想・IMAGINE」

- ・情報を文脈で捉え、視野と信頼性を確保
 - 発信元の異なる情報源を動的に関連づける
出版物、新聞、写真、美術館・博物館、Web、...
 - 不確かな情報を高信頼な情報に関連づける
- ・主体的な情報収集と発見を支援
 - 記憶は個性的な言葉の文脈、「連想」は指紋
- ・分散管理された多数の情報源を動的融合
 - コンテンツ群の動的マッシュアップ技術
 - 世界的に見ても類似技術なし

11

大規模異種情報の収集・結合技術



12

想 IMAGINE

想いを連ね、想いを映し、その響きに耳を澄まそう

The screenshot shows the homepage of the 'Imagine Book Search' website. At the top, there is a search bar with the word 'Imagine' and a 'Web Search' button. Below the search bar, there are five main categories of books displayed as thumbnails:

- pop.新書マップ-テーマ
- pop.新書マップ-本
- pop.text.岩波 生物学辞典 第4版
- pop.text.岩波 理化学辞典 第5版
- pop.text.広辞苑 第五版

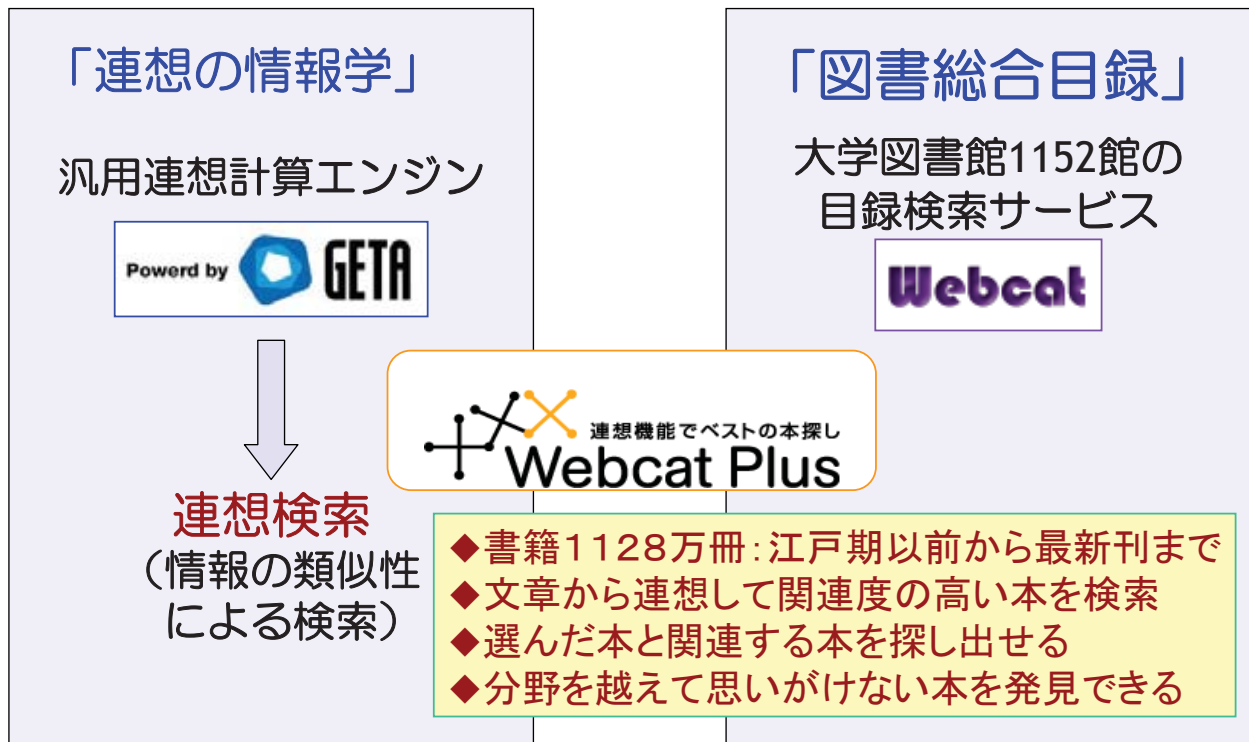
Each category has a corresponding book cover image and a brief description. To the right of the book thumbnails, there is a 'Web Search' section with the Google logo and a 'goo' search engine link. The website has a blue header and footer, and the overall design is clean and professional.

