エッジ型情報収集支援における個人クエリ相互利用の利用者視野拡大効果 - 実験報告 -

金沢工業大学情報工学科　小林　暉 (3年)
佐治寿一 (3年)
風尾勇侑 (3年)
芦川拓実 (3年)
指導教官　金道敏樹

はじめに

　他人のクエリを利用することで、情報アクセスへの利便性と視野拡大を並行できるかもしれない。

目次

- 構想について
- 作ったシステム
- 使用した技術
- クエリ利用とは
- 実験結果
- まとめ

構想の前提

- 各個人にカスタマイズされた個人情報フィルタ
  その中にある個人クエリは、ユーザの個人視野を表す。

構想-１

- １．ユーザの興味(ユーザの視野)を広げるには、他者のクエリを利用すると便利なのではないか。

- ２．あるコンテンツに対して、興味の似通った他者のクエリを利用した検索が行えれば、ユーザの興味範囲が広がる（かも）。

他者のプライバシーにどう配慮するか

構想– ２　プライバシーを保つために

- プライバシーをユーザがコントロールするには、エッジ型の方が良い。
  - 今回はサーバクライアント（≠エッジ型）で実装

- 個人クエリをユーザ端末に置ける技術で実現しよう
  - 大きな辞書を使わない情報フィルタINSOPを活用

作ったシステム

おもしろ？映画検索支援サービス

まず、いくつかの映画にチェックをつけてください

名前：

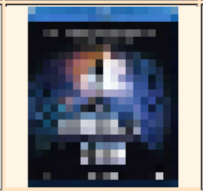
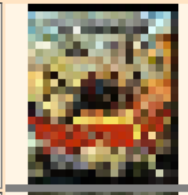
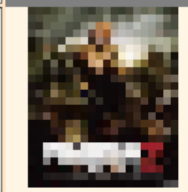
学習　シャッフル

CheckBox	Image	Title	Abst
○ ×		ID:197「アンティーク西洋骨董洋菓子店」	
○ ×		ID:215「ウタモノガタリ CINEMAFIGHTERSproject「ファンキー」」	
○ ×		ID:20「パニックインベガスエイリアン襲来」	
○ ×		ID:132「フローズンライター」	

