

情報探索と検索学期末レポート課題

201821636 村松直哉

2018 年 6 月 28 日

1 はじめに

Googleの検索エンジンには、Knowledge Panel(KP) [1]という機能がある。図1に、KPの表示例を示す。

これは様々な情報源から収集したセマンティック検索情報を用いて、検索結果を拡張する。 [3]より、スマートフォン画面での配置を工夫されたKPの表示は、ユーザ検索作業にとって有効であることがわかる。しかし、PC画面に表示された場合（図1）右端に表示される。従来の表示手法になれたユーザは、検索結果を上から見ていくと考えられる [2]。そのため、KPに十分に注意を引くことができない。

私は、KP情報を音声情報に変換し自動的に再生するGoogle++を提案する。対象クエリに関する基礎的な情報を含んだKPを、最初にユーザに提示する。これにより、ユーザは検索対象情報を得るまでの時間を減らすことを目的とする。

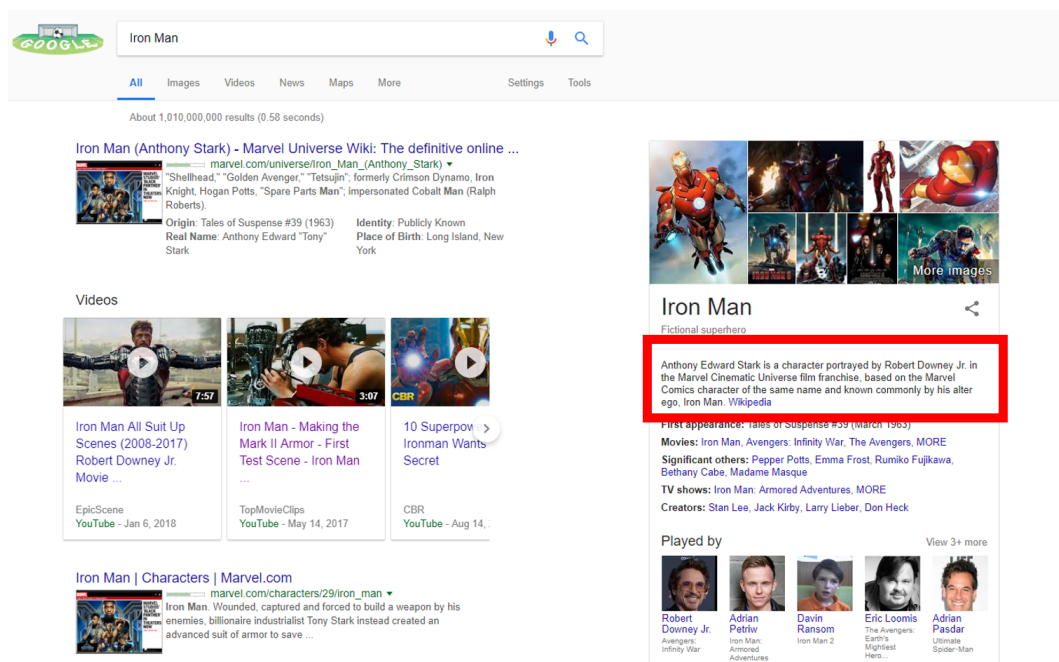


図1 PC画面でのGoogle Knowledge Panelの表示例

2 システム構成

ユーザは従来の検索エンジン同様に、作業関連する情報ニーズを表した検索質問（クエリ）をシステムに

入力する。クエリはまずGoogleの検索エンジンによって検索結果を得る。ここでKPが表示される場合、本システムはこれを解析し、text-to-speech(TTS)アルゴリズムにより音声情報を、検索結果とともにユーザに渡す。

これにより、ユーザが検索対象情報を得るまでの時間を減らすことを目的とする。言語対応については、英語を対象にしたシステムを想定とする。

音声読み上げについては、TTSによりテキスト情報から変換を行い、生成された音声を再生する。使用するアルゴリズムは、自然性の高い読み上げが可能なWaveNet [5]を用いる。これは高速動作のため、学習済みモデルをダウンロードして、JavaScriptによりブラウザ側で処理を行う。

