

学術情報システム総合ワークショップ

調査計画書

テーマ: 学術情報データベースでの検索行動の分析

2班 亀崎、渡邊

内容

- 背景
- 動機(問題意識)
- 課題・目的
- 課題を変更した理由
- 調査について
- 調査方法
- 調査備忘
- スケジュール/作業分担

背景

- 情報環境の変化

- インターネット、携帯電話、スマートフォン、タブレット端末等の普及

- 情報検索者の変化、広がり

- 老若男女誰もが知識を得るために検索を行う

- 検索に知識/技術が必要であるという意識は希薄

⇒ 情報検索は様々な層の人が知識を得るために行う生活行動の一部となっている

背景

- 学術情報データベースの現状

- 網羅性の高いDBは基本的な検索機能のみで、絞り込みや新しい提案をする機能まで持たない
- 高い検索機能や絞り込み機能を持つDBは検索対象がきわめて限定的であったり、有料のサービスだったりする
- DB内のコンテンツはそのDB内に限定されてしまっていて、その情報を求める利用者に認知されない
- Google ScholarなどからDBに誘導されても、使い方がわからない

⇒ 学術情報が手軽に検索できない状況が、非学術研究者や学生などの初学者を学術情報から遠ざけてしまっている

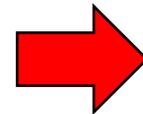
動機（問題意識）

「一般的なインターネットサービスから手軽に検索することを日常としているユーザ」が学術情報を必要としたとき、その日常体験から十分な検索を行わず、求める学術情報にたどり着けない、もしくは偶然発見した学術情報のみで満足し、本来たどり着くべき多くの学術情報と出会えていない可能性がある。

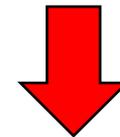
動機（問題意識）

「十分な検索を行なっていない」「最適な学術文献にたどり着いていない」原因

- 本人の意欲
- ユーザーインターフェースのわかりにくさ
- 適切な文献検索を行う方法知らない
- etc.



ユーザの検索行動を詳しく知ることができれば、解決のための方法を提案できるのではないか



最適な文献検索ができないユーザの行動を把握することが必要

課題・目的

7月の打ち合わせでは

私たちは、初学者のユーザが自らえらんだ最適な文献に案内するために、ユーザが検索結果の一覧を検討したのか知りたいです。



次のように変更した

私たちは、求める文献を的確に発見できないユーザが適切な文献検索を行う方法を提案するために、該当ユーザの検索行動や傾向を知りたいです。

課題を変更した理由

7月の段階では、調査方法を次のように計画していた。

1. アクセスログから、検索結果一覧の1件目の詳細画面あるいは本文を1回閲覧し検索を終了したユーザが一定数いることを立証し、該当ユーザを、検索結果一覧を検討していない初学者ユーザとみなす。
2. 初学者ユーザの1.の検索行動に影響を与えた要因が何であるかを調査によって特定する。
3. ユーザが2.の阻害要因に影響されず、最適な文献にたどり着く方法を提案する。
4. 1.で該当ユーザが一定数存在しないことが立証された場合は、2.の調査や3.の提案の対象をユーザ全員とする。

課題を変更した理由

実証をアクセスログの分析、検索行動に影響を与えた要因の2通りで行う予定であったため、実質課題が2つあることになり、調査スケジュールを考えると、どちらについても深く掘り下げて調査ができない



メインとなる課題を求める文献を的確に入手できないユーザの検索行動とし、適切な文献検索を行う方法については提案にとどめることにした

調査について

概要

- 学術情報データベースのアクセスログを元にユーザの検索行動や傾向を多角的に把握し、それらの結果から解決のための方法を考える。

調査対象

- CiNii Articlesで何らかの検索行動をとったユーザ(アクセスデータ)