作ったシステム

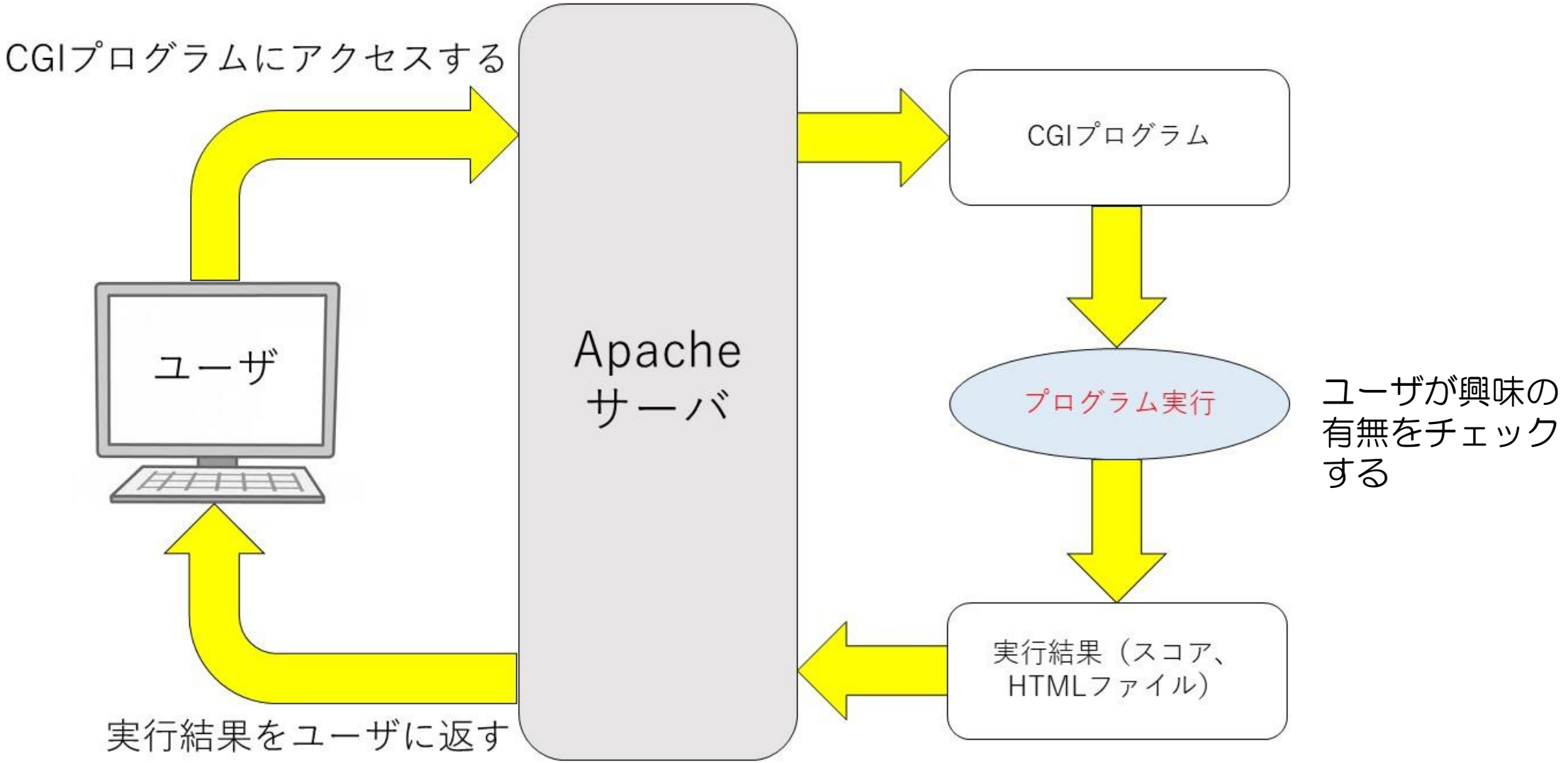
名前： User1

学習 （済）表示 （済）非表示 シャッフル

CheckBox	Image	Title	Abst	Score
○ × (済)		ID : 28「アンドロメダ… 」		117.6
○ × (済)		ID : 24「ディープインセクト 」		85.91
○ × (済)		ID : 266「ウォーキングZ 」		52.88
○ ×		ID:1「プラネットオブロボット 」		15.88

スコアの大きい順に並べている

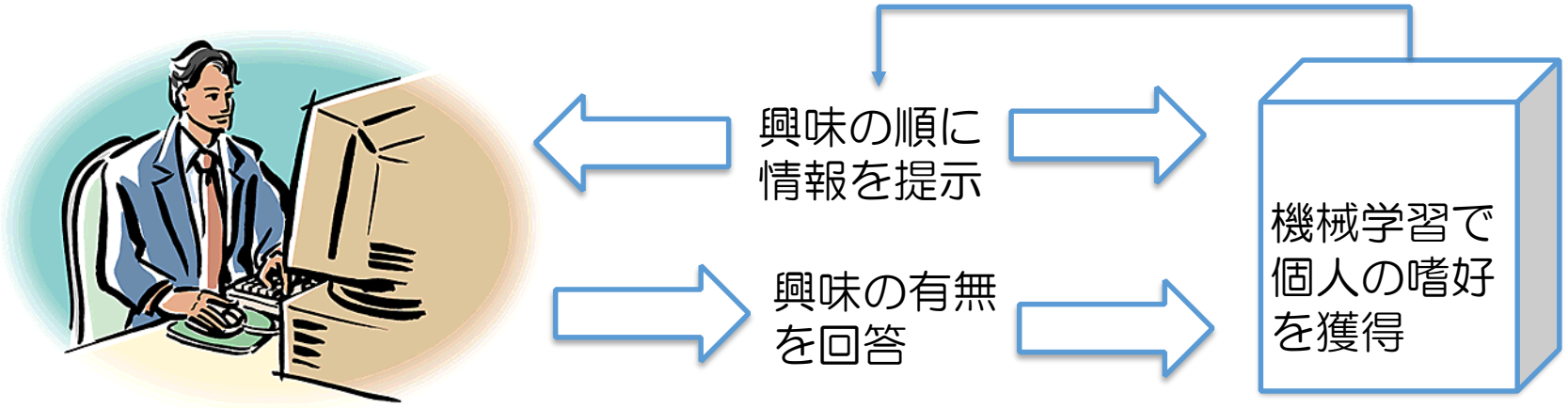
システム構成



ユーザが興味の有無をチェックする

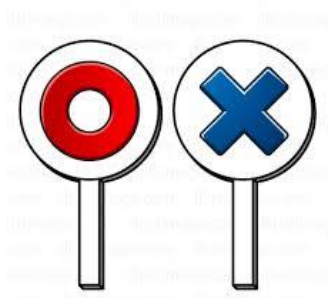
使用した技術１－プロファイリング技術：INSOP

- INSOP
  - ユーザプロファイルを生成する個人情報フィルタ



使用した技術１－INSOP

Yes/No入力



INSOP

Personal Profile

含まれるKW	ユーザ特有の値
KW1	SKC（KW1）
KW2	SKC（KW2）
KW3	SKC（KW3）

Contents Score

Score

使用した技術１－INSOP

ユーザごとに小さな辞書を自動生成し、「価値」のあるキーワードを登録

Positive key words: q(wi) >> p
Negative key words: q(wi) << p
Neutral key words: q(wi) ≒ p

