

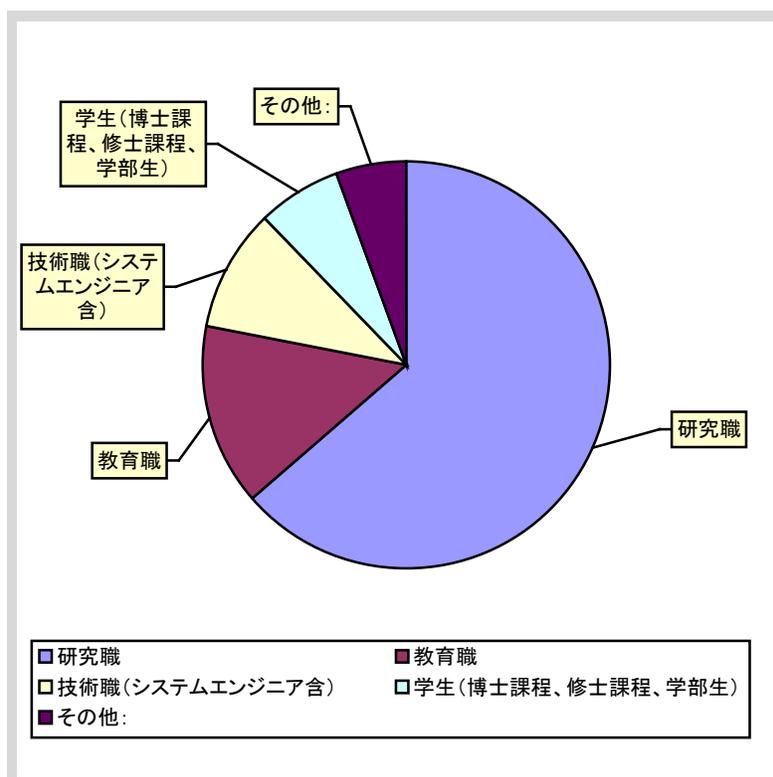
国内の医学／生命科学データと生物資源をさがす

## Sagace

アンケート集計結果

2014/04/21

Q1. 現在の職業について、最も近いものを1つお選びください。

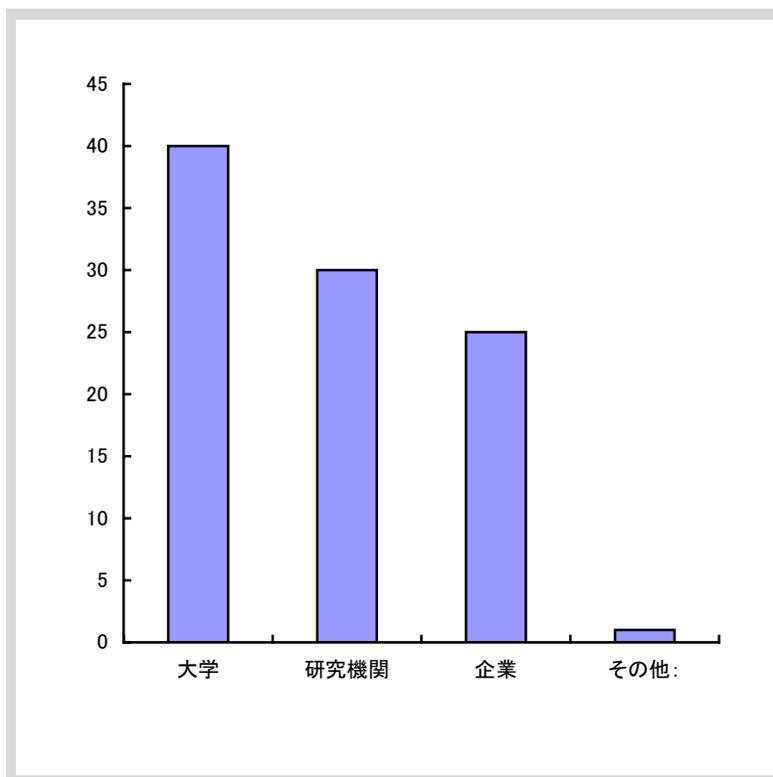


回答	件数	割合
研究職	58	63.7%
教育職	13	14.3%
技術職(システムエンジニア含)	9	9.9%
学生(博士課程、修士課程、学部生)	6	6.6%
その他:	5	5.5%

その他 内訳

回答	件数
研究員	1
開発営業	1
テクニカルサポート	1

Q2. 現在の所属機関の種別をお選びください。(複数回答可)

