



# 第19回統合DBミーティング

---

Integrated MTG in NIBIO

# 本日の予定

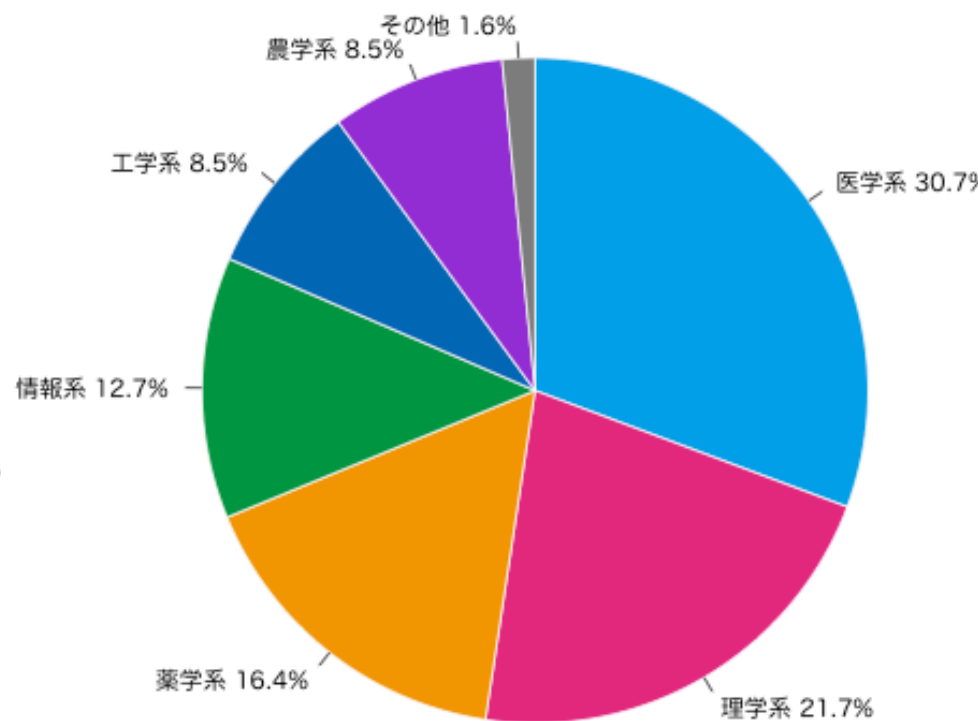
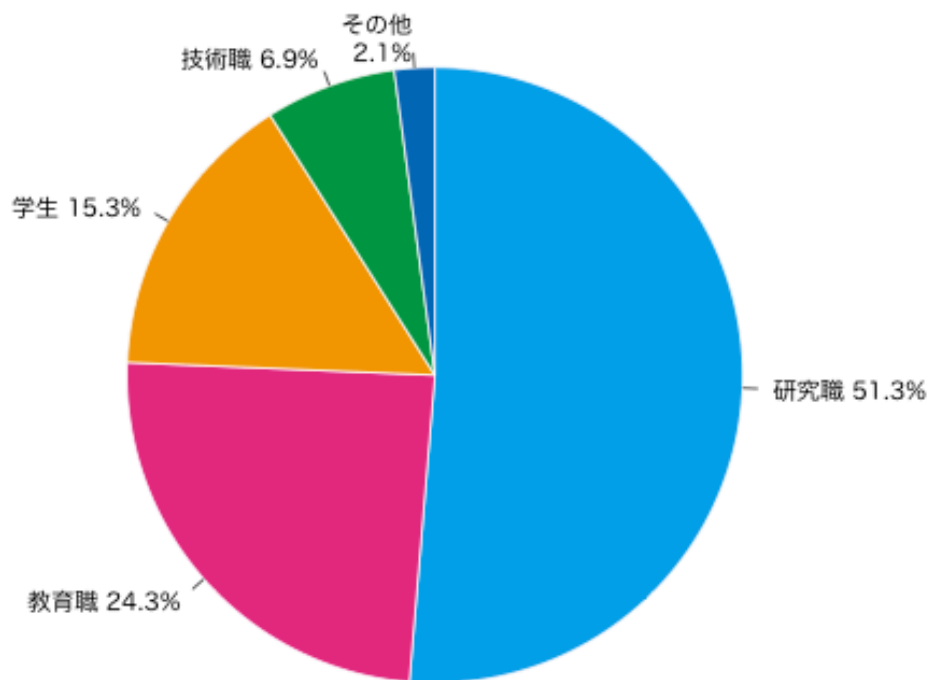
- アンケートの集計結果
- 今年度の予定
- アクセス解析