興味有り＝正
興味なし＝負

p:ユーザが興味ありとする確率

q(wi):キーワードwiが付いた情報をユーザーが興味ありとする確率

個人辞書へは、キーワードが運ぶ個人の興味の情報量の多い
Positiveキーワードと Negativeキーワードが登録される。

使用した技術１－INSOP

興味あり

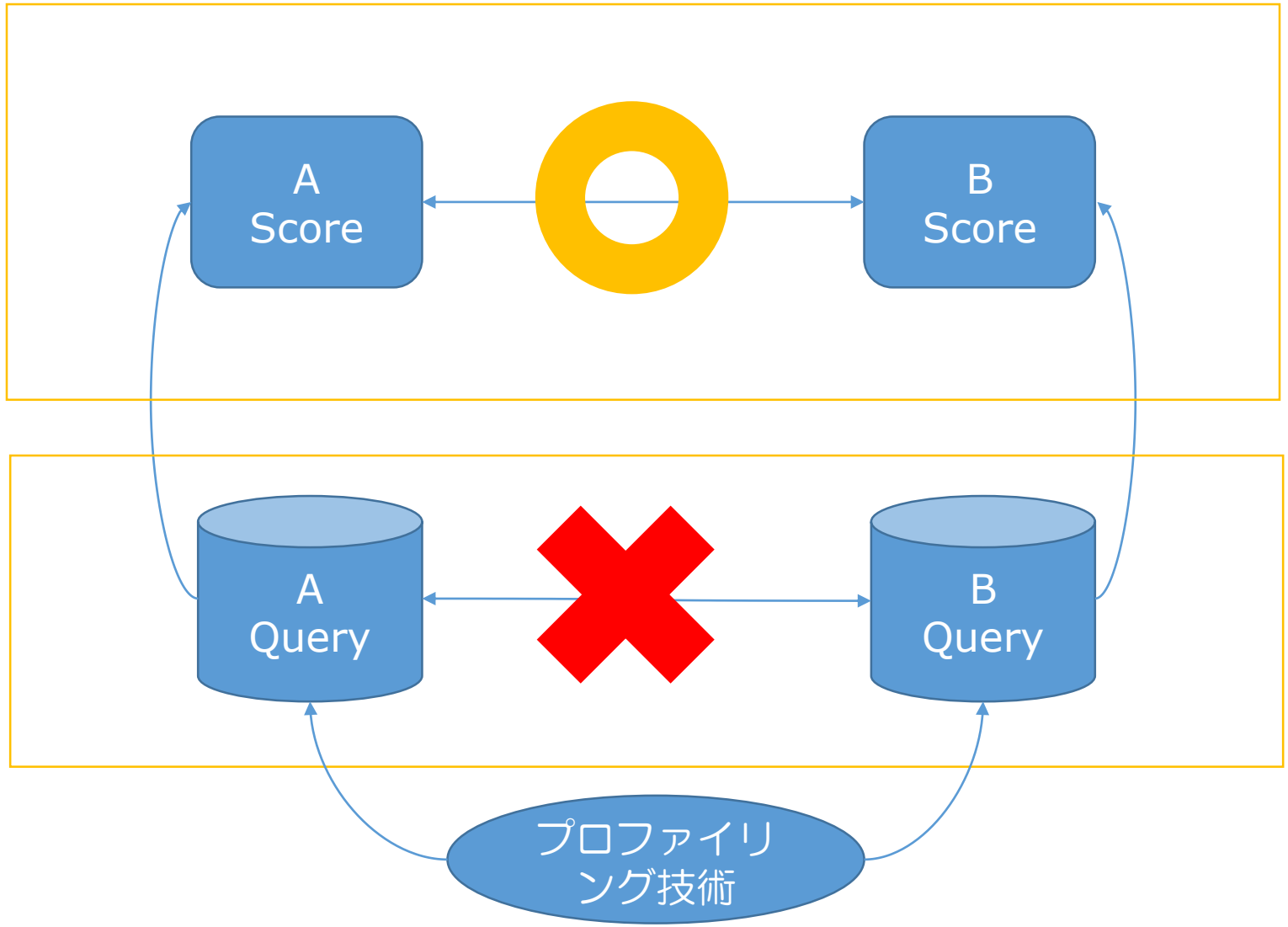
興味なし

キーワードが運ぶ「興味度」は統計的
キーワード評価値(SKC)で計算される。

各情報に対する「興味度」は
SKCの総和で推定できる。

SKCの総和の大きい順に情報を並べる。

クエリ利用：プライバシーへの配慮

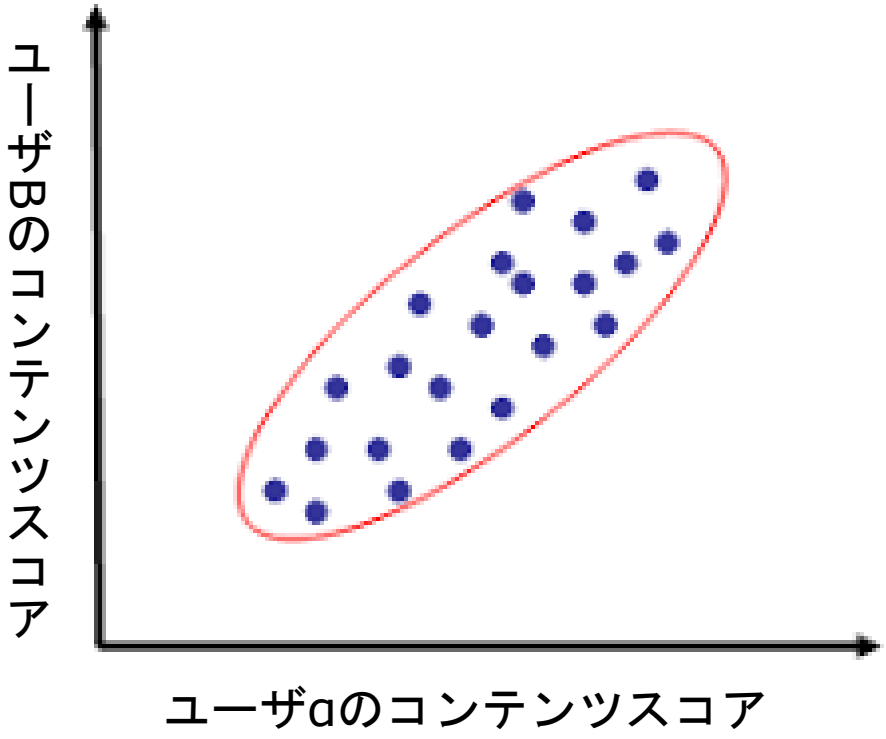


個人クエリの中身を比較するのではなく、スコアの類似性の計算をしているだけ！

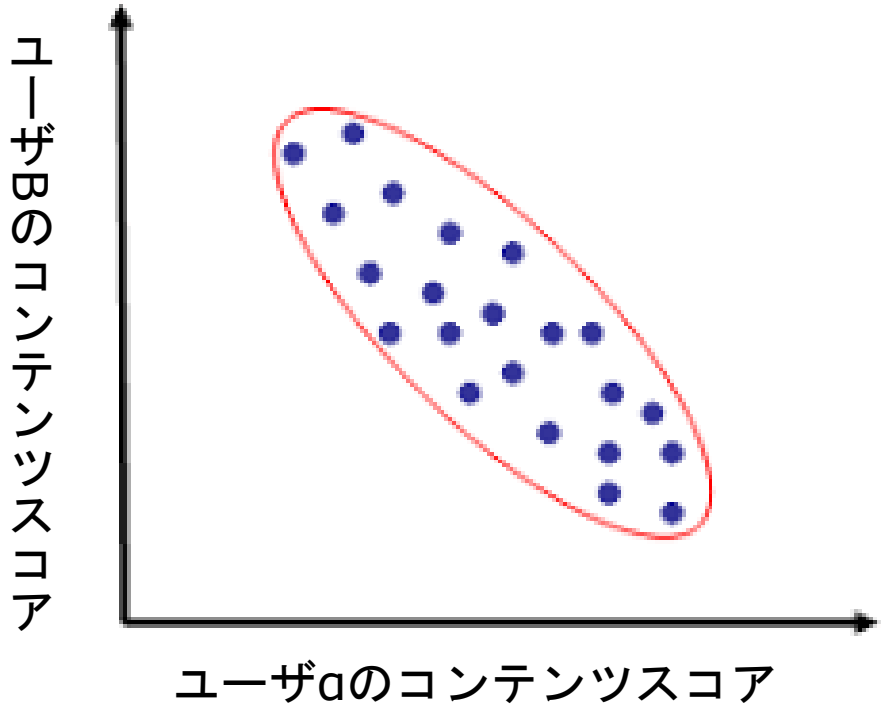
直接比較はユーザのプライバシー保護が難しい

使用した技術2 - コンテンツスコアによる類似性の推定

正の相関

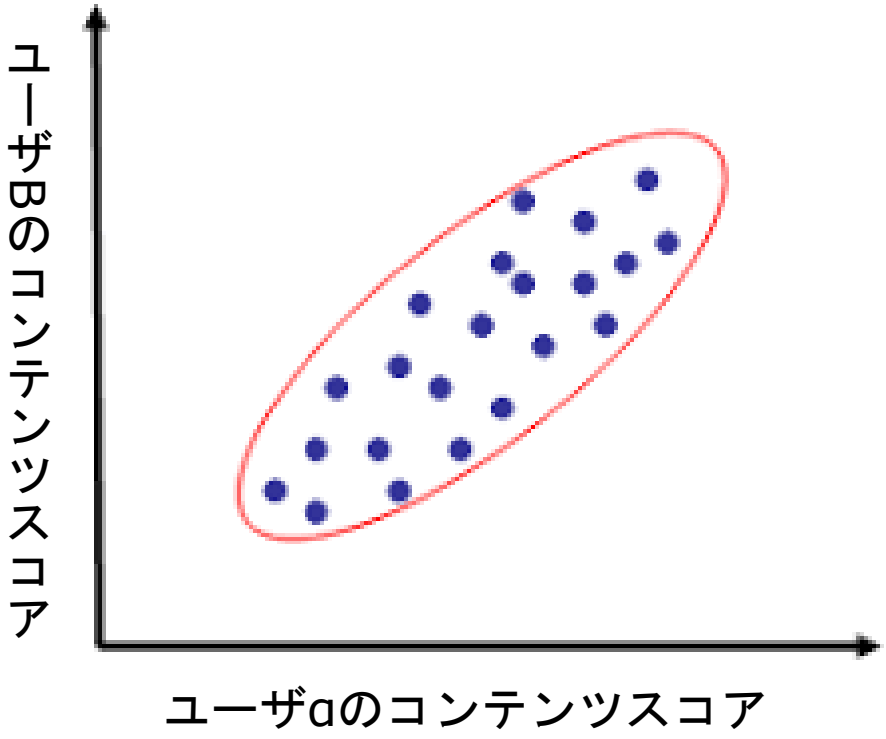


負の相関

