

Assisting the search

佃 洸掇 (産総研)

An Eye-Tracking Study of Query Reformulation

Carsten Eickhoff, ETH Zurich

Sebastian Dungs, Univ. Duisburg-Essen

Vu Tran , Univ. Duisburg-Essen

何をした論文？

検索中のユーザが注目している単語とユーザがクエリの修正に使用する単語の関係を分析してクエリ推薦に応用

先行研究との違いは？

検索中のユーザの視線やマウスカーソルの動きから様々な分析が行われてきたが、クエリの修正に使用する語との関係を調べた研究は少ない。

パラグラフ単位の視線情報からクエリ推薦をした研究はあったが、単語単位に着目したのは初めて。

- **検索中に見る語とクエリ修正に使用される語の関係分析**
 - (次以降のスライドで詳細を紹介)
- **マウスカーソルが視線の代わりに使えることを検証**
 - 視線検出は高コストなのでカーソルをトラックしましょう
- **単語単位での視線情報をもとにクエリを推薦**
 - パラグラフ単位の視線情報を使うよりも高い推薦精度を記録

- ・クエリ入力
- ・検索結果閲覧

- ・ページ閲覧

- ・クエリ修正
- ・検索結果閲覧



- 被験者が見たページの中に「金閣寺」は含まれる？
- 被験者が実際に見た語の中に「金閣寺」は含まれる？



- クエリ内の語の**43%**は被験者が見たページ内の語
- クエリ内の語の**21%**は被験者が実際に見た語

- 仮説：見た語の類義語や反義語を入力しているのでは？
- 検証：被験者が見た単語とクエリの語の類似度を計算して、類似度ごとの注視時間を比較

		fixation duration
金閣寺・階段 →	$LCH < 0.25$	30ms
	$0.25 \leq LCH < 0.5$	36ms
	$0.5 \leq LCH < 0.75$	39ms
	$0.75 \leq LCH < 1.0$	34ms
	$1.0 \leq LCH < 1.25$	32ms
	$1.25 \leq LCH < 1.5$	38ms
	$1.5 \leq LCH < 1.75$	45ms*
	$1.75 \leq LCH < 2.0$	53ms*
金閣寺・清水寺 →	$LCH \geq 2.0$	86ms*

クエリの語と類似している語ほど注視時間が長くなる

Differences in the Use of Search Assistance for Tasks of Varying Complexity

Robert Capra, Univ. North Carolina

Jaime Arguello, Univ. North Carolina

Anita Crescenzi, Univ. North Carolina

Emily Vardell, Univ. North Carolina

何をした論文？

48名の被験者によるユーザスタディを行い
他人の検索行動を表示する検索支援ツールの有用性を検証

先行研究との違いは？

ユーザスタディや検索ログの分析によって
ユーザの検索行動を分析する研究は行われてきたが
他人の検索行動を表示する検索支援ツールの有用性を
検証する研究はほとんど行われていない