生成された音声は、十分に短い時間内で読み上げられる必要がある。KP内の文章が長すぎる場合には、さらに情報を要約する。この手法については、2.2で詳しく述べる。

2.1 Knowledge Panel抽出手法

KPの有無やテキスト抜き出しには、通常のスクレイピング手法を用いる。

KPが検索結果に表示される際には、表示ページのHTMLに、クラス名kp-blkを含むdivタグが出現する。これにより、KPの存在を確認できる。また、KPに記述される短い説明文も同様に、クラス名kno-rdescを基に抽出できる。

2.2 文章要約手法

読み上げ速度は、生成された音声データにより依存する。これを計測したところ、約2.3 words/sであった。情報をユーザに提示する際に、検索結果を確認する時間は、約8秒である [2]。つまり、8秒以内に読み上げる事のできる単語数は、約18 wordsである。

英語の1文あたりの単語数は15-20 wordsである [4]。そのため、本システムで読み上げられる単語数は、基本的に20 words以下とする。

KPには、対象情報に関する短い説明文である、knowledge description(KD)が付く（図1における赤枠の部分）。本システムでは、KDの文章を利用して、音声情報を生成する。

アルゴリズム1は、KDの文章が長すぎる場合の要約手法を示している。基本的には、最初の文を利用する。それが長い文の場合、カンマを基準に文を分割する。分割した各テキストから、対象クエリを最も多く含むテキストを利用する。それでも長いテキストになる場合には、音声の読み上げスピード上げることで、8秒以内に情報提示を終えるようにする。

3 検証実験

本システムの有効性を検証するために、様々な被験者実験を設計し、実施する。そのとき、ベースラインとして既存のGoogleの検索エンジンを用いる。被験者実験は、KPを基にした短い音声情報提示により、ユーザの情報探索方法や要する時間に、どのように影響するかを調べる。音声情報は、KPを基にしているため有名な人物や場所などの、ある実在するもの（[louvre]や[angelina jolie]など）に関連するクエリ情報を提供する。音声情報により素早く情報を概要をつかめるため、実験でのタスクでは、一般的に知られていないクエリを対象とした。

検証実験には、以下のプロトコルで行った。参加者には、20の検索タスクのリストが載ったウェブページを提示する。リスト内の各エントリは、タスクの説明、2つのハイパーリンク—1つは検索結果ページ（タスクに関連する事前定義されたクエリを指す）、2つ目はタスク後のアンケートページである。参加者は、タスクの説

Algorithm 1 文章要約アルゴリズム

```
D ← the knowledge description
d ← Dの最初の文
q ← stemming処理したクエリ
speechspeed ← 1.0
th ← 20
if dの単語数 ≤ th then
  return d
else
  c ← dをカンマで分割した最初の要素
  d ← cでqを最も多く含む要素
  if dの単語数 > th then
    speechspeed ← d/th
    return d
  end if
end if
```

明を読んで、検索エンジンを利用してタスクの答えを見つけて、タスク後のアンケートをしてもらった。

各タスクの難易度が同じであることを確認するために、著者は各タスクについて、対応する1つ検索結果ページに答えが含まれていることを確認する。したがって、参加者に行ってもらうタスクは非常に簡単で（1分未満で済む）、参加者は1タスクにつき3分を超えないように指示する。答えを見つけたら、参加者はブラウザの「戻る」ボタンを使って実験トップページに戻り、2番目のハイパーリンクをたどってアンケートを完了してもらう。アンケートのページでは、参加者に7段階のリッカート尺度で検索結果全体の満足度を評価するよう依頼する（1は不満であり、7件は満足である）。クエリはタスクごとにあらかじめ定義されており、クエリの再定義は許可しない。