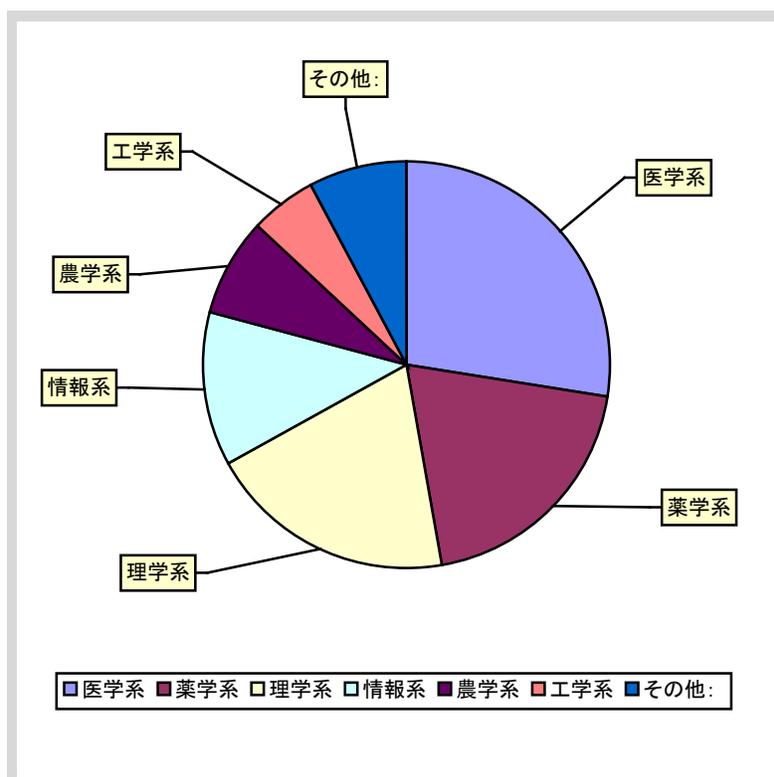


回答	件数
大学	40
研究機関	30
企業	25
その他:	1

その他 内訳

回答	件数
独立行政法人	1

Q3. 現在の専門分野について、最も近いものを1つお選びください。

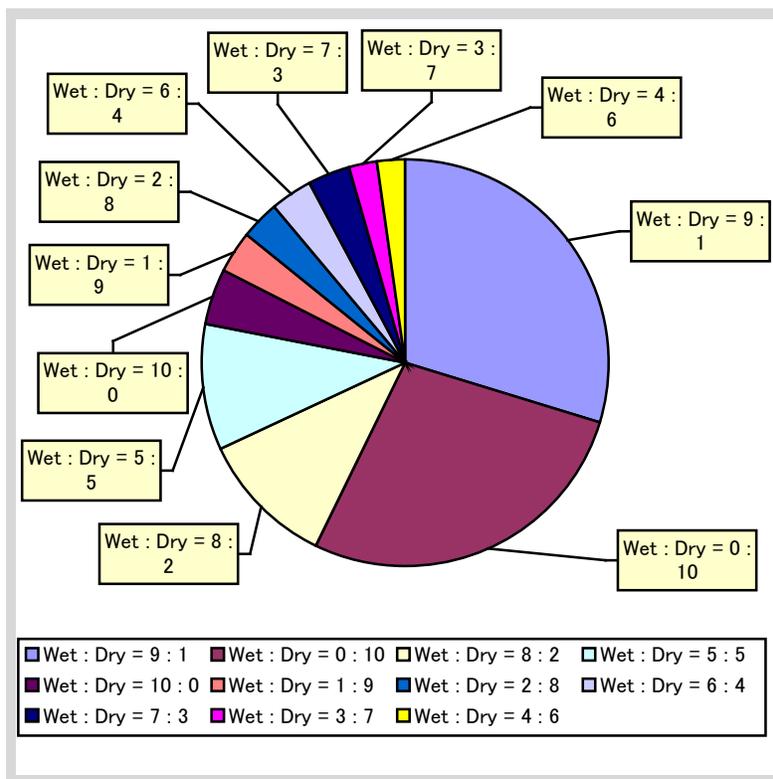


回答	件数	割合
医学系	25	27.5%
薬学系	18	19.8%
理学系	18	19.8%
情報系	11	12.1%
農学系	7	7.7%
工学系	5	5.5%
その他:	7	7.7%

その他 内訳

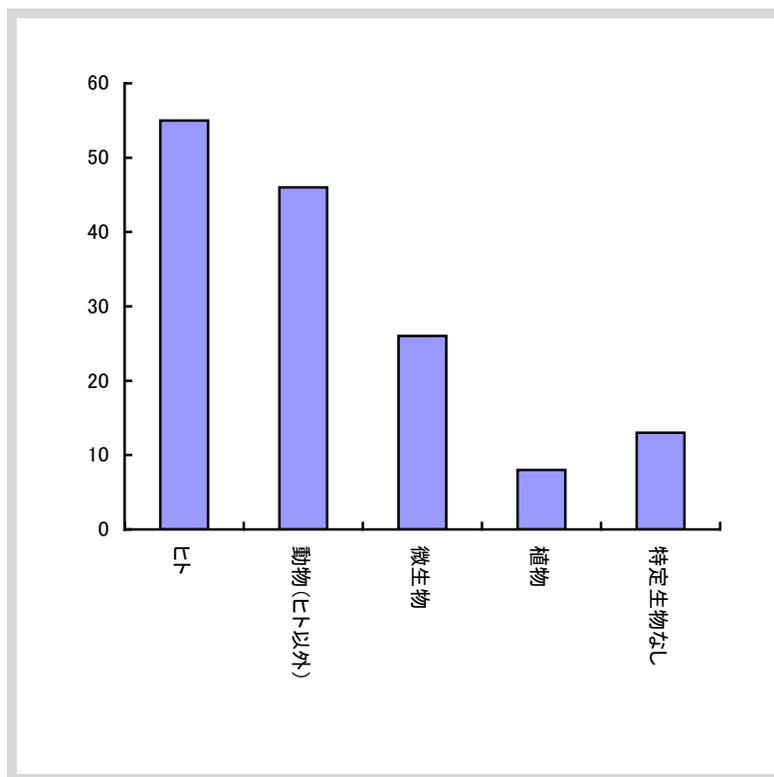
回答	件数
電気系	1
歯学系	1
バイオ全般(農・理・薬・医)	1
生命工学系	1
統計	1

Q4. あなたの研究において実験 (Wet) とデータベース解析・ツール利用の割合 (Dry)はどれくらいでしょうか。最も当てはまるものをお選びください。



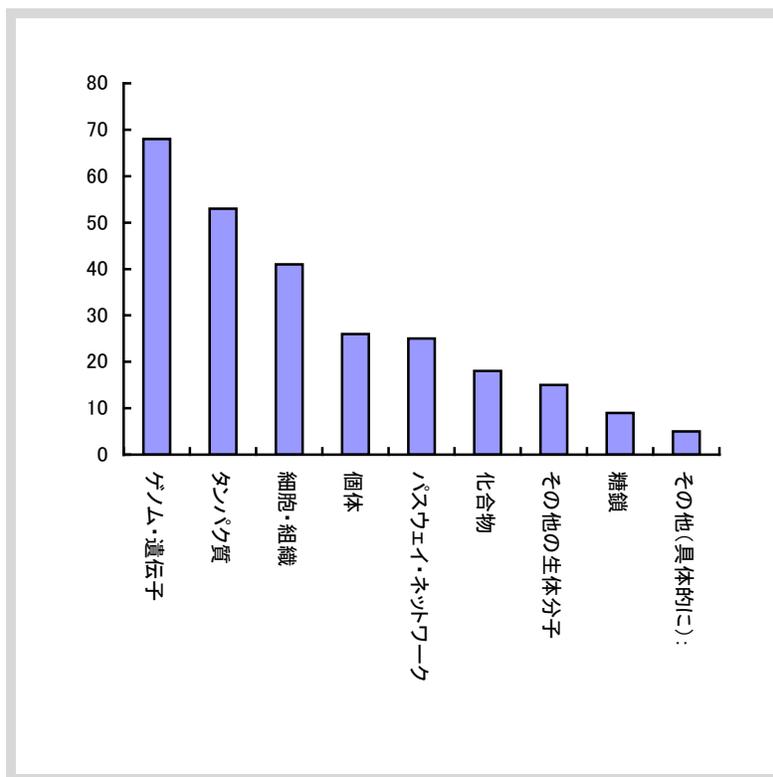
回答	件数	割合
Wet : Dry = 9 : 1	27	29.7%
Wet : Dry = 0 : 10	25	27.5%
Wet : Dry = 8 : 2	10	11.0%
Wet : Dry = 5 : 5	9	9.9%
Wet : Dry = 10 : 0	4	4.4%
Wet : Dry = 1 : 9	3	3.3%
Wet : Dry = 2 : 8	3	3.3%
Wet : Dry = 6 : 4	3	3.3%
Wet : Dry = 7 : 3	3	3.3%
Wet : Dry = 3 : 7	2	2.2%
Wet : Dry = 4 : 6	2	2.2%

Q5. あなたの研究で主に扱っている対象の生物を以下の項目の中からお選びください。（複数回答可）



回答	件数
ヒト	55
動物(ヒト以外)	46
微生物	26
植物	8
特定生物なし	13

Q6. あなたの研究テーマに該当する分野を以下の項目の中からお選びください。  
(複数回答可)