# 職業・専門

概ねSagaceの想定ユーザー

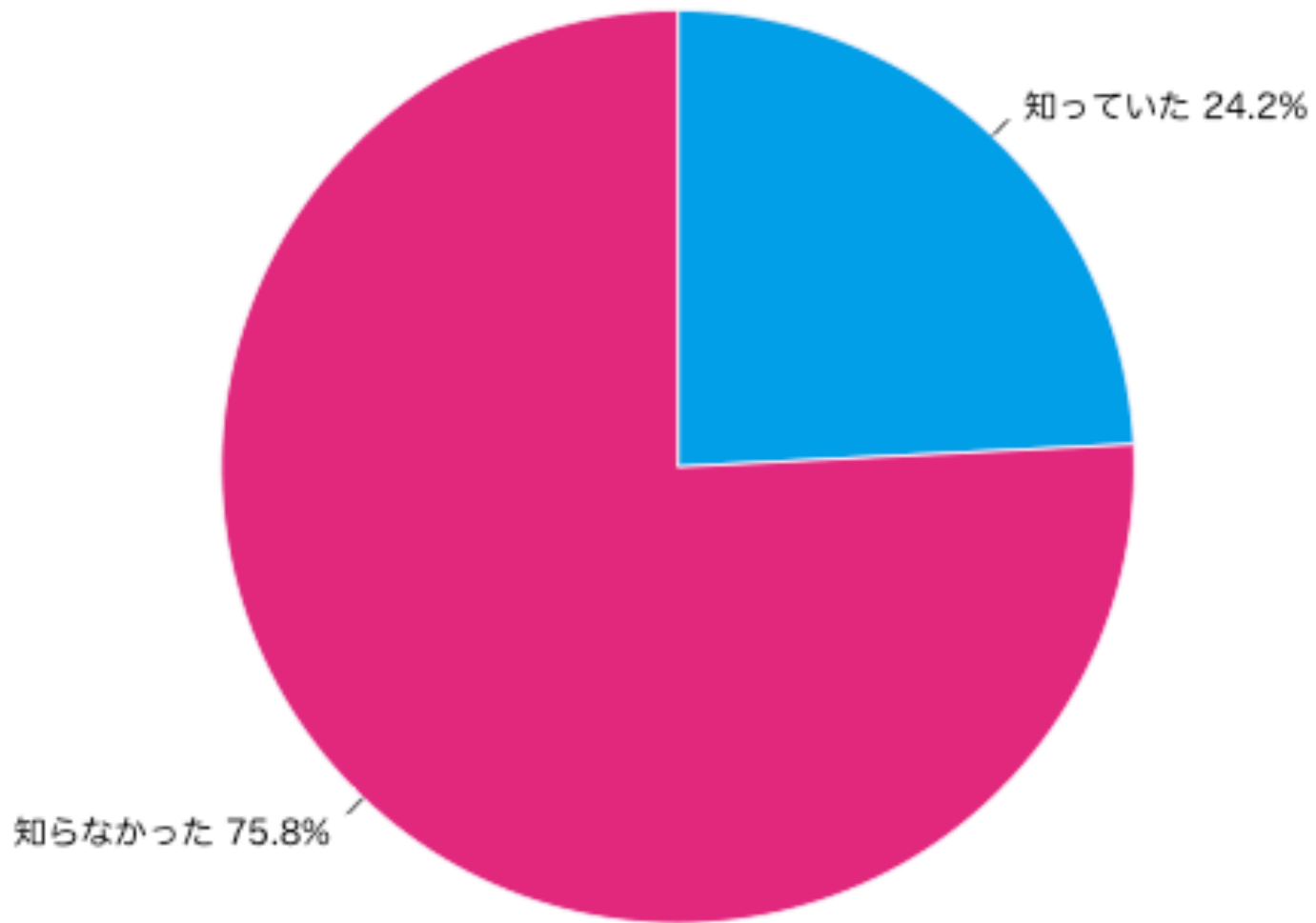
専門

職業

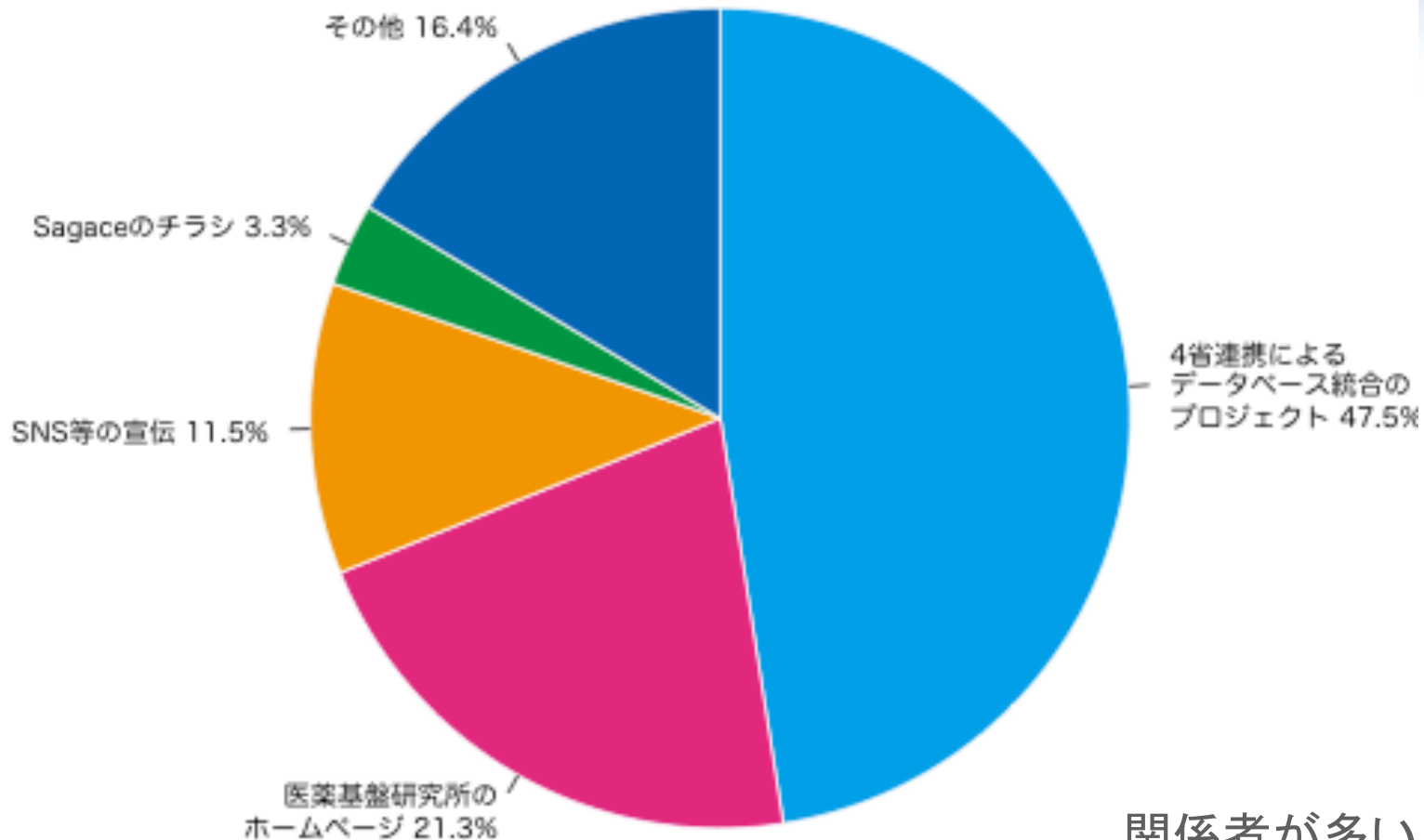


# Sagaceの知名度

まだまだ宣伝が必要



# Sagaceを知った理由



関係者が多い。  
見る人が多いページにリンクがあると良い？

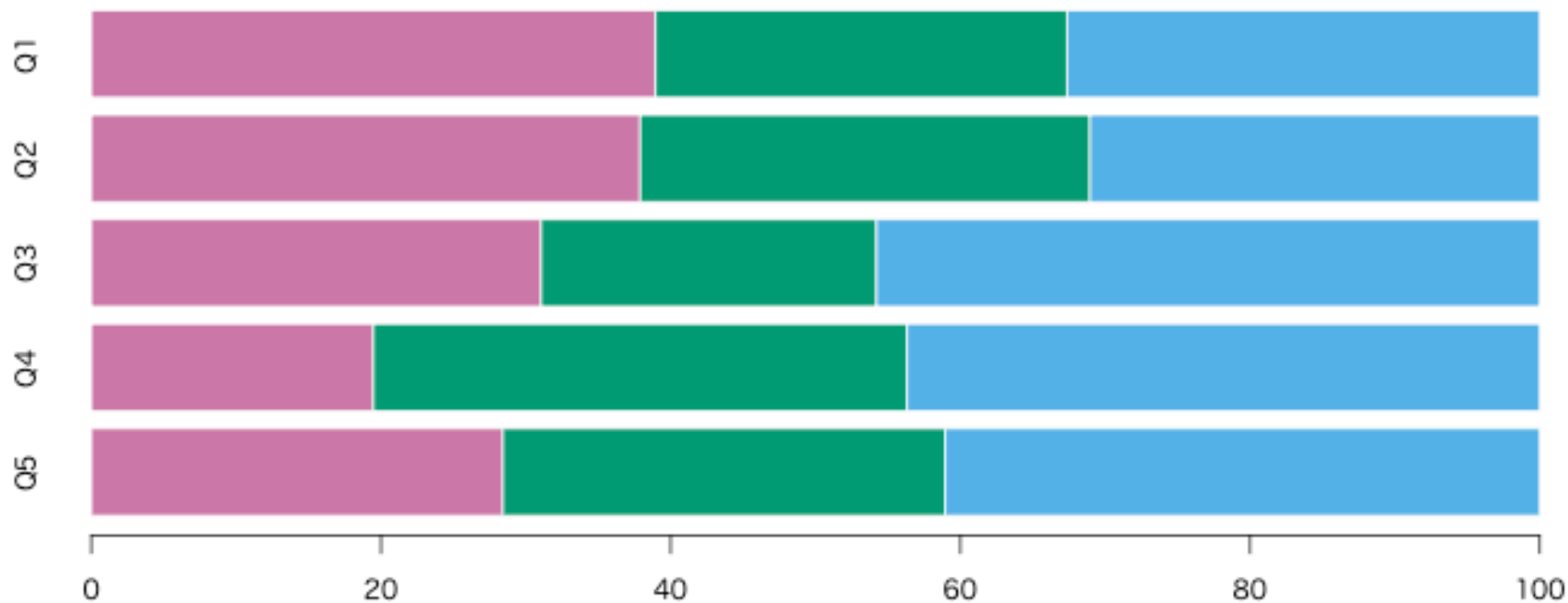
# 満足度について

- Q1. 検索速度
  - Q2. 欲しい情報が得られたかどうか
  - Q3. ファセットによる絞り込み機能の使い心地
  - Q4. ファセットの分類項目は適切か
  - Q5. リッチスニペットによる絞り込み機能の使い心地
- 
- 5段階評価 … 使いやすい (5,4,3,2,1) 使いにくい
  - ・不満(1,2),普通(3),満足(4,5)の割合を%で示してあります。



