

# リサーチ & プランニング 第3回

## 検索の未来

デジタルハリウッド大学  
橋本大也

# 本日の内容

- 基本編に続く検索の応用編
- 検索技術の多様性とトレンドを理解する
- デスクトップ検索を使いこなす
- 検索とは何か、その未来を考える

# 世界の先端検索サービスを見る

- YAHOO!やGoogle以外にどんな検索エンジンがあるのか
- 検索技術の多様性と可能性を眺めてみる

# 検索エンジンの事例 1/2

- メタ検索と文書クラスタリング
- Vivisimo
- <http://www.vivisimo.com/>
- メタ検索結果をビジュアルな地図で表示
- Kartoo
- <http://www.kartoo.com/>
- 過去のWebを検索する
- Internet Wayback Machine
- <http://www.archive.org/>
- 徹底的にパーソナライズする検索エンジン
- SurfWax
- <http://www.surfwax.com/>
- A is Bに着目した検索エンジン
- Googlism
- <http://www.googlism.com/>
- 関連やアクセス数グラフがわかる
- Alexa
- <http://www.alexa.com/>
- 雑誌の記事を検索
- FindArticles.com
- <http://www.findarticles.com/PI/index.jhtml>
- ブログを検索
- Feedstar
- <http://www.feedster.com/>
- 質問の答えとなるWebを検索
- AskJeeves
- <http://www.ask.com/>
- オープンソースの検索サーバ
- htdig
- <http://www.htdig.org/>
- 1分前に更新されたページも検索
- GigaBlast
- <http://www.gigablast.com/>
- 使いやすさを追求するメタ検索
- Fazzle
- <http://www.fazzle.com/>

# 検索エンジンの事例 2/2

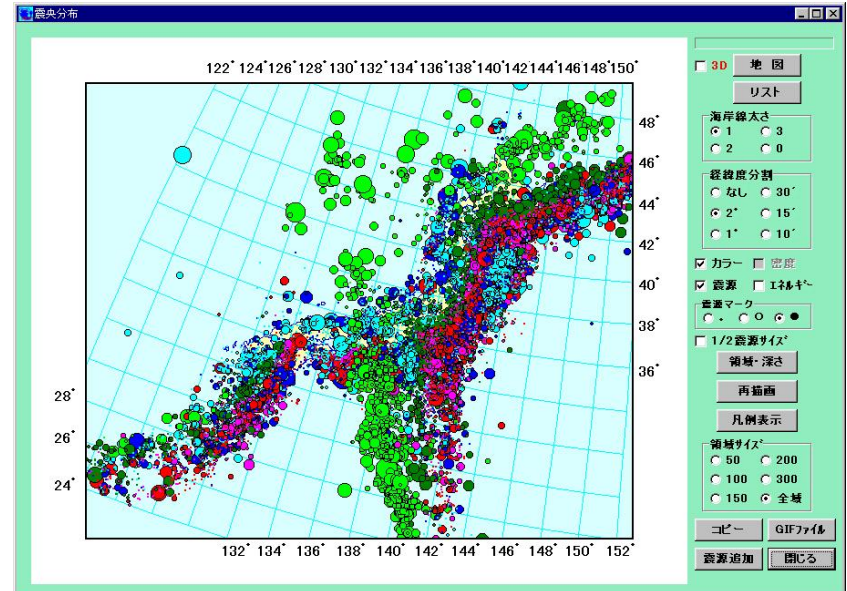
- 世界最大のショッピング比較検索
- MySimon
- <http://www.mysimon.com/>
- ラジオ番組のテキスト検索
- Speechbot
- <http://speechbot.research.compaq.com/>
- 数字データの分布の検索と可視化
- The Secret lives of Numbers
- <http://www.turbulence.org/Works/nums/>
- P2Pでインデックスを作成
- HyperBee
- <http://www.hyperbee.com/>
- ストリーム放送を検索
- SingingFish
- <http://www.singingfish.com/>
- PageRankアルゴリズムを拡張
- Teoma
- <http://www.teoma.com/>
- Googleによる商品検索
- Froogle
- <http://froogle.google.com/>
- リンクの関係をネットワーク図表示
- Anacubis
- <http://www.anacubis.com/>
- 検索結果にサムネイルを付加
- Girafa
- <http://www.girafa.com/vsearch/index.acr>
- お絵かきで形状を検索する
- 3d models search engine
- <http://shape.cs.princeton.edu/search.html>