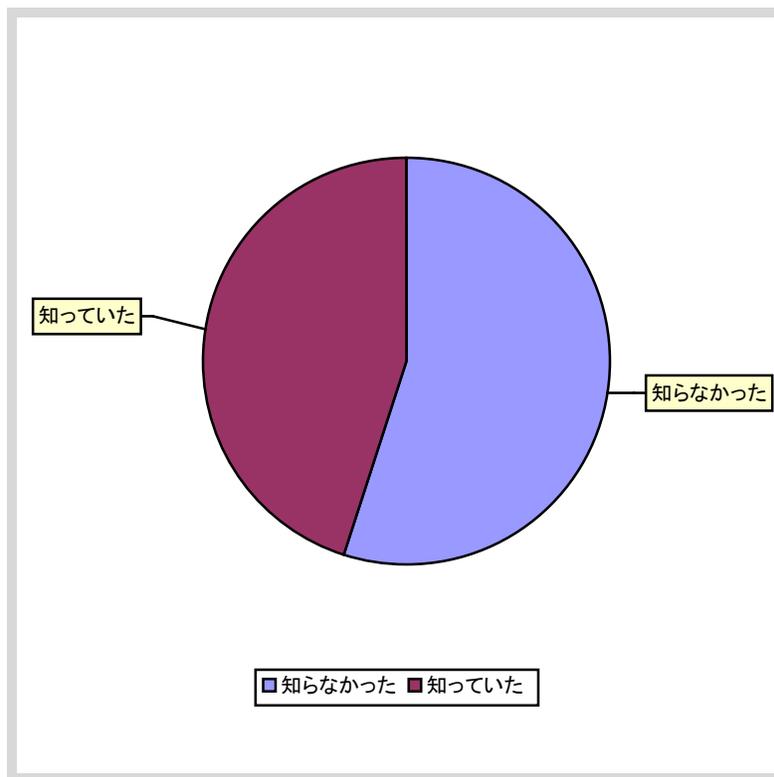


回答	件数
ゲノム・遺伝子	68
タンパク質	53
細胞・組織	41
個体	26
パスウェイ・ネットワーク	25
化合物	18
その他の生体分子	15
糖鎖	9
その他(具体的に):	5

その他 内訳

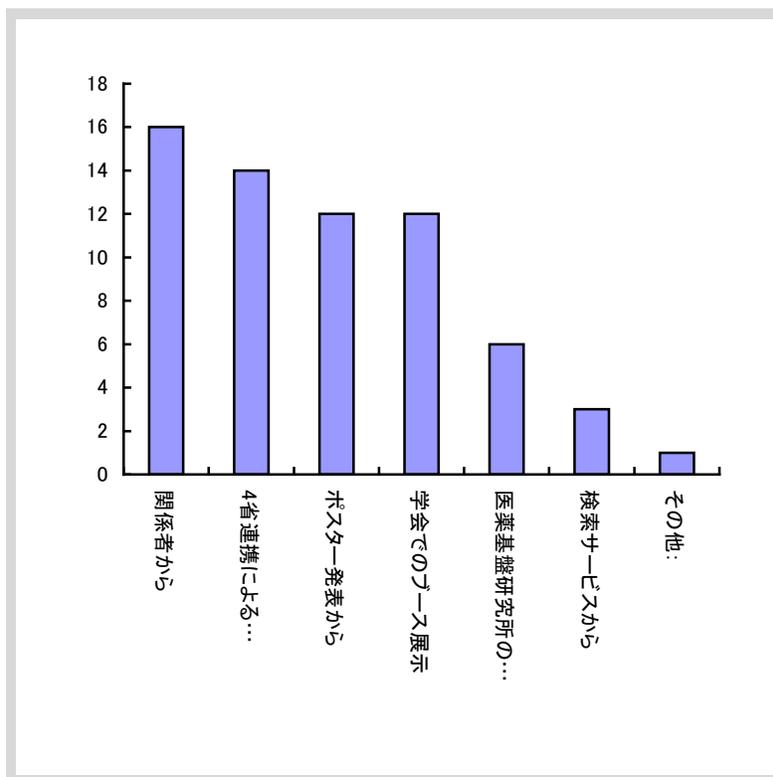
回答	件数
エピジェネティクス	1
特許	1
細菌	1
教育	1
社会医学	1

Q7. Sagaceについてご存知でしたか



回答	件数	割合
知らなかった	50	54.9%
知っていた	41	45.1%

Q8. "知っていた"とお答えの方にお聞きします。sagaceを何で知りましたか。  
(複数回答可)

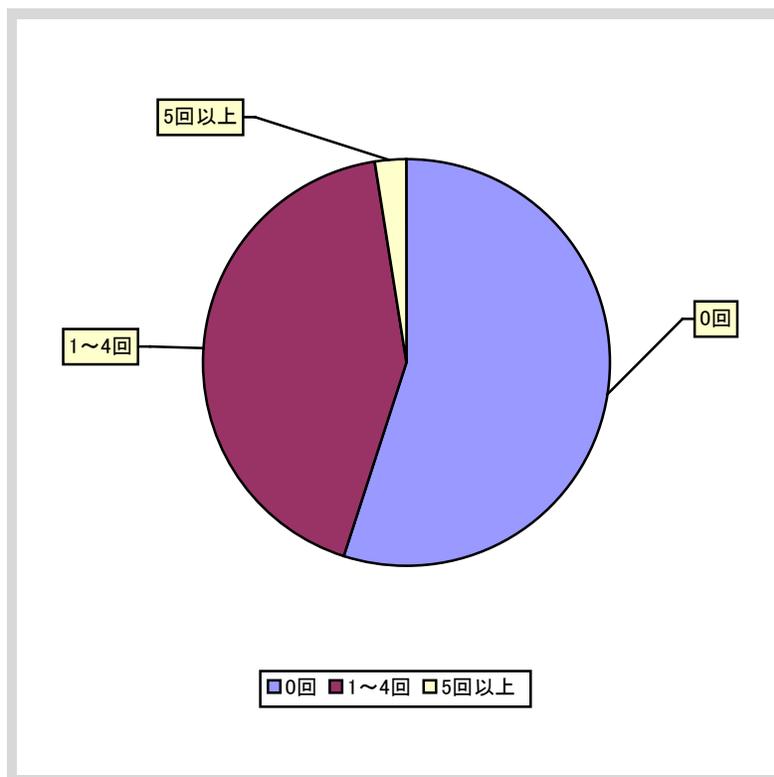


回答	件数
関係者から	16
4省連携によるデータベース統合のプロジェクト	14
ポスター発表から	12
学会でのブース展示	12
医薬基盤研究所のウェブページから	6
検索サービス(Google、Yahoo!など)から	3
その他:	1