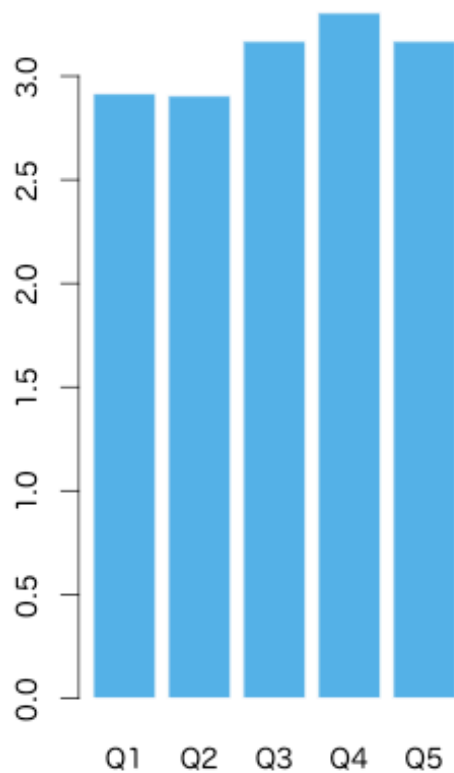
# 滿足度

滿足度 <不滿(1,2),普通(3),滿足(4,5)>

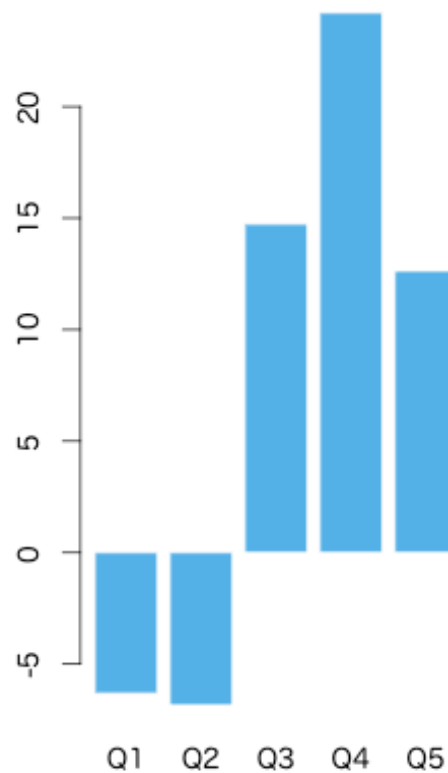


# 質問ごとの比較

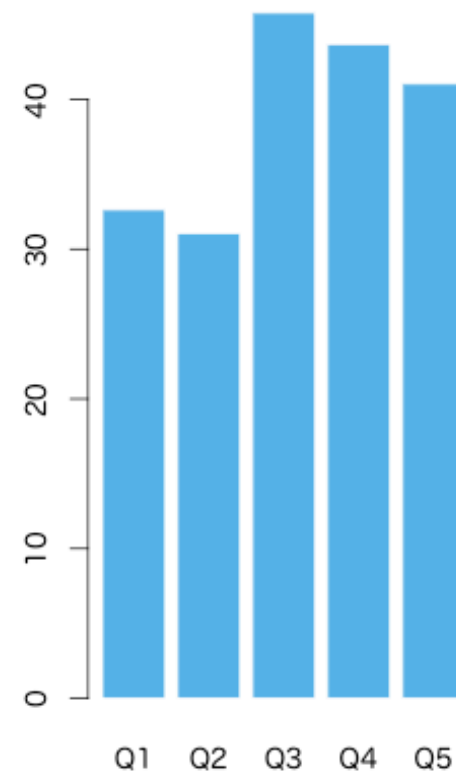
平均値



良い(%)-悪い(%)



満足度(%)