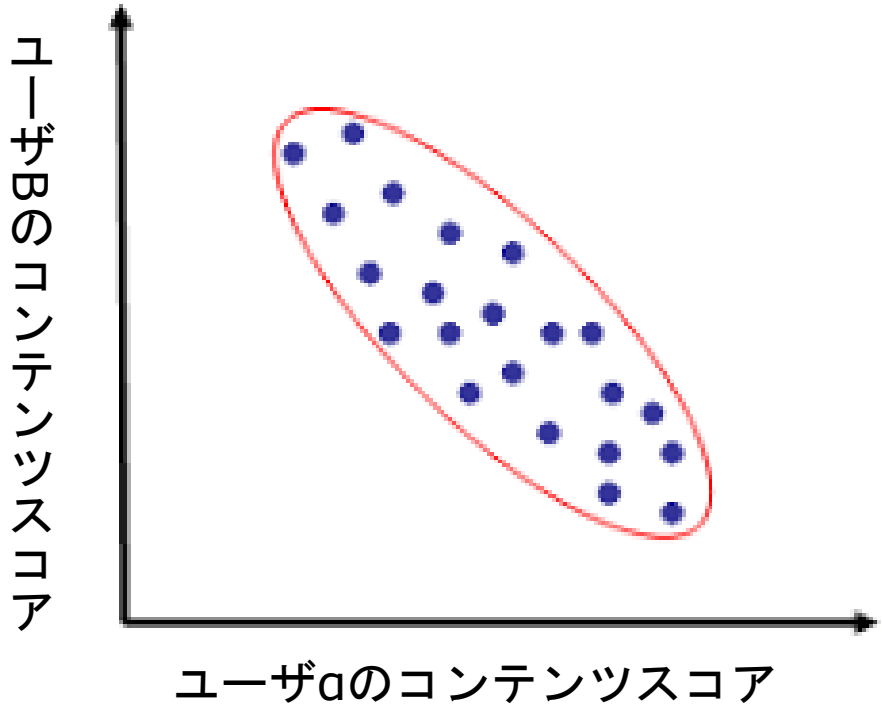


使用した技術2 - コンテンツスコアによる類似性の推定

正の相関



負の相関



あるユーザ α とユーザ β の各コンテンツに対する興味の類似性を、相関により表す。

実験結果 - 1

- 金沢工業大学の学生並びに職員にアプリケーションを利用してもらい、個人クエリを収集した。

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S15	S16	S18	S19	S21
S1	10	-1	2	4	1	2	0	2	1	0	0	4	1	1	0	2	1	2
S2	-1	10	1	2	1	0	1	-1	1	2	-1	-1	0	2	3	-1	0	-2
S3	2	1	10	2	-1	0	-2	0	0	4	-1	-1	3	-1	0	-2	-2	2
S4	4	2	2	10	2	1	1	2	2	1	1	3	1	1	2	2	0	3
S5	1	1	-1	2	10	2	0	0	1	0	3	-1	1	0	1	2	0	1
S6	2	0	0	1	2	10	1	2	2	1	2	0	1	3	2	2	2	3
S7	0	1	-2	1	0	1	10	1	0	-1	1	2	0	2	2	1	2	-1