その他 内訳

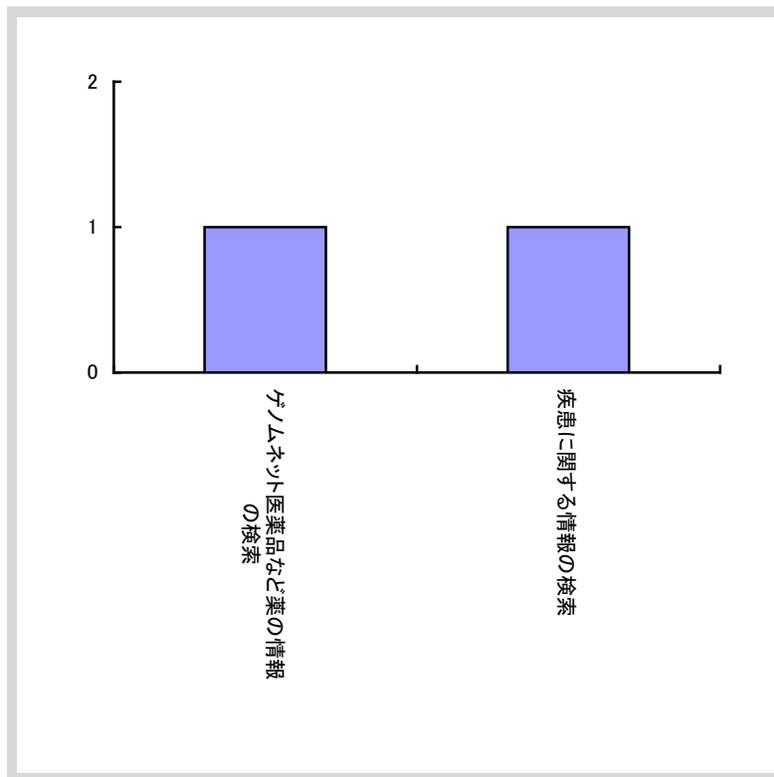
回答	件数
Twitter	1

Q9. "知っていた"とお答えの方にお聞きします。sagaceを過去1ヶ月以内に何回使ったことがありますか。



回答	件数	割合
0回	22	55.0%
1~4回	17	42.5%
5回以上	1	2.5%

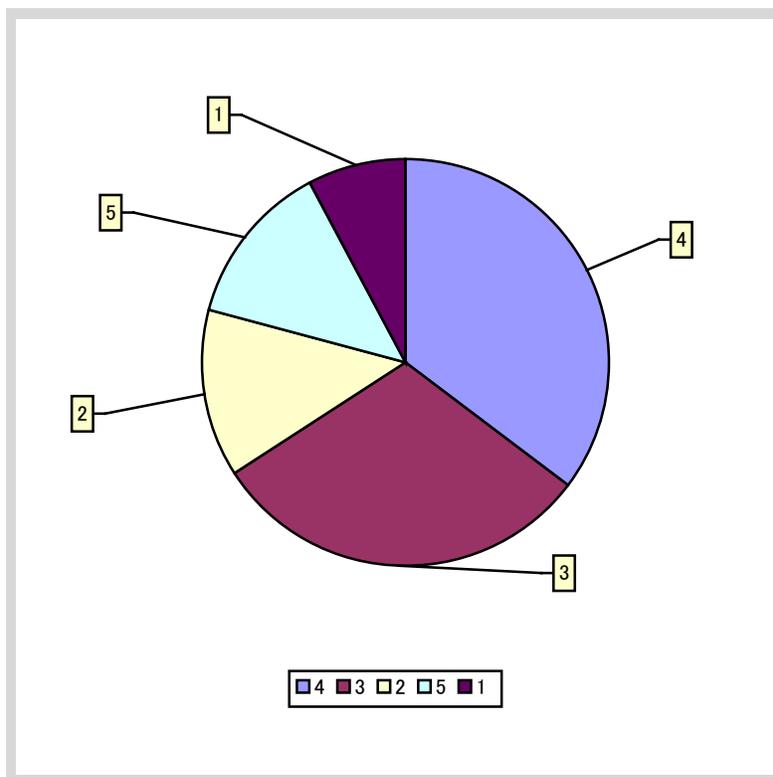
Q10. 上記の質問で、“5回以上”とお答えの方にお聞きします。Sagaceをどのような時に使いますか。（複数回答可）



回答	件数
ゲノムネット医薬品など薬の情報の検索	1
疾患に関する情報の検索	1

### Q11. 検索速度について

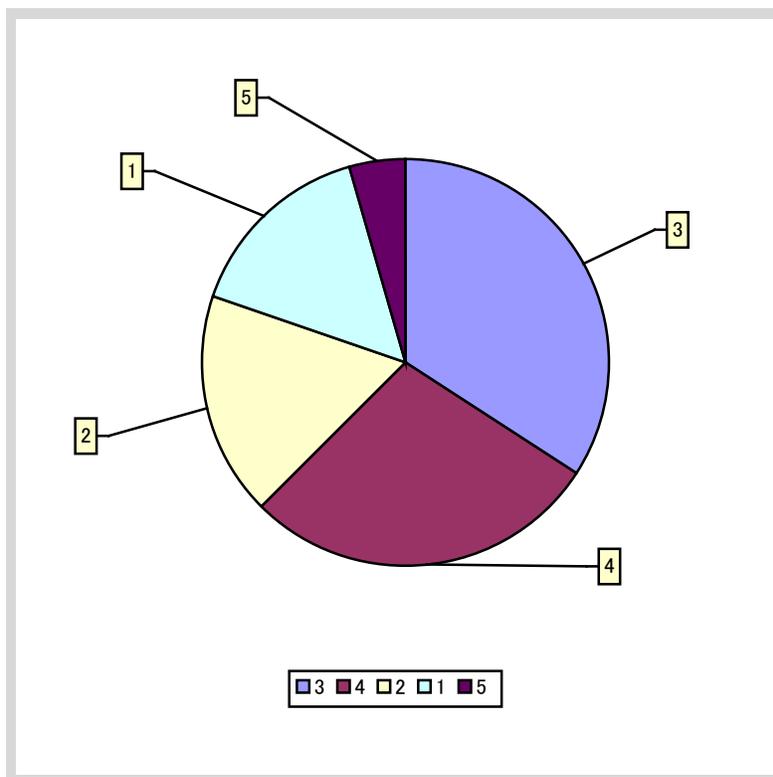
速い 5 4 3 2 1 遅い



回答	件数	割合
1	7	7.7%
2	12	13.2%
3	28	30.8%
4	32	35.2%
5	12	13.2%

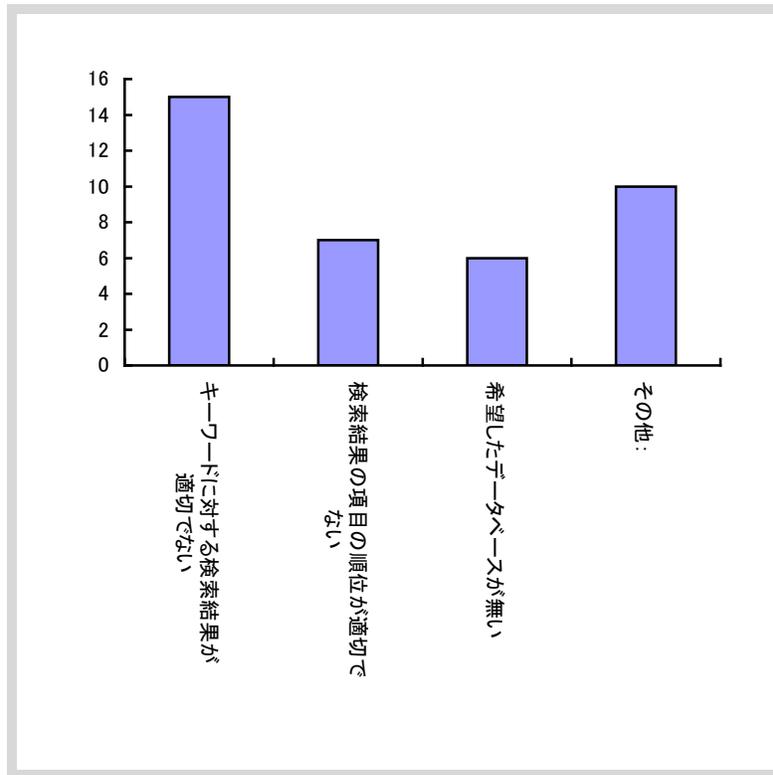
Q12. 欲しい情報が得られた。

得られた 5 4 3 2 1 得られない



回答	件数	割合
1	14	15.4%
2	16	17.6%
3	31	34.1%
4	26	28.6%
5	4	4.4%

Q13. 得られなかった (+, ++) とお答えした方に伺います。どういったところが適切でないと感じられたでしょうか。以下の項目からお選びください。(複数回答可)