# ユニークな検索技術5

キーワード検索ではない検索

# 1 地震情報の検索

- EQLIST
  - <http://www5b.biglobe.ne.jp/~t-kamada/CBuilder/eqlist.htm>
- 1885年以降のマグニチュード4以上の地震3万件を検索できるフリーソフト
- 地域と期間でリスト表示と地図表示



番号	地震番号	年月日	時分	緯度	経度	深さ	マグニチュード	震源地
1	416/08/23	00:00	0'00"	0'00"	0	0.0	0.0	遠飛島宮付近(大和)
2	599/05/28	00:00	0'00"	0'00"	0	7.0	7.0	大和
3	679/00/00	00:00	32'42"	130'24"	0	6.7	6.7	筑紫
4	684/11/29	00:00	32'30"	134'00"	0	8.4	8.4	土佐その他南海・東海・西海地方
5	701/05/12	00:00	35'36"	135'24"	0	7.0	7.0	丹波
6	715/07/04	00:00	35'06"	137'54"	0	6.4	6.4	湯江
7	715/07/05	00:00	34'42"	137'24"	0	6.7	6.7	三河
8	734/05/18	00:00	34'18"	136'06"	0	7.0	7.0	畿内・七道諸国
9	744/06/30	00:00	0'00"	0'00"	0	6.4	6.4	肥後
10	745/06/05	00:00	35'30"	136'36"	0	7.9	7.9	美濃
11	762/06/09	00:00	35'30"	137'18"	0	7.4	7.4	美濃・飛騨・信濃
12	799/09/18	00:00	0'00"	0'00"	0	0.0	0.0	茶陵

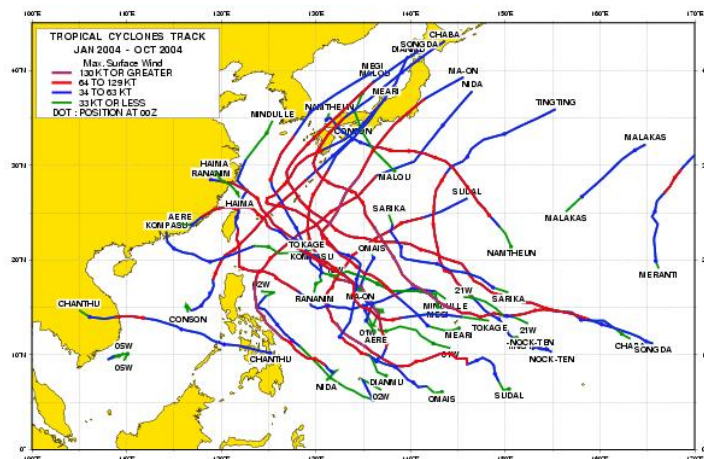
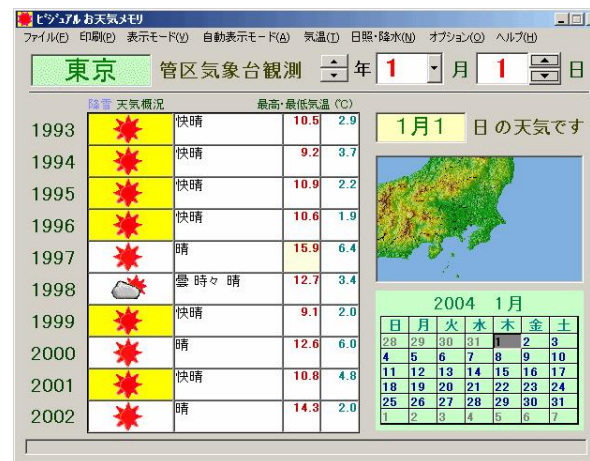
# 考察1 天変地異情報も個人で分析

- 13年間の気象庁データから天気を検索できる「ビジュアルお天気メモリ」があれば夏休みの絵日記も後書きOK

- <http://www.vector.co.jp/magazine/softnews/030322/n0303224.html>

- TRMM台風データベースで衛星データから過去の台風の規模と経路を一覧できる

- [http://www.eorc.jaxa.jp/TRMM/typhoon/index\\_j.htm](http://www.eorc.jaxa.jp/TRMM/typhoon/index_j.htm)