S8	2	-1	0	2	0	2	1	10	3	1	2	3	0	2	0	3	1	2
S9	1	1	0	2	1	2	0	3	10	3	2	0	0	3	2	1	1	2
S10	0	2	4	1	0	1	-1	1	3	10	1	0	2	0	1	-2	1	2
S11	0	-1	-1	1	3	2	1	2	2	1	10	2	1	1	1	4	3	2
S12	4	-1	-1	3	-1	0	2	3	0	0	2	10	-1	0	-1	3	3	2
S13	1	0	3	1	1	1	0	0	0	2	1	-1	10	0	1	-1	-1	1
S15	1	2	-1	1	0	3	2	2	3	0	1	0	0	10	1	2	1	0
S16	0	3	0	2	1	2	2	0	2	1	1	-1	1	1	10	-1	1	1
S18	2	-1	-2	2	2	2	1	3	1	-2	4	3	-1	2	-1	10	2	1
S19	1	0	-2	0	0	2	2	1	1	1	3	3	-1	1	1	2	10	1
S21	2	-2	2	3	1	3	-1	2	2	2	2	2	1	0	1	1	1	10

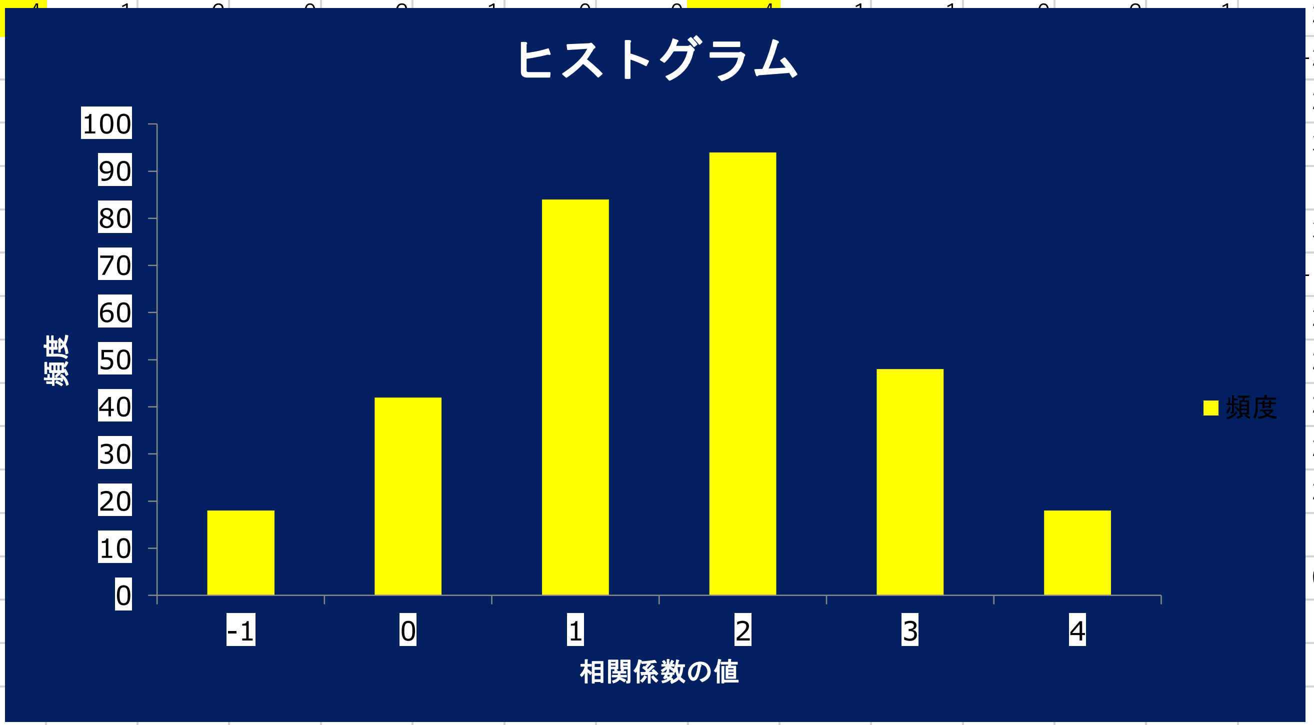
相関係数が最も大きい値（黄色）を利用することで、ユーザにとって便利な推薦になる。