本実験では、2つの要素（ユーザーの情報ニーズに対する音声情報の妥当性、および音声情報の存在）を含む2 × 2のwithin-subjectを使用する。両方の要素には2つのレベルがある。つまり、関連性（関連性または無関係性）と存在（表示または非表示）である。各参加者は20の検索タスクを実施する（条件ごとに5つのタスク）。タスク提示の順序は、学習効果と比較印象の影響を排除するためにランダムに行う。参加者は、使用するPCと実験フローに慣れるために、本実験を開始する前に4つの練習課題を行ってもらう。最初の実験で10のタスクを完了した後、参加者は5分間の休憩をとってから、残りの10のタスクを行ってもらう。

表1 被験者実験で使用する質問タスクの例

Query	Task Description
Suspension bridge	In what state is the bridge?
Tony Stark	Who made him?
university of cambridge	What was the enrollment of the University of Cambridge in 2012?

3.1 被験者

インフォームド・コンセントを得た18–65歳の様々な職種の30人（男性15人、女性15人）を対象として、被

験者を募集する。被験者らは、それぞれいくつかの検索エンジン使用経験があり、利用サービスを回答してもらう。被験者らは、正常または矯正視力（メガネやコンタクトレンズ）を有し、PC画面を操作できることを確認する。

3.2 装置

実験には、Dell ディスプレイモニター SE2717H（27インチ）を使用し、Chromeブラウザをフルスクリーン表示の状態で行う。

3.3 操作の記録方法

ある時点でスクリーンに表示された正確な情報を記録するために、ビューポートロギングを使用する。これにより、画面に現在表示されているWebページの部分だけでなく、ページに表示されているすべての検索結果のバウンディングボックスを記録することができる。ビューポートロギングはJavaScriptによって設置され、ユーザーに表示されるすべての検索表示結果に挿入される。このスクリプトは、ページがブラウザでレンダリングされた直後に検索結果の境界ボックスを記録し、スクロールやズームなどのログビューポート変更イベントを記録する。すべてのビューポートイベントはバッファされ、その後HTTPリクエストとともにユーザー調査サーバーに送信され、後で分析するために保存する。このログにより、ユーザが画面上で見たものをいつでも再構成することができる。

4 結果の分析

設計したwithin-subjectに従い、回答の正誤と、そのタスク終了時間をまとめる。正解のものだけを対象に統計分析を行う。このとき対象とするのは、「音声情報提示が回答の満足感に与える影響」、「音声情報提示の関連性の効果」である。以下のような仮説のもと検定を行う。

- H1 平均的に音声情報提示がある方が、回答の満足度が上がる
- H2 音声情報提示があるとき、関連性がある方が、回答の満足度が上がる
- H3 音声情報提示の関連性がないときは、ある場合に比べタスク完了時間が長くなる。

5 結論

本レポートでは、まさにGoogleの検索エンジンを拡張したGoogle++を提案した。これは音声により、情報をすばやく提供する事によって、人間の検索結果に対する分析速度の向上を狙ったものである。そのアルゴリズムは非常に単純であるため、高速に動作することが可能であると考えられる。実験計画については、音声による情報提供の有効性を確かめるために設計されており、特にタスク完了までの時間に注目したものになっている。

この提案は音声による情報提供の有効性を明らかにすることで、視覚と聴覚を組み合わせた新しい情報探索方法において、重要な結果を示すことができる。

参考文献

- [1] S. Amit. Official Google Blog: Introducing the Knowledge Graph: things, not strings, 2012.
- [2] L. A. Granka, T. Joachims, and G. Gay. Eye-tracking analysis of user behavior in WWW search. In *Proceedings of the 27th annual international conference on Research and development in information retrieval - SIGIR '04*, page 478, New York, New York, USA, 2004. ACM Press.
- [3] D. Lagun, C.-H. Hsieh, D. Webster, and V. Navalpakkam. Towards Better Measurement of Attention and Satisfaction in Mobile Search.
- [4] C. Mesa and W. H. DuBay. The Principles of Readability. page 24.
- [5] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu. WaveNet: A Generative Model for Raw Audio. sep 2016.