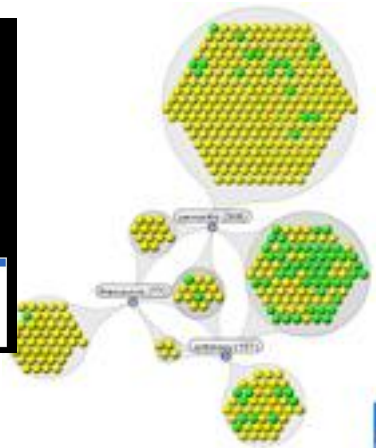
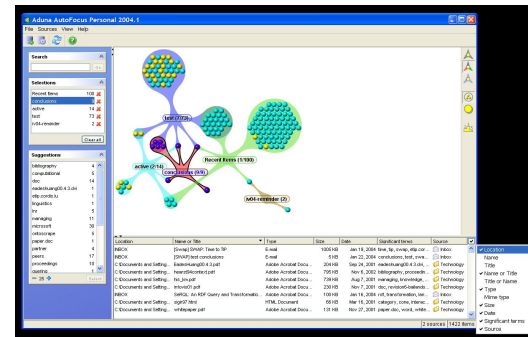
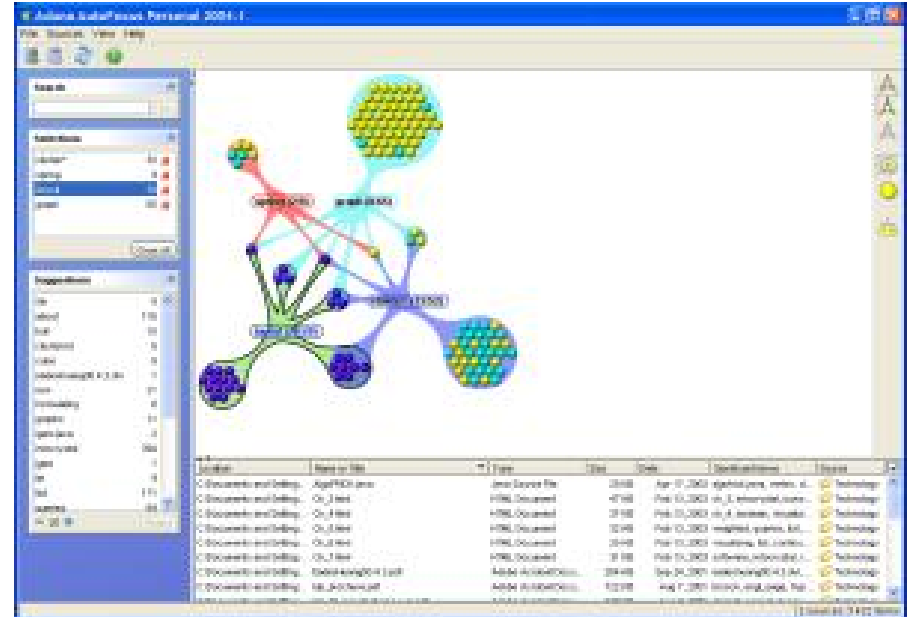


# 2 デスクトップ検索の未来系

## AdunautoFocusPersonal

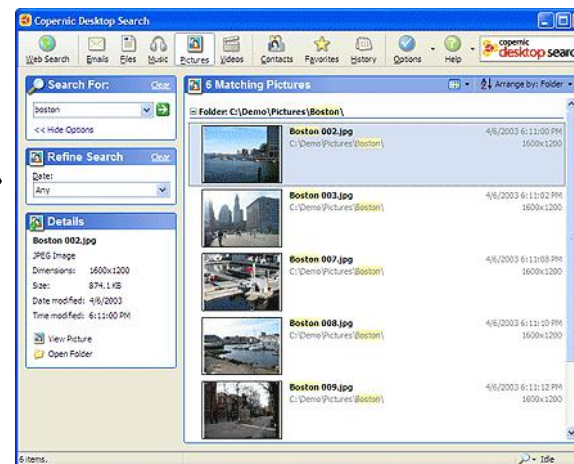
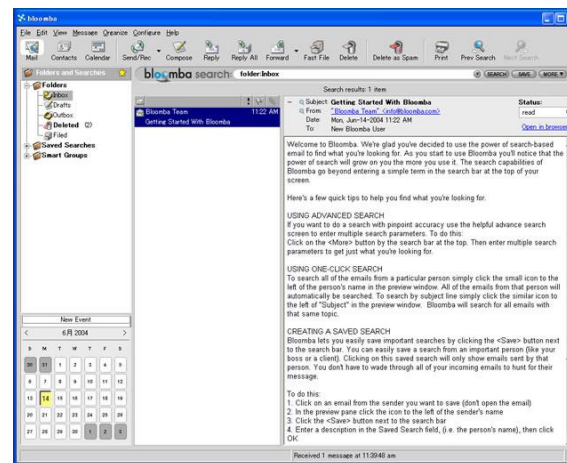
– <http://aduna.biz/products/autofocus/personal/index.html>