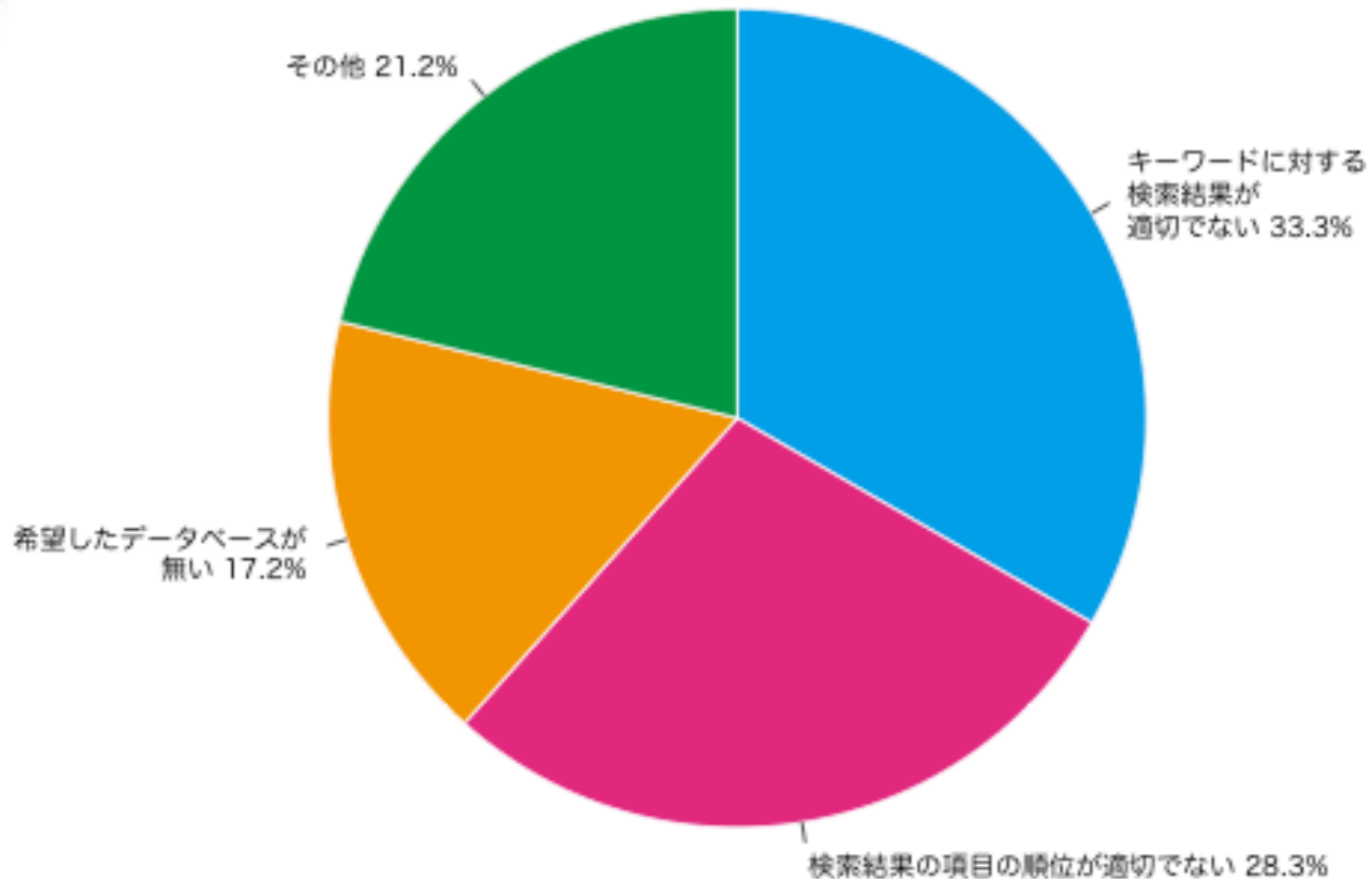




# 満足度のグラフより

- 基本的な速度, 精度に不満を持っている人が多い。
- ファセットの絞り込み機能, 分類項目は概ね好評
- リッチスニペットはファセットより評価がやや落ちる。

# Sagaceで検索して欲しい情報が 得られない理由



# Sagaceで検索して欲しい情報が得られない理由

- 適切でない検索結果, 順位に対する不満, 欲しいデータベースの不足と続く。
- 対策
  - 独自に新規追加するDBのスニペットの精査
  - 連携機関のインデックスをSagaceに追加する際にはより厳しく事前にチェック
  - ログを使った順位の変動
  - 新規検索エンジンの開拓

# Sagaceで検索して欲しい情報が 得られない理由

- その他の意見より
  - 特許が多い(3)
  - 同義語が引っかからない
- 対策
  - 最初のページに出る特定のDBからのヒット件数や、エントリが似ているものを削除する
  - 同義語は増強中