回答	件数
キーワードに対する検索結果が適切でない	15
検索結果の項目の順位が適切でない	7
希望したデータベースが無い	6
その他:	10

その他 内訳

回答	件数
データベース選択が難しい	1
ヒット数97で「クリックすると一致する情報は見つかりませんでした」	1
データベースの選択欄が見にくかった	1
データベースの分類が分かりにくい	1
概念について検索し、その概念について関連したページの表示を期待したが、単にそのキーワードが含まれているだけの概念そのものとは関係の薄いページが表示された。	1
絞り込み方法が十分でない	1
データベースの選択が必要なところ	1
検索結果のリストが見づらい。タイトルが“Search Result”だったりよくわかりません。	1
どのデータベースに行けばほしい情報が得られるのか知っているのなら直接そのデータベースに行くはずなので、なぜ最初にデータベースを選択させるのかよくわからない。また、各データベースがどんな領域に詳しいかを名前だけで判断するのは限界がある。	1

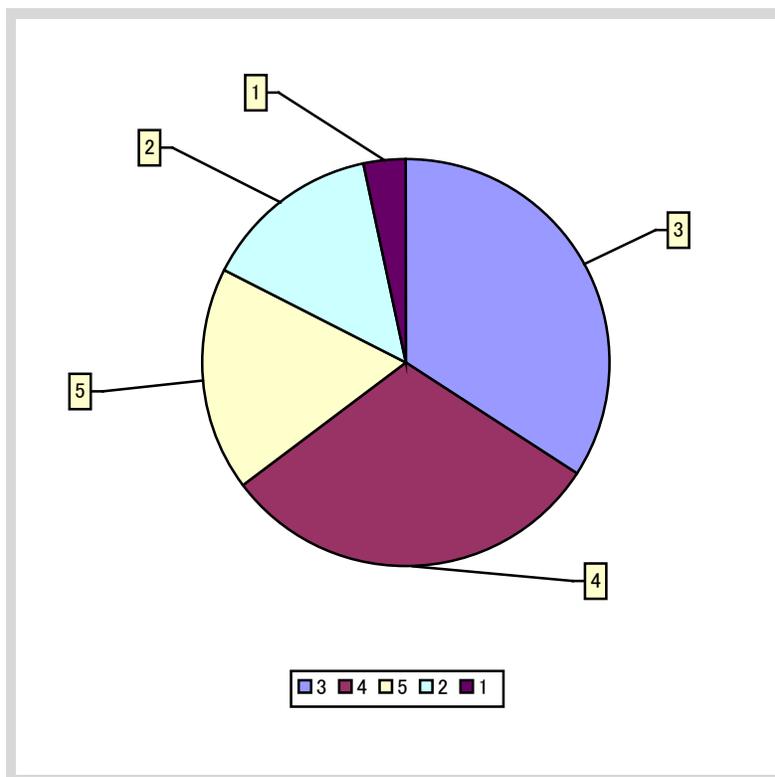
Q14. sagaceでは検索対象となるデータベースの増強をしたいと考えています。  
 検索対象として欲しいデータベースがありましたら、教えてください。

回答	件数
特になし	3
博士論文書誌データベース	1
addgene, The Human Gene Mutation Database	1
Biomodels	1
BioPortal	1
DrugBank, ChEMBL, Drug Data (FDB)	1
Methodからの文献検索	1
MGI	1
Nextprot <a href="http://www.nextprot.org/db/">http://www.nextprot.org/db/</a> また、各タンパク質がもつmotif情報がデータベースごとに異なるので、それをまとめたような情報が得られるとうれしい。	1
PDB	1
UCSC genome browser, 電子版がアクセスできない学会予稿集、 国際会議 アブスト(特にCSHL関係)	1
いまいち何ができるかわからない・・・	1
エピゲノム	1
ケモタキシス 細胞株 セルライン	1
コクランライブラリーのようなシステマティックレビューがあれば利用したいです。	1
セルラインに関するデータベース(遺伝子発現その他)	1
タンパク質の機能	1
デフォルトで検索結果が国内に限られていたように思いますが、国内DBのみ限定で検索したいシチュエーション自体が稀なのではないかと思えます。このためマンパワーがかかるとは思いますが、海外DBも含めた方が良いでしょうと思います。	1
ヒト遺伝疾患DB(Orphanetや、権利上むずかしいかもしれませんが、OMIM)やマウス表現型(MGIなど)	1
科研費データベース	1
海外のpatent情報(USpatent、中国など)	1

回答	件数
学会、セミナー、特許、和雑誌	1
学会情報は検索しにくいので有効	1
厚生労働科学研究成果データベース	1
酵素等に対する阻害剤や活性化剤のDB。ただし、DB増強よりも精度向上の方が必要。	1
国際出願された特許公報も検索できたら嬉しいです。	1
市販の抗体情報	1
疾患との関連情報, 薬との関連情報	1
特にありません。 現在登録されているDBについて、最新のVersionに対応していればよいと思います。	1
日化辞Web	1
論文も直接サーチ出来るようにしてもらいたい	1

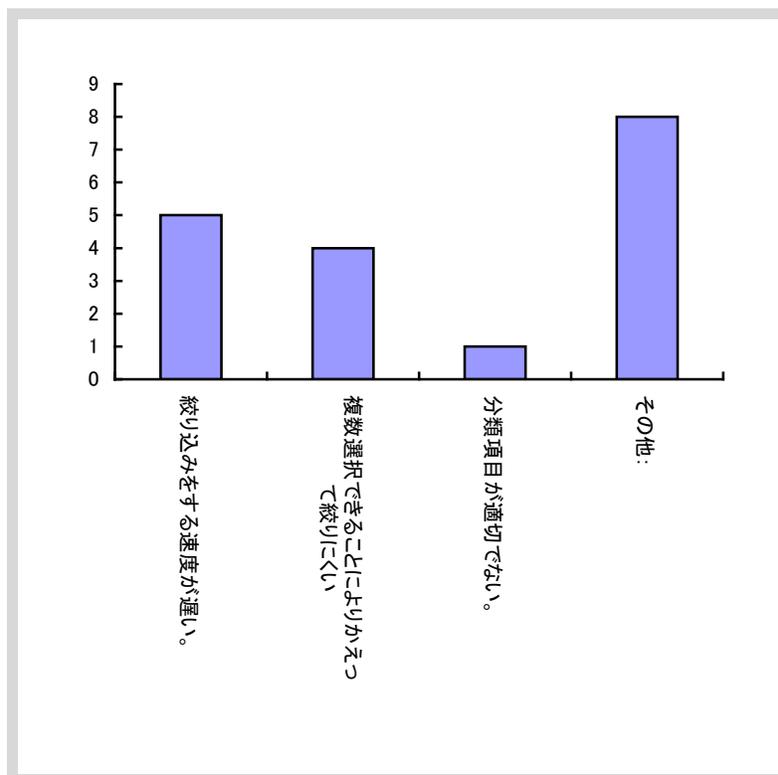
Q15. この絞り込み機能について、使い心地はいかがですか。

使いやすい 5 4 3 2 1 使いにくい



回答	件数	割合
1	3	3.3%
2	13	14.3%
3	31	34.1%
4	28	30.8%
5	16	17.6%

Q16. 使いにくい (+, ++) とお答えした方に伺います。どのような点が使いにく  
いでしょうか。以下の項目からお選びください。(複数回答可)