- デスクトップファイル、メール、Web履歴に含まれる単語でファイルを群れとして表示する。



# 考察2 仮想机上検索は机上の空論 ではなくなった

- YAHOO!は整理不要のメール検索ソフト Bloomba (Statalab) を買収したと10月21日発表、衝撃。聞いてみよう。
- GDS、Lookout (M\$ が買収)、Copernic (CNET Editorレビューで高評価) など、デスクトップの検索が技術トレンドに。



# 3 音楽の波形で検索

- 11月下旬に発売。アナログ音源の波形データを元にCDDDBから楽曲情報を取得する“MusicID”を搭載したCarryOn Music。3500円くらい。
- グレースノート(CDDDB)の音楽認識技術MusicIDを採用した最初の国内プロダクト。
  - [http://www.gracenote.com/gn\\_products/music\\_id.html](http://www.gracenote.com/gn_products/music_id.html)



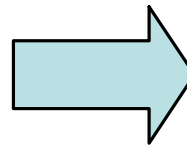
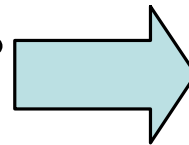
# 考察3 ノリで検索、ノリ検

- Shazam社の楽曲認識技術を使って、ボーダフォンで「携帯楽曲認識」がスタート。携帯にBGMを聞かせると楽曲情報&購入

– <http://www.shazam.com/uk/do/home>

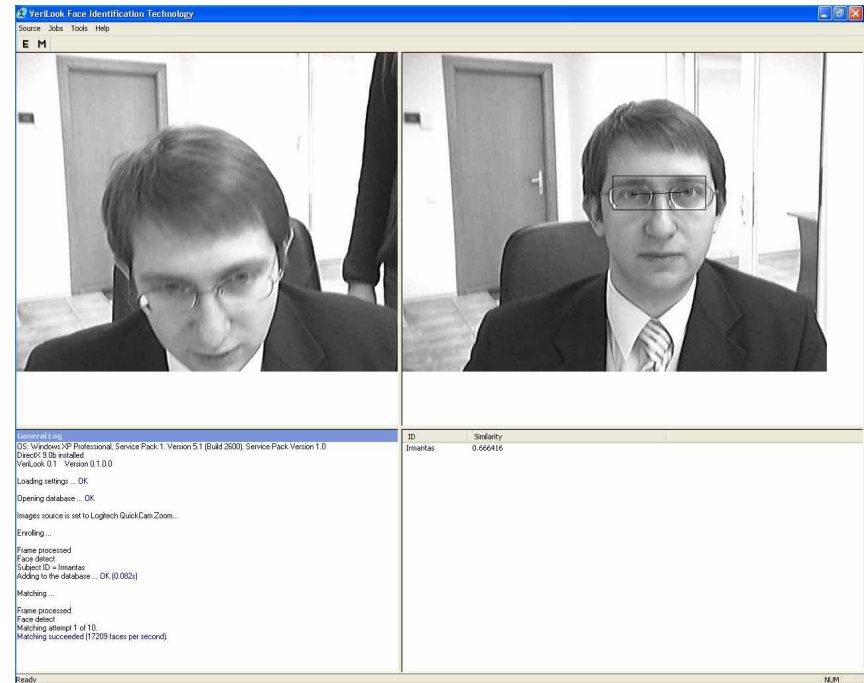
- 音楽の雰囲気 handheld のMP3を自動カテゴリ分類。“ロマンチックな曲”、“アップテンポな曲”で検索が可能になるデスクトップ音楽検索ソフトMoodLogic