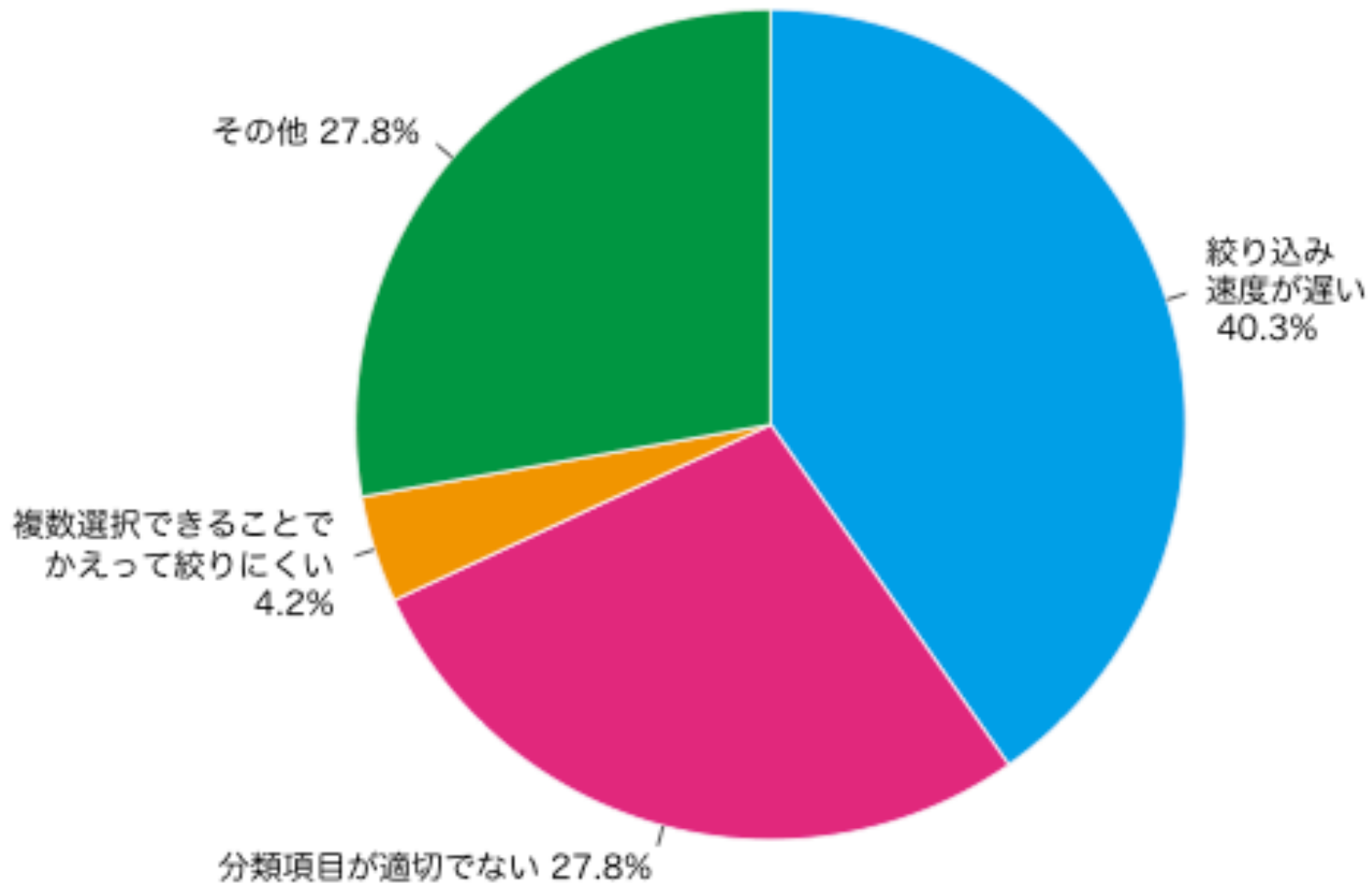
# 検索対象DBの増強について

- 多様な要望
- 他の検索サービスのリクエストもある
- 有料のものは難しい
- 来月以降に候補と絞り込みを行う

- 医薬品 (2)
  - メディカルオンライン, 中医協の記録, 臨床試験のDB, 副作用情報, 添付文書情報
- 論文 (2)
  - J-Stageの英語論文, 医中誌, JGlobal関連サービス, CiNii(2)
- 医薬関連の法律
  - 薬事法, 医師法, 化審法など
- 研究費(KAKEN)
- 研究資源
  - 疾患モデル動物のリソースデータベース
- ミューテーション, アセチル化・メチル化など、DNA、タンパク質の化学修飾などの情報
- 厚生労働科学研究成果データベース
- 理研のデータベース



# ファセット機能が使いにくい理由



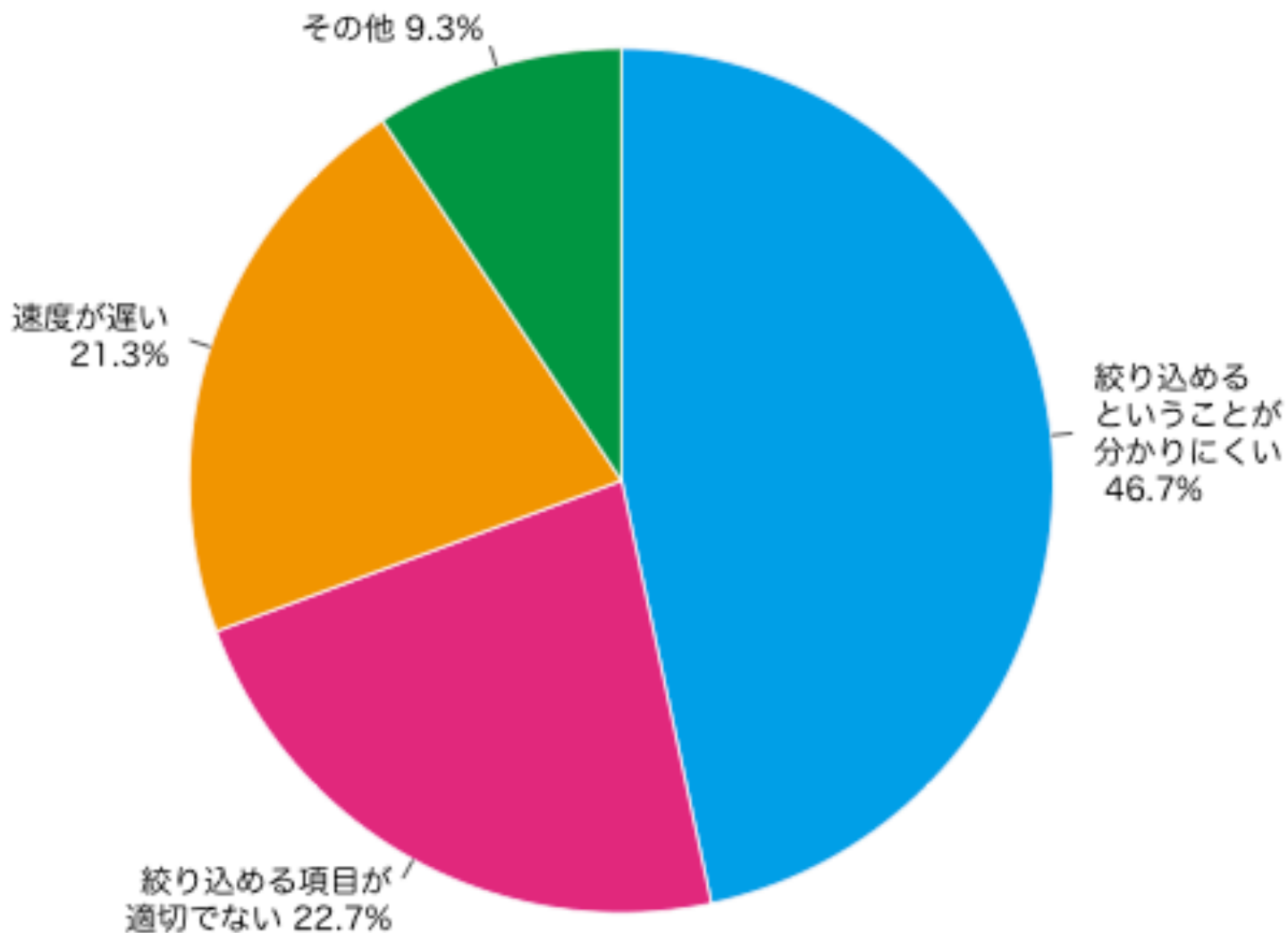
# ファセット機能が使いにくい理由

- 速度に関する不満が多い。
- ファセット分類項目自体への不満, 複数選択への不満と続く。
- 対策
  - より効率的に速度を出せる方法を模索する。
  - HE以外の検索エンジンで解決できるかテストする。