This screenshot shows the search results for the term 'CO2' on the 'Imagine Book Search' website. The search results are displayed in a grid format, with each result including a book cover, a title, and a brief description. The search results are as follows:

- 持続可能な開発** [英 sustainable development]
人類の発展を続けようとして自然環境の保全と調和や環境の保全のキーワードとされる。開発は現在の社会や社会を改良し、そこから利益を導出して、人間が長年、大規模かつ急激な開発が人間の環境を非可逆的に劣化させる状況、例えば熱帯林の破壊、砂漠の拡大、二酸化炭素の増加にもとも地球温暖化、オゾン層の破壊、酸性雨、海面上昇汚染、野生生物の絶滅などの点から環境において人間の発展を継続可能とする観点になった。
- 地球温暖化** [英 global warming]
二酸化炭素、メタンなどの大気中の温室効果ガスの濃度の増加によって地球大気の温室効果が増加し、その結果、地球全体の地上気温が高くなるというシナリオ、1989年に北アメリカで発生した干ばつの原因が化石燃料の燃焼による人為的な地球温暖化であるとする見解が広がり、大きな社会問題となり、二酸化炭素の排出量規制の議論が展開した。二酸化炭素の温室効果は、氷河期から間氷期に移る際に、気候変動のメカニズムの一つとして、地球温暖化を引き起こす。
- 環境科学** [英 environmental science (仏 science du milieu 独 Umwelt-Wissenschaft 露 наука об окружающей среде)]
1950年代の後半、DOTを中心とする自然環境保全運動がきっかけとなり、化学的におよんで主に学際領域健康リスクなど、レポートの公害問題として、多分野の学際領域を含む無機物、有機物の人間への影響、光、騒音、振動などの物理的現象が、公害衛生学、物理化学、気象学、土壌学、衛生学、経済学、社会学など人文科学の中で人間の生存を中心課題の調和を保つべきかという点が
- 砂漠化** [英 desertification]
植生のある土地が砂漠になること。地球上の陸地の約1/3は砂漠である。砂漠の中二古代都市の遺跡や岩に彫られた動物の絵が残されているから、かつては緑が豊かだった。人や動物が住んで、土を踏み固め、気候が変化して、降水の分布が変化すると、自然の状態で砂漠化が生じる。しかし、近年、社会問題となっているのは、人間の活動によって砂漠化が進行する問題である。現在、全世界の人口の約1/3の人々が乾燥地帯で生活している。農業のため灌漑を行なると、塩害が地上に供給され、土地がアルカリ性になって植物が育たなくなる。牧畜を行なると、家畜が草を根こそぎ食べるために、表層土が雨で流出されやすくなり、植物が育たなくなる。燃料にするために木を伐採することも砂漠化を助長する。このような人間の作用で、毎年、約77万km²の乾燥地帯が形成されている。特に、サハラ砂漠の周辺部のサヘル地方、カラハリ砂漠の南東、タウラマタ、砂漠の北東で砂漠化が進行している。

The search results are displayed in a grid format, with each result including a book cover, a title, and a brief description. The search results are as follows:

連想機能をもつ図書検索システム



文化遺産オンライン(試験公開版:2004.4公開、正式版:2008.3公開)

文化遺産 Cultural Heritage Online **オンライン** <http://bunka.nii.ac.jp/>



- 文化遺産オンライン構想 (文化庁)
全国の博物館、美術館が所蔵する文化遺産を
ひろく国民に向けて情報発信
- 国立情報学研究所は文化遺産のポータルサイト
構築に連携協力して、共同運営
- 専門知識をもっていない一般利用者を
文化遺産情報へ適切にナビゲート

時代から探す

分野から探す

地域から探す



平成20年度参加館 90館
収録文化財 16000件

安土・桃山で検索



茶釜、志野茶碗から連想検索



Webcat Plusへ連想検索

最も近い5作品





21



2

http://www.itmedia.co.jp - ITmedia News: 本を置くだけで情報検索 千代田図書館で - Micr...
ファイル(F) 編集(E) 表示(V) お気に入り(I) ツール(T) ヘルプ(H)

ニュース

本を置くだけで情報検索 千代田図書館で

台の上に本を置くと、その本に関連する書籍やWikipediaの情報などを検索し、ディスプレイに表示する。図書館の蔵書だけでなく、図書館近くの神保町古書店の在庫情報も検索できる。

2007年04月25日 20時34分 更新

5月7日にリニューアルオープンする千代田図書館(東京都千代田区)に、新しい情報検索システムが導入される。専用の台の上に本を置くと、その本に関連する書籍やWikipediaの情報などを検索し、ディスプレイに表示する仕組み。図書館の蔵書だけでなく、近所の古書店の在庫情報も検索できる。見知らぬ本に出会うことで発想を広げてもらいつつ、古書店に誘導して地域活性化にもつなげる狙いだ。



本の台の上に本を置くと、その本に関連する情報がディスプレイに表示される

23



自発的な学びを育む 連想的情報アクセス技術の研究

・「連想の情報学」を基礎づける連想計算機構

- 連想計算＝情報空間における関連情報の探索・分析・提示
- 検索から連想へ：文字列検索と根本的に異なる情報処理原理
- 「人の連想」と「電子情報空間における連想」の創造的相互作用
- スケーラブルな分散処理(文書1億件)、規模・性能とも世界一

・「連想する情報サービス」の構築方式

- コンテンツ(情報の蓄積)に相互運用性を付与
 - 連想検索、関連語抽出、関連性フィードバック
- 基本システム一式をオープンソース形態で公開予定

・「連想する場」としてのコンテンツ利用環境「想・IMAGINE」

- 多数の異なる情報源を動的に融合して、情報に文脈を与える技術
- ユーザの主体的な情報収集・発見を支援、世界的に類似技術なし