– <http://www.moodlogic.com/>



# 4 顔で検索

- 顔で検索する技術例：  
Verilookは1秒間に27000人の顔を認識することができる。開発者向けのソフトがダウンロード可能。  
[http://www.neurotechnology.com/vl\\_sdk.html](http://www.neurotechnology.com/vl_sdk.html)
- Wired News - スーパーボウルで観客全員の顔がスキャンされていた  
<http://hotwired.goo.ne.jp/news/news/culture/story/20011228201.html>



プライバシーの危険もあるが写真アルバムの整理、テレビ自動録画に画期的な技術になりそうな予感がするので、**3へえ**あげます。

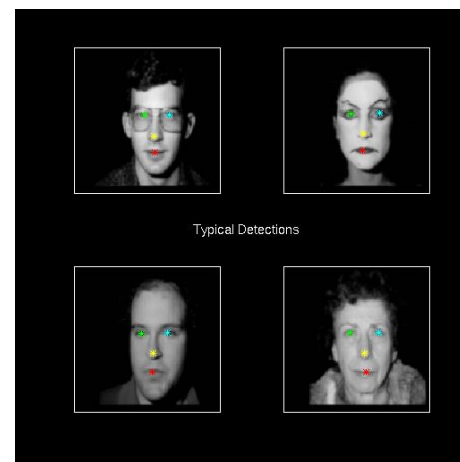
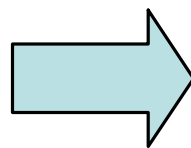
# 考察4 顔面検索の危険と可能性

- 写真の顔から人物を特定する検索技術大会 FaceRecognitionVendorTest (顔面認識業者試験！) がある(5へえ)

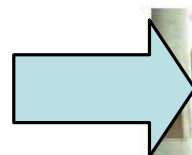
– <http://www.frvt.org/>

- デジタルドアマンArgusはドア上部に取り付けられたカメラで入室者を自動検出してしまふ監視技術。

– <http://www.computer.org/intelligent/ex2001/pdf/x2014.pdf>



<http://vismod.media.mit.edu/vismod/demos/facerec/system.html>



–製品情報

–[http://www.identix.com/products/pro\\_security\\_bnp\\_argus.html](http://www.identix.com/products/pro_security_bnp_argus.html)

# クリックで検索

- iPodとmp3とmusicは同じページで同時に使われる言葉(共起性)
- 検索語と一緒によく使われる言葉のリストが100個表示されるAeiwi
  - <http://www.aeiwi.com/>
- クリックで選ぶとその言葉と一緒によく使われるリストがまた100個芋づる式で入力要らず
- 何がキーワードか強調するアプリ。PopupPrism

yahoo Search

Don't panic ! Just read the [introduction](#).