# ファセット機能が使いにくい理由

- その他より抜粋
- 医薬的な分類項目が欲しい(2)
  - DBを増やした後に対応を検討
- 検索前に絞り込みたい, 複数選択チェックしたい, 解除が分かりにくい
  - ラジオボタンの検討

# リッチスニペットが使いにくい理由



# リッチスニペットが使いにくい理由

- 絞り込めるということが分かりにくい。
- 対策
  - 画像ロゴ (Species, 更新日など) もう少しアピールをした実装をしてテストをする。

# リッチスニペットの希望

例: 病名, 文献ID, 細胞の画像, データ作成日, データ作成者の名前

回答	件数(2件以上)
文献の ID (PMID含む)	37
病名, 疾患名	35
データ作成日	33
データ作成者の名前	14
データ更新日	12
画像	11
細胞の画像	7
文献情報(著者名と発表年月など)	6
データ作成者の機関名、組織名	3
遺伝子名	3
化合物名	3
細胞種(例 HEK293)	2
試薬番号 ID	2
組織画像	2
著者	2
薬剤名	2

文献, データ作成者,  
画像の実装を検討

\* 実装済



# 今年度の計画

- 横断検索システムの拡張
  - 検索インデックスの拡張
  - 検索ユーザビリティの改善
  - メタデータの利活用
- 新規データ統合へ向けた調査と研究
  - 厚労省関連データベースの調査
  - 既存のデータベースのRDF化と拡張

# 横断検索システム

- インデックスの拡張
  - 医薬品に関するDBの追加
    - 文書系のDBを優先, アンケートにも考慮
- ユーザビリティの改善
  - 速度
  - 精度
    - 新規追加のDBは精度を高める
    - DBを追加する際には個々のエントリまで厳しくチェック
    - 同義語の追加

# 今後の予定

- メタデータの普及と利活用
  - Schema.orgの拡張
  - メタデータを埋め込んだDBを増やす
  - 埋め込み先のDBで恩恵を受けられる機能
  - 文献と画像のメタデータの対応
  - RDFaクローラーと実装とSagaceへの反映
- 新規検索システムの提案
  - HEに代わる検索システムの調査とテスト

# 新規データ統合へ向けた調査と研究

- 厚労省関連のDBの調査と英語化
- Toxygatesの拡張
- 基盤研内部のデータ(希少疾病用医薬品・希少疾病用医療機器開発支援品目, 実験動物バンク)のRDF化

	4	5	6	7	8	9	10	11	12	1	2	3
<b>インデックスの拡張</b>												
医薬に関するDBを追加 (実験プロトコル, 生物資源)												
速度の改善												
同義語の追加												
UIの改善												
<b>メタデータの普及と利活用</b>												
Schema.orgの拡張												
メタデータが対象となるDBを増やす												
新規検索システムの提案												
RDFaクローラーの実装とSagaceへの対応												
<b>厚労省関連のDBの調査</b>												
カタログ作成												
<b>新規データ統合へ向けた調査と研究</b>												
Toxygatesの拡張												
基盤研内部のDBのRDF化												

- ・6月23日(日)～28日(金) BioHackathon2013 @ 東京
- ・10月5日(土)
- ・12月3日(火)～6日(金) 分子生物学会 @神戸