回答	件数
絞り込みをする速度が遅い。	5
複数選択できることによりかえって絞りにくい	4
分類項目が適切でない。	1
その他:	8

その他 内訳

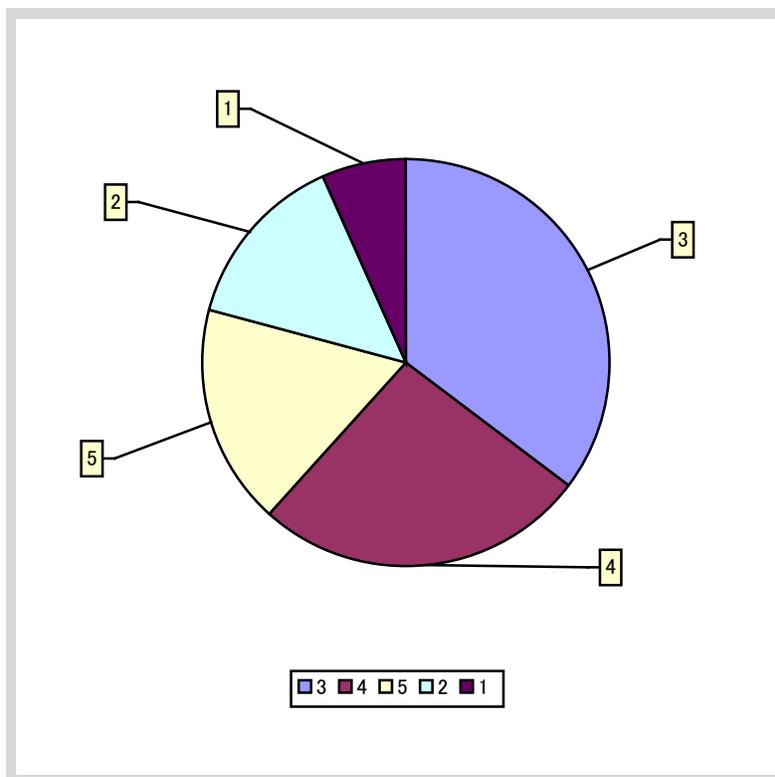
回答	件数
http://sagace.nibio.go.jp のトップ画面の検索の際にあらかじめ絞り込んでおきたい	1
あまり複数で絞り込まない	1
検索結果のタイトルでどのDBの結果なのかわかりにくい。例)ASTRAがSearch Resultとしか表示されていない	1
絞り込みができるのはうれしいが、絞り込みを簡単にリセットできるようなボタンが欲しい。	1
絞り込み機能の検索結果ができるまでに時間がかかるから	1
絞り込んだ後のヒット数が連動しない	1
生物種と生体の階層で絞り込めない	1
欲しい情報に行きつけない	1

Q17. 上記の質問で、“分類項目が適切でない”とお答えした方にお聞きします。どういったところが適切でないと感じられたでしょうか。他にどんな項目があると良いと思われませんか。

3つの分類のタイトルを載せた方が良いと思いますし、分類軸はたくさん合った方が良いと思います。疾患分類もそうですし、他にも現在どれくらい使われている薬なのか、など産業的な情報もあるとありがたいと思います。	1
---	---

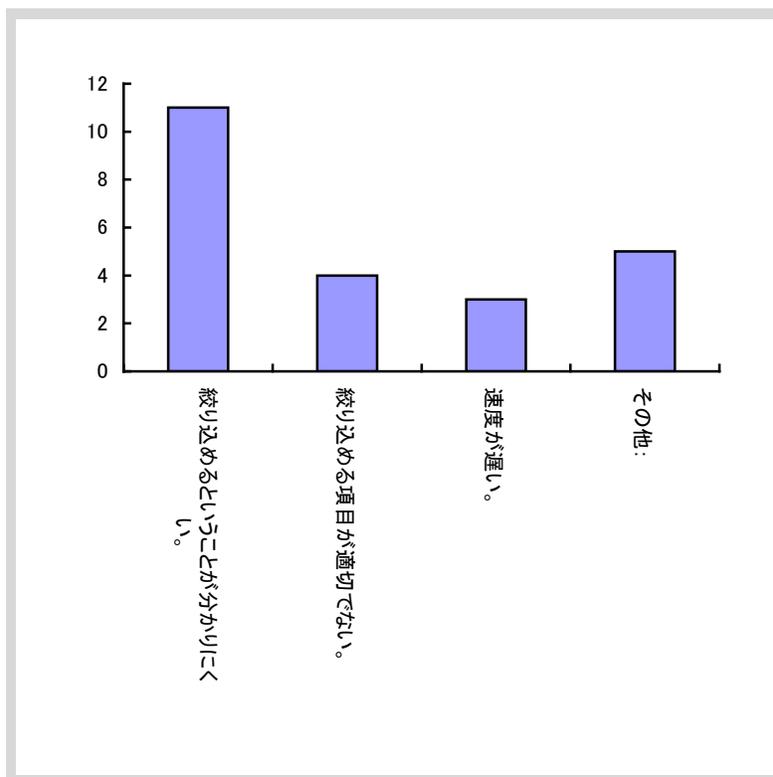
Q18. この絞り込み機能について、使い心地はいかがですか。

使いやすい 5 4 3 2 1 使いにくい



回答	件数	割合
1	6	6.6%
2	13	14.3%
3	32	35.2%
4	24	26.4%
5	16	17.6%

Q19. 使いにくい (+, ++) とお答えした方に伺います。どのような点が使いにくいでしょうか。以下の項目からお選びください。(複数回答可)



回答	件数
絞り込めるということが分かりにくい。	11
絞り込める項目が適切でない。	4
速度が遅い。	3
その他:	5

その他 内訳

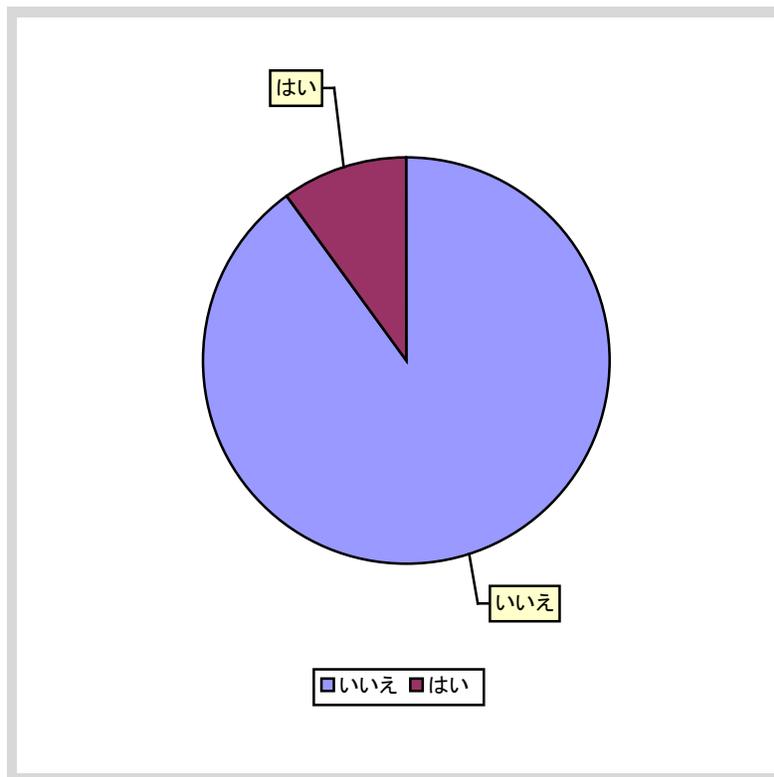
回答	件数
その他のデータベース   動物(ヒト以外), 植物, 微生物   個体 という分類が出て、微生物をクリックするとノーヒットになってしまった。期待としては、最低1つは残ると思ったのですが。	1
ある検索をした際に「文献・用語集・辞書   ヒト   細胞・組織」というリッチスニペットが表示された検索ヒット(Aとします)があり、そこで「細胞・組織」をクリックすると「情報は見つかりませんでした」となる。こちらの理解が間違っているのかもしれないが、少なくともAは再度ヒットすべきのように思います。	1
「ヒト」なのにマウスが検索結果に出てくる。	1
絞り込みの解除の仕方がよくわからない。	1
ファセット検索との違いがわからない。同じなら不要では？	1
もっと目立つ色と太字で表示してほしいです。	1