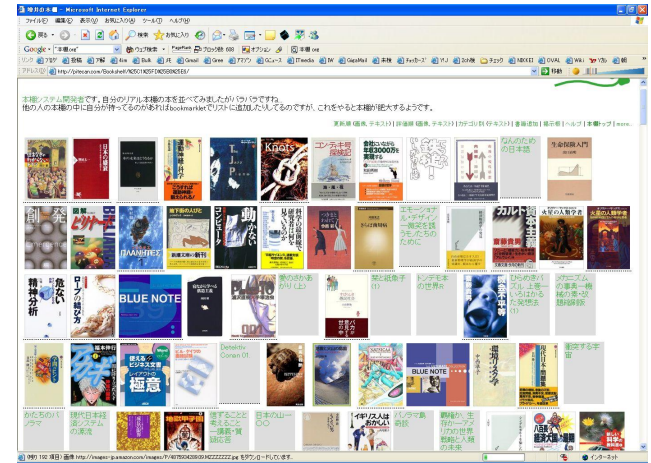
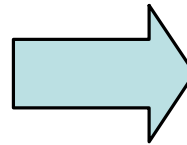
<a href="#">amateur</a> (1454)	<a href="#"> ficken</a> (1143)	<a href="#">masturbation</a> (1321)	<a href="#">service</a> (2058)
<a href="#">art</a> (1823)	<a href="#">forum</a> (2161)	<a href="#">minolta</a> (1429)	<a href="#">sex</a> (4495)
<a href="#">auto</a> (1183)	<a href="#">fotos</a> (1247)	<a href="#">money</a> (1687)	<a href="#">shop</a> (1897)
<a href="#">bilder</a> (2165)	<a href="#">free</a> (3062)	<a href="#">movies</a> (1532)	<a href="#">shopping</a> (1908)
<a href="#">binocular</a> (1368)	<a href="#">fuji</a> (1406)	<a href="#">mp3</a> (1484)	<a href="#">slr cameras</a> (1355)
<a href="#">binoculars</a> (1475)	<a href="#">games</a> (2746)	<a href="#">mp3 players</a> (1471)	<a href="#">software</a> (2986)
<a href="#">business</a> (2784)	<a href="#">gifts</a> (1705)	<a href="#">mp3 review</a> (1355)	<a href="#">sony</a> (1777)
<a href="#">camcorder review</a> (1355)	<a href="#">girls</a> (1889)	<a href="#">music</a> (2332)	<a href="#">sport</a> (1421)
<a href="#">camcorders</a> (1491)	<a href="#">gps</a> (1570)	<a href="#">news</a> (1618)	<a href="#">teen</a> (1834)
<a href="#">canada</a> (1144)	<a href="#">gps devices</a> (1356)	<a href="#">nikon</a> (1429)	<a href="#">teens</a> (2294)
<a href="#">canon</a> (1596)	<a href="#">gps review</a> (1356)	<a href="#">olympus</a> (1427)	<a href="#">telescope review</a> (1355)
<a href="#">cd</a> (1547)	<a href="#">gratis</a> (2254)	<a href="#">online</a> (3014)	<a href="#">telescopes</a> (1378)
<a href="#">chat</a> (1999)	<a href="#">health</a> (1740)	<a href="#">panties</a> (1732)	<a href="#">training</a> (1173)
<a href="#">cheap</a> (1151)	<a href="#">holiday</a> (1179)	<a href="#">pentax</a> (1384)	<a href="#">travel</a> (2230)
<a href="#">children</a> (1177)	<a href="#">home</a> (1929)	<a href="#">photo</a> (1190)	<a href="#">tv</a> (1597)
<a href="#">computer</a> (1679)	<a href="#">hosting</a> (1770)	<a href="#">photos</a> (1284)	<a href="#">uk</a> (1415)
<a href="#">design</a> (2634)	<a href="#">hotel</a> (2486)	<a href="#">php</a> (1169)	<a href="#">urlaub</a> (1750)
<a href="#">digital camera</a> (1501)	<a href="#">hotels</a> (1701)	<a href="#">pics</a> (1411)	<a href="#">video</a> (2756)
<a href="#">digital camera review</a> (1357)	<a href="#">information</a> (1200)	<a href="#">pictures</a> (1778)	<a href="#">videos</a> (1706)
<a href="#">digital cameras</a> (1593)	<a href="#">internet</a> (3883)	<a href="#">portable stereo</a> (1355)	<a href="#">voice recorders</a> (1360)
<a href="#">download</a> (1753)	<a href="#">ipod</a> (1417)	<a href="#">private</a> (1223)	<a href="#">web</a> (2327)
<a href="#">dvd</a> (2361)	<a href="#">kostenlos</a> (1596)	<a href="#">projectors</a> (1385)	<a href="#">web design</a> (1542)
<a href="#">email</a> (1159)	<a href="#">links</a> (1400)	<a href="#">real estate</a> (2074)	<a href="#">web hosting</a> (1188)
<a href="#">entertainment</a> (1345)	<a href="#">love</a> (1187)	<a href="#">review</a> (1526)	<a href="#">webdesign</a> (1205)
<a href="#">fetish</a> (1363)	<a href="#">marketing</a> (1914)	<a href="#">reviews</a> (1623)	<a href="#">website</a> (1594)

[Top Referees](#) [Add Link](#) [Aeiwi Info](#) [Feedback](#)