検索結果に加えて、過去に同じタスクの検索を行った3名のユーザのクエリ、クリックした検索結果、ブックマークしたページ*を表示

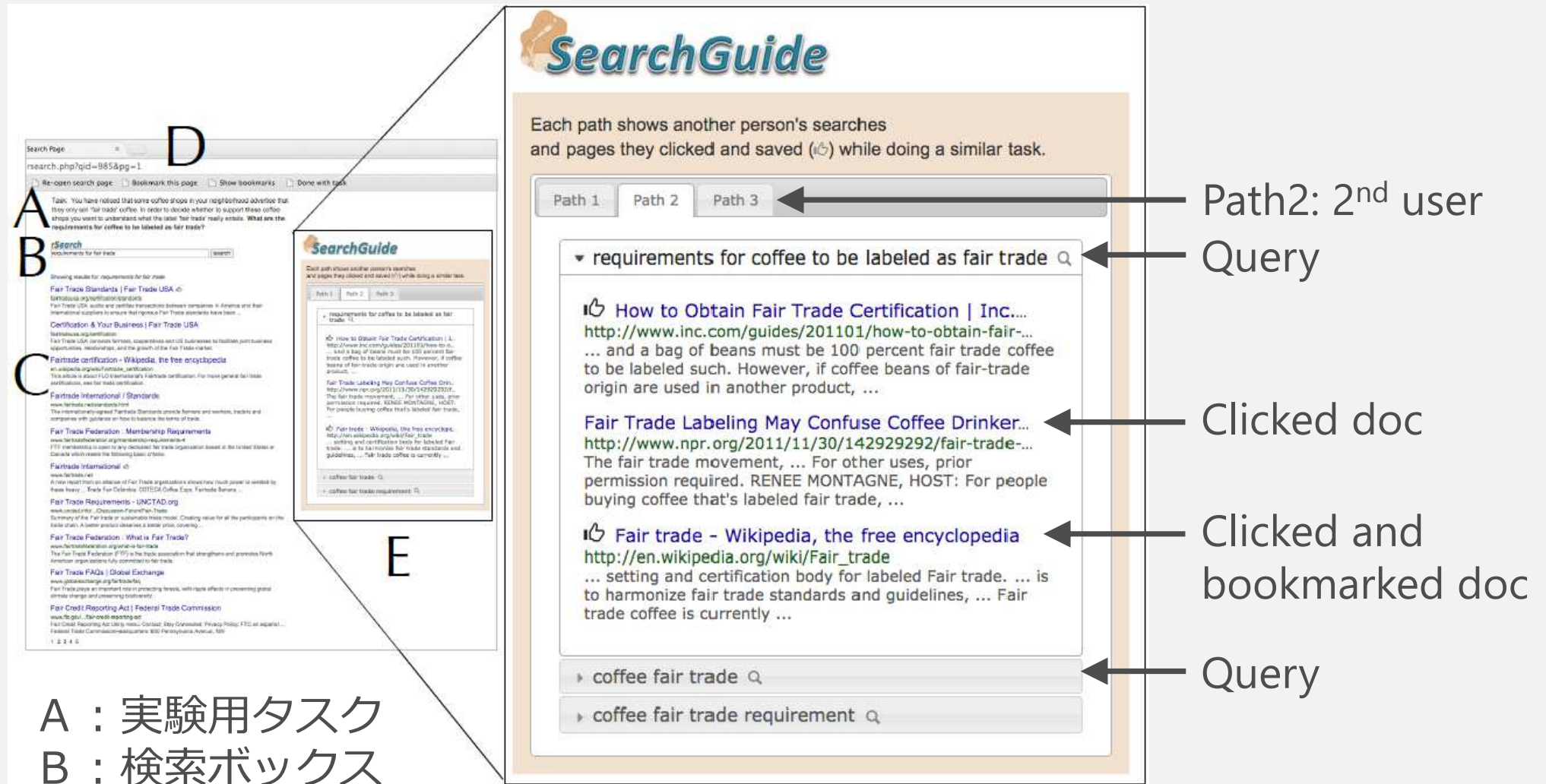


Figure 1

* 表示するデータには著者の過去の研究で得られたログを使用

- **タスクの難しさと支援ツールへの依存度の関係は？**
 - タスクが難しくなるほど支援ツールを頻繁に使用
- **タスクの性質と支援ツールへの依存度の関係は？**
 - どう検索すれば良いか見通しの立てやすいタスクほど他人の検索行動の内容も理解しやすいのでツールを頻繁に使用
- **支援ツールによってタスク達成後の感想は変わる？**
 - 集中度や満足度をはじめとする大部分で変化なし
- **支援ツールがユーザの体験に与える影響は？**
 - 支援ツールを使ったのに良いページが見つけれないと評価最悪、ツールを使って良いページが見つかりると評価最高
- **支援ツールはいつ何のために使われる？**
 - タスクが優しいときは答え合わせのために使用し、タスクが難しいときは新しい情報源を見つけるために使用

タスクのcognitive complexityには4段階のレベルが存在

タスクを達成するために必要とされる学習量とcognitive effort

● Lv. 1: Remember (長期記憶から関連情報を想起)

- 過去にアメリカでHIV患者のテレビ番組を見た。最近になってHIVの予防に興味湧いてきたけど、アメリカでは現状何人がHIVに感染している？

● Lv. 2: Understand (解説記事を見て知識を獲得)

- 甥がサッカーを始めると言っているけどサッカーは危なそう。長期間サッカーをすることで生じる健康上の害は何？

● Lv. 3: Analyze (事象を細かく分割して分析)

- 人口舌の話題を最近よくきくけど、人口舌にはどんな種類がある？それぞれにはどんな健康上のリスクがある？

● Lv. 4: Evaluate (物事の判断を下す)

- 友達がタトゥーを入れたことを後悔している。タトゥーを消すために現状使用できる方法とその有効度は？どの方法がベストだろうか、またそれがベストだと思う理由は？

Exploring Session Context using Distributed Representations of Queries and Reformulations

Bhaskar Mitra, Microsoft

何をした論文？

検索セッションのcontextを考慮したDeep Learning
を使ってquery auto-completionの精度を向上

先行研究との違いは？

検索セッションのcontextを考慮した検索精度の向上
は行われてきたが、query auto-completionは少ない。
クエリ修正をベクトルで表現したのは初めて

$vector(\text{new york}) + vector(\text{newspaper}) = vector(\text{new york times})$

$v(\text{beyonce}) + v(\text{pictures}) = v(\text{beyonce images})$

$v(\text{www.facebook.com}) - v(\text{facebook}) + v(\text{twitter}) = v(\text{www.twitter.com})$

$v(\text{university of washington}) - v(\text{seattle}) + v(\text{chicago}) = v(\text{chicago state university})$

2007 toyota yaris

us

us elections 2014 predictions

usaa

used cars ✓

←←← 直前に入力したクエリ

←←← 入力途中のクエリ

←←← ベースライン 1

←←← ベースライン 2

←←← 提案手法

airline tickets

amer

american express

american express

american airlines ✓

pace university

amer

american express

american girl

american university ✓

the fighter

amer

american express

american express

american psycho movie ✓

usairways

us

us elections 2014 predictions

usps.com

usairways.com ✓