調査前提

- 本調査では、「求める文献を的確に発見できない」原因の代表的なものとして、ユーザの「データベースや検索自体に対する知識不足」や「締め切りなどの時間的制約」「不本意な動機による意欲低下」「ユーザインターフェースの難解さ」を想定し、結果として「網羅的な検索を行わない」「検索結果を熟慮しない」「検索語の選定を疎かにする」などの様々なユーザ行動に反映されているものとする。

調査方法

概要

- アクセスデータに対し、一番妥当性が高いと思われる基準で「求める文献を的確に発見できないユーザ」と「それ以外」の2者に仕分け、それぞれの性質（傾向）を比較する。また、上述の基準に限定せず別の基準を模索し、考案した基準で「求める文献を的確に発見できないユーザ」と「それ以外」の2者に仕分け、それぞれの性質（傾向）の比較を繰り返しながら、その差異を把握していく。

調査方法

1. 「求める文献を的確に発見できないユーザ」と「それ以外」を仕分ける基準として、「DBを1回検索し1回詳細画面を見て検索を終了(あるいは、「検索するだけで終わり」「本文リンクを見て終わり)」という行動」を採用する。その基準に合致するログとしないログを仕分け、「検索語」「エージェント」「滞在時間」「プロバイダ」「複雑な検索方法(前方一致等)」「詳細検索・簡易検索・全文検索」「参照元」などを視点に分析を行う。Google Analyticsを主に使用するが、分析内容によってはアクセスログデータそのものを分析に使用する。分析結果を考察するとともに、分析方法を定型化し、次工程以降に活用できるように整備する。

DBを1回検索し1回詳細画面を見て検索を終了
(求める文献を的確に発見できないユーザ)

区分

それ以外

以下の視点で比較

検索語(キーワード)

エージェント(PC、タブレット...)

滞在時間

プロバイダ(大学、政府機関...)

複雑な検索方法(前方一致、完全一致...)

簡易/詳細/全文検索、参照元、etc...

調査方法

2. 「求める文献を的確に発見できないユーザ」を仕分けるための基準を探するため、当該ユーザの類型・行動例を列挙し、基準となりうる行動を選定する。(新たな基準作り)
3. 2. で決められた基準を元に、再度調査と分析を行い、1. を補強する。(1. の補強調査)
4. 全ての分析結果を元に当該サービスの評価を行い、「適切な文献検索を行う方法」を提案する。

調査備忘

- ツール

- Google Analytics及びRubyでプログラムを組んで処理をする。
- 「Google Analyticsで確認・検証可能な項目の洗い出し」と、「Rubyによる前処理方法の確立」を最優先で行う。

- 生ログ範囲と受け渡し

- 調査の妥当性を確保するには、20～40日分の生ログデータを使用することになる。
- 受け渡し方法は、Google Drive or フラッシュメモリ。要相談。

	9月 一週目	9月 二週目	9月下旬	10月	11月	12月
亀 崎 渡 邊 備 考	GAを触って実施可能な調査の模索 (-9/6) 調査計画書確認 (9/3-4)	生ログ処理方法の模索	GA・生ログによる調査1の実施	GA・生ログによる調査2・3の実施	調査まとめ (調査4) 報告会準備	調査報告書作成
	GAを触って実施可能な調査の模索 (-9/6) 調査計画書確認 (9/3-4)	GAで可能な調査対象・項目の確定 (-9/13)	GA・生ログによる調査1の実施	GA・生ログによる調査2・3の実施		
	調査計画書作成(-9/2) GAを触って実施可能な調査の模索(-9/6)	生ログ処理方法の模索	GA・生ログによる調査1の実施	GA・生ログによる調査2・3の実施		
	調査報告書提出(9/4) GAのみで工程1が実施可能か、あたりをつける	生ログ調査対象期間 (日数)を決める GAのみで調査1が実施可能か確定		できれば各人別々の調査対象を持つ		
	毎週木曜日に19:00～定例会を、Googleハンガアウトを使って行う					