# 考察5 検索というより探索

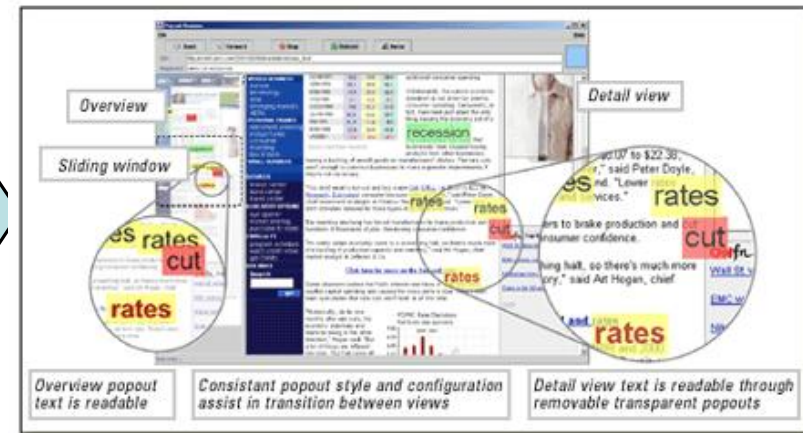
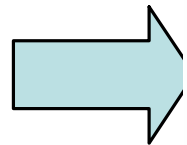
- 超便利な本棚.org。本の表紙でビジュアル検索する。自分の本棚をWebで簡単公開。

- <http://pitecan.com/Bookshelf/>



- 閲覧中のページの重要語を強調表示する PopoutPrism

- <http://www2.parc.com/csl/projects/popoutprism/default.html>

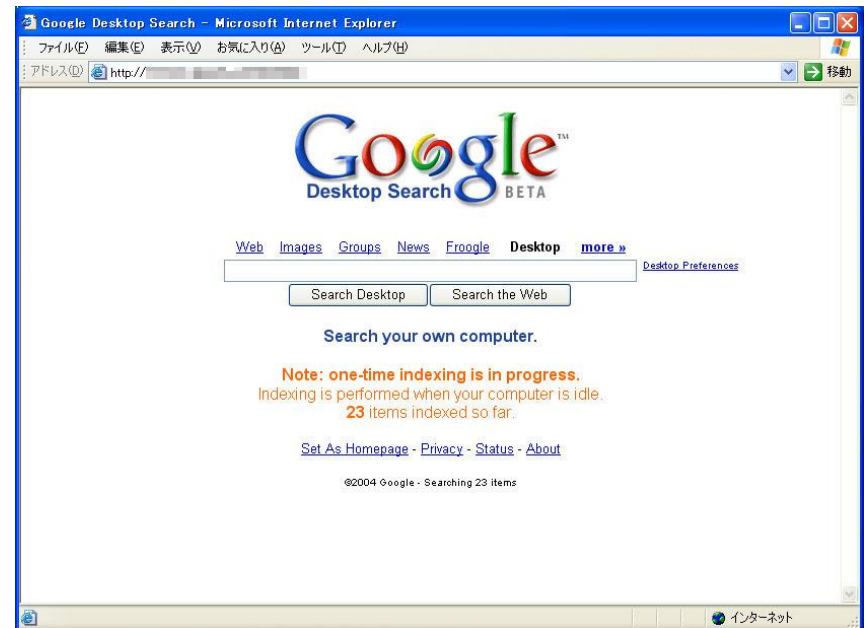




# デスクトップ検索を活用する

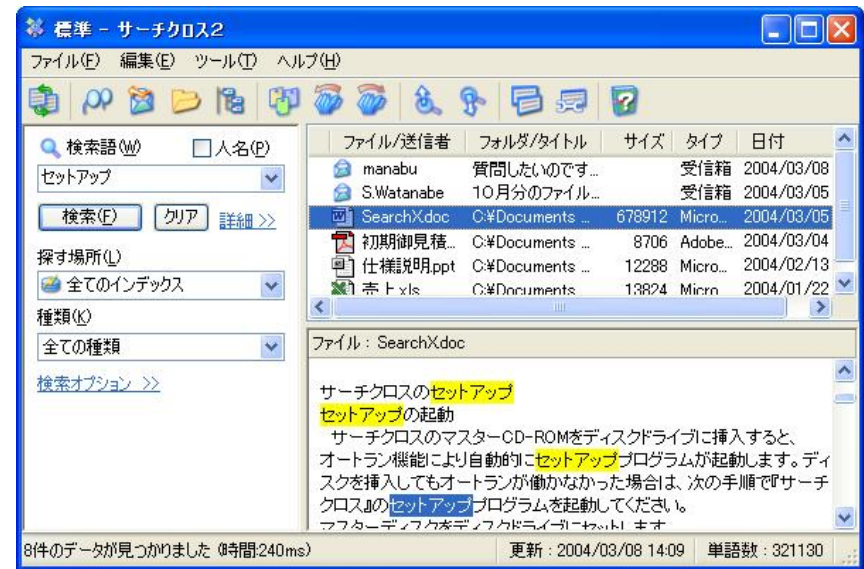
# Google Desktop Search

- Googleが2004年11月にリリース
- ローカルWebサーバ上で動作し、ブラウザで操作する
- Google Web検索と連動してリモートとローカルを横断検索
- オフィス文書、オフィス文書、メール(一部のメールソフト)、ブラウザーWeb履歴