# アクセス解析

## トラフィック サマリー

2013/03/25 - 2013/04/24

比較: 2013/02/22 - 2013/03/24

アドバンス セグメント | メール エクスポート ▾ マイレポートに追加 ショートカット

訪問数の割合の変化: +0.00%

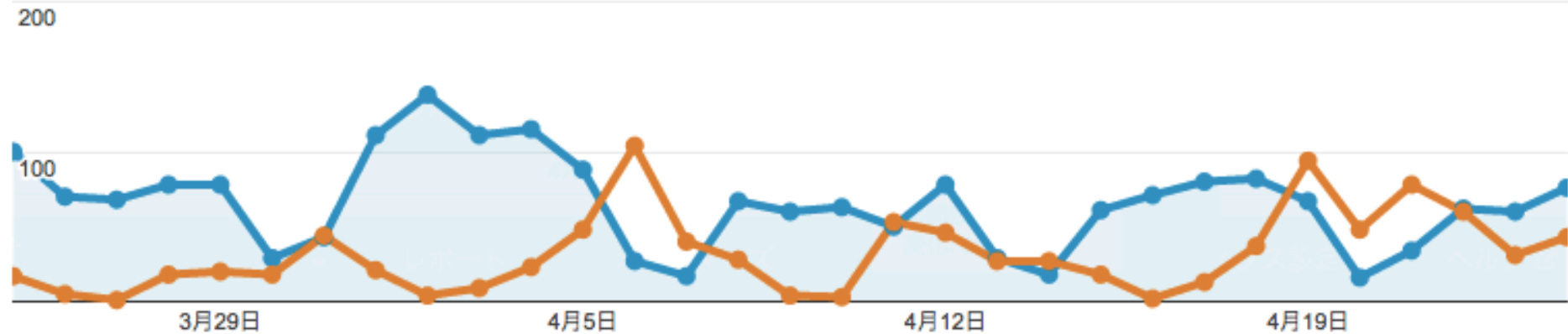
増加傾向

サマリー

時間別 日 週 月

2013/03/25 - 2013/04/24: ● 訪問数

2013/02/22 - 2013/03/24: ● 訪問数





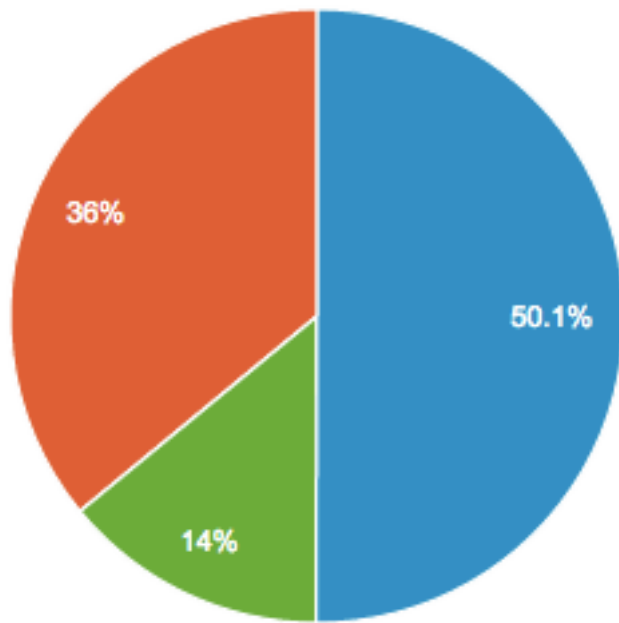
# アクセス解析

検索サイト

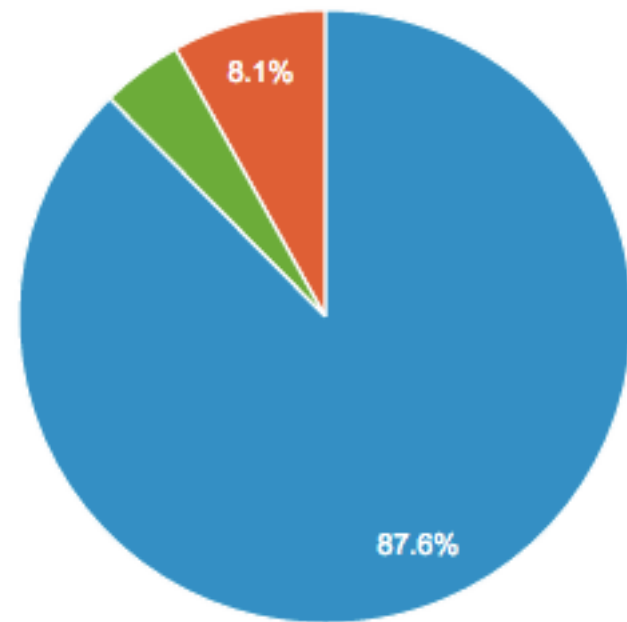
貼られたURL

■ 検索トラフィック ■ 参照トラフィック ■ ノーリファラー





2013/02/22 - 2013/03/24



2013/03/25 - 2013/04/24



# アクセス解析-参照トラフィック

参照元	訪問数	訪問数 (%)
1. <a href="#">google</a>		
2013/03/25 - 2013/04/24	1,006	 57.22%
2013/02/22 - 2013/03/24	256	 66.15%
変化率	<b>292.97%</b>	<b>-13.49%</b>
2. <a href="#">yahoo</a>		
2013/03/25 - 2013/04/24	739	 42.04%
2013/02/22 - 2013/03/24	128	 33.07%
変化率	<b>477.34%</b>	<b>27.09%</b>
3. <a href="#">bing</a>		
2013/03/25 - 2013/04/24	11	0.63%
2013/02/22 - 2013/03/24	0	0.00%

# アクセス解析-検索エンジン参照元

名古屋大の廣明秀一さんが講義資料のリンク先で紹介

<input type="checkbox"/> 参照元	訪問数	↓	訪問別ページビュー	訪問時の平均滞在時間	新規訪問の割合
<input type="checkbox"/> 1. <a href="http://presat-vector.org">presat-vector.org</a>					
2013/03/25 - 2013/04/24	25		5.60	00:06:17	92.00%
2013/02/22 - 2013/03/24	0		0.00	00:00:00	0.00%
変化率	∞%		∞%	∞%	∞%
<input type="checkbox"/> 2. <a href="http://integbio.jp">integbio.jp</a>					
2013/03/25 - 2013/04/24	9		4.78	00:02:52	88.89%
2013/02/22 - 2013/03/24	3		2.67	00:00:14	100.00%
変化率	200.00%		79.17%	1,155.28%	-11.11%
<input type="checkbox"/> 3. <a href="http://192.168.1.7">192.168.1.7</a>					
2013/03/25 - 2013/04/24	6		1.67	00:00:09	0.00%
2013/02/22 - 2013/03/24	16		1.81	00:00:55	12.50%
変化率	-62.50%		-8.05%	-84.35%	-100.00%

# 次回

- 5月24日(金)?
- 31日(金)?
  
- 新規追加DBについて
- (連携機関, Sagaceへの追加)