Q20. 他にどういった情報がリッチスニペットとして得られると良いと思われま  
すか。(例：細胞の画像、タンパク質の立体構造の画像、データ作成者の名  
前等)

回答	件数
細胞の画像	3
タンパク質の立体構造の画像	2
:細胞の画像、タンパク質の立体構造の画像	1
「その他のデータベース」になることが多いのもう少し分類しては？ メタデータの項目をアイコン化するのも視認性が良くなるかも。	1
「大学・研究機関」など情報源の信頼性を測れるキーワード検索	1
GeneIDなど	1
IDがわかれば十分。	1
MHCで検索を行い、植物が表示されます。クリックすると、当然、該当 するデータはないと表示されます。関連する言葉のみにしたほうが良 いと思われました。	1
NGSのデータ解析をやっている立場なので、SRAデータベースのIDが あると嬉しい。	1
PDBID	1
PDB情報、スプライスバリエント情報、GO	1
オントロジー	1
グーグルの検索結果みたいに画像も一部出てくるようであれば直感的 にわかりやすい	1
これというのではなく、統一された分類項目の適当な部分。 クリックするかを決めるための情報提示なのですが、対象データ全体 に何があるのかわからない状況では、提示される情報がユーザーの 求めるフィルター情報ではない可能性が高いのでは。ページタイトルの ようなそのページを本当に要約するような1行などがほしい。 といってもタイトルそのままでは他のデータベース由来のページと比較 するのは適当でないなど、問題がある。 データ作成者の名前などは、10以上の項目からユーザーがあらかじめ 選択したものを表示するならいいかもしれない。	1
コンテンツの使用言語、信頼度(品質)	1
すみません、思い当たりません。	1
タンパク質の立体構造の画像 組織、細胞の局財性の画像 電気泳 動やWestern blotting像	1
タンパク質の立体構造の画像、細胞の画像	1
タンパク質立体構造の構造データ	1
データの出所が明確にできること	1
データもしくはページが作成された日時	1

回答	件数
データ更新日	1
データ作者の名前と所属機関	1
データ作成者の名前	1
データ作成者の名前 形態形質の分かる画像	1
レコードの更新日や作成者など信頼性を高めることができるパラメータがあると良いと思う。	1
遺伝子のヒト組織発現	1
遺伝子検索をした際に、その遺伝子が関連する疾患(日本人で)が横断的に検索できると便利です。	1
医薬品情報の副作用	1
閲覧数:注目度がわかると信頼性にも繋がり、参考になります。	1
画像イメージ, 画像情報	1
画像は欲しいと思います	1
機能に関連するのか、特許に関連するのか、どのDBの情報があるのかが、直接わかるといいと思います。	1
項目に関連した参照文献	1
左のpaneにそのまま絞りこみを表示させれば良いのでは？	1
最終更新日が欲しい	1
細胞などの画像データ	1
細胞の画像、タンパク質の立体構造の画像	1
細胞の画像、タンパク質の立体構造の画像など絵的なものが増えると良いですね。最近の化学系の論文誌ではVisual Abstractが増えてい	1
作成責任者の情報	1
疾患と変異の理論的關係性の説明(仮説も可)	1
疾患名	1
写真	1
写真系は有ると嬉しい。求めているデータか否か、判断が早くなる。	1
症例画像 治療薬剤	1
情報がありすぎると画面が混雑してくるので、シンプルの方が良いです。(最低限の情報表示)	1
情報の数は今のままでよいと思う。	1
特になし	1
日付(ページが更新された年など)	1
発現組織、関連疾患、その分子に紐づいて頻繁に検索されているワー	1
分子の機能分類名	1
変異体	1
報告されている論文の画像データがあると良い	1
胞の画像、タンパク質の立体構造の画像、データ作成者の名前	1

Q21. あなたはデータベース管理者ですか。

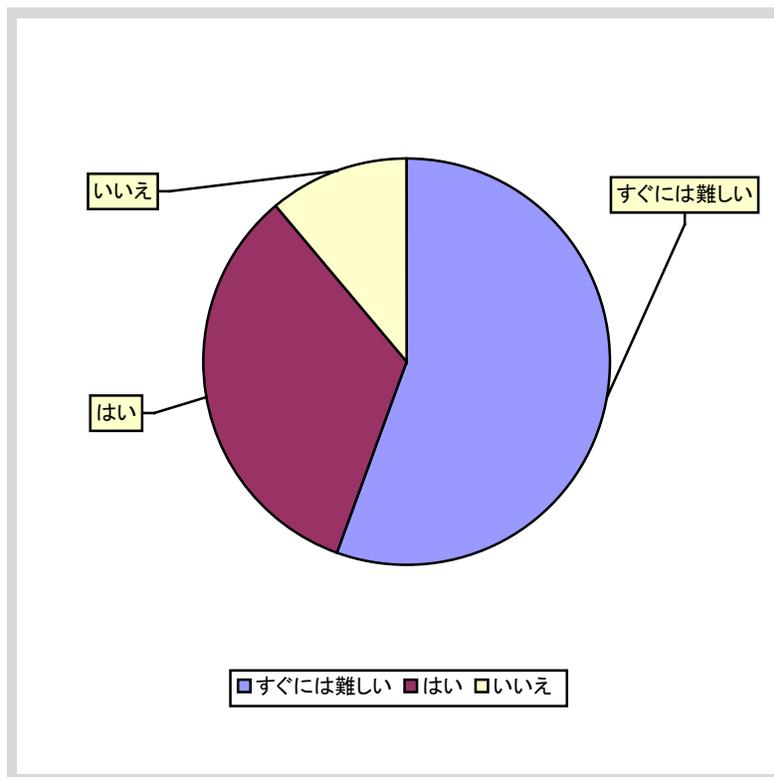


回答	件数	割合
いいえ	82	90.1%
はい	9	9.9%

Q22. “はい”と答えた方に質問します。リッチスニペットを検索結果に反映させるためには、検索対象となるデータベースにメタデータを入れていただくことが必要です。以下のリンク先をご覧ください。