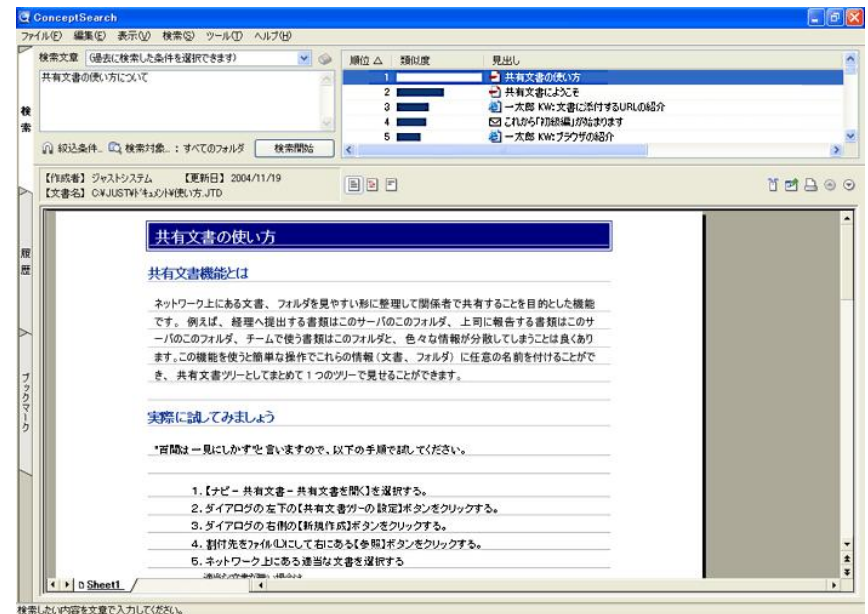
# サーチクロス (ビレッジセンター)

- ヴィレッジセンターが2000年から販売。専門化には高評価。3000円。
- 現時点で最も高機能で使いやすいと評価。
- 検索対象ファイルの種類、キーワードの切り出しの細かさなど秀逸。
- 多数のメールソフトや2ちゃんねるログなどにも対応している。
- デスクトップ検索の機能ベンチマーク



# ConceptSearch (ジャストシステム)

- 高度な自然言語処理によるデスクトップ検索
- 2005年3月リリース予定。1ライセンス5000円。
- 技術的には国内最高レベルのインデクス、アルゴリズムを保有。
- 概念で検索する



# 検索の未来のキーワード

- インデックスとアルゴリズムの時代から
- インタフェースとコミュニケーションへ
- 情報を「引き出す」技術が次世代検索の鍵

# 検索＝情報を引き出す技術

- 情報をどこから引き出すか？
    - 1 過去から
      - 記録 紙、画像、音声、映像、Web、石版
- インデックスとアルゴリズム

これに加えてこれからは、

- 2 現在から
  - 想起 記憶、推論、連想、発想
- 3 未来から
  - 創発 対話、対立、予測

インタフェースとコミュニケーション

# インタフェースとコミュニケーション

- インタフェースとは
  - 自分の中から引き出す技術
- コミュニケーションとは
  - 他者の中から引き出す技術
- ページランクを超えて
- 検索力 =
  - アルゴリズム × インタフェース × コミュニケーション

# まとめ

- 世界にはさまざまな検索技術と応用サービスがあること
- 技術のトレンドを事例から把握した
- 特にデスクトップ検索は実用レベル、使いこなそう
  
- 検索とは「引き出す」技術である
  - 過去: 記録から引き出す
  - 現在: 記憶から引き出す(思い出す)
  - 未来: 創造する(対話や創発)
  
- コミュニケーションは一種の検索になる
- インタフェースも検索の効率を左右する要素
  
- 新しい検索サービスを積極的に使いこなす