※本来の数値を10倍し、整数で表示している

実験結果 - 1

- 金沢工業大学の学生並びに職員にアプリケーションを利用してもらい、個人クエリを収集した。

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S15	S16	S18	S19	S21
S1	10	-1	2															2
S2	-1	10	1															2
S3	2	1	10															2
S4	4	2	2															3
S5	1	1	-1															1
S6	2	0	0															3
S7	0	1	-2															-1
S8	2	-1	0															2
S9	1	1	0															2
S10	0	2	4															2
S11	0	-1	-1															2
S12	4	-1	-1															2
S13	1	0	3															1
S15	1	2	-1															0
S16	0	3	0															1
S18	2	-1	-2															1
S19	1	0	-2															1
S21	2	-2	2	3	1	3	-1	2	2	2	2	2	1	0	1	1	1	10



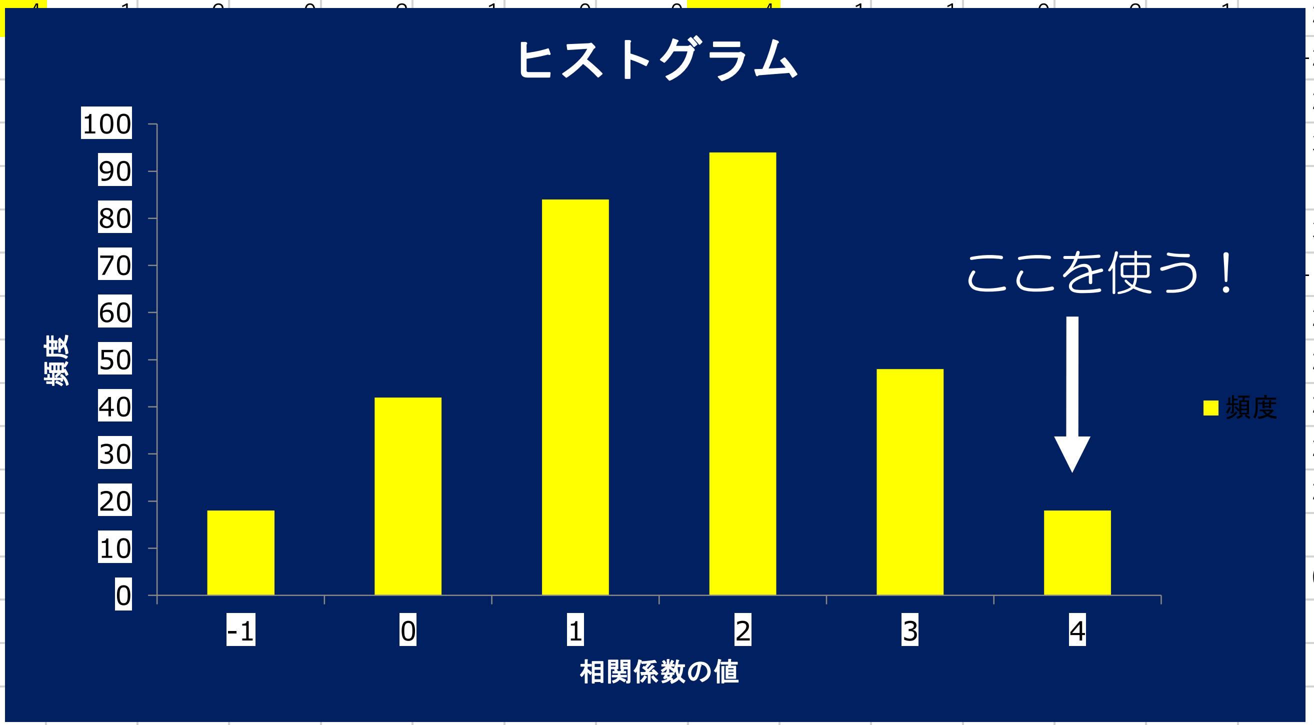
相関係数が最も大きい値（黄色）を利用することで、ユーザにとって便利な推薦になる。

※本来の数値を10倍し、整数で表示している

実験結果 - 1

- 金沢工業大学の学生並びに職員にアプリケーションを利用してもらい、個人クエリを収集した。

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S15	S16	S18	S19	S21
S1	10	-1	2															2
S2	-1	10	1															2
S3	2	1	10															2
S4	4	2	2															3
S5	1	1	-1															1
S6	2	0	0															3
S7	0	1	-2															-1
S8	2	-1	0															2
S9	1	1	0															2
S10	0	2	4															2
S11	0	-1	-1															2
S12	4	-1	-1															2
S13	1	0	3															1
S15	1	2	-1															0
S16	0	3	0															1
S18	2	-1	-2															1
S19	1	0	-2															1
S21	2	-2	2	3	1	3	-1	2	2	2	2	2	1	0	1	1	1	10



相関係数が最も大きい値（黄色）を利用することで、ユーザにとって便利な推薦になる。

※本来の数値を10倍し、整数で表示している

実験結果 - 2　ユーザαの推薦結果

他人のスコア利用前

○ ×		ID:60「明日に向って撃て！」
○ ×		ID:187「夕陽の群盗」
○ ×		ID:184「GiveandGo-ギブアンド ゴー-」
○ ×		ID:294「オリーブの林をぬけて」
○ ×		ID:30「ゴジラ キング・オブ・モンスターズ」
○ ×		ID:34「劇場版仮面ライダーエグゼイドトゥルーエンディング」
○ ×		ID:109「マーダーオブキャット」
○ ×		ID:191「バルトの楽園」
○ ×		ID:85「コラインとボタンの魔女」
○ ×		ID:24「ディープインセクト」
○ ×		ID:283「最低で最高のサリー」
○ ×		ID:173「ひろしま」
○ ×		ID:297「ファントム・スレッド」
○ ×		ID:293「ういらぶ」

ユーザαのスコアを見ると、ミステリー要素、少しのラブコメ要素を含むコンテンツに興味があることが分かった。

実験結果 - 3　推薦結果の変化

自分のスコアのみ

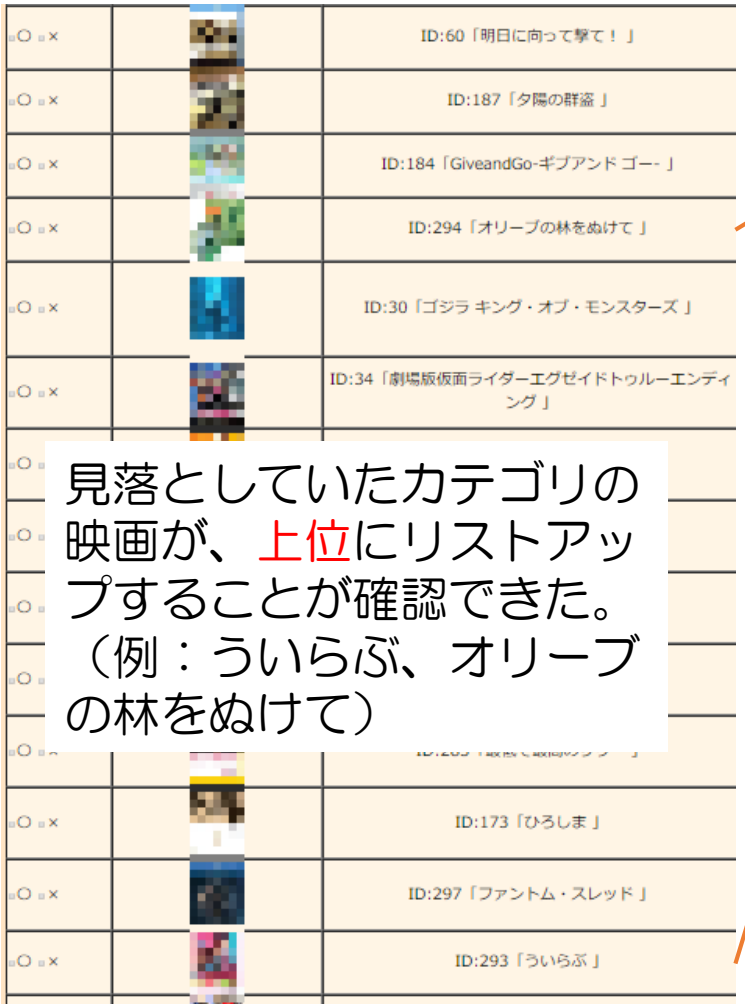
○ ×		ID:60「明日に向って撃て！」
○ ×		ID:187「夕陽の群盗」
○ ×		ID:184「GiveandGo-ギブアンドゴー-」
○ ×		ID:294「オリーブの林をぬけて」
○ ×		ID:30「ゴジラ キング・オブ・モンスターズ」
○ ×		ID:34「劇場版仮面ライダーエグゼイドトゥルーエンディング」
○ ×		ID:109「マーダーオブキャット」
○ ×		ID:191「バルトの楽園」
○ ×		ID:85「コララインとボタンの魔女」
○ ×		ID:24「ディープインセクト」
○ ×		ID:283「最低で最高のサリー」
○ ×		ID:173「ひろしま」
○ ×		ID:297「ファントム・スレッド」
○ ×		ID:293「ういらぶ」

他人のスコア利用後

○ ×		ID:286「セカンドバージン」
○ ×		ID:294「オリーブの林をぬけて」
○ ×		ID:89「きみと、波にのれたら」
○ ×		ID:293「ういらぶ」
○ ×		ID:96「ONE PIECE FILM GOLD」
○ ×		ID:287「おとなの恋はまわり道」
○ ×		ID:290「夏ノ日、君ノ声」
○ ×		ID:217「心が叫びたがってるんだ。(2017)」
○ ×		ID:53「烈車戦隊トッキュウジャーTHEMOVIEギャラクシーラインSOS」
○ ×		ID:87「君の膵臓を食べたい(2018)」
○ ×		ID:296「想影」
○ ×		ID:125「マスカレードホテル」
○ ×		ID:75「ディリリとパリの時間旅行」
○ ×		ID:78「夏目友人帳うつせみに結ぶ」