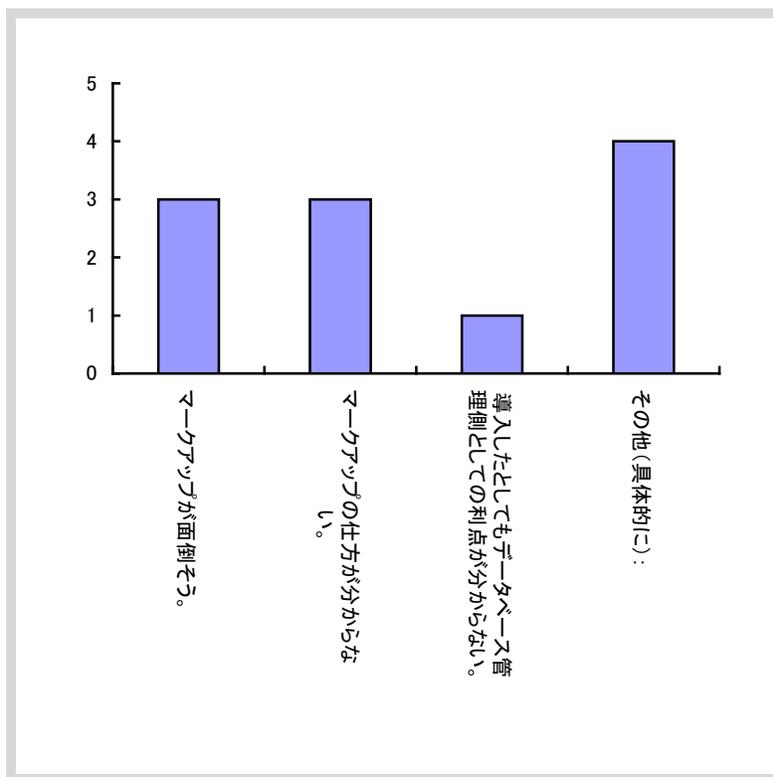
<http://sagace.nibio.go.jp/press/metadata/markup/>

リッチスニペットの導入について、あなたのデータベースに導入して頂くことは可能でしょうか。



回答	件数	割合
はい	3	33.3%
いいえ	1	11.1%
すぐには難しい	5	55.6%

Q23. 上記の質問で、“すぐには難しい”、“いいえ”とお答えの方にお聞きします。導入が難しい理由は何でしょうか。（複数回答可）

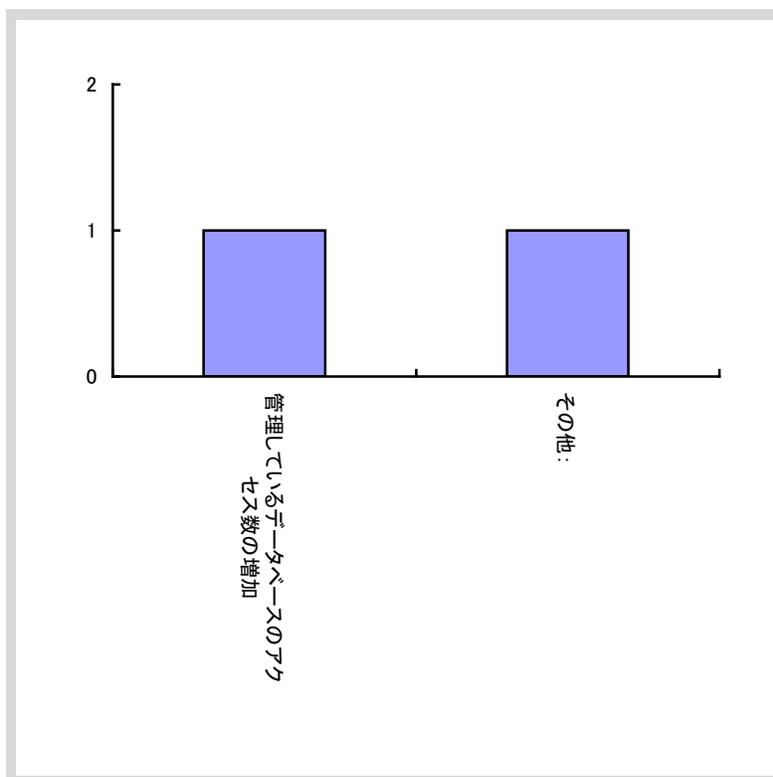


回答	件数
マークアップが面倒そう。	3
マークアップの仕方が分からない。	3
導入したとしてもデータベース管理側としての利点 が分からない。	1
その他(具体的に):	4

その他 内訳

回答	件数
何をマークする対象にするのが適当か、すぐには分 からない	1
技術内容とその影響をまだ深く理解できていないか	1
togodbの環境を利用しているため	1
イントラネットのものだからセキュリティ的に無理	1

Q24. 上記の質問で、“導入したとしてもデータベース管理側としての利点がない。”とお答えした方にお聞きします。どうしたら利点があるといえるでしょうか。（複数回答可）



回答	件数
管理しているデータベースのアクセス数の増加	1
その他:	1

その他 内訳

回答	件数
更新、データ整合性チェックなどの効率化(あるかどうか知りませんが)	1

Q25. その他、ご質問やお問合せがございましたら、ご記入をお願いします。

- マークアップについては技術ありきではないか。この試みを否定するわけではないですが、もしかしたらユーザーは他の問題で困っているのではないのでしょうか。よく会話をされているのでしたらすみません。
- 最終更新日のリッチスニペットは検索に反映されてないようですが、取得が難しいのでしょうか？
- 使い勝手が悪くて、意見できるレベルに達していない。有償のDBを参考にして、大幅な改善が必要。
- データベース自体の検索は必要性を感じる時もあり、システム的に対応しなければならぬほど膨大な数のDBがあるのは存じていますが、Sagaceに限らずMEDALS、生命科学データベース横断検索その他、何度も試したことはあるのですが、満足のいく検索結果を得たことがありません。  
このため、googleで調べてなければ、pubmedでreviewを探したり、NARのDB issueを探したり、それでもなければTwitter等コミュニティに聞いてみないと、  
結局のところ、網羅性が確保されているか、重大な見落としがないかの安心が得られないので、その他の検索手段をメインで使わざるを得ません。  
自分のシステム開発経験から言えば、ニーズが一番先にあり、そのニーズを満たす結果を得るための処理手順があり、それを繰り返すことでシステムティックにニーズを満たすことのできるシステムが出来上がっていくのだと思いますが、  
これら横断検索の現在の開発スキームに汎用的なニーズを満たせるゴールがあるとは期待できなくなってしまうのが正直な印象です。  
個人的にはメタ情報を細分化して構造化して充実させていくことで検索可能性を増やす方向とは別に、このような横断検索エンジンを利用した「成功例」をいくつも作り、統合TVなり、サイトの広報ページなり、担当者のTwitterでのアナウンスなりで広報していくことで、開発者側も「成功体験」を重視するようなシステム開発作りにつながるでしょうし、開発者とSNS上での交流も生まれ、良く使われるシステムになっていくように思います。最初から全てのニーズを満たすのは当然不可能でしょうし、現行のスキームで実現しやすい手のつけやすいものから対応し、場合によっては開発担当が足りない部分をSPARQLやSQL等の直叩きで探し、それと同様の手順をシステム上でも可能にしていくという手順を踏めば、近いうちには、特定の検索の仕方なら信頼できる結果が返るということがユーザーやコミュニティに認知され、その範囲を徐々に拡大していけるのではないかと思います。  
そもそも取り扱い範囲が広く、政治的にもいろいろな組織の利害に関係しそうなシステムではあるので、大変難しいとは思いますがコミュニティと繋がりのあるご担当者関わっている点でご期待申し上げたいと思っております。

- 同義語展開した時に、同義語がたまたま違う用語の略称の場合にもヒットしてしまっている。ADAM17→Tace (tumor necrosis factor- $\alpha$ -converting enzyme) はOKだが (Transcatheter Arterial Chemoembolization) はNGなど。
- 国内の DB だけを対象にしているので欲しい情報に辿りつきにくい？
- 普段利用していませんが、すごく使いやすくなったと思います。
- 良くできていると思います。