実験結果 - 3　推薦結果の変化

自分のスコアのみ

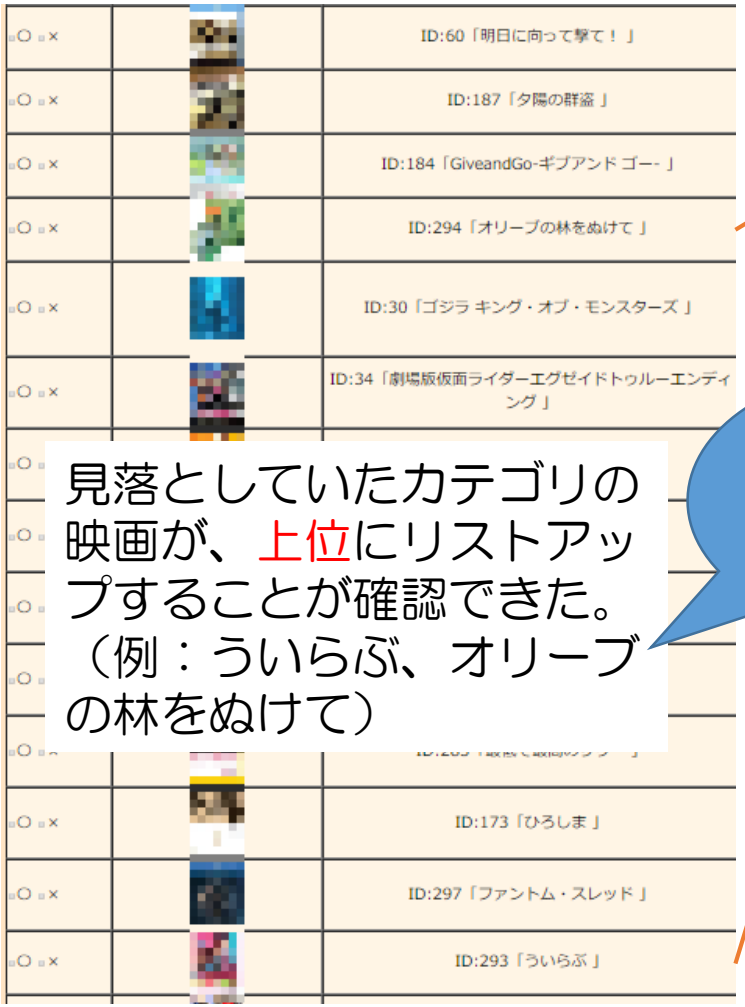


見落としていたカテゴリの映画が、上位にリストアップすることが確認できた。（例：ういらぶ、オリーブの林をぬけて）

他人のスコア利用後

実験結果 - 3　推薦結果の変化

自分のスコアのみ



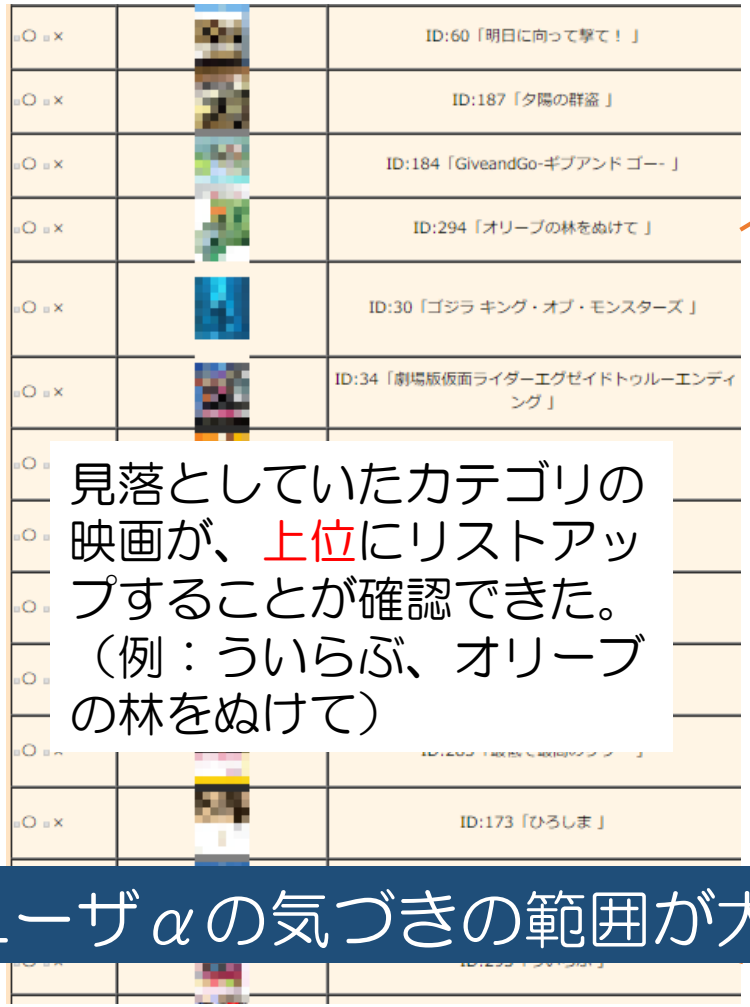
見落としていたカテゴリの映画が、上位にリストアップすることが確認できた。（例：ういらぶ、オリーブの林をぬけて）

いい感じ！

他人のスコア利用後

実験結果 - 3　推薦結果の変化

自分のスコアのみ



見落としていたカテゴリの映画が、上位にリストアップすることが確認できた。（例：ういらぶ、オリーブの林をぬけて）

他人のスコア利用後

ユーザαの気づきの範囲が大きくなる方向に可能性が見られた

まとめ

ユーザβのクエリをユーザαが利用することで、
ユーザαが見落としていた映画をすくい上げることができた。

↓

新しいコンテンツを目に入れることができた。
これは視野拡大につながるはず

↓

今後、視野拡大への影響を調べたい。
個人クエリの変化で計測できると考えている。

お願い

- 具体的な計算方法については次の芦川が説明しますので、よろしくお願いします。

Fin

ご清聴